INFORMATION SOCIETY TECHNOLOGIES
(IST)
PROGRAMME

Contract for: **Shared-cost RTD**

Annex 1 - "Description of Work"

Project acronym:	CROSSGRID
Project full title:	Development of Grid Environment for Interactive Applications
Proposal/Contract no.:	IST-2001-32243
Action line:	IST-2001-5.1.9
Type of action:	RTD Project

Date of preparation of Annex 1: 15 November, 2001

Operative commencement date of contract: 1 January, 2002

Document identifier: CROSSGRIDANNEX1_V3.1.DOC

1 PROJECT SUMMARY

1.1 Abstract

The Cross Grid project will develop, implement and exploit new Grid components for interactive compute and data intensive applications like simulation and visualisation for surgical procedures, flooding crisis team decision support systems, distributed data analysis in high-energy physics, air pollution combined with weather forecasting. The elaborated methodology, generic application architecture, programming environment, and new Grid services will be validated and tested thoroughly on the CrossGrid testbed, with an emphasis on a user friendly environment. The work will be done in close collaboration with the Grid Forum and the DataGrid project to profit from their results and experience, and to obtain full interoperability. This will result in the further extension of the Grid across eleven European countries.

2 OBJECTIVES

2.1 Introduction

The primary objective of this Project is to further extend the Grid environment to a new category of applications of great practical importance, and into 11 new European countries.

The applications we are interested in are characterised by the interaction with a person in a processing loop. They require a response from the computer system to an action by the person in different time scales; from real through intermediate to long time, and they are simultaneously compute- as well as data-intensive. Examples of these applications are: interactive simulation and visualisation for surgical procedures, flooding crisis team decision support systems, distributed data analysis in high-energy physics, air pollution combined with weather forecasting. A visualisation engine should be developed and optimised for these applications.

To enable efficient development of this category of applications for the Grid environment, new tools for verification of parallel source code, performance prediction, performance evaluation and monitoring are needed. This, in turn, requires extension of the Grid by new components for application-performance monitoring, efficient distributed data access, and specific resource management. Users should be able to run their applications on the Grid in an easy and transparent way, without needing to know details of the Grid structure and operation. CrossGrid will develop user-friendly portals and mobile personalised environments and will integrate new components into the Grid and application development tools.

The elaborated methodology, generic application architecture, programming environment, and new Grid services will be validated and tested thoroughly on the CrossGrid testbeds. This will result in further extension of the Grid across Europe. In our CrossGrid development we will exploit all the available achievements of DataGrid, EuroGrid and other related projects in a way which enables their interoperability. CrossGrid will closely collaborate with DataGrid.

2.2 CrossGrid Application Development

The main objective of this workpackage is to elaborate the methodology for development of, and generic architecture of, large-scale grid-enabled applications for simulation and visualisation that require real time response. The challenges are the distribution of source data, simulation and visualisation tasks, virtual time management, simulation/visualisation rollback due to user actions, and platform independent VR visualisation. This workpackage will provide a representative collection of sample applications from various fields that can exploit the specific interactive functionalities to be developed in the CrossGrid project. Their need for specific services and their performance characteristics will provide a driving force for the technology-oriented workpackages. They will serve as benchmarks for CrossGrid performance, and crucially, they will be demonstrators for their respective fields of the added value provided by the Grid in general, and the technology developed by the CrossGrid project in particular.

As the first application, we will develop a **Grid-based prototype system for pre-treatment planning in vascular interventional and surgical procedures** through real-time interactive simulation of vascular structure and flow. The system consists of a distributed real-time simulation environment, in which a user interacts in Virtual Reality (VR). A 3D model of the arteries, derived using medical imaging techniques, will serve as input to a real-time simulation environment for blood flow calculations. The user will be allowed to change the structure of the arteries, thus mimicking a surgical procedure and the effects of this adaptation will be analysed in real time while the results are presented to the user in a virtual environment.

The second application will be a **Grid-based Support System for flood prevention and protection**. The kernel of this system is numerical flood modelling that requires an appropriate physical model and robust numerical schemes for a good representation of reality, as well as grid distributed supercomputing aspects for realistic simulations when dealing with large problem sizes.

As the third application we will develop **final user applications for physics analysis** running in a distributed mode in a Grid-aware environment using large distributed databases for High-Energy Physics (HEP). The challenging points are: seamless access to large distributed databases in the Grid environment, development of distributed data-mining techniques suited to the HEP field, and integration in a user-friendly interactive way, including specific portal tools. As indicated in the CERN LHC Computing Review Report, interactive data analysis tools is one of the areas where joint efforts amongst one or more experiments, resulting in common projects and products, might lead to cost savings, or decreased risk, or both. The proposed applications explore and develop the use of advanced simulation and interactive data mining techniques.

The fourth application is **weather forecast and air pollution modelling**. An important component will be a data mining system for the analysis of the archive of operational data from a mesoscale model and meteorological reanalysis DBs which include homogeneous meteorological information from a single numerical weather prediction model integrated over decades. The system will also include an atmospheric pollution chemistry module.

All these applications will heavily rely on the performance tools, resource and network services, and management tools developed in the other CrossGrid workpackages. Their deployment on the CrossGrid testbed will test these applications in the final user environment and provide feedback to application developers. In this way they can fully exploit the possibilities of the Grid. The generic software developed in this workpackage will support Grid-distributed interactive simulation and visualisation.

2.3 Grid Application Programming Environment

For end users and program developers, one of the critical aspects of the Grid is its complexity. The development, debugging and tuning of parallel and distributed applications on the Grid is difficult and tedious due to the permanently changing and heterogeneous nature of the Grid infrastructure, which incorporates numerous different hardware and software components. The aim of this workpackage is to specify, develop, integrate and test tools that facilitate the development and tuning of parallel distributed, compute- and data-intensive, interactive applications on the Grid.

Verifying that user applications comply with the MPI standard will reduce the need for debugging sessions on the inhomogeneous Grid environment. In order to make parallel applications portable and reproducible, this workpackage will develop a debugging and verification tool for MPI programs. Efficiently using the Grid as an environment for large-scale interactive applications requires the end-users to monitor and analyse the performance of their jobs. For this purpose tools will be developed that automatically extract high-level performance properties of Grid applications. Also, on-line performance monitoring tools will be developed to analyse performance data for detecting bottlenecks of applications. Benchmarks will be developed that are sensitive to data transfer, synchronisation and I/O delay, and CPU utilisation. Taken together with the monitoring tools, they will allow analysis of the system at any level of granularity, from the Grid level down to the process level.

The tools developed in this workpackage will be integrated into the testbed and will be promoted by and tested with the real end-user applications of WP1.

2.4 New Grid Services and Tools

As a lot of Grid services and tools are already available and ready to be used, we will only focus on those which are necessary for development of interactive compute- and data-intensive applications, and which have not been addressed in other Grid projects, as well as on services necessary for the end users.

At first this WP will address user-friendly Grid environments. Portal access to the Grid infrastructure and user applications, independent from the user location, is very important from the practical point of view. According to the questionnaire on Grid user requirements analysed by the ENACTS project, more users would be interested in Grids as soon as such services become more easily accessible. Portal technologies, such as the Grid Portal Development Toolkit (GPDK), and iPlanet, have already started to address this issue, however, no practical implementations exist so far. One of the portal features that we propose to implement is roaming access, i.e. the mobile personalised environment. Secondly, we will address the construction of new resource management techniques based on self-adaptive scheduling agents for scheduling a particular parallel application submitted to the Grid. The goal is to achieve a reasonable trade-off between resource usage efficiency and application speedup, according to the user's optimisation preferences. As the third objective we propose to develop a prototype infrastructure for the needs of monitoring-related activities for automatic extraction of high-level performance properties and for tool support for performance analysis (WP2). A Grid monitoring service to support management decisions will be implemented with Jiro - a Java language based implementation of the Federated Management Architecture. Finally, we are going to significantly extend the data management system that is under development in the DataGrid Project. In order to help users to interact with data stores, a kind of expert system is proposed that will operate as an advisor for selection of migration/replication policies in user-defined circumstances. This heuristic tool will be supported by a subsystem that will estimate the time needed to make the file available. We also plan to design and implement middleware that is independent of the mass storage system for faster access (lower start-up latency time) to large tape resident files. This will be done by using a sub-filing strategy.

Deployment of these new Grid tools and services on the CrossGrid testbed will allow for testing them in the final user environment, and so provide feedback to the tool developers.

2.5 International Testbed Organisation

The role of the international testbed is to assure the integration of all the applications, programming tools, and new Grid services developed in this Project. Organisational issues, performance and security aspects, including network support, can only be evaluated with a testbed relying on a high-performance network (which will be provided by the Géant project), with participation of an adequate number of computing and data resources distributed around Europe. The main objectives of the testbed organisation are to:

- provide distributed computer resources where the Project developments can be tested in a Grid environment,
- integrate national Grids,
- integrate the software developed within the Project,
- co-ordinate the software releases,
- assure interoperability with other Grids, in particular with DataGrid,
- support the construction of Grid testbed sites across Europe.

Testbed sites will be placed in 10 different institutions distributed across 9 different European countries, significantly expanding the Grid community.

2.6 CrossGrid Architecture

Being aware of other Grid projects, like DataGrid or EuroGrid, we need to assure compatibility with those projects. As previously stated, a lot of Grid services and tools already exist, ready to be used, so we will focus only on those that are indispensable for our applications.

The technical co-ordination will focus on:

- definition of the overall CrossGrid architecture,
- establishment of rules for component transfer from other Grids and integration into CrossGrid,
- selection of appropriate software development methodologies,
- definition of measures necessary for interoperability with other Grids.

The above-mentioned actions will be elaborated by the Architecture Team (Task 5.2).

2.7 Dissemination and Exploitation

The innovative character of the Project requires a systematic distribution of information and acquired knowledge through a variety of routes to potential users or beneficiaries. Planned dissemination methods include written information, electronic media and person-to-person contact. Potential users and customer groups outside the CrossGrid consortium will be identified and their needs and requirements should be considered in the development of the Project. In this way a catalogue of intensive computing applications and institutions potentially involved in Grid exploitation will be created. Potential users will get assistance, information and ideas from the CrossGrid Consortium through the dissemination task leader. We have already started to establish close contacts with other European and international Grid projects.

3 PARTICIPANT LIST

Partic. Role*	Partic. no.	Participant name	Participant short name	Country	Date enter project**	Date exit project**
CO	1	Academic Computer Centre CYFRONET AGH	CYFRONET	Poland	Start	End
AC	2	Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw	ICM	Poland	Start	End
AC	3	The Henryk Niewodniczanski Institute of Nuclear Physics, High Energy Physics Department	INP	Poland	Start	End
AC	4	The Andrzej Soltan Institute for Nuclear Studies, Laboratory for High Energy Physics	INS	Poland	Start	End
CR	5	Universiteit van Amsterdam, Faculty of Science	UvA	Netherlands	Start	End
AC	6	Ustav Informatiky, Slovenska Akademia Vied, Departement of Parallel and Distributed Processing	II SAS	Slovakia	Start	End
AC	7	Institut für Technische Informatik und Telematik, Johannes Kepler Universität Linz, Abteilung für Graphische und Parallele Datenverarbeitung	Univ. Linz	Austria	Start	End
CR	8	Forschungszentrum Karlsruhe GmbH. Central Information and Communication Technologies Department	FZK	Germany	Start	End
AC	9	Universität Stuttgart, Rechenzentrum Universität Stuttgart	USTUTT	Germany	Start	End
AC	10	Technische Universität München, Lehrstuhl für Rechnertechnik und Organisation/Parallelrechner Architektur, Fakultaet für Informatik	TUM	Germany	Start	End
CR	11	Poznan Supercomputing and Networking Center Affiliated to the Institute of Bioorganic Chemistry of PAN	PSNC	Poland	Start	End
AC	12	University of Cyprus, Department of Computer Science, University of Cyprus	UCY	Cyprus	Start	End
AC	13	DATAMAT S.p.A.	DATAMAT	Italy	Start	End
AC	14	Department of Computer Science, Trinity College Dublin	TCD	Ireland	Start	End
CR	15	Consejo Superior de Investigaciones Cientificas	CSIC	Spain	Start	End

AC	16	Universitat Autonoma de Barcelona	UAB	Spain	Start	End
AC	17	Universidade de Santiago de Compostela	U.S.C.	Spain	Start	End
AC	18	National Centre for Scientific Research 'Demokritos', Institute of Nuclear Physics	Demo	Greece	Start	End
AC	19	Aristole University of Thessaloniki, Division of Nuclear and Particle Physics, Department of Physics	A.U.Th.	Greece	Start	End
AC	20	Laboratorio de Instrumentacao e Fisica Experimental de Particulas LIP Computer Centre	LIP	Portugal	Start	End
AC	21	Algosystems S.A., Applied Research Department	ALGO	Greece	Start	End

*CO – Coordinator

AC – Assistant Contractor

CR – Contractor

** Normally insert “Start of project” and “End of project”. These columns are needed for possible later contract revisions caused by joining/leaving participants

4 CONTRIBUTION TO PROGRAMME/KEY ACTION OBJECTIVES

The Information Society Technologies (IST) Programme places the needs of the user at the centre of future development. The "2001 Workprogramme" document includes this vision statement summarising the programme orientations:

"Start creating the ambient intelligence landscape for seamless delivery of services and applications in Europe relying also upon testbeds and open source software, develop user-friendliness, and develop and converge the networking infrastructure in Europe to world class".

The priorities for WP2001, focused on the challenges of realising this vision, include:

"Support the development of large scale demonstrations and trials for the adoption and development of IST products and services that involve citizens and businesses of all sizes across Europe".

In this context, the programme supports Cross-Programme actions that focus on a limited number of specific themes relevant to the entire IST programme, and the CrossGrid project clearly addresses the objectives of *Cross Programme Action V.1.9 CPA9: Grid testbeds, deployment and technologies: "To foster and encourage community-wide (research and industry) development, deployment, experimentation and integration of the Grid".*

The Project is also focused on the issues proposed in this Cross-Programme action:

- The CrossGrid testbed will be conducted in the context of full-scale applications, compute- and data-intensive, from scientific and industrial fields: health, physics and environment management. The deployment on many testbed sites distributed across Europe will expand the Grid community and provide a realistic framework to assess organisational issues.
- The distributed and interactive nature of these applications motivates the use of a Grid-specific software architecture. Their development needs enabling technologies and new services that will be implemented and integrated in the form of middleware and toolkits, including verification of parallelisation, resource scheduling, performance prediction and monitoring. User access to the CrossGrid resources via specific portals will ensure their user-friendliness, and ubiquitous computing and access to information will be improved with the use of roaming techniques.
- Interoperability with other Grid testbeds is a key point in the CrossGrid project. Complementarity will be considered from the architecture definition to the testbed set-up. In particular, co-ordination with the ongoing DataGrid project will be promoted. Moreover, as Cross-programme Action Lines should seek to work closely with the most relevant projects in the Key Actions, grouping into cross programme clusters could be expected. This will also result in a common infrastructure.

The Project will make available the developed middleware to the Grid community, and promote corresponding standards in international fora. It will establish membership in the Global Grid Forum. The consortium keeps active contacts and constructive relations with major US centres involved in Grid initiatives, sharing also a very active user community.

Participation of industrial partners will promote the incorporation of the Grid as a common instrument in industry and business, assuring a wider applicability of Project results. Explicit plans for exploitation and dissemination are being handled by a SME participant in the consortium with previous experience, who will also participate to provide user friendliness for the applications. The other large industrial partners will collaborate in key parts of the Project: performance monitoring and resource scheduling.

In summary, the Project addresses all the points made explicit in CPA9, matches the priority criteria for the Workprogramme 2001, and will contribute to the programme objectives by providing an answer to the interactive compute- and data-intensive needs of a wide user's community, promoting the use of the Grid as an instrument for European competitiveness in research, industry and business.

5 INNOVATION

5.1 Introduction

The essential novelty of this Project consists in extension of the Grid to a completely new and societally important category of applications. The main features of these are the presence of a person in a processing loop, and a requirement for real-time response from the computer system. Moreover, these applications are at the same time compute- and data-intensive. Generic applications were selected: interactive simulation and visualisation for surgical procedures, flooding crisis team decision support systems, distributed data analysis in high-energy physics, and air pollution combined with weather forecasting. A visualisation engine will be developed and optimised for these applications. The components of the CrossGrid Project and their relation to other Grid projects are presented in Fig.1.

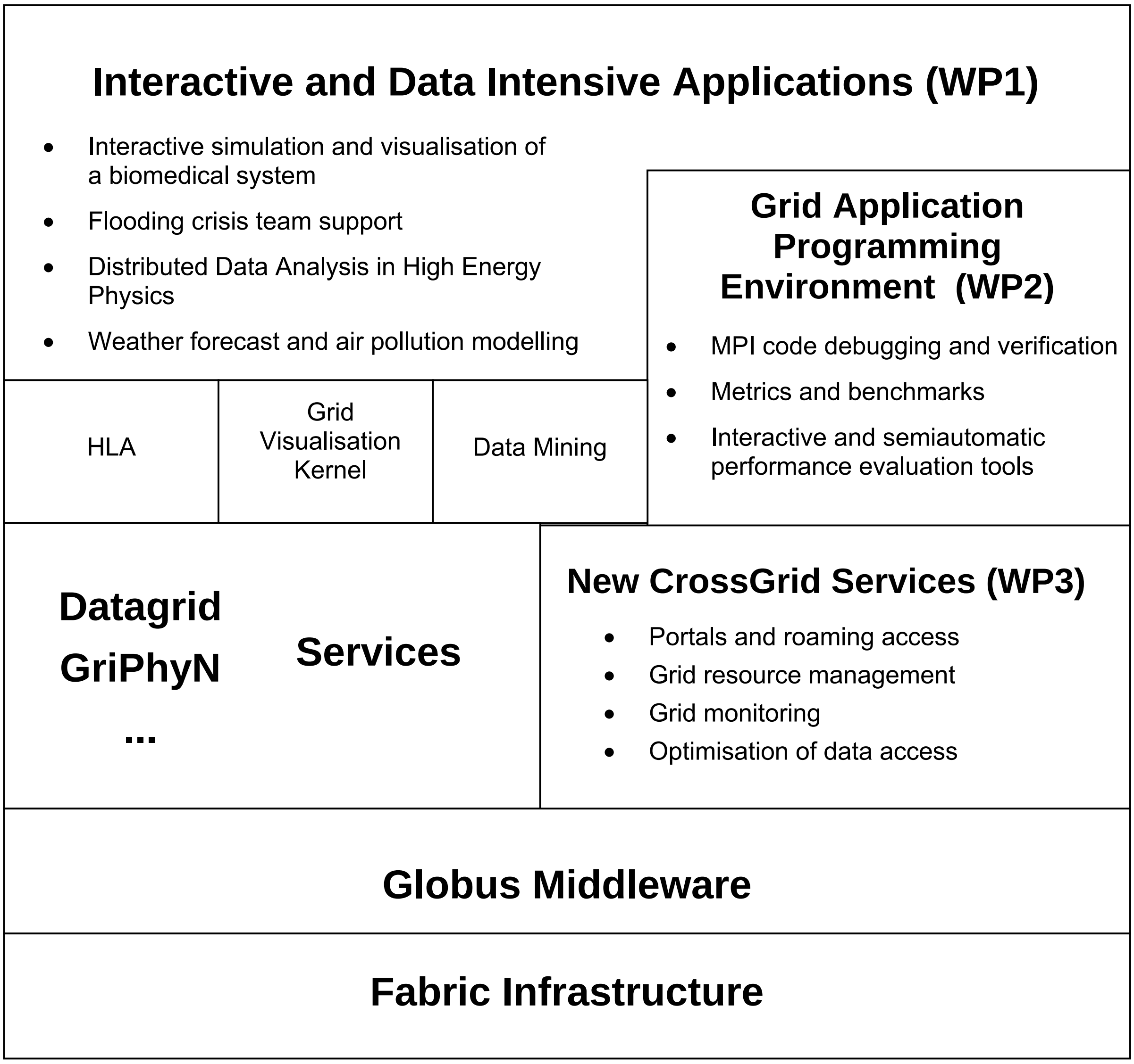


Fig. 1. Basic components of the CrossGrid Project.

The new methodology, generic application architecture, programming environment, and new Grid services will be validated and tested thoroughly on the CrossGrid testbeds. This will result in further extension of the Grid across Europe into 11 new countries.

5.2 Enabling New Applications for Grid

We will develop enabling technologies for visualisation, interaction and performance optimisation for Grid-based applications that will allow such applications to be applied to a variety of areas. These application areas often address problems of international or supra-national import. The developed technology will be tested and its usefulness will be verified in WP1 by a diverse collection of prototype applications. These include a bio-medical application, a decision support application for flood control, a large physics application, and an application addressing weather forecasting and air pollution.

In comparison to the applications in other Grid projects (EuroGrid, DataGrid and GriPhyN), the applications in the CrossGrid project emphasise the use of interaction and visualisation, technologies that will be vital to a broad acceptance of Grid-based applications. Together with data-mining, the third recurring theme in our applications, these technologies put very stringent demands on the performance of the Grid. The development and refining of technology that supports such highly demanding applications is a core issue in the project. In this manner a coherent package of software supporting a wide range of important applications will be realised.

All applications in WP1 have a relevance that transcends national boundaries. The actual research is always embedded into large international collaborations, even if only a single collaboration partner is responsible for the task within this Project. All applications represent a wider range of similar applications with similar demands. The biomedical application integrates a number of new and existing technologies (blood-flow simulation, medical imaging, interactive simulation and virtual reality) into a system that allows surgeons to better predict the outcome of a common surgical procedure. The flood control decision support system explores the use of Grid technology to access widely distributed data and computing resources in decision support for international crisis management. The high-energy physics application explores data mining on huge distributed databases. The relevance of these techniques is further explored in the very different application area of weather forecasting and air pollution modelling. The development of advanced visualisation techniques in a Grid environment will be vital to the acceptance of the Grid as an environment conducive to a wide range of highly demanding applications.

Atherosclerosis is a widespread disease that particularly manifests itself in the developed countries and is of major concern in Europe. A prototype system for pre-treatment planning in vascular interventional and surgical procedures will therefore be highly relevant on a European scale. A system for flood prediction should result in improvements in both the long-term and the short-term protection against floods due to the high rain-fall and snow melts that frequently plague large areas of Europe and often affect several countries simultaneously. The air pollution and weather forecasting work will improve the understanding of these crucial problems. The visualization kernel will contribute to the usefulness of the Grid for a wide range of users.

5.3 New Tool Environment for Grid Application Development

The results of this workpackage will promote an efficient and easy use of the Grid as a programming environment and computing resource. A debugging and code verification tool for Grid-enabled MPI programs will be developed, to check that an application adheres to the MPI standard. This will provide a significant step towards code portability and reproducibility by facilitating the execution of applications on any platform on the Grid in a smooth and seamless way. Verifying that user applications comply with the MPI standard will reduce the need for debugging sessions in a heterogeneous Grid environment. Current approaches to parallel debugging breed either classical debuggers where the debugger is attached to multiple processes, or debug versions of the MPI library which catch some incorrect usage and internal errors in the library. Neither of the existing approaches addresses portability or reproducibility, two of the major problems when programming MPI. The use of automatic techniques that do not need user intervention would also allow one to debug programs executing on hundreds or thousands of processors.

Using the Grid as an environment for large-scale interactive distributed applications requires performance monitoring and analysis. In the framework of this Project, tools will be designed and developed for automatic extraction of high-level performance properties of Grid applications. The performance information will be made available by visualisation tools. Since the analysis process already takes into account the structural and architectural details of the Grid, information on performance properties and/or bottlenecks in applications can be provided without the need for the user to know the infrastructure architecture. This global, highly pre-processed performance information will significantly simplify program optimisation and tuning, making the work of programmers and scientists more efficient, and thus eliminating possible barriers to work on the Grid.

Based on performance data of the Grid measured by benchmarks, sensitive to data transfer, synchronisation and I/O delay, and CPU utilisation, models will be developed to predict the performance of Grid applications without really executing them. Besides providing the user with a simple tool for suitable planning of resources, the results of this work are seen as a first step to future accounting and billing systems on the Grid. Furthermore, on-line performance monitoring tools will be developed to analyse performance data during application execution.

5.4 Extension of Grid Services

The CrossGrid portal will securely authenticate users to remote resources and help them make better decisions for scheduling jobs by allowing them to view pertinent resource information obtained and stored on a remote database. In addition, profiles created and stored for portal users will allow them to monitor submitted jobs and view results.

One of the new portal features we propose to implement is roaming access, i.e. the mobile personalised environment. CrossGrid will pursue the design and implementation of architectures and interfaces that will allow mobile users to access the Grid from different locations and different platforms (including mobile terminals). Roaming access to Grid applications and an infrastructure with personalised user environments independent of the location of this access is a real innovation in Grid environments. It will enable users to access their data and applications with the same environment settings everywhere.

The management system will support inter-organisational and international Grids. Organisations that contribute to a Grid system usually differ in management schemes and policies. A unified management framework will harness and synchronise the resources and employ them for common goals. At the same time, the hierarchical structure of management data will be flexible enough to allow individual institutions to restrict the use of their resources by other parties and control the outgoing management information.

Another new contribution to Grid technology will be a monitoring system oriented towards large distributed applications-performance analysis. It will be built in compliance with standard tool/monitor interface requirements, which imply de-coupling the monitoring services and tool functionality to enable simplified porting of tools to different platforms.

A further contribution will significantly extend the data management system under development within the DataGrid Project. In order to help users interact with a data store, a kind of expert system is proposed. This will improve the performance of existing MSMS's used in various applications fields.

5.5 Spreading the Grid

The development of CrossGrid will be fully co-ordinated with DataGrid, and it will aim to be a catalyst for the unification of several Grid initiatives, both in Europe and USA, that could be integrated into a single world-wide Grid. This would result in a much more effective use of resources, both in manpower requirements and common infrastructure use.

The European Grid community will benefit from this Grid size increase, which will have a positive impact on the promotion of new Grid technology standards and deployment of common middleware. Moreover, the extension to several European countries will promote national interest in, and support for, Grid initiatives. As this testbed will support a wide range of user applications, this could help enormously in establishing the necessary critical mass for the extension of the Grid to other user communities, thereby establishing the Grid as a standard computational framework.

The novelty of this Project is in the organisation of a testbed that will be able to support the use of distributed interactive applications. The new requirements in terms of quality of the whole testbed will be evaluated.

As indicated above, one of the main innovative aspects in this Project is the strong intention to provide interoperability with other Grid initiatives. This effort will be co-ordinated by the Architecture Team of the Project.

Another innovative issue is the broad range of final user applications to be supported. This will test the flexibility of the testbed, and will also require solutions to practical organisational issues.

6 COMMUNITY ADDED VALUE AND CONTRIBUTION TO EU POLICIES

6.1 Background

The development of information technology brings new opportunities to the society. The progress on networking allows for integration of computational resources on a large scale, which opens new possibilities for challenging applications. The recent Global Grid Forum and DataGrid conferences presented some of these new possibilities and addressed technological and organizational problems. The accompanying presentations of the EU coordinators called for a concentrated effort towards the future e-Europe, including development and implementation of the Grid technologies and new applications. The CrossGrid project responds to such calls.

6.2 European level

Many areas of education and communication, pure and applied sciences, society and environment, industry and commerce, cross the national boundaries and require a broad international cooperation. There are a number of problems that cannot be solved on a local, national level, since they require larger infrastructures and a top-level expertise. The CrossGrid project concentrates on a few selected problems from the field of environment, medicine and physics, which require access to widely distributed data, high performance computing, and which often have to be solved in a timely fashion (in *real-time*). To be able to attack these problems a proper programming environment has to be created and additional software tools need to be developed. Any individual problem requires the integration of several scientific centres of one or more countries (for biomedical and/or environmental applications), and the integration of many European centres or even trans-continental integration, in the extreme case of multi-national collaborations for high-energy physics. The CrossGrid project will integrate the computing infrastructures of all the members of the project, and this will be demonstrated with the above applications.

The main thrust of the CrossGrid project lies with the applications. However the integration effort is also very important. There is a need to exceed a critical mass of transnational grid traffic and activity. Hence clusters of computers in Cyprus, Germany, Greece, Poland, Ireland, Portugal, the Netherlands, Slovakia, and Spain will be connected together as a testbed, running common middleware, and used for the evaluation of the CrossGrid applications. Close attention will be paid to interoperability with other Grid initiatives to assure that this infrastructure will be available to the widest possible user community.

The CrossGrid project involves two industrial European partners. We believe that their presence demonstrate the European commercial interest in the prepared programme of research and development. But it will also help us to prepare the end-products according to the industrial standards, and matching the expectations of potential customers. We propose a special task, oriented towards the active promotion of the project results – it should help in the transfer of the project results to a broader community.

The CrossGrid project integrates 21 partners from 11 European countries with an appropriate mix of expertise. At the time of unification of Europe it is important to stress that the project is going to stimulate the development of local and national infrastructures towards common European solutions. We are convinced that the close cooperation of all these institutions and individuals from so many countries, over several years, will foster a common technical and human culture.

6.3 International dimensions

Close cooperation is planned between the CrossGrid project and the other European and international Grid projects (EuroGrid, DataGrid, GryPhyN, PPDG). A special *architecture integration team*, has been defined within the CrossGrid project, with its goal to track other developments and to assure that integration with the DataGrid infrastructure can be realised across the whole Europe. Already the connections to the DataGrid project are very good, as several members of the Consortium are involved in the work at CERN, and in addition there are personal contacts at the managerial levels. The Consortium has also a representative in the Global Grid Forum steering group. Still, in the long-term a *cluster* project bridging DataGrid and CrossGrid would be desirable, as it would lead to full compatibility, unification and standarisation.

The European Grid initiatives are already recognised as strong partners for worldwide cooperation, precipitating the creation of the Global Grid Forum. Another European project in this area will enhance the role of Europe in this programme, and the GridStart project will result in close cooperation between EU Grid initiatives.

7 CONTRIBUTION TO COMMUNITY SOCIAL OBJECTIVES

We believe that the development and spreading of the Grid technologies across the borders is essential for the future society. Access to the Grid, as with access to the Web, will help in the equalisation and the unification of communities. It will open new opportunities for education and learning, for industry, tourism and commerce. Furthermore it has a very important political dimension, as we head towards a unified Europe.

7.1 End-users

There are many large-scale problems which require new approaches to computing, such as earth observation, environmental management, biomedicine, industrial and scientific modelling. The CrossGrid project addresses realistic problems in medicine, environmental protection, flood prediction, and physics analysis. These applications are oriented towards specific end-users:

- medical doctors, who could obtain new tools to help them to obtain correct diagnoses and to guide them during operations,
- industries, which could be advised on the best timing for some critical operations involving risk of pollution,
- flood crisis teams, which could predict the risk of a flood on the basis of historical records and actual hydrological and meteorological data,
- physicists, who could optimise the analysis of massive volumes of data distributed across countries and continents.

Each of these efforts will contribute to improving the quality of life and environment.

7.2 User-friendly environment

These new developments and applications could be complex and difficult to use, even by experienced users. This problem is recognised, and the CrossGrid project plans to develop several tools which will make the Grid more friendly for average users. *Portals* for specific applications will be designed, which should allow for easy connection to the Grid, create a customised work environment, and provide users with all necessary information to run their jobs.

8 ECONOMIC DEVELOPMENT AND S&T PROSPECTS

8.1 Introduction

Objectives and goals

The information community has discovered the potential of networking and now plans are afoot to create a new global computing infrastructure, the Grid. The members of the CrossGrid Consortium want to join in this effort and plan to extend the Grid to a number of new locations, as well as merge it with the evolving global Grid. However, as its prime goal, the CrossGrid project would like to demonstrate the use of the Grid for a new kind of applications – interactive applications. These are applications in which, apart from intense computations and large data volumes, the results are needed in "*real-time*". Before such applications can be installed on the Grid, new programming tools and services will need to be developed.

The CrossGrid Consortium consists of European academic centers, research institutes and industry. Those involved include top experts and specialists in computing and other disciplines, as well as an exciting mix of university teachers, young researchers, PhD students and industry staff. We believe that such a mixture of knowledge and cultures, combined with the vitality of the younger generations, will foster development of a new Grid environment, as well as the development and deployment of ambitious application tasks.

The objective of the CrossGrid is to extend the Grid environment to new categories of applications, such as visualization for surgical procedures and decision support services for flooding management and air pollution combined with weather forecasting. High-energy physics applications will also be addressed, in cooperation with the DataGrid project already in progress.

For these purposes, new tools, components and services will be implemented. Interoperability with previous Grid development will be addressed, in order to exploit development and software engineering approaches defined in previous relevant projects (mainly DataGrid). The methodology, architecture, programming environment and new Grid services will be tested in the context of the project test beds.

Developments planned within the CrossGrid project are of general interest to society, as they manifest global progress in the domain of e-sciences. They are also of specific interest to the communities of computer scientists and other fields concerned with applications proposed. One of the global objectives of the project is to provide these communities with sufficient information on general and specific developments. The specific objectives are:

- mantain contacts with the communities involved in the Grid developments to obtain the synergy of actions
- disseminate the information on ongoing technical developments to the community of computer science professionals
- work closely with potential end-users, to get information concerning their needs and expectations and provide them with information on ongoing developments
- introduce the new generation of students and researchers to the technology of distributed computing and provide them with an opportunity to train in this domain
- provide broad information concerning the potentials of distributed computing, including the demonstration of specific developments and applications.

The actions planned are summarised in the next few chapters; specific centralised activities are part of workpackage WP5 on Management.

Clustering with Grid projects

Several projects on Grid developments are already in progress. The exchange of information between different Grid projects is of prime importance when planning for a common Grid. It will be realised mainly in technical presentations at dedicated workshops and conferences.

However, to obtain the synergy of actions, there is also a need for “clustering” of the Grid projects. In such a context, the CrossGrid will collaborate closely with the DataGrid project. Some members of the CrossGrid are also partners of the EuroGrid, GridLab and GRIP projects. An accompanying measure project, GridStart, in which CrossGrid participates, has recently been submitted to the EU office.

International standardization bodies and dissemination entities in the European Union, as well as worldwide, provide a well-defined context for the CrossGrid development. The project will ensure links with all these schemes in order to exploit their resources and their established user base.

Clustering with national Grid initiatives will also be implemented as a bilateral rapport between CrossGrid and national R&D activities in the countries of the participants in the consortium.

Dissemination of information

Participation of the CrossGrid members in all major Grid events is projected. CrossGrid will disseminate the project results in technical and scientific presentations in the context of conferences, workshops and discussion fora.

Dissemination will be considered on the technical level and will be supported through the participation of the partners in international Grid fora as well as in National Grid initiatives. On the other hand, exploitation will be considered on the application level and will address specific user groups from the CrossGrid application fields (medical visualization, flood management, air pollution etc).

Annual CrossGrid workshops will be organized on a national level by a number of partners, including international participation. These workshops will be organized mainly in the facilities of the academic partners of the consortium and will be scheduled every year. The partners of the consortium will organize lectures and seminars on CrossGrid developments and applications.

Apart from many working meetings, the CrossGrid consortium plans on holding workshops every year to review progress and prepare material for status reports.

Exploitation of technical results

Within the CrossGrid project we will develop several applications. The range of potential customers extends from medical clinics (bio-medical task), via research centers (physics analysis), to factories and governmental agencies (air pollution combined with meteo and flood predictions). The middleware developed will be open-source.

The developers of CrossGrid applications are already working in close contact with a number of institutions concerned – this applies to all four application tasks. These communities will be constantly informed about the progress of ongoing work, they will get first hand demonstration of results and they will provide first assessments of the products. We believe that such direct contacts present the best way for dissemination and also provide feedback which will be beneficial for the developers.

The developments of the CrossGrid are limited to a few applications and the range of interaction of the developers is limited. However, the technology developed by CrossGrid has a lot of potential, while specific applications could also be of interest to many institutions and agencies. To facilitate the dissemination of the CrossGrid results, we plan to centrally prepare the necessary documentation and PR materials, which could be used by all partners of the consortium, as well as organise dedicated seminars and/or demonstrations to the communities of potential users.

Research, teaching and training

The topics of the CrossGrid project concern computer sciences and several other disciplines: biomedicine, meteorology and air pollution, hydrology and particle physics. For the computer sciences, Grid technology problems are highly interesting on their own, as they present new challenges. The development of Grid testbeds, as well as new Grid tools and services for interactive applications, poses a number of interesting research problems. These will be addressed by the members of the collaboration. We are sure that these topics will generate many PhD theses, thus implanting new technological concepts in the minds of computing experts of the future, but also enhancing the Grid middleware and tools.

For other disciplines, the CrossGrid project proposes several challenging applications, which require synergy between computer sciences and the other fields. We believe that the CrossGrid project will demonstrate the power of the Grid in solving complex problems of biomedicine, earth sciences and physics, in real-time. We hope that these applications will find customers, but we also think that such demonstration will provide an important inherent value for future research.

Many members of the Consortium have regular teaching duties. This is a unique opportunity for the promotion of new scientific ideas and technologies, including the Grid. There is no doubt that access to younger generations is the best way to introduce new technologies and habits, and many industrial companies invest in this. We see teaching as an important vehicle in the promotion of new ideas and we are sure that it greatly amplifies the message.

Direct contacts to students give many of us an opportunity to expose them to new ideas, but also to select promising young researchers to work on the Grid technologies, leading to PhD theses.

Public relations

There is a great public interest in the “information society” and related technologies. We see it as our duty to provide information on ongoing developments, trends and prospects. The members of the consortium will fulfill their obligations in contacts with national and international media; existing material will be distributed and some more, concerning new developments, will be prepared.

8.2 Specific exploitation and dissemination plans for each Partner

Specific exploitation and dissemination plans for the CrossGrid partners are reported below.

Academic Computer Centre CYFRONET AGH (CO1)

Dissemination and exploitation

The Academic Computer Centre CYFRONET is the leading unit in the field of high performance computing and networking. The mission of ACC is to offer access to computational facilities and network services for universities and research institutes, as well as serve as the centre of competence in this field and promote new developments in efficient exploitation of computers and computer networks.

Since 1999, ACC is taking a very active part in the development of the Polish cluster of SGI computers and this experience will be used in establishing the national Grid network in Poland. On the other hand, participation in the CrossGrid project will be a very important factor for this activity.

The ACC will promote Grid computing among the Polish scientific community. The ACC staff members have started to take part in the Grid Forum and this will be a natural opportunity to disseminate results of the CrossGrid. ACC is also going to work with partners from software companies and industrial research centres (e.g. Motorola and ABB divisions in Cracow).

The daily promotion of the CrossGrid activity will go through the ACC web page dedicated to this project, where we plan to put information about the project objectives, participants, results obtained with international testbeds, as well as multimedia presentations.

ACC has already decided to organise each year the Cracow Grid Workshop to bring together people from Poland and Central Europe interested in the development and application of Grid technologies; the first event of this series was held on November 5 – 6, 2001 (http://www.cyfronet.krakow.pl/cgw01).

The Grid technology is already included in the advanced courses for students of computer science at the University of Mining and Metallurgy. In the last two years, about ten MSc theses were related to Grid computing and this activity will be continued in the following years. About ten MSc and six PhD students from the Institute of Computer Science (which closely collaborates with ACC) will be involved in the CrossGrid Project. At ACC, research is performed in an entirely open way and the methods and software developed are available to the scientific community.

Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw (AC2)

Dissemination

Distributed data processing and computations are already a must in a couple of activities performed at ICM, viz. particle physics, meteorology and physics of atmosphere, astrophysics, biomolecular physics and genomics. All these disciplines will profit from technologies and experience gained from the testbeds for particle physics and for the weather forecast distribution realized in the framework of this project at ICM. It is going to be a preliminary step before setting up a larger scale national grid infrastructure in Poland, where ICM is playing the leading role. Data repositories and software are going to be exploited and gradually modified over the whole period of currently running large grid and scientific projects that are using them (LHC experiments, astrophysics projects, meteorology and Eurogrid). As ICM is already involved in the BioGRID project, which is part of the Eurogrid project, it is taking an active part in the Grid Forums.

Special Web pages devoted to the CrossGrid project will be maintained at ICM (http://crossgrid.icm.edu.pl) in analogy to already exiting BioGRID Web page http://biogrid.icm.edu.pl/ and will be part of the ICM web site (http://www.icm.edu.pl/). These will present and promote the CrossGrid as a whole and ICM's involvement in it, as well as tasks descriptions and their status.

One of the CrossGrid testbeds is to be built at ICM. It will be part of the larger structure built in Poland (in Krakow, Poznan and Warsaw). The ICM testbed will be devoted to testing general CrossGrid middleware (especially Grid monitoring software which will be developed at ICM) and meteorological and physics applications (which are to be developed at ICM).

A special seminar concerning Grid technologies is going to be organized at ICM (possibly with the cooperation of the Soltan Institute for Nuclear Studies) for specialists, students and a broader audience. Its main goal will be to promote GRID initiatives held at ICM (CrossGrid, Eurogrid, GRIP) and to adress technical aspects of the Grid.

We plan to have M.Sc. and/or Ph.D. students involved in building this testbed site and developing meteorological and physical applications.

Exploitation

The data, the software and the documentation produced and analyzed at ICM will be available to all members of scientific and technological projects involved in them. In view of the academic nature of this institute, these will also be available to academic groups from other fields.

The Henryk Niewodniczanski Institute of Nuclear Physics (AC3)

Dissemination and exploitation

The Polish Data Grid Consortium has been established in 2001. It includes seven Polish institutions (four physics institutes/universities - including INP Krakow, and three computing centers). The Consortium has been established largely due to the initiative of the INP Krakow. Professor Michal Turala (of INP Krakow) has been asked to lead the Consortium.

Most partners of the Consortium are participating in the CrossGrid project, therefore it is natural that CrossGrid activity will play an essential role in the Consortium operations. Members of the Consortium are the partners in the POLGRID proposal, which has been submitted to the Polish State Committee for Scientific Research in spring 2001, proposing the development of GRID computing across a broad scientific community in Poland and awaiting final discussions and decisions. Experience with the CrossGrid will be transferred to the POLGRID project, if approved.

Until now, the INP Krakow has been involved in the large international collaboration around the LHC experiments at CERN Geneva, SLD experiments in Fermilab, USA, and Belle experiments at KEK, Japan - all of them working intensively on GRID projects for physics. In these experiments, the INP personnel is involved in "on-line" filtering of data and data analysis, working hand-to-hand with other laboratories. One of the tasks (subtasks) of the CrossGrid project will perform the study of the use of GRID technologies in "on-line" filtering of physics data. We believe that, if successful, GRID technology could help substantially to reduce the size of local computing farms, necessary for "interactive" data processing.

Actually, the INP staff is participating in the GRID meetings of the physics community at CERN and elsewhere. Some members participate in the Global Grid Forums (e.g. Amsterdam).

The Web address of INP is http://www.ifj.edu.pl/. and the CrossGrid project will be presented on its Web pages. INP will not develop its own testbed, however the testbed installed at Cyfronet will be used extensively. INP will develop the task (sub-task) of using the GRID for "on-line" filtering of the physics data which will be tested using the Cyfronet testbed.

INP Krakow is interested in promoting GRID technology, as it provides opportunities for future scientific computing. In November 2000, we organized a one day workshop attended by four people from CERN DataGrid and about 20-30 people from INP Krakow, the Krakow Academy of Mining and Metallurgy, INS Warsaw, ICM Warsaw, IPSNC Poznan, Warsaw University, the Krakow Branch of Hewlett-Packard. In May 2001, a seminar entitled "Time for GRID?" was organized with the participation of about 20 people from INP Krakow. The results obtained in the field of "on-line" filtering of the physics events using GRID technology (one of the CrossGrid tasks/sub-tasks) will be presented at several large international fora of LHC collaborations (several meetings yearly, each meeting attended by 20-30 people). In addition, these results will be presented at specialized conferences, such as IEEE Real Time Symposium (attended by a few hundred people), Workshops on LHC Electronics (attended by 150-200 people), etc.

The CrossGrid activity will be described in the Annual Report of the INP Krakow - printed in a quantity of a few hundred and distributed to more than a hundred scientific libraries and some leading scientists in the field of physics.

The staff of INP Krakow gradually introduces GRID technology in its lectures to students of physics, computational physics and other specializations at the Jagiellonian University and Technical Universities of Krakow; participation varies from a few to a few dozens of students per lecture. We plan to get 1-2 students involved at the level of their M.Sc. and/or Ph.D. thesis. The work will be focused mainly on the application of GRID technologies for "on-line" filtering.

INP has good relations with the local press and will keep them informed about the CrossGrid project and the role of Polish institutions (including INP Krakow) in the project.

The Andrzej Soltan Institute for Nuclear Studies (AC4)

Dissemination

The group of high energy physics at INS is taking part in two data intensive experiments at CERN, COMPASS and ICARUS, becoming operational before LHC, and three LHC experiments at CERN: ALICE, CMS and LHCB. The principal goals of the tasks realized at SINS are to setup the environment and tools for distributed data production and data analysis in these experiments and use them for the real data and simulated data production and analysis, first at the COMPASS and later at the ICARUS and LHC experiments. In particular, first data production at COMPASS is projected immediately after the completion of the first run in September 2001. For LHC, the test distributed data production for Monte Carlo simulated events in CMS is planned for the second half of 2001. The same scenario will be repeated every year.

Special Web pages devoted to the CrossGrid project will be mantained at SINS (http://crossgrid.fuw.edu.pl) and will be part of the Warsaw High Energy Physics Group web site (http://hep.fuw.edu.pl). These will present and promote the CrossGrid as a whole and SINS's involvement in it as well as tasks descriptions and their status.

One of the CrossGrid testbeds will be built at SINS. It will be part of the larger structure built in Poland (in Cracow, Poznan and Warsaw). The SINS testbed will be devoted to testing general CrossGrid middleware and physics applications, which are to be developed and implemented at SINS.

The Grid community at SINS will take an active role in organizing a special seminar concerning Grid technologies at the Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, for specialists, students and a broader audience. Its main goal will be to promote GRID initiatives held at SINS and ICM (CrossGrid, Eurogrid, GRIP) and to adress technical aspects of the Grid.

We plan to have Ph.D. student(s) involved in building this CrossGrid testbed site and developing physics applications.

Exploitation

The data, software and documentation produced and analysed at SINS will be available to all members of collaborations mentioned above and these include all major laboratories in the field in Europe, the US and the Far East.

Universiteit van Amsterdam (CR5)

Dissemination

The Universiteit van Amsterdam (UvA) participates in the CrossGrid project primarily through the development of an interactive, grid-based biomedical application, aimed at pre-operative analysis and simulation of vascular problems. The UvA closely collaborates with various medical institutes, nationally and internationally on the development of this system. Partner institutes include the Leiden University Medical Centre (LUMC), the Amsterdam Medical Centre (AMC) and Stanford (Md C. Zarins and Prof. C Taylor). This close collaboration will ensure that the system will find acceptance in the medical world.

Aside from the targeted promotion of our results through our contacts in the medical world, the normal channels for dissemination and publication for scientific institutes – journal publications, conferences and our group web pages – will be used extensively. As a case in point, at the ICCS 2002 conference (http://www.science.uva.nl/events/ICCS2002/), three workshops on medical topics will be organised.

Exploitation

The group at the UvA has extensive contacts in the world of medical research that will be employed to validate and promote the technology developed within the CrossGrid project. These include connections with the e-Science program (Tony Hey, U.K), a program in the area of medical visualisation (Prof. N. Avis, UK), A. Marsh's company Virtual Medical Worlds, and Medis, a software company in Leiden (NL). In the course of developing this application further spin-off results will be produced. These include a parallelised version of the lattice-Boltzmann method, suitable for interactive simulations. Another important spin-off will be in the area of interactive, immersive visualisation of 3D medical scan data.

Ustav Informatiky, Slovenska Akademia Vied (AC6)

Dissemination

SAS participates in the creation of a national grid for flood alert systems in Slovakia within the framework of POVAPSYS project, which is supported by the Slovak government. Besides II SAS, Slovak Hydro-meteorological Institute (SHI) and Slovak Water Enterprise (which is responsible for the management of Slovak River Authorities) will participate in this project. Building and managing a Slovak national grid will be the role of SAS in the project dissemination context.

Lectures about Flood Alert Systems based on Grid technology will be organized by SAS. Contacts with the Slovak Water Enterprise will be ensured with SAS contacts. The SAS will participate in hydraulic-hydrological modeling of the project test beds.

The CrossGrid project will be advertised in the Journal of Computing and Informatics (bimonthly) of the Institute.

Ph.D students will work in the project in the topic of hydrological modeling based on Grid technology.

Exploitation

The flood modeling modules and other grid tools in CrossGrid could be used in POVAPSYS. During CrossGrid, SAS would like to activate the building of the International Flood Alert Grid for international rivers providing hydraulical-hydrological modeling modules. The flood modeling modules and other grid tools in CrossGrid could be used in the International Flood Alert Grid.

GUP, Johannes Kepler Universität Linz (AC7)

Dissemination

GUP-Linz is involved in the Austrian Grid Consortium AGRID. As an impulse for the project, the AGRID Consortium is attempting the Austrian Medical Grid (AMG) project, which is currently being established.

The University of Linz is project leader of the AMG project and additionally covers the area of distributed visualization in Virtual Reality environments, as well as high performance distributed and parallel computing.

The Austrian Medical Grid (AMG) project is intended as a major Austrian research initiative for grid computing, with the focus on its extension to a more international grid. Several prospective partners of European Union and Candidate Countries are already involved in the discussion and establishment of the AMG. Since GUP-Linz is actively involved as a leader in these efforts, an extension of this role towards the international level is intended. It is possible that partners from the CrossGrid consortium will be invited for participation, as soon as AMG reaches an international level. Some initial discussions are already underway.

The University of Linz is currently not actively involved in Grid Forums, but is taking strong interest in these efforts. However, there is some bilateral cooperation with partners engaged in Grid Forum initiatives.

The University of Linz is engaged in several projects concerning Virtual Reality visualization. One example is the SAVE (Safety Training in Virtual Environments - http://www.gup.uni-linz.ac.at/save) project for the Austrian OMV petrol company. Past activities include visualization of numerous applications in computational science and engineering, and parallel and distributed high performance computing.

The website of GUP-Linz can be reached via http://www.gup.uni-linz.ac.at/ . GUP will devote some pages to the CrossGrid project and especially the intended Task 1.5 of WP 1. An icon for the CrossGrid involvement of the Institute will be placed on the organization homepage.

The University of Linz has the appropriate hardware for project test beds available. One example is the SGI Origin 3800 installed at the Johannes Kepler University Linz, which hosts 128 processors and is currently the strongest academic supercomputer of Austria (see http://www.top500.org). Another example is the available CAVE Automatic Virtual Environment, a 3-dimensional cube-sized projection system for immersive Virtual Reality applications, which has been used during several projects of partner AC7. Some other Virtual Reality gadgets, such as Head Mounted Displays (HMDs) and the like are also available. In addition, partner GUP is in constant cooperation with the Ars Electronica Center (AEC) in Linz, which is widely known for its activities in Virtual Reality.

At present, the University of Linz is organizing the Special Session on Parallel and Distributed Programming Tools for Grids at the 10th Euromicro Workshop on Parallel, Distributed and Network-based Computing (PDP 2002) in Las Palmas de Gran Canaria, Canary Islands, Spain in January 2002 (see http://www.gup.uni-linz.ac.at/pdp2002). This session is established together with the Laboratory of Parallel and Distributed Systems of the Hungarian Academy of Sciences, which is also involved in the DataGrid project. In addition, it is involved in the organization of the International Conference on Computational Science (ICCS 2002) to take place in Amsterdam, The Netherlands in April 2002 (see http://www.gup.uni-linz.ac.at/iccs). In September 2002, partner AC7 will host the 9th EUROPVMMPI Conference together with the 4th Austrian Hungarian Workshop on Distributed and Parallel Systems at the Johannes Kepler University of Linz. For this conference, partner AC7 plans to establish a special session on Grid computing in the scope of PVM and MPI. Other conference and workshop activities are planned for the future.

The area of visualization is a very good candidate for both, CD and Web multimedia presentations, and therefore also important for CrossGrid activities.

The Johannes Kepler University produces several periodic publications, which are intended to describe the work of its members. The monthly magazine "News from the Campus" describes research activities taking place at the Johannes Kepler University. The quarterly magazine "Forschungsmedienservice" of the Johannes Kepler University describes the latest research results of the Technical Faculty, where partner AC7 is a member. In addition, there is the "Supercomputing and Visualization" brochure of the Central Computing department, which is suitable for the research conducted during the CrossGrid project. The University of Linz plans to present its involvement in each of these official publications, as well as in its own information folder.

Currently, the University of Linz offers the following courses and lectures at the Johannes Kepler University Linz.

- Parallel Computing
- Computer Graphics
- Virtual Reality in the CAVE
- Applied Computer Graphics
- Parallel Computer Graphics
- Software Engineering for Parallel Systems

Each of these courses features a theoretical part as well as practical work.

In addition, the University of Linz reserves a space for intermittent lectures concerning

- Selected Chapters on Parallel Computing
- Selected Chapters on Computer Graphics

These lectures are intended for visiting professors at Univ. Linz. During the course of the CrossGrid project, we intend to invite some specific partners for each of these lectures.

Students of partner AC7 (Univ. Linz) have always participated in research projects. On the one hand, the Computer Science curriculum at the Johannes Kepler University Linz requires the students to conduct such a project during their studies. On the other hand, they are free to continue this project during their diploma thesis, or even their PhD thesis. The number of students currently involved in scientific research projects at partner AC7 is 5 project students, 7 diploma students, and 6 PhD students.

The "Forschungsmedienservice" magazine intended for distribution to the press will be used for advertising CrossGrid. The University of Linz plans to advertise its participation in CrossGrid with the appropriate means since it is one of the first involvements of Austria in an international grid project.

Exploitation

The GVK (Grid Visualization Kernel) developed in Task 1.5 of CrossGrid may also be adapted as a core element of the AMG project. However, specific changes related to medical computing will have to be applied.

Forschungszentrum Karlsruhe GmbH (CR8)

Dissemination

The research program of FZK is concentrated on the five main areas of energy, environment, structure of matter, health, and key technologies. Within these fields FZK is involved in numerous computing intensive domains like meteorology and climatology, nuclear fission and fusion technologies, materials research, nanotechnology, solid-state physics and astrophysics. It is envisioned to introduce the developed CrossGrid technologies and tools into these numerical sciences by direct user support. Dissemination and information exchange about the new CrossGrid technologies is thus guaranteed through numerous international collaborations and contributions to scientific workshops, conferences and publications. In 1996, the Computing Centres of FZK and the University of Karlsruhe founded the Virtual Computing Centre Karlsruhe to establish a common IT-infrastructure. End-users of both institutions not only share different computing facilities but are also provided with complementary information systems like Web- or News-services. Developments of the CrossGrid technology will be promoted to both user groups via these information systems and published in the monthly Newsletters of both Computing Centres. Together with three German universities FZK is presently building a German Tier-B computing centre for high-energy physics data from the BaBar experiment at the Stanford Linear Accelerator Centre SLAC. This centre will be linked to RAL, CCIN2P3 at Lyon and SLAC to form an international Grid-infrastructure for BaBar computing. Data-intensive applications from the experiment analysis will participate in the Cross Grid test bed so that the results are directly made available to respective end-users.

FZK has a division devoted to the dissemination of scientific and engineering developments to the general public, as well as a unit devoted to technology transfer and marketing addressed to the industry. It will publish official press releases and is planning on disseminating Grid technologies during its next “open day for the general public” in 2002.

An official Web page about FZK activities in Grid Computing is under development and will be available by the end of October 2001. Links to CrossGrid activities will be included.

Exploitation

A much larger “Regional Data and Computing Centre Germany” (RDCCG) at FZK is presently under discussion. The RDCCG is expected to provide a complete Grid computing, Grid R&D and tools support infrastructure for LHC- as well as non-LHC experiments with hundreds or even thousands of users world-wide. FZK is a member of the international “LHC Computing Grid Project” to be launched at the end of 2001 and will support the linking of CrossGrid developments with these activities.

Inside FZK, IT-coordinators of the different scientific institutes exchange information and decide about strategic investments in hardware and software technologies through periodic meetings with the computing centre. End-users are informed about current IT-developments in the Centre through monthly newsletters and through Web- and News-services. CrossGrid technologies will be included in these information systems.

Universität Stuttgart (AC9)

Dissemination

In addition to the dissemination from the technical experts to the end–users in the project, the CrossGrid results will be presented to a target audience of ISVs, hardware vendors and interested users of metacomputing/coupling technology. In addition, the University of Stuttgart will present CrossGrid using proven methods and channels:

- Presentation of papers at conferences and other meetings
- Attendance at relevant exhibitions (exhibitors, marketing talks)
- Distribution of CrossGrid brochures at conferences/exhibitions
- Announcement of CrossGrid results in the relevant printed/electronic publications

The CrossGrid dissemination material will be constantly adapted to show the latest results.

Exploitation

Business activities relevant to this project

The developments of the CrossGrid project will provide a direct benefit and improvement for the RUS core business to consult and help its users. Distributed computing and Grid computing are becoming a more and more important technology to do more complex simulations.

Market opportunities

RUS/HLRS together with debis SH and Porsche AG has formed a joint company to foster the use of Supercomputing by small and medium sized enterprises. The aim of this group is to provide easy access and use of distributed resources. It is quite obvious that there is a market for such technical computing and that this market is steadily growing. To become an important provider will require being able to provide computing power at one's finger tips without all the overhead that is prohibitively high today. CrossGrid will develop tools that help to overcome some of the limitations in that field. Especially, the increased productivity of parallel program development using the MPI debugging and verification tool will strengthen the value of the scientific simulations, because it will allow faster creation of more reliable programs faster.

Exploitation Plan

RUS will exploit the CrossGrid results by giving access to the software tools to its user community. It will further use the tool in the training courses on MPI, teach how to use the MPI standard in a portable way and help users solve their problems.

Technische Universität München (AC10)

Dissemination

The Lehrstuhl für Rechnertechnik und Rechnerorganisation / Parallelrechner at TUM (LRR-TUM) has close contacts with industrial users of high performance computing technology and consortia like the Global Grid Forum. These contacts, as well as conferences, workshops and seminars will be used to disseminate the results of the project. Another main path of dissemination will be the KONWIHR competence network that supports the use of high performance computers, especially the Hitachi SR8000 at the Leibniz Computing Center, the most powerful computer in Europe. The South Bavarian office of KONWIHR is housed by LRR-TUM. Results related to performance analysis and tools will also be propagated in the APART2 working group. Finally, the project’s results will immediately be incorporated into teaching activities, especially a lecture on grid computing.

Exploitation

The computer science department of TUM will build an *active campus* at its new site in Garching, where it will move in 2002. The active campus, which will be developed within a projected special research grant provides an electronic information and service infrastructure for mobile users, especially students and university teachers. The possible scenarios include: A ubiquitous, personalized information room for students and guests; spontaneous use of services and equipment without a need for precedent configuration or cabling; sharing of information, common documents, (remote) experiments, etc. As the computing infrastructure of the active campus consists of a loosely coupled, frequently changing collection of heterogeneous computer resources without a centralized administration, it shares most features and problems with a Grid environment. The techniques for programming and managing this infrastructure will therefore be fundamentally based on the techniques, services and tools developed in the CrossGrid project (esp. in WP 2 and 3).

Together with an industrial partner from previous cooperations, LRR-TUM currently arranges a research project on remote visualization and steering for applications in medicine and computational prototyping. This project will directly profit from the results of WP 1 and the grid infrastructure developed in the other workpackages.

Poznan Supercomputing and Networking Center (CR11)

Dissemination and exploitation

CrossGrid middleware and applications will be used in the Polish National Grid and locally in the metacomputer of PSNC. It was created by PSNC in co-operation with other Polish computing centres (Krakow, Gdansk, Wroclaw, Lodz, Szczecin, Torun). Polish supercomputer centres provide computing power for all scientific communities, therefore supercomputers are used by scientists and students from different fields of science such as physics, chemistry, engineering and biology. Polish scientists can use applications developed in the CrossGrid project. Especially, scientists from Medical Universities can benefit from interactive simulation and visualisation of a biomedical system (WP 1.1). Most users can also find results of visualisation in a grid environment (WP 1.3) interesting, especially in the context of the planned creation of the Polish High Performance Visualisation Grid. PSNC is starting the development of a virtual laboratory in which there are plans to make the NMR spectroscope available for a grid community. This project can also use experiences and visualisation tools from the CrossGrid.

Currently, the Polish National Grid is constructed with the use of the Load Sharing Facility job management system from Platform Computing and the Virtual Users' Accounts System developed in the Poznan Supercomputing and Networking Center. There are plans to evolve the Polish National Grid toward a more grid-oriented environment. The CrossGrid experiences and middleware will be used to create a grid environment that will form a basis for scientific grid applications. Most of the Polish scientists run third party applications, but there is also a large group of users who write and run their own programs. Debugging, performance and monitoring tools (WP 2.3, 2.3, 2.4) will help them write programs that can efficiently use grid resources. This tool will also be valuable to staff of supercomputer centres, because they can provide better assistance to users in parallelizing their applications and develop new, distributed tools for a grid environment.

Grid resource management and monitoring tools (WP 3.2, 3.3) will be applied to improve the Polish Grid management. It will help in better resource utilisation and faster problem location. These tools will also be used for other computers in PSNC that will not be included in the grid. Portal and roaming access tools (WP 3.1) will also be put to use in PSNC. Portal access to a specific application will be used for some bioinformatic applications. A project for writing this kind of application is starting now. PSNC will also dedicate a special server for roaming access to the grid.

PSNC also participates in some international projects: GridLab, ENACTS, SEQUIN and is involved in Polish-German border regional co-operation. These projects are all connected with network or grid activity, therefore experience gained during CrossGrid development can contribute to these projects as well.

PSNC is one of the initiators of the European Grid Forum. One of the EGRID meetings was hosted in Poznan. PSNC has a representative on the GGF Steering Committee. PSNC is also active in the GGF working group, especially in the Security Workgroup and Scheduling and Resource Management Workgroup.

Exploitation of the CrossGrid testbed by the Poznan Supercomputing and Networking Center is planned through the integration of their resources with CrossGrid infrastructures and tools.

University of Cyprus (AC12)

Dissemination

CrossGrid will provide the basic framework for building a national Grid infrastructure in Cyprus. Therefore, the full benefits of the Grid resources must be available to larger segments of the academic, research and educational communities. To this end, the project will undertake the diffusion of the Grid technologies in other Cypriot scientific and research sectors (other Departments of the University of Cyprus, private institutions of tertiary education, the Cyprus Meteorological Service, the Institute of Neurology and Genetics, the Agronomical Institute, the Cyprus Telecommunications Authority, etc.) and the private sector.

In the quest for establishing and promoting the Grid infrastructure in Cyprus, a number of activities are planned in view of the CrossGrid project:

- Within six months from the beginning of the project: workshop on Grid initiatives in Europe, to be held at the University of Cyprus, with invited speakers from the European Commission, the CrossGrid consortium, and other large Grid-oriented projects like DataGrid.
- Within 24 months from the beginning of the project: workshop on exploitation potential of local testbed, and update on CrossGrid activities and tools.
- Development and deployment of a Web-site with information about resources available to the national Grid community, as CrossGrid evolves.
- Development of Web-based helpdesk to support users of CrossGrid in Cyprus.

A number of small workshops may be developed, if needed, to support user groups planning to use the national testbed.

Exploitation

Exploitation and diffusion efforts will be along the following axes:

1. **Exploitation of the project test bed**

- Exploitation of the CrossGrid test bed by the Department of Computer Science and the Nuclear and High-Energy Physics Laboratories of the University of Cyprus, through the integration of their resources with CrossGrid infrastructures and tools.
- Support for participating at CERNS's LHC and other large-scale international experiments, like HADES.

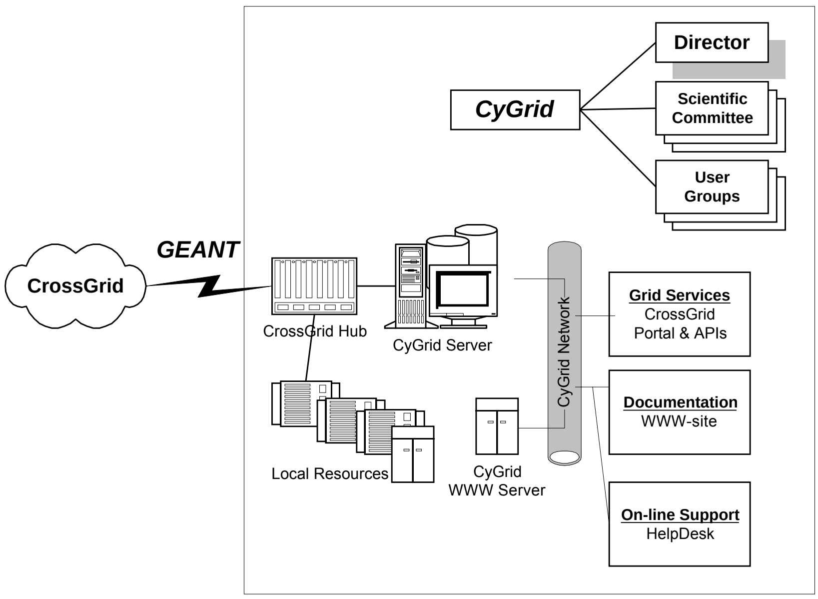
2. **Formation of a National Grid Infrastructure**

- Integration of national research computing resources into a national Grid and connection with CrossGrid. Attraction of new users and applications to the Grid.
- Promotion of local and international collaborations in the field of computational sciences and Grid Computing.
- Better exploitation of the Géant network. Support of and active participation in the upcoming initiatives for upgrading the network-connectivity in the Euro-Mediterranean region in the context of the Eumedis programme.

3. **Management & Co-ordination of National CrossGrid Activities**

- Establishment of a ***CyGrid***, a Coordination Center for Grid-services provision, under the auspices of the University of Cyprus.
- Formulation of rules and regulations for the provision of Grid-services at a national level. Support for the formation and running of user-groups. Exploration of schemes for further funding and for pricing of Grid-based services to safeguard continuous support and upgrade of Grid infrastructures, past the lifetime of the CrossGrid.
- Development and deployment of Web server with Web-based manuals on CrossGrid services and tools, and an interactive Web-based helpdesk for user support.

The structure of CyGrid is depicted in the following picture. Establishment of the CyGrid's structure will be submitted for approval to the Senate of the University of Cyprus, and to the Research Promotion Foundation of Cyprus, which coordinates and fund research activities in the country.



DATAMAT S.p.A. (AC13)

Dissemination and exploitation

DATAMAT is already strongly involved in the Grid activities at a continental level, being one the three industrial partners of DataGrid and being the prime contractor of SpaceGrid, a large project on application of Grid technology to the space field (earth observation, solar terrestrial physics, spacecraft engineering, space science), funded by the European Space Agency.

For this reason, DATAMAT is already actively participating in Global Grid Forums and is, in particular, within the steering committee of the DataGrid Industry and Research Forum.

Further than those specific events, generally involving at most Grid-aware organisations, DATAMAT is also, for the time being, actively promoting Grid towards the general public through the following means:

- press conferences / articles on national newspapers, either on its own or in conjunction with the main Italian scientific institutions involved in the Grid (namely INFN and CNR);
- a specific brochure on rationale behind DATAMAT commitment in GRID projects (including DataGrid, SpaceGrid - on going - and CrossGrid - under negotiation), which has been printed and distributed at recent GGF3 and at the last press conference

DATAMAT is also planning to have a specific section of its web site devoted to Grid related projects, in particular, once it has achieved significant results from its internal investment project.

In conclusion, DATAMAT will disseminate CrossGrid-related information through the same means and events, making it complementary to the already on going involvement in the other projects.

Trinity College Dublin (AC14)

Dissemination and Exploitation

The Department of Computer Science at Trinity College Dublin will establish a testbed node for CrossGrid and will actively promote CrossGrid within the Grid-Ireland community, in which TCD is a founding partner. CrossGrid will greatly enhance the national grid activities, both systemically and in its international dimension.

TCD will also participate in the development of the monitoring facilities that will support application-level performance evaluation. These nicely complement the non-invasive trace tools and databases that TCD has been developing over several years, and substantively extend the work in system-level monitoring that TCD is involved in within DataGrid.

TCD will publish its results via the normal journal, conference, workshop and web channels, but will also introduce relevant material into undergraduate and postgraduate teaching.

Consejo Superior de Investigaciones Científicas (CR15)

Dissemination

CSIC will lead a national Grid network in Spain as Coordinator. CrossGrid will be integrated in this network as the first multidisciplinary Grid.

CSIC is also involved in international Grid projects in progress and participates actively in GGF1 and GGF3. The CrossGrid web site can be promoted through the Institute's site (http://www.grid.ifca.unican.es/).

CSIC is already involved in DataGrid and will participate in the test beds of the CrossGrid project.

The contribution of CSIC to the creation of the CD-ROM and the CrossGrid e-zine is projected with articles and material concerning the project achievements in Spain. CrossGrid will be presented in the CSIC Annual

Review and in the national press. A number of students working for their PhD in Physics with CSIC will be involved in the CrossGrid development. Several lectures and talks have been organized by CSIC concerning various Grid issues in the past. CrossGrid will be the subject of similar talks and lectures along the project duration.

Exploitation

The Institute has an agreement with a Spanish electrical company for exploring the Industrial application of Grid simulation of electrical loads. New CrossGrid services can be considered around this topic.

Universitat Autonoma de Barcelona (AC16)

Dissemination and exploitation

Researchers from the UAB participating in the CrossGrid project belong mainly to two groups: the High Energy Physics Institute (IFAE) and the Computer Architecture and Operating Systems (CAOS) group.

The IFAE group is the coordinator of Spanish HEP groups participating in the DataGrid project. They are involved in the testbed workpackage and, as a consequence, they will work in the deployment of a national grid network linked to the european DataGrid network.

The CAOS group has been collaborating in scheduling problems since 1999 with the researchers of the University of Wisconsin-Madison (UWM) that are responsible for the Condor system. The computing resources of the CAOS Condor pool are planned to be linked in the near future with the Condor pool of the UWM, which has more than 700 machines. Further discussions will be undertaken to link the Spanish HEP grid with the UWM Condor pool in order to establish an international grid of sites that use the Condor system.

On the other hand, both the IFAE and the CAOS groups, have submitted a proposal to the Catalan regional government to create a network of groups interested in distributed computing. The network will focus mainly on the dissemination of issues related to distributed computing and, eventually, will facilitate access to existing grid environments deployed by members of the network.

UAB groups are involved in the testbed workpackage and will participate in the framework of HEP applications.

Members of the CAOS group are currently participating in the discussions of the Scheduling Group of the Grid Forum. CrossGrid activities relating to scheduling will be disseminated to this discussion group as long as results of the CrossGrid consortium are produced.

A common CrossGrid web page will be set up by the IFAE and the CAOS groups. This page will be linked to other particular pages from both groups and will also be linked to institutional web pages of the UAB devoted to dissemination of research projects that are carried out by UAB groups. A CrossGrid presentation could also be included in a newsletter of the UAB (named l'Autonoma) that is published every month and is addressed to the university community.

Finally, the CAOS group will include references to the CrossGrid project in some of their graduate courses devoted to parallel and distributed computing. Some M.Sc. and Ph.D. students (two or three) are also planned to be involved in CrossGrid activities. In particular, they will work in scheduling topics, grid testbed set-up and development of grid HEP applications.

Universidade de Santiago de Compostella (AC17)

The USC will work close with University of A Coruña (UDC) and Supercomputer Centre of Galicia (Centro de Supercomputación de Galicia, CESGA). In fact, the main testbed facilities in Galicia (a region in the north west of Spain) will be at CESGA, improving the computer facilities for the R&D existing actually. They can be used free of charge by researches of universities of Galicia, public R&D centers belonging to the local goverment (more than 25) and centres of CSIC around Spain (one of the partners of the CrossGrid consortium).

USC, UDC and CESGA are strongly involved in projects related to distributed computing. They have developed their own model, named Virtual Supercomputer of Galicia, that runs in three diferent locations spread around Galicia. In this distributed cluster recently the first 3D european annimation film (el bosque animado) was rendered (115000 pictures, 1.2 TB. See http://www.cesga.es/dixitos/2001/xullo01.pdf).

A variety of high performance computing servers are available at CESGA allowing researchers to always choose the most appropriate architecture to meet their computing needs. CESGA's users can choose from vector, parallel vector, scalar, shared or distributed memory parallel scalar architectures according to the needs dictated by the research they are conducting. So, the Grid testbed will be a new architecture available for these researches.

CESGA's main equipment is made up of a Fujitsu VPP300E parallel-vector processing server, a Fujitsu AP3000 massive parallel-scalar server and a SUN Microsystems Enterprise 4500 shared memory processing server. Both the AP 3000 and the Enterprise 4500 are equipped with UltraSPARC-2 processors.

Dissemination

A number of disseminations activities are planned by the University of Santiago and CESGA. Among these activities are included:

- publication of several articles concerning the project in the CESGA newsletter, DIXITOS (http://www.cesga.es/dixitos). It is distributed in paper to the most important R&D public and private centers of Galicia and to the private IT companies. So, the activities of CrossGrid and, in general, Grid computing will be well disseminated around Galicia.
- The advances of CrossGrid will be presented in the annual workshop about HPC that CESGA organizes every year. It will include one or several presentations about CrossGrid in 2002 and 2003 (see 2001 at http://www.cesga.es/Cursos/default.htm?2001/hpcn_axenda.html&2).
- Press brieftings and releases concerning the project achievements will be published in local and national newspapers.

Exploitation

The explotation of the results of the project is direct. CESGA will try to include the Grid model in the computing resources available in the center and in the universities, allowing the researches of universities and public R&D centers to join the global GRID. Also, it has an agreement with CESCA (Cataluña, Spain) and CICA (Andalucía, Spain) for sharing resources. CESGA has proposed the Grid model as the solution for administering it. Finally, the researches will have a distributed platform for testing new applications and architectures adapted to the GRID.

National Centre for Scientific Research 'Demokritos' (AC18)

Dissemination and Exploitation

The Institute of Nuclear Physics of NCSR Demokritos is placing special emphasis on the dissemination of the results and the technological know-how and experience acquired in the context of the CrossGrid project.

To this respect, the following actions will be implemented:

- Organization of seminars on Grid technologies, tools and applications. These will be scheduled within the current program of seminars and advanced courses which the Institute is offering to postgraduate students and young researchers in the field of High Energy Physics.
- Demonstration of the capabilities of the Grid network to interested parties in the Greek scientific community. The multi-disciplinary character of NCSR 'Demokritos', as well as that of the National Technical University of Athens and the University of Athens (both sub-contactors within CrossGrid) will facilitate communication with scientists and researchers in other scientific fields (e.g. biology, chemistry, computer science, meteorology etc).
- Presentation of our results and experience at national and international scientific conferences and fora, as well as through publication of these results in international scientific journals.

- In collaboration with the General Secretariat of Science and Technology of the Greek Ministry of Development, the INP will organize a daily presentation with the participation of leading authorities in the field of Grid technology. The aim will be to expose the Greek scientific community, as well as the private technological sector, to the capabilities of Grid-related technologies and applications.

It is our belief, that these actions will greatly contribute towards the exposition and dissemination of our results and experience to interested parties and potential users of Grid technologies in Greece.

Aristotle University of Thessaloniki (AC19)

Dissemination and exploitation

The Physics Department of the Aristotle University of Thessaloniki will play an active role in establishing collaboration with a number of departments and divisions with strong interests to benefit from the outcome of the CrossGrid project. In particular, the Computer Science Department is interested in the development of databases and visualisation in medical applications. Therefore they have a direct interest in the various WP's of the project.

The Department of Biology is active on DNA mapping. The Division of Environmental Physics is interested in handling large volumes of meteorological and atmospheric data. The Division of Astrophysics is also dealing with large volumes of data and are potential users of the testbed that will be developed by the HEP group within the context of the project. In close collaboration with the University and, in particular, with the Department of Computing of the Polytechnic School of A.U.Th. is the Technological Park of Thessaloniki, which is a collaboration of various software and network-specialized companies with the University and a number of young researchers and graduate students. Some companies involved in the Technological Park are Intrasoft, Datalogic, Sigular etc. The prospects of the CrossGrid project will be presented to the above interested departments and companies and discussed via local workshops, newsletters and appropriately formulated web pages. The usual way is short - one day long - workshops, closely spaced in time - three to six months apart - where working groups will present results and get feedback from the companies and the various research groups. In appropriate web pages, which will be developed under the project, highlights from all the WP's of the CrossGrid project will be regularly presented and updated and the various communication tools will be accessible for trying and feedback.

Finally, interesting results on applications will appear regularly in magazines and monthly journals of the University and the Research Committee like: ANALECTA, POLYMICHANOS, PHENOMENON etc.

Last but not least, the above mentioned divisions and departments have common projects with other Balkan countries like Turkey, Bulgaria, Rumania etc, with exchange of researchers and organization of common workshops and conferences on computing and networking. The aim is to establish contacts with these countries, participating in the above activities in order to develop ties with the CrossGrid project and the Research Institutes in the above mentioned countries. A CrossGrid conference, where researchers from the above countries will be invited to participate, will be organized. Finally, one should not underestimate the educational role that the CrossGrid project can play in the Physics Department. In particular, a close collaboration will be established between the local team and the Graduate Course in Computational Physics, which is currently under way.

Laboratorio de Instrumentacao e Fisica Experimental de Particulas (AC20)

Dissemination

LIP has three research centers in different Portuguese cities. These centers will be integrated. For the time being, there are no plans to build a Portuguese national Grid, the main reason for this is insufficient bandwidth in the major Universities and research centers. However LIP will have an important role in the divulgation of this technology near universities and research centers namely through its membership in the Portuguese Foundation for the National Scientific Computing (FCCN). LIP will push the deployment of GRID technologies within workshops, conferences and member meetings.

LIP is a research laboratory, which is involved in several research activities with CERN, ESA and other international research institutes in the fields of High Energy Physics, Astrophysics and Nuclear Medicine.

Currently LIP is involved in research activities within the CERN LHC programme, which will require the usage of grid technology. LIP is currently participating in the development and usage of analysis and simulation tools, which will be grid aware. LIP is also interested in using resulting grid technologies in other research activities in the domain of astrophysics and nuclear medicine. LIP is extremely interested in promoting and, if possible, linking CrossGrid activities with these current and future research activities. Some presentations on grid technologies have already been carried out in the framework of an international astrophysics project showing the potential of the technology. LIP is involved in the DataGrid Industry and Research forum and has strong interest in contributing to the bridging of these two projects, contributing in different ways for the same fundamental goals.

Contacts will be used for technology transfer and dissemination of the project results. An example of divulgation will be the LIP/ESA workshop being organized in the context of the EUSO experiment, which will be held in November 2001 in Lisbon. In this workshop, Portuguese companies in the field of Aerospace and fault tolerant software will be present. We plan to make a presentation on grid technologies.

LIP has a web site and intends to build pages dedicated to the project at two levels:

1) For the general public through the LIP outreach group web site

2) More technical pages through the LIP computer center web site

LIP will ensure the presence of a crossgrid banner in its main web pages.

LIP will participate in the testbed activities, mainly supporting HEP applications, but we are open to helpin to test and demonstrate any other crossgrid applications.

LIP will participate in the deployment and support of the crossgrid international testbed.

Plans are made for the presentation of the work that will be developed in the context of the CrossGrid project within national workshops and conferences on computing. These presentations will be directed to technical audiences from industry and universities. The major events of this kind in Portugal are the FCCN conference on “Computer Networks” and the workshop on “Computer Centers”. The conference on “Computer Networks” usually gathers around 200 engineers.

LIP will try to involve students in its grid activities. Although difficult we will try to involve students to work on fabric management technologies and authentication issues.

LIP has good relations with the press and plans to release to the press information about the involvement of the institute in CrossGrid and its developments.

Exploitation

The integration of LIP research centers within the CrossGrid test bed will show the advantages of the grid technology. The participation of LIP in grid activities has been well received by FCCN, which is also responsible for the Portuguese National Research Network. Currently, LIP and FCCN are cooperating in order to prepare the Portuguese NREN to support the connection of the LIP research centers internally and externally. This collaboration is focused on providing a quality-of-service aware network and bandwidth capable of supporting grid requirements.

LIP is a member of the Portuguese Association of Electrical and Electronic Equipment Industrials (ANIMEE), which joins the main constructors of electrical, electronic and computing devices in the Portuguese market. This membership provides a communication channel between LIP and the Portuguese industry.

Algosystems S.A. (AC21)

Dissemination

Algosystems will be in charge of organizing the CrossGrid dissemination and exploitation plan with the collaboration of all partners and the support of the Project Management Board.

Dissemination issues that will be addressed by Algosystems include the establishment of the Greek User Groups in collaboration with A.U.Th., the preparation of the project CD, the brochure and the information package.

The design and development of the CrossGrid web site, with the approach of a Grid vortal, in order to support the interaction and communication of Grid-interested communities is very important to the dissemination activity.

The contribution of Algosystems also covers the design of the project Newsletter and the project dissemination area of the web site. The advertisement of CrossGrid in the national press and in technical and scientific reviews of the academic and industrial sector is projected.

Exploitation

The exploitation of the CrossGrid services and the middleware, that will be developed, towards industrial deployment will be addressed through the Industrial Forum and the Application User Groups anticipated by the dissemination and exploitation plan of the project.

In addition, Algosystems will investigate the possibility of exploiting the CrossGrid output to application domains that are of interest to the company such as forest fire behaviour assessment, ground water management, portal services and network traffic engineering.

9 WORKPLAN

In this section we describe the overall workplan that implements the technical system described above.

9.1 Methodology of Work and Structure of Workplan

The work that will be done within the CrossGrid project is based on two principles:

- Clear definition of responsibilities: A single partner will be identified who is responsible for each major workpackage. Thus, the responsibility will cover the overall workpackage. We try to avoid as much overlap in joint work within one workpackage as possible to make the responsibilities very clear and to ease the co-operation within the project. The co-operation between partners within the project is arranged by linking the different workpackages together and merging the different technological components into one common solution.
- Rapid prototyping and iterative engineering: The technical development within the CrossGrid project will be done in various phases and various parallel tracks. The common approach for the development is to focus on small prototypes that are available in a very early phase of the project and can be enlarged with more and more functionality throughout the project. This iterative approach has been shown to be very successful and will be followed within the Project.

The overall structure of the work plan is closely related to the software modules defined previously. This ensures that each module can be thoroughly tested before it is added to the integrated system. Hence, it is possible to monitor the progress more continuously and detect possible problems at an early stage. The consortium can then react quickly by shifting resources to the critical workpackages. Experience has shown that in most cases this is a feasible approach that helps overcome unforeseen difficulties provided that they are detected early enough.

9.2 General Description

The work of this project will research, design, develop, implement, and test the technology and components essential for the implementation of new tools and services that will allow one to deploy a large range of scientific and engineering applications across a Grid. This is a complex project involving many organisations, software engineers and scientists. It builds upon many national initiatives in this area and, as indicated in each of the subsequent workpackage descriptions, to ensure success many of the contributing partners are making available considerably more effort to the project than they are requesting EU funding for.

The Project is split into 5 workpackages:

- **WP1** CrossGrid Application Development, **WP2** Grid Application Programming Environments, **WP3** New Grid Services and Tools will each develop specific well-defined parts of different applications and GRID middleware.
- **WP4** International Testbed Organisation will collect all of the developments from the workpackages WP1-3 and integrate them into successive software releases. It will also gather and transmit all feedback from the end-to-end application experiments back to the developers, thereby linking development, testing, and user experience.
- **WP5** Project Management will ensure the active dissemination of the project results and its professional management.

Each of the development workpackages will start with a user-requirement gathering phase, followed by an initial development phase before delivering early prototypes to the testbed workpackage. Following the delivery of these prototypes a testing and refinement phase will continue for each component until the end of the Project.

9.3 Workpackage list

Workpackage list

Work-package No[1]	Workpackage title	Lead contractor No[2]	Person-months[3]	Start month[4]	End month[5]	Phase[6]	Deliverable No[7]
WP1	CrossGrid Application Development	CR5	537 (365)	0	36	–	D1.0.1 – D1.4.7
WP2	Grid Application Programming Environment	CR8	233 (156)	0	36	–	D2.1 – D2.7
WP3	New Grid Services and Tools	CR11	421 (258)	0	36	–	D3.1 – D3.9
WP4	International Testbed Organisation	CR15	567 (435)	0	36	–	D4.1 – D4.9
WP5	Project Management	CO1	168 (102)	0	36	–	D5.1.1 – D5.3.18
			1926 (1316)				

[1] Workpackage number: WP 1 – WP n.

[2] Number of the contractor leading the work in this workpackage.

[3] The total and (funded) number of person-months allocated to each workpackage.

[4] Relative start date for the work in the specific workpackages, month 0 marking the start of the project, and all other start dates being relative to this start date.

[5] Relative end date, month 0 marking the start of the project, and all end dates being relative to this start date.

[6] Only for combined research and demonstration projects: Please indicate R for research and D for demonstration.

[7] Deliverable number: Number for the deliverable(s)/result(s) mentioned in the workpackage: D1 - Dn.

9.4 Effort per Partner in PM distributed over Workpackages

		WP1 T	WP1 F	WP2 T	WP2 F	WP3 T	WP3 F	WP4 T	WP4 F	WP5 T	WP5 F	TOTAL T	TOTAL F
CO1	CYFRONET	12	6	38	19	98	49	40	28	132	66	320	168
AC2	ICM	68	34			28	14	24	12			120	60
AC3	INP	46	23									46	23
AC4	INS	26	13					20	10			46	23
CR5	UvA	128	72					12	6			140	78
AC6	II SAS	76	76					12	12			88	88
AC7	Univ. Linz	46	23									46	23
CR8	FZK	6	6	13	13			73	73			92	92
AC9	USTUTT			31	31							31	31
AC10	TUM			68	34							68	34
CR11	PSNC	12	6			138	61	18	18			168	85
AC12	UCY			36	22	14	7	30	15			80	44
AC13	DATAMAT					34	34			3	3	37	37
AC14	TCD					32	16	24	12			56	28
CR15	CSIC	74	74	20	20	16	16	72	72			182	182
AC16	UAB	15	15			40	40	19	19			74	74
AC17	U.S.C.	28	17	27	17			41	26			96	60
AC18	Demo							59	40			59	40
AC19	A.U.Th.							74	43			74	43
AC20	LIP							49	49			49	49
AC21	ALGO					21	21			33	33	54	54
	TOTAL	537	365	233	156	421	258	567	435	168	102	1926	1316

T – Total

F – Funded

9.5 Long Description of Workpackages

WP1 CrossGrid Application Development

Workpackage manager: Peter M.A. Sloot, UvA (CR5)

Objectives

Data gathering, processing and interpretation in health and the environment is driven by modern technology. The geographical distribution of data generators (Medical scanners, Patient Databases, Environment-Actuators and Monitors) requires an integrated approach to access and process these data. With the advent of Grid-based technology many new possibilities for data presentation and fast decision making come within reach. It is the goal of the CrossGrid project to explore, incorporate adapt and validate this technology for those application areas that require interactive access to resources, be it Databases, Super Computers, Visualisation Engines, Medical Scanners, or environmental data input devices.

Decision making processes stemming from many health and environmental problems require integration of distributed data and near real-time interaction. The objectives of this workpackage are to provide a representative collection of sample applications from various fields that will drive and exploit the specific (interactive) functionalities to be developed in the CrossGrid project. Firstly, through their need for specific services and their performance characteristics, they will provide a driving force for the technology-oriented workpackages. Secondly, they will serve as a benchmark for the performance of these workpackages, and finally, they will serve as demonstrators, for their respective fields, of the added value provided by the Grid in general, and the technology developed by the CrossGrid project in particular. Each task in this workpackage will focus on the development of one of these applications. Together, they cover a wide range of final user communities, from health care to environment management, and basic research.

All tasks will heavily rely on the performance tools, resource and network services, and management tools developed within the other CrossGrid workpackages. Their deployment on the CrossGrid testbed will test these applications in the final user environment and provide feedback to the application developers. In this way they can fully exploit the possibilities of the Grid.

Though the tasks cover a wide range of application areas, they share many common requirements. Two tasks depend very explicitly on near-real time performance across the network (Tasks 1.1. and 1.2). All tasks depend on efficient access to large databases; in three of the four tasks the data will be distributed over various locations. All four tasks depend heavily on large-scale parallel simulations. These simulations will all make use of MPI for their internal communication.

Many of these common requirements will be addressed by the technology workpackages (WP2, 3, and 4). Other common requirements will be addressed in collaboration between the tasks. This includes visualisation (primarily developed in Task 1.1, but also used in the other tasks), data mining (common to Tasks 1.3 and 1.4), and data discovery (common to Tasks 1.2 and 1.4).

Task descriptions

The tasks in this workpackage map directly to the various health and environment applications, and as such do not strongly depend on each other, so there is no temporal sequencing of tasks. However, at the technology level, tasks will frequently interact and cross-fertilise. In order to drive and to utilise the technology workpackages in a consistent and coherent manner, co-ordination will be needed. All tasks run for the entire duration of the project. Four of the tasks aim at the CrossGrid-embedding of applications, the fifth one is responsible for the overall management of the workpackage and the coherence within and between the tasks.

Task 1.0 Co-ordination and management (Month 1 – 36)

Task leader: Peter M.A. Sloot, UvA (CR5)

This task will be responsible for the co-ordination between the tasks and the coherence in their interaction with the other workpackages in the project. In order to monitor the progress of each task in this workpackage, and to mutually exchange results, the task leaders will be required to submit a formal report to the WP management twice a year. WP co-ordination meetings will be held with the same frequency.

During various phases of the project and the development of the applications, activities will be undertaken to ensure coherency between the various tasks, genericity of the adopted solutions and a focussed collaboration on common issues such as the application deployment on the grid.

In the first phase of the project, the definition of common and coherent detailed requirements to WP 2 and WP 3 is one of the prime objectives of this WP. As part of the kick-off meeting a one-day workshop on the applications will be organised, specifically aimed at identifying common interests. After that, beside the WP co-ordination meetings, technical meetings will be organised to further identify and discuss common requirements between tasks and to come to a coherent first deliverable for the tasks.

Further workshops will be held at least semi-annually. Identified common interests and requirements will be followed up. If deemed effective, further collaborations between the tasks will be initiated. Working groups of experts from each task will be assembled to monitor and guide the parallel developments within the tasks. These issues will include, but will not be limited to, implementation of the testbeds, and application deployment on the grid, adoption of technology developed in the other CrossGrid workpackages and outside the project. In this way it will be ensured that to the maximum extent possible, a common grid environment will be developed, capable of supporting the fairly wide range of data and compute intensive, interactive applications in this workpackage.

Subtasks

1.0.1 General WP management (UvA). Organisation of meetings, collection of status reports, progress monitoring, co-ordination of dissemination.

1.0.2 Requirements definition (UvA with II SAS, CSIC and ICM, M 1-3). In this subtask representatives of all workpackage managers will together produce a coherent document detailing the requirements imposed by the applications on the other workpackages. This task will interact closely with the CrossGrid Architecture team.

1.0.3 Coherency (UvA with II SAS, CSIC and ICM, M 4 - 36). This subtask will be responsible for the continued monitoring and tuning of the work on joint issues, within and outside WP 1. This task will interact closely with the CrossGrid Architecture team.

Number	Description	Responsible partner	Sub-task	Delivery date	Nature	Dissemina-tion level
D1.0.1	Joint requirements definition document	UvA	1.0.2	M3	R	CO[8]
D1.0.2	Semi-annual managerial report for WP1	UvA	1.0.1	M6	R	CO
D1.0.3	Extended yearly managerial report for WP1	UvA	1.0.1	M12	R	CO
D1.0.4	Intermediate report on coherency	UvA	1.0.3	M18	R	CO
D1.0.5	Semi-annual managerial report for WP1	UvA	1.0.1	M18	R	CO
D1.0.6	Extended yearly managerial report for WP1	UvA	1.0.1	M24	R	CO
D1.0.7	Semi-annual managerial report for WP1	UvA	1.0.1	M30	R	CO
D1.0.8	Final report on coherency	UvA	1.0.3	M36	R	CO
D1.0.9	Final managerial report for WP1	UvA	1.0.1	M36	R	CO

[8] Abbreviations used in tables are explained in Appendix B.

Task 1.1 Interactive simulation and visualisation of a biomedical system (Month 1 – 36)

Task leader: G. Dick van Albada, UvA (CR5)

Objectives

In Task 1.1 we propose to develop a Grid-based prototype system for pre-treatment planning in vascular interventional and surgical procedures through real-time interactive simulation of vascular structure and flow. The system will consist of a distributed real-time simulation environment, in which a user interacts in Virtual Reality (VR). A 3D model of the arteries, derived using medical imaging techniques, will serve as input to a real-time simulation environment for blood flow calculations. The results will be presented in a specially designed virtual reality environment. The user will also be allowed to change the structure of the arteries, thus mimicking an interventional or surgical procedure. The effects of this adaptation will be analysed in real time and the results will be presented to the user in the virtual environment (Fig. 2). The work in this task is embedded in the research on medical applications at the UvA and will be performed in close collaboration with the Leiden University Medical Centre (LUMC).

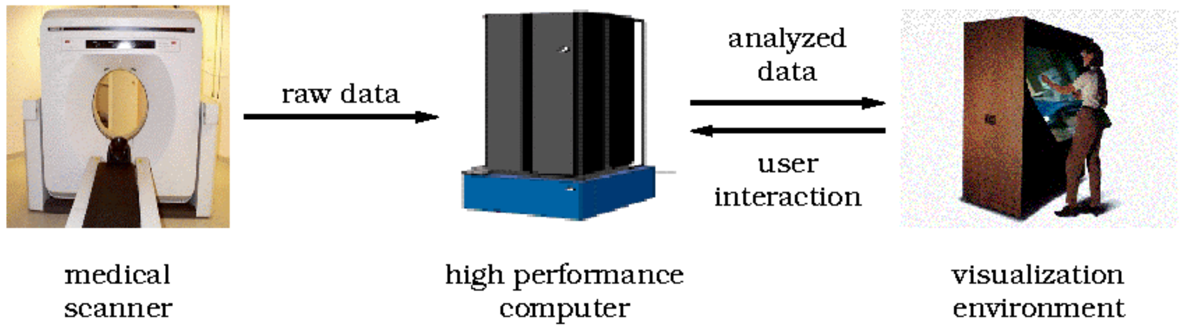


Fig. 2. Experimental set-up of interactive bio-medical simulation and visualisation.

An important aspect of this work will also be the support for distributed near real-time interaction, where data, simulation, and visualisation all are at different locations. Partitioning and scheduling of the scientific application may dynamically change in the course of time. Simultaneous visualisation clients may be activated anywhere on the Grid, with the possibility of multiple clients operating on different kinds of output devices. Therefore, the connection between the server application and the visualisation client faces several problems, such as a limited and varying bandwidth, the possibility of communication failures, and the dynamics of the Grid environment itself. These core aspects of the visualisation system will primarily be investigated at the Univ. Linz and included in the Grid Visualisation Kernel (GVK).

In designing a grid-enabled visualisation system, aside from the heterogeneity, there are two other issues that must be addressed. These are throughput and latency. Throughput is important in generating a good quality display of time-evolving data as generated by the simulation. Latency is important when interaction and feedback are involved.

The throughput as measured in bytes per second is ultimately limited by the network capacity. The problem is to make the most effective use of this capacity. Issues here are data selection (what data is needed when), display rate, location of the (rendering) components of the system and data compression. Data selection is influenced not only by the generated data set (unchanged data need not be retransmitted), but also by the user focus (parts that are not currently being displayed need not be transmitted at the same level of detail as those that are). Especially in interactive systems, the user focus can often be known rather accurately, even as a subset of the displayed data.

When the system is interactive, the user will also experience latency effects between his actions and the response of the system. As long as the interaction primarily is concerned with exploration of the data, much of it can be performed locally on the data already present (like rotating the image). Changes in user focus may result in a brief delay, while the selection of the transmitted data is being adjusted, but the effects could be kept rather small. Maximising the effective throughput, and minimising these latency effects are important issues both for the UvA and for Univ. Linz (GVK).

When the user interacts with the simulator, latency effects become important in two ways: firstly, because the simulator will have proceeded while the image is being displayed, and the user specifies his actions, secondly, because the simulator and visualisation system will need some time to show the effects of the interaction to the user. The delay of the user response to the simulator may make it necessary for the simulator to roll back to a previous state (viz. the state that the user interacted with). The delay in the simulator's response to the user will make it necessary to provide suitable intermediate information to the user. In both instances the additional delays caused by the use of a Grid based system will usually not be the dominant effect. Coping with these effects is a central issue in the research at the UvA. Their experience with optimistic discrete event systems and checkpointing will be put to use here. HLA was chosen as middleware as it was designed for distributed interactive simulation (viz. battlefield simulations) and includes provisions for virtual time management.

As described, the challenges in this task will be: distribution of source data, distributed simulation and visualisation tasks, near real-time response, virtual time management, simulation/visualisation rollback due to user actions, and VR visualisation.

In the area of performance management, use will be made of the results of the Dynamite project [Iskra2001] and of the performance tools developed in WP2. To support the distributed interactive simulations, use will be made of the concepts and techniques provided by the High Level Architecture (HLA) [HLA1999].

The parallel simulation code will internally use MPI.

The system will build upon current work at the UvA on interactive visualisation and simulation and on lattice-Boltzmann simulations of vascular blood flow [Belleman2000]. This work is performed in the context of various national and international projects, including ICES/KIS.

Subtasks:

1.1.1 Simulation kernel (UvA, M 1-36). In this subtask a lattice-Boltzmann simulation kernel will be developed that supports interactive work. This implies the addition of state-save and roll-back capabilities, as well as on-the-fly modifications to the computational grid. Performance enhancements for the required soft real-time behaviour will include improved load balancing and load redistribution. The tools developed in WP 2 will provide important support in this area.

1.1.2 VR environments (UvA, M 1-36). In this subtask specific presentation and interaction capabilities for this application will be developed, for use in a wide range of VR systems, including CAVE, Immersa-desk and PC-based systems.

1.1.3 Grid Visualisation Kernel (Univ. Linz, M 1-36). The Grid Visualisation Kernel (GVK) developed within this task will address the problems of distributed visualisation on heterogeneous devices. It will allow one to easily and efficiently interconnect Grid applications with existing visualisation tools (such as AVS, OpenDX, or VTK) by providing a transparent interface. The GVK will handle multiple concurrent input data streams from arbitrary scientific applications in the Grid, and multiplex compressed data and images efficiently across long-distance networks with limited bandwidth to any desired graphical toolkit and output device, such as standard 2-D displays or sophisticated Virtual Reality devices.

1.1.4 Integration and deployment (UvA, Univ. Linz and Cyfronet, M 6-36). In this subtask a testbed will be realised and the various components of the system will be integrated and deployed on it. This involves the coupling of the simulation to the VR environment, integrated time-management for simulation and visualisation, management of simulation checkpoints, management of selected simulation results, management and configuration of grid resources, and integration testing obtaining medically relevant examples. The integrated application will be tested; the test results will be made available to the other subtasks and the technology workpackages.

1.1.5 Dissemination (UvA and Univ. Linz, M 1 - 36).

Deliverables

In order to limit the total number of deliverables, all deliverables will be joint deliverables, containing chapters contributed by the various subtasks. The partner leading a subtask will be responsible for each chapter.

Number	Description	Responsible partner	Subtasks	Delivery date	Nature	Dissemination level
D1.1.1	Application description, including use cases for Task 1.1	UvA	1.1.1, 1.1.2, 1.1.3	M3	Report	PU
D1.1.2a D1.1.2b	Internal progress report First internal software release	UvA	all	M12	Report SW	CO
D1.1.3	Report on application prototype 0 run	UvA	1.1.4	M18	Report	CO
D1.1.4a D1.1.4b	Internal progress report Second internal SW release	UvA	all	M24	Report SW	CO
D1.1.5	Report on application prototype 1 run	UvA	1.1.1 – 1.1.4	M30	Report	CO
D1.1.6	Application final demonstration and report	UvA	All	M36	D, R	PU

Task 1.2 Flooding crisis team support (Month 1 – 36)

Task leader: Ladislav Hluchy, II SAS (AC 6)

Objectives

In Task 1.2 we will develop a support system for establishment of a Virtual Organisation (VO) associating a set of individuals and institutions involved in flood prevention and protection, and integrating the most advanced techniques in data sampling, data processing, data transfer, data management and Grid technology (Fig. 3). Grid technology will enable co-operation between flood crisis teams on international rivers.

The kernel of this task is numerical flood modelling that uses an appropriate physical model and robust numerical schemes for a good representation of reality. Realistic simulations of large problem sizes are computationally challenging but need to be performed on short time-scales in crisis situations [Hluchy2001a, Hluchy2001b, Hluchy2001c].

The system to be developed in this task will employ a Grid technology to seamlessly connect together the experts, data and computing resources needed for quick and correct flood management decisions. The final aim is to be able to work in real-time. This will also require the timely transfer of individual new precipitation events (e.g. evaluated from the meteorological radar monitoring).

The need to access widely distributed data and computational resources makes the use of Grid technology essential for this application. The functionality of the system is based on the hydro-meteorological (snowmelt) rainfall-runoff modelling of a relatively extended orographic area. In such an area the hydrological modelling of a flood response on precipitation is probably the only solution.

The challenges in this task will be: acquisition of significant resources at short notice, near real-time response, the combination of distributed data management and distributed computing, computational requirements for the combination of hydrological (snowmelt- rainfall-runoff) and hydraulic (water surface elevation, velocity, dam breaking, damage assessment, etc.) models, and eventually, mobile access under adverse conditions. 2D visualisation of the simulation results is currently possible, but needs adaptation to the grid environment; the possibilities for 3D visualisation will be studied in collaboration with Task 1.1.

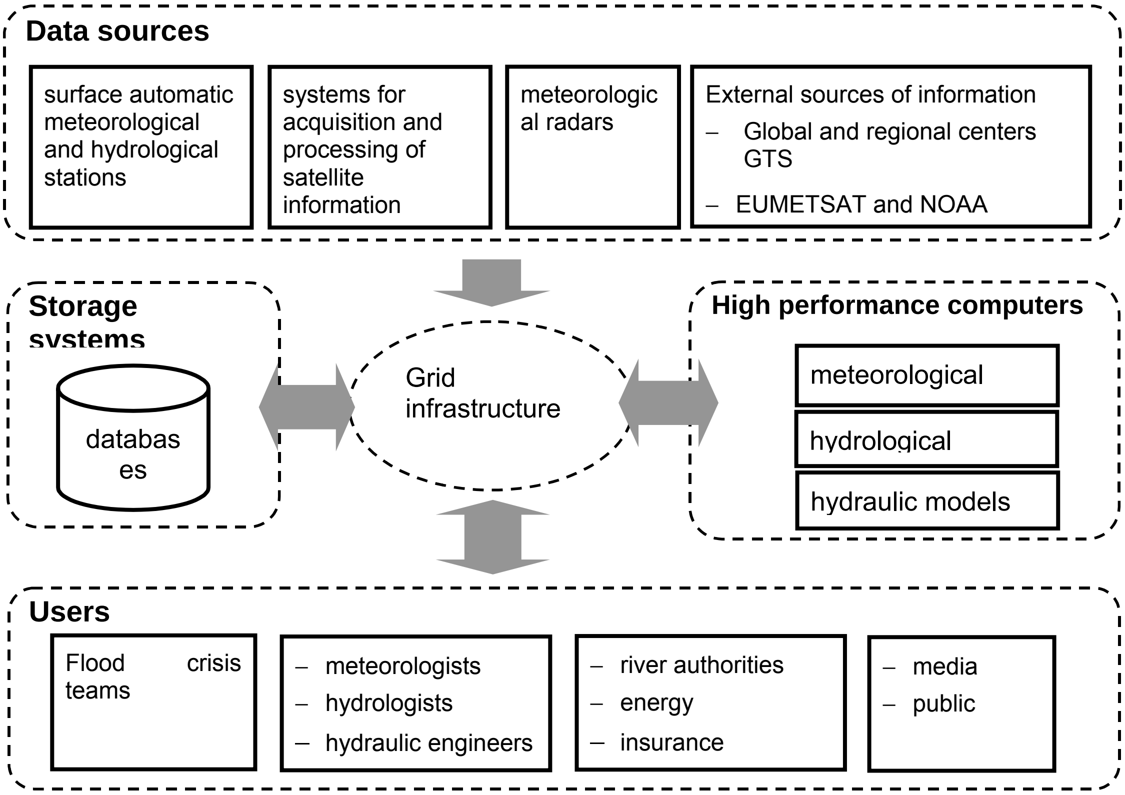


Fig. 3. Support system for Virtual Organisation for Flood Prevention and Protection.

In flood modelling we will adapt advanced meteorological (ALLADIN/LACE, MM5), hydrological (HEC-1, TR-20, HSPF, NLC, HBV) and hydraulic (FESWMS, RMA2, NLN) models to Grid environment. Models with unsatisfactory response time will be optimised for HPCN systems using tools developed in WP2. For allocation of resources for time critical parts of simulations and their efficient use, as well as for the required roaming access, the tools from WP3 will be needed.

This task will build upon the results of the IST project (2000-2002) Data Fusion for Flood Analysis and Decision Support – ANFAS - in which II SAS participates. In this project our focus is oriented on preparing data for real scenarios for hydraulic simulation. It is appropriate to use these results, that are now available, for the planed interactive near real time flood simulation in the CrossGrid project. Meanwhile, II SAS has started collaboration with the Slovak Hydrometeorological Institute regarding the more effective use of hydrological models.

Subtasks

1.2.1 Distributed data collection. In this subtask the available and necessary sources of data will be further identified and techniques will be developed to allow a timely and distributed access to these necessarily distributed data. (II SAS, M 1-36)

1.2.2 Distributed simulation and data analysis. In this subtask hydro meteorological and hydraulic models will be developed and implemented that can make use of the computational resources available through the grid, and that will be able to work with possibly incomplete data. The performance of these models will be an important issue. (II SAS, M 1-36)

1.2.3 Distributed access, support for the virtual organisation. In this subtask the problems related to the distributed and roaming access to the data and models will be addressed. (II SAS, M 1-36)

1.2.4 System integration, testing and demonstration. In this subtask, the technologies developed in the preceding subtasks will be implemented on the testbed (Task 4.1). The performance of the integrated system will be tested and the results be coupled back to subtasks 1.2.1 to 1.2.3 and to the technology workpackages.(II SAS, M 6 - 36)

1.2.5 Dissemination and industrial deployment. Dissemination will be performed and industrial deployment promoted as described in the section on dissemination and exploitation. (II SAS, M 1 - 36)

Deliverables:

In order to limit the total number of deliverables, all deliverables will be joint deliverables, containing chapters contributed by the various subtasks.

Number	Description	Responsible partner	Subtasks	Delivery date	Nature	Dissemination level
D1.2.1	Application description, including use cases for Task 1.2	II SAS	1.2.1, 1.2.2, 1.2.3	M3	Report	PU
D1.2.2a D1.2.2b	Internal progress report First internal software release	II SAS	all	M12	Report SW	CO
D1.2.3	Report on application prototype 0 run	II SAS	1.2.4	M18	Report	CO
D1.2.4a D1.2.4b	Internal progress report Second internal SW release	II SAS	all	M24	Report SW	CO
D1.2.5	Report on application prototype 1 run	II SAS	1.2.1 -1.2.4	M30	Report	CO
D1.2.6	Application final demonstration and report	II SAS	1.2.1 -1.2.4	M36	Report Demo	PU
D1.2.7	Final report on dissemination and exploitation	II SAS	1.2.5	M36	Report	CO

Task 1.3 Distributed Data Analysis in High Energy Physics (Month 1 – 36)

Task leader: C. Martinez-Rivero, CSIC (CR15)

Objectives

In Task 1.3 we aim to develop final user applications for physics analysis running in a distributed mode in a GRID-aware environment using large distributed databases. They will be used in the high-energy physics field, the main focus being future LHC experiments (ALICE, ATLAS, CMS and LHCb).

Next generation experiments in HEP will require unprecedented computing resources for physics analysis [LHCC2001]. Hundreds of physicists around the world will collaborate in the analysis of Peta bytes of data stored in distributed databases. Most of the experiments are adopting a model for distribution of these databases, such as the MONARC model proposed for LHC computing, proposing tiered data and computing centres. The DataGrid project will provide a file-level service, based on replication, for access to these databases, including transparent user access that accommodates distributed simulation and reconstruction requirements. For the final interactive user analysis, however, an object-level service is also required, to optimise the use of the resources.

This can be implemented in a classical three-tier model, accessing either a pure OODBMS (like Objectivity) or an O/R DBMS (commercial like Oracle or IBM DMS or an open source product like mySQL), via a middleware server employing XML as a format for query and data retrieval. Security can be achieved through the use of http over SSL/TLS. It can be also implemented on a more specific full HEP solution, like ROOT, offering file service replication (ROOTD). Both approaches will be tried for LHC experiments, and tested in production with running experiments (such as BABAR or CDF). This task will nicely complement those being developed in the DataGrid WP2.

The task will address several challenging points: access to large distributed databases in the Grid environment, development of distributed data-mining techniques suited to the HEP field, definition of a layered application structure, flexible enough to adapt to different experimental set-ups, and integration of user-friendly interactive access, including specific portal tools.

These strategies will be developed as open-source code, platform independent, with specific portal tools developed in such a way that all users can profit, including those with HEP expertise who are not computing experts, as in the case of previous portals developed within large HEP collaborations [IES2000]. This middleware will be based on existing physics analysis packages like ROOT or JAS, reusing their components as applets. Other Grid specific portal tools will be incorporated from WP3.

Subtasks:

1.3.1 Interactive Distributed Data Access (CSIC, M 1-36). In this subtask the best strategies for interactive use of distributed data access in the GRID framework will be developed. The use of replication strategies in O/R DBMS for HEP data at the DST level, characterised by its write-once/read-many nature from the server point of view, will be studied. The client side requires a light-weight interface based on XML both for query and data-transfer, but with local persistence capabilities, and should accept integration of complex distributed data-mining techniques.

The first prototype is likely to be based on Oracle 9i or IBM DB2, accessed in a three-tier scheme, where data partitioning implemented through replication will assure load balancing.

This subtask will benefit from WP2, in particular its optimisation will need the performance and monitoring tools developed in it.

1.3.2 Data Mining Techniques on GRID (CSIC, M 1-36). Data-mining techniques in the HEP field are driven by the use of neural networks, both for classification and parameterisation. The classical interactive analysis based on sequential cuts that can be translated into simple filtering is being replaced by the above techniques, which are much more time consuming and therefore preclude interactive use. The distributed implementation of data-mining can be done either in an agent-based way where the user job is performed at each database server, returning an index collection for later use, or in a parallelised way where the task is distributed between several GRID machines, querying on the different database partitions, running on the corresponding query output, and integrating back the result to the user. Which strategy is best depends on the physics analysis problem, from the low level filtering to sophisticated neural networks.

The tools developed in WP2, in particular those related to MPI or similar parallelisation tools for applications, will be an essential ingredient to this subtask, together with the basic Grid middleware. The improvement in response time for the user thanks to the use of the amassed GRID power, should allow reasonably interactive use, ranging from seconds to minutes depending upon the complexity of the analysis, and so optimising the manpower used for this purpose in large collaborations.

The results of this subtask will be applied in Task 1.4, to trigger a similar work on data mining algorithms specific for environmental applications that so will start not from scratch but from the basic solutions found in this subtask. Close collaboration is assured thanks to the presence of a common partner (CSIC).

1.3.3 Integration and Deployment (FZK with CSIC, M 6-36)

Final user applications should run on simulated data from the LHC collaborations (ATLAS, CMS, ALICE, LHCb). The integration with the corresponding experiment framework, including a common user interface through the use of Portal tools addressed in WP3, will benefit from the use of a XML Schema for data modelling. Deployment will take place at the testbed sites in FZK and CSIC, and then will be extended to the complete testbed.

1.3.4 Application to LHC Physics TDR and High Level Trigger (INP with INS, CSIC, UAB, M 1-36)

Final user applications in HEP run on a filtered set of events; the LHC collider will operate at an extremely high rate of almost 10^9 proton-proton collisions per second. Most of these correspond to well-known processes and are, therefore, of low interest; only a small number of these interactions will be induced by new phenomena. A fundamental discovery is possible only if these rare processes are efficiently extracted from this huge background of ordinary interactions. The allowed event rate written to tape is about 100 Hz so that a reduction of 10^7 has to be done by the multilevel trigger system of each experiment. The low-level triggers are based on fast hardware solutions. Accepted events are processed subsequently by the high level triggers (HLT) where sophisticated algorithms have to be used in order to attain high efficiency for events showing the desired features. The quality of these algorithms has to be almost as good as the quality of the off-line analysis chain. It implies that a kind of simplified event reconstruction has to be applied at the trigger level before writing the event onto tape. The time to make the decision at the last trigger levels is of the order of a few hundredths of milliseconds. Optimisation of these sophisticated algorithms is a mandatory step prior to the study of the final physics expectations from LHC experiments employing simulated data, the objective of the Physics Technical Design Report that should be completed by 2004. These results will require the input of DataGrid simulated and processed MC samples, while the corresponding final user analysis is a real challenge for our CrossGrid project. The results of this challenge will be the main dissemination topic inside the HEP community. This task will also take care of the elaboration of well-defined short and longer-term use cases.

1.3.5 Dissemination of the results of the HEP application.

Deliverables

In order to limit the total number of deliverables, all deliverables will be joint deliverables, containing chapters contributed by the various subtasks. The partner leading a subtask will be responsible for each chapter.

Number	Description	Responsible partner	Subtasks	Delivery date	Nature	Dissemination level
D1.3.1	Application description, including use cases for Task 1.3	CSIC	1.3.1, 1.3.2, 1.3.4	M3	Report	PU
D1.3.2a D1.3.2b	Internal progress report First internal software release	CSIC	All	M12	Report SW	CO
D1.3.3	Report on application prototype 0 run	CSIC	1.3.3	M18	Report	CO
D1.3.4a D1.3.4b	Internal progress report Second internal SW release	CSIC	All	M24	Report SW	CO
D1.3.5	Report on application prototype 1 run	CSIC	1.3.1 - 1.3.4	M30	Report	CO
D1.3.6	Application final demonstration and report	CSIC	All	M36	D, R	PU

Task 1.4 Weather forecast and air pollution modelling (Month 1 – 36)

Task leader: Bogumil Jakubiak, ICM (AC2)

Objectives

The objectives of this task are to provide a representative collection of sample applications of Grid tools, for use by the atmospheric/oceanographic community. Atmospheric applications require large computational power and fast access to large data sets stored in geographically distributed archives. We plan to develop a data mining system for the analysis of the archive of operational data from a mesoscale model (in volume 32 GB) and meteorological reanalysis DBs which includes homogeneous meteorological information from a single numerical weather prediction model integrated over decades (in the order of Tera-bytes). Data mining techniques, including association rules, linear and non-linear correlation methods, and neural nets (Self-Organised Maps SOM), will be developed for extracting interesting patterns (clusters) or trends (spatial and temporal correlations and tele-connections) within the DB.

The interactive use and the scalability of Grid technology, in order to meet atmospheric research and application user community requirements, will be investigated. Specification of requirements typical for Atmospheric Research, their development and their implementation, will be the core of the activity. These components will be validated through prototyping activity and consolidated by running specific testbeds. A complete application proposed by ICM involves the implementation and development of Grid tools enabling distant, co-ordinated feedback between atmospheric models of different resolution and wave models based on local coastal data and forced by wind fields generated by atmospheric components of the system. Atmospheric pollution chemistry studies, conducted by U.S.C., will be also included.

Easy and fast access to actual and reliable atmospheric and marine forecasts, covering basic parameters of the atmosphere and sea state, is very important for the safety of human activities in the sea, as navigation, harbour operations, fisheries, combating hazards of the marine environment, dredging activities, realisation of hydro-technical investments etc. The most important forecast parameters are visibility, wind, waves and surface currents. For effectiveness of activities in some special cases such as rescue or combating oil spills, it may also be necessary to know sea water temperature, air temperature and water salinity. Lack of such forecast when immediate decision-concerning navigation is needed, during extreme and highly variable hydro-meteorological situations, may cause a large threat to the life of people in the sea and floating equipment losses. For such purposes a system of dissemination of meteorological and oceanographic information have been developed. The main partners of this system are ICM, Warsaw University, producing atmospheric forecasts, Institute of Oceanography, University in Gdansk, producing oceanographic forecasts forced by atmospheric models, Maritime Institute in Gdansk and Institute of Hydro-Engineering, Polish Academy of Sciences both producing wave forecasts driven by winds simulated by atmospheric model. Models used in a system are at the European level. The atmosphere is modelled by the mesoscale version of the UKMO Unified Model [Niezgodka1998], the ocean is modelled by the POM (Princeton Ocean Model [Mellor1998]) adopted to Baltic Sea and Gulf of Gdansk conditions; for waves prediction a third generation of WAM model is used [WAMDIGroup1988, Kalas1998, Paplinska1999]. The weakest part of the actual system is the one-way interaction between the data producers and data users - the ocean and wave models are able only to read data produced by atmospheric models according a constant schedule. Our proposal is concentrated in an application of the modern GRID tools to enhance the operational possibilities of such system.

Air pollution chemistry models need to solve tens to hundreds of gas-phase chemical reactions, coupled to the air pollution transport. Gas-phase chemistry is a heavy task for personal computers and aqueous chemistry and scavenging processes add more complexity to the problem, so parallel systems should be applied for a reasonable response time. We propose to consider the Sulphur Transport Eulerian Model 2 (STEM-II [Carmichael, 1986]) as a first approach to develop a GRID based code, although we can focus on other models. The STEM-II is an air quality model used for simulating the behaviour of pollutants in the air. It includes different pollutant emissions, transport by advection, convection and turbulence, with dynamic meteorological conditions; chemical transformation by gas-phase and aqueous-phase, and pollutants removal by dry and wet deposition.

This model is currently used together with meteorological prediction programs to simulate local and regional acid deposition from the As Pontes Power Plant (A Coruña, Spain) environment. This power plant is placed in a EMEP 17,6 cell (150x150 km^2) considering 15 height levels from ground level and reaching 4200m. It considers 56 chemical species, 16 long-lived species, 40 short-lived species, and 28 radicals, such as OH and HO_2. The chemical mechanism includes 176 gas-phase reactions, 31 aqueous-phase reactions and 12 aqueous-phase solution equilibria. Model equations are integrated by using the locally one-dimensional finite element method (LOD-FEM) [Carmichael, 1986], with the resulting transport equations solved by the Petrov-Crank-Nicolson-Galerkin FEM [Berge1997]. Chemistry and mass transfer terms are integrated by using a semi-implicit Euler's [Preussner1981] and the pseudo-analytic method [Carmichael, 1986]. Therefore, several integration time steps are used in the model, depending on the different physical and chemical processes solved, and the atmospheric conditions.

In our previous work [Singh2000, Mouriño2001a], several studies on this model were developed. We have carried out an analysis of the computational costs of each module of the code and we concluded that the chemical-simulation module is the most costly part of the program. With the aim of reducing the execution time, a set of sequential optimisations has been applied mainly to the chemical phase obtaining a 13% improvement in the execution time of the sequential program.

Next, we focused on the analysis of spatial and temporal locality of memory references as well as on data dependencies [Singh2000] on the most costly parts of the program. Once this analysis was completed we were able to parallelise them. Initially we considered two main approaches in order to execute the STEM-II code on a high performance computer: vector processors and scalar multiprocessors. For vector systems, the performance was low due to data dependencies in the loop body as well as several calls to external functions. So, we concluded that STEM-II is not suitable for being executed on a vector processor. In further studies [Mouriño2001a], we demonstrated that the structure of the STEM-II code is very appropriate for an execution on a scalar multiprocessor.

We have developed an initial parallel implementation of the selected codes using the message passing paradigm. Specifically we have used the MPI (Message Passing Interface) library [Gropp1994]. Different tests were executed on the Fujitsu AP3000 system [Mouriño2001a] obtaining encouraging results. The achieved speed-ups are due to both the parallelisation itself and the data locality (each processor accesses a smaller data volume, taking better advantage of the memory hierarchy).

We propose to further pursue the parallel implementation of this model, adapting it to the heterogeneous GRID infrastructure, paying special attention to the communication costs and load balancing. Moreover, we propose to include more chemical mechanisms into the process to improve the results.

The proposal consists of adapting the SPMD parallel versions of this application (photochemical grid model), to the GRID. This work has to cope with heterogeneous components and communications, paying special attention to good dynamic load balances [Mouriño2001b]. We will use the MPI library to minimise the communication overheads, for which we will depend on the MPI tools developed within WP 2.

For the visualisation of the results, meteorologically oriented visualisation tools (NCAR Graphics, GRADS, VIS5D) are available in the public domain. In order to realise Grid-support for these libraries, use will be made of the GVK developed in Task 1.1.

The development of the required distributed data-base access and data mining technology will done in close collaboration with Task 1.3.

The challenges in this task will be:

1. Integration of distributed databases into GRID structure (ICM, CSIC, and U.S.C.).
2. Meet the data mining and network requirements for atmospheric research (CSIC).
3. Migration of data mining algorithms to GRID (ICM and CSIC).
4. Integration of distributed atmospheric and wave models into the Grid structure (ICM).
5. Integration of the parallel codes for air quality models into GRID (U.S.C.).
6. Validation and improvement of performance of the codes.

Subtasks

1.4.1. Integration of distributed databases into GRID structure (ICM, CSIC with U.S.C.,M 1- 36)

1.4.2. Migration of data mining algorithms to GRID (CSIC with ICM, M 1 - 36)

1.4.3. Integration of distributed atm. & wave models into the GRID structure (ICM, M 1 - 36)

1.4.4. Integration of parallel codes for air quality models into GRID structure (U.S.C., M 1 - 36)

1.4.5. Integration, testing and demonstration of the application in the testbed environment. (PSNC with ICM, CSIC, U.S.C., M 6 - 36)

1.4.6. Dissemination. (ICM, CSIC, U.S.C., M 1 - 36).

Deliverables

Number	Description	Responsible partner	Subtasks	Delivery date	Nature	Dissemination level
D1.4.1	Application description including use cases for case 1.4	ICM CSIC U.S.C	1.4.1 – 1.4.4	M3	Report	PU
D1.4.2	Internal progress report on migration of data mining algorithms to GRID structure	ICM CSIC	1.4.2	M12	Report	CO
D1.4.3	Integration of parallel codes for air quality models into GRID structure Prototype A phase 0	U.S.C.	1.4.4 1.4.5	M18	Report + Prototype	CO
D1.4.4	Integration of distributed atmosphere and wave models into GRID structure). Prototype B phase	ICM	1.4.3 1.4.5	M24	Report + Prototype	CO
D1.4.5	Prototype A running (phase 1) (integration of parallel codes for air quality models into GRID structure).	U.S.C	1.4.4 1.4.5	M30	Report + Prototype	CO
D1.4.6	Prototype B running (phase 1) (integration of distributed atmosphere and wave models into GRID structure).	ICM	1.4.1 – 1.4.5	M30	Report + Prototype	CO
D1.4.7	Application final report	ICM CSIC U.S.C	All	M36	Demo, Report	PU

Workflow and interfaces to other WP

M1-3: Definition phase

For all partners: Refine the existing application descriptions. Define use cases. Define common infrastructure (hardware, software) requirements for applications; describe where specific applications need a non-standard infrastructure. This work will result in a deliverable providing a detailed description plus use cases for each application, and a common deliverable specifying the characteristics of the software and hardware components required for the final application in sufficient detail. In the definition phase the various tasks will closely collaborate with the CrossGrid Architecture team and the initial tasks in WP 2, 3, and 4. Provides output to WP 2, 3 and 4 in month 3.

Number	Description	Responsible partner	Subtasks	Delivery date	Nature
D1.0.1	Joint requirements definition document	UvA	1.0.2	M3	Report
D1.1.1	Application description, including use cases for Task 1.1	UvA	1.1.1, 1.1.2, 1.1.3	M3	Report
D1.2.1	Application description, including use cases for Task 1.2	II SAS	1.2.1, 1.2.2, 1.2.3	M3	Report
D1.3.1	Application description, including use cases for Task 1.3	CSIC	1.3.1, 1.3.2, 1.3.4	M3	Report
D1.4.1	Application description, including use cases for Task 1.4	ICM; CSIC; U.S.C	1.4.1 –; 1.4.4	M3	Report

M4-12: Initial implementation and development

Create initial Grid-enabled implementations based on existing Grid software. Deploy on initial testbeds. These versions of the applications will not yet depend on the code developed in WP2 and WP3, but where necessary, will be prepared for integration with this software. Requires input from WP 4 (M 6). Provides to other WP: Initial implementations of all applications, ready for integration and testing (M12).

Number	Description	Responsible partner	Subtasks	Delivery date	Nature
D1.0.2	Semi-annual managerial report for WP1	UvA	1.0.1	6	Report
D1.0.3	Extended yearly managerial report for WP1	UvA	1.0.1	M12	Report
D1.1.2a	Internal progress report	UvA	1.1.1 - 1.1.4	M12	Report
D1.1.2b	First internal software release	UvA	1.1.1 - 1.1.4	M12	SW

The software delivered in D1.1.2b will be concerned with the modifications required in the visualisation and simulation modules required for later integration. For the simulator the primary concerns will be the mechanisms needed for user feedback and simulation rollback. For the visualisation this will be the implementation of user interaction and data transfer. The level of integration in the software in D1.1.2b will be limited to non-interactive, on-the-fly display of simulation data. The software delivered in D1.2.2b will be grid-aware prototypes of individual applications. The software delivered in D1.3.2b will be grid-aware prototypes of individual applications.

Number	Description	Responsible partner	Subtasks	Delivery date	Nature
D1.2.2a	Internal progress report	II SAS	1.2.1 - 1.2.4	M12	Report
D1.2.2b	First internal software release	II SAS	1.2.1 - 1.2.4	M12	SW
D1.3.2a	Internal progress report	CSIC	1.3.1, 1.3.4	M12	Report
D1.3.2b	First internal software release	CSIC	1.3.1, 1.3.4	M12	SW
D1.4.2	Results of migration of data mining algorithms to GRID structure	ICM; CSIC	1.4.2	M12	Report

M13-18: First testing and refinement phase.

During this phase the performance of the applications will be tested in the first version of the CrossGrid environment. Especially during the initial testing period intensive support will be required from developers in other WP.

Requires: prompt support from WP 2 - 4 developers

Produces: Test and bug reports to other WP (regularly). First refined version of application SW utilising SW from other CrossGrid WP. This period will be closed with the production of reports describing the "prototype 0" performance for all applications. Development on the application will continue during this phase, but results need not fully integrated immediately with the prototype.

Month 18

Number	Description	Responsible partner	Subtasks	Delivery date	Nature
D1.0.4	Intermediate report on coherency	UvA	1.0.3	M18	Report
D1.0.5	Semi-annual managerial report for WP1	UvA	1.0.1	M18	Report
D1.1.3	Report on application prototype 0 run	UvA	1.1.4	M18	Report
D1.2.3	Report on application prototype 0 run	II SAS	1.2.4	M18	Report
D1.3.3	Report on application prototype 0 run	CSIC	1.3.3	M18	Report
D1.4.3	Integration of parallel codes for air quality models into GRID structure Prototype A phase 0	U.S.C.	1.4.4; 1.4.5	M18	Report + Proto-type

M19-24: Second development phase.

Application functionality will be extended and an initial integration of application components will be realised. Tasks will be deployed on all suitable testbeds, but not all testbeds may be suitable for all applications. Applications will approach, but will not yet realise full functionality.

Requires: prompt updates on design documents and prototypes from other WP.

Provides: M 24: Applications ready for integration and testing on all applicable testbeds.

Month 24

Number	Description	Responsible partner	Subtasks	Delivery date	Nature
D1.0.6	Extended yearly managerial report for WP1	UvA	1.0.1	M24	Report
D1.1.4a	Internal progress report	UvA	1.1.1 - 1.1.4	M24	Report
D1.1.4b	Second internal SW release	UvA	1.1.1 - 1.1.4	M24	SW
D1.2.4a	Internal progress report	II SAS	1.2.1 - 1.2.4	M24	Report
D1.2.4b	Second internal SW release	II SAS	1.2.1 - 1.2.4	M24	SW
D1.3.4a	Internal progress report	CSIC	1.3.1 - 1.3.4	M24	Report
D1.3.4b	Second internal SW release	CSIC	1.3.1 - 1.3.4	M24	SW
D1.4.4	Integration of distributed atmosphere and wave models into GRID structure. Prototype B phase	ICM	1.4.3; 1.4.5	M24	Report + Proto-type

M25-29: Second testing and refinement phase

During this phase the performance of the applications will be tested in the second version of the CrossGrid environment. Especially during the initial testing period intensive support will be required from developers in other WP.

Requires: prompt support from WP 2 - 4 developers

Produces: Test and bug reports to other WP (regularly). Second refined version of application SW utilising SW from other CrossGrid WP. This period will be closed with the production of reports describing the "prototype 1" performance for all applications. Development on the application will continue during this phase, but results need not but fully integrated immediately with the prototype.

Month 30

Number	Description	Responsible partner	Subtasks	Delivery date	Nature
D1.0.7	Semi-annual managerial report for WP1	UvA	1.0.1	M30	Report
D1.1.5	Report on application prototype 1 run	UvA	1.1.4	M30	Report
D1.2.5	Report on application prototype 1 run	II SAS	1.2.4	M30	Report
D1.3.5	Report on application prototype 1 run	CSIC	1.3.3	M30	Report
D1.4.5	Prototype A runs (phase 1) (integration of parallel codes for air quality models into GRID structure).	U.S.C	1.4.4; 1.4.5	M30	Report + Proto-type
D1.4.6	Prototype B runs (phase 1) (integration of distributed atmosphere and wave models into GRID structure).	ICM	1.4.3; 1.4.5	M30	Report + Proto-type

M30 - 33 Third development phase.

During this phase the application demonstrators will be developed and integrated with the final versions of the CrossGrid infrastucture SW. Final integrated versions of the applications will be realised and deployed on applicable testbeds.

M34 - 36 Finishing phase

In this phase demonstrators will be prepared for the various applications and final reports will be written.

Month 36

Number	Description	Responsible partner	Subtasks	Delivery date	Nature
D1.0.8	Final report on coherency	UvA	1.0.3	M36	Report
D1.0.9	Final managerial report for WP1	UvA	1.0.1	M36	Report
D1.1.6	Application final demonstration and report	UvA	1.1.4, 1.1.5	M36	Report, Demo
D1.2.6	Application final demonstration and report	II SAS	1.2.4	M36	Report, Demo
D1.2.7	Final report on dissemination and exploitation	II SAS	1.2.5	M36	Report
D1.3.6	Application final demonstration and report	CSIC	1.3.3, 1.3.4	M36	Report, Demo
D1.4.7	Application final demonstration and report	ICM; CSIC; U.S.C	1.4.5	M36	Report, Demo

Technology required from other WP

The application tasks developed in WP 1 all depend on very good performance of the underlying infrastructure. These requirements can only be realised in the following way:

All tasks have a strong need for the MPI tools and the various performance tools developed in WP 2. As the MPI tools will strongly interact with the SW development in the application, an early availability is necessary (M 12).

Roaming access is vital to Task 1.2 and very useful for the other tasks (WP3). Specifications (M 3) and prototypes (M 12) will be needed early.

The performance evaluation tools from Task 2.4 will be strongly needed during the integration and refinement phases (from M 12 on) for all applications.

Grid resource management and grid monitoring (WP 3) are essential in different ways to all tasks. Tasks 1.1 and 1.2 primarily address the High Performance Computing and interactive aspects of the use of the grid; Tasks 1.3 and 1.4 tend more to High Throughput Computing, though short response times remain important. The scheduling and monitoring capabilities realised in Tasks 3.2 and 3.3 are vital in this respect. The availability of prototypes is required around month 12 for the development of the appropriate interfaces. Fully functional versions are required by M 24.

Optimisation of data access (WP 3) is important primarily to Tasks 1.2, 1.3 and 1.4, where distributed, heterogeneous data are to be used.

All applications require access to one or more testbed sites.

Technology required from outside sources

Tasks 1.2, 1.3 and 1.4 will make use of the data management software developed in DataGrid.

Task 1.1 will make use of HLA, developed by the DoD.

All tasks will use suitable software components developed elsewhere to the maximum extent possible (Globus toolkit).

Technology generated

Besides the applications, a variety of high-level grid-enabled tools will be developed within WP 1. These include tools in the area of visualisation and interaction, data mining, and OODB.

Resources

Total and (funded) Person Months (PM).

Task	Task PM	CYFRONET	ICM	INP	INS	UvA	II SAS	Univ. Linz	FZK	PSNC	CSIC	UAB	U.S.C.
1.0	20 (10)		3 (1)			11 (3)	3 (3)				3 (3)		
1.1	175 (98)	12 (6)				117 (69)		46 (23)					
1.2	73 (73)						73 (73)						
1.3	142 (106)			46 (23)	26 (13)				6 (6)		49 (49)	15 (15)	
1.4	127 (78)		65 (33)							12 (6)	22 (22)		28 (17)
Total PM	537	12	68	46	26	128	76	46	6	12	74	15	28
Funded PM	365	6	34	23	13	72	76	23	6	6	74	15	17

WP2 Grid Application Programming Environment

Workpackage manager: Holger Marten, FZK (CR8)

Objectives

The aim of this workpackage is to specify, develop, integrate and test tools that facilitate the development and tuning of parallel distributed high-performance and high-throughput computing applications on Grid infrastructures.

Verifying that user applications comply with the MPI standards will reduce the need for debugging sessions on the Grid environment. This workpackage will develop a debugging and verification tool for MPI programs.

Efficiently using the Grid as an environment for large applications requires job performance measurement and analysis [Bubak, 2001]. In this workpackage, an on-line tool will be developed that analyses the performance of Grid applications and graphically presents the results. Besides a pure performance measurement module, the tool will include an automatic analysis component [Fahringer, 2000; Dikaiakos, 1998] that assesses the measured performance w.r.t. the performance capabilities of the Grid infrastructure. In addition, the tool will support performance prediction for selected application kernels based on analytical performance models.

The workpackage will develop benchmarks that model workloads typical of Grid applications of interest. These benchmarks should capture and isolate performance characteristics of the Grid, with respect to processing power, data throughput, synchronisation, communication and I/O overhead, etc.. Benchmarks will allow identification of important factors that affect application performance, and provide application developers with initial performance estimates. In addition, benchmark results are an input to the automatic analysis component in the performance analysis tool, which will help application developers to analyse the system at various granularity levels, from the Grid level down to the process level which finally allows them to tune their applications for optimal Grid performance.

The tools developed in this workpackage will interoperate [Wismueller, 2000] and be integrated into the testbed and will be promoted by and tested with the real end-user applications of WP1.

The component structure of the programming environment, as well as its relations to the other workpackages in the project is presented in Fig. 4. Solid arrows denote interactions between software components, dashed ones indicate manual or semiautomatic steps. Numbers in parenthesis identify the workpackage and task related to a component. Run-time data on an application executing on the Grid is collected by the Grid monitoring system. It is used both for MPI verification and performance analysis. The verification tool additionally uses static information derived from the applications' source code. The performance analysis tool can visualize three supplementary kinds of performance data: the measured raw data, higher-level performance information obtained by an automatic analysis, and results of performance predictions based on an analytical model. Automatic analysis takes into account the performance characteristics of the Grid as determined by benchmarks. Analytical models will be available for selected kernels extracted from the applications' source code. Using the results of verification and performance analysis, the application developer finally can correct and optimise the applications' code.

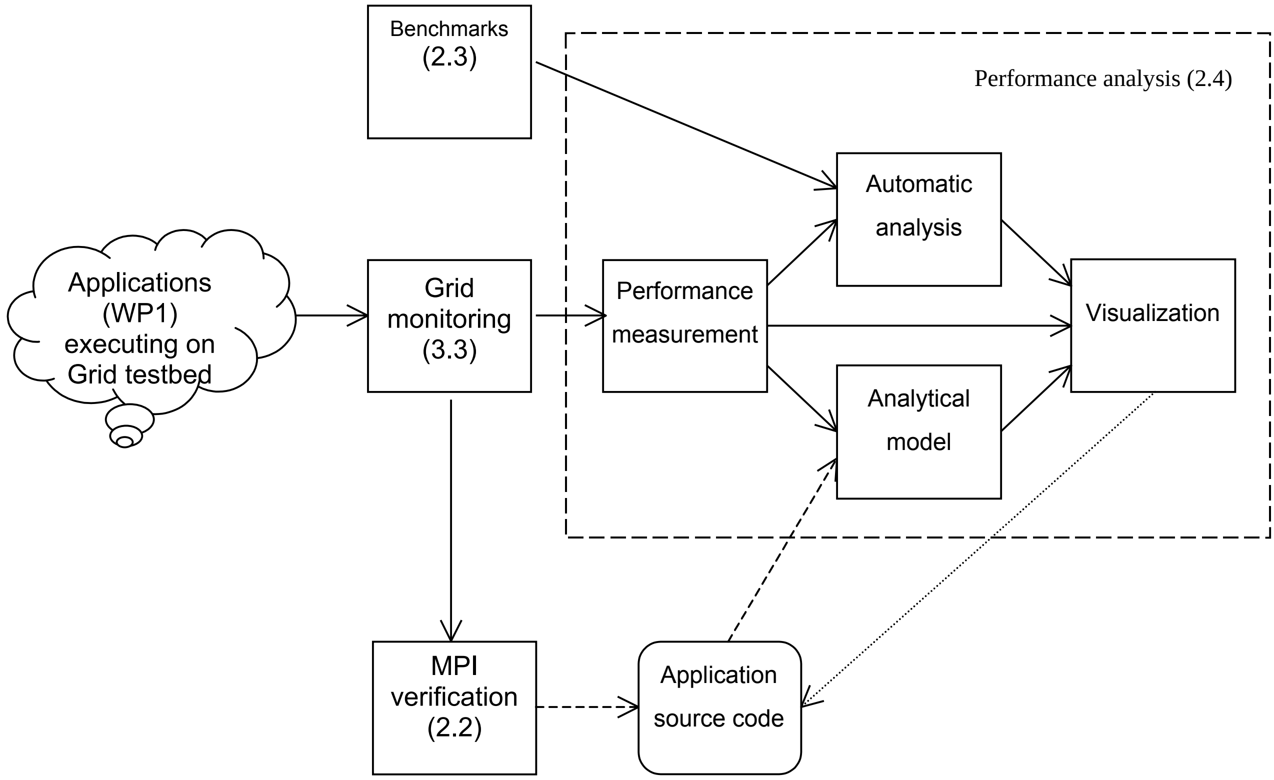


Fig. 4. Relations between tasks of WP2 and other WPs of the CrossGrid Project.

Task descriptions

Task 2.0 Co-ordination and management (Month 1-36)

Task leader: Holger Marten, FZK (CR8)

Co-ordination is required to ensure the seamless integration of the interfaces and software components, the timely flow of information and deliverables between WP partners, other WPs and the CrossGrid Architecture Team. FZK will coordinate this WP2.

Task 2.1 Tools requirements definition (Month 1-3)

Task leader: Roland Wismüller, TUM (AC10)

A full requirements analysis will be performed to evaluate the needs of all program developers in WP1 w.r.t. development tools for interactive Grid applications. A focus will be put on the support for the following activities:

- verifying the correctness of critical applications, e.g., in crisis support or medicine,
- prediction of application response time in early development phases,
- detection of performance bottlenecks in applications.

These requirements address the functionality of the tools, their integration, and the programming languages, middleware and hardware to support. They will be determined in close interaction between the WP2 partners, the program developers in WP1, and the CrossGrid Architecture Team.

In addition, a review of existing technologies will be performed. Currently, there are no tools directly comparable to those to be developed in WP2: verification and performance prediction tools do not yet support Grid environments; performance analysis tools for the Grid, e.g. NetLogger, mainly provide low level information on the Grid infrastructure, rather than application related data. However, concepts and data provided by these tools can and will be reused as far as possible. In particular, the DataGrid project plans to provide services that create a directory of static and dynamic information on the Grid infrastructure and a platform for application instrumentation (DataGrid WP3, deliverables D3.1 – D3.3). We will analyse the types of run time data needed by the tools to be developed in WP2 and examine which of the needed data is already available via the DataGrid services. Since the access to all run time data will occur through the unified monitoring interface developed in Task 3.3, the results of the analysis will define the major requirements for this.

Task 2.2 MPI code debugging and verification (Month 4-36)

Task leader: Matthias Müller, USTUTT (AC9)

The objective of this Task is to develop a tool that verifies the correctness of parallel, distributed Grid applications using the MPI paradigm. The primary issues are how to make end-user applications portable, reproducible and reliable on any platform of the Grid. Another goal is to enable the debugging of applications executing on hundreds of processors. Simple (NxN) matrix algorithms, benchmarks developed in Task T2.3 as well as applications from Tasks T1.2, T1.3 and T1.4 will be used during the test and refinement phase.

The technical basis of this development is the MPI profiling interface which allows a detailed analysis of the MPI application at runtime. Existing tools like commercial debuggers (e.g. TotalView from Etnus) only cover a very limited subset of the planned functionality (e.g. TotalView does not address portability), or they are targeted for shared memory machines (e.g. Assure from KAI, or "umpire" under development at LLNL) and are thus neither portable to distributed memory platforms nor suitable for the Grid.

The main components of this tools will be developed by USTUTT which is also leading this task. The CSIC team involved in this task will provide Grid-enabled (NxN) matrix algorithms and collaborate in analysing and optimising the use of MPI in the most time consuming applications (Tasks 1.3 and 1.4), namely data-mining techniques based on Neural Network algorithms. The debugging tools previously cited will be used to verify the distributed algorithms, and trace possible conflicts. New strategies on the use of MPI on the Grid taking into account availability and heterogeneity properties will be proposed. In particular, explicit algorithms to optimise the global Grid answer time will be developed, in the framework proposed by the following tasks.

Task 2.3: Metrics and benchmarks (Month 4-36)

Task leader: Marios Dikaiakos, UCY (AC12)

This task will propose a set of performance metrics to describe concisely the performance capacity of Grid configurations and application performance, and it will develop and implement benchmarks that are representative of typical Grid workloads. Such benchmarks will be used to estimate the values of performance-metrics for different Grid configurations, to identify important factors that affect end-to-end application performance, and to provide application developers with initial estimates of expected application performance. A suite of benchmarks will be deployed and validated in the CrossGrid testbed.

The main work of this task will be performed by UCY. TUM will participate in the definition of metrics, which are also important for the assessment component developed in Task 2.4.

Task 2.4: Interactive and semiautomatic performance evaluation tools (Month 4-36)

Task leader: Wlodzimierz Funika, CYFRONET (CO1)

Task 2.4 will develop on-line tools that allow application developers to measure, evaluate, and visualise the performance of Grid applications with respect to data transfer, synchronisation and I/O delay as well as CPU, network and storage utilisation. The tools will support any level of granularity, from the Grid level down to the process level. Acquisition of raw performance data will be based on the monitoring infrastructure developed in WP3. The performance evaluation tools will exhibit four distinctive and novel features:

1. The tools operate on-line. This means that the user can select the performance data to be measured at run-time, based on the results of previous measurements. Thus, the need to acquire, store and transmit large amounts of trace data is avoided.
2. Based on the work done by the APART working group, the tools include an automatic component that can extract high-level performance properties from the measured raw performance data. This includes an assessment of the application's performance w.r.t. the performance characteristics of the currently used Grid configuration. The extractable performance properties at the application level, as well as the data model for the raw performance data will be modelled using the APART specification language ASL. The Grid's performance characteristics will be determined using the benchmark results and metrics delivered by Task 2.3.

3. A tool will be provided to extract the relationships between the application execution features, i.e. the performance of the grid (network speed, CPU speed, memory bandwidth), the problem sizes (vectors and matrices) and the real execution time for some selected kernels in the applications, in a form of an analytical model. After having selected the relevant kernels from the applications of WP1, the above relationships are extracted on a statistical basis, to build an analytical model. When applying the kernels to different Grid configurations, a visualisation module will be used to present the results of the analytical model. They comprise the number of FLOPs, the load balance, the volume of communication, the prediction of the communication times, the amount of memory needed, and the prediction of the total runtime, w.r.t. the grid and problem features.
4. The tools will use a standardised tool/monitor interface, which will be an extension of the OMIS specification for Grid applications. This will enable the interoperability of tools, whose simultaneous use is intended to give a synergistic effect for performance analysis of applications.

CYFRONET will define performance measurements as well as develop a tool/monitor interface and GUI. TUM will develop the component for automatic extraction of performance properties, and, in addition, it will support CYFRONET in the definition of the tool/monitor interface. U.S.C. will collaborate on the automatic extraction of performance features and on the development of analytical performance models. CSIC will analyse the impact of the underlying network infrastructure on the final performance of the Grid.

Task 2.5 Integration, testing and refinement (Month 9-36)

Task leader: Roland Wismüller, TUM (AC10)

The components developed in WP2 must be integrated to build a uniform tool environment for Grid programmers. At the beginning of the project, this task will refine the design of this environment, which is sketched in Fig. 4. In particular, the interfaces and interactions between the different components will be fully specified and design rules for the component's user interfaces will be set up.

In order to guarantee a timely availablilty of the tool environment for its use in WP1, a prototyping approach will be used. For each prototype version created by Tasks 2.2, 2.3, and 2.4, this task will handle the integration, testing and refinement of the developed software components. The prototypes will first be installed at the testbed site FZK. Since FZK will have a lot of experience in the field of HEP, the application from Task 1.3 will be used for the first real-world tests of the tool prototypes, which will be performed jointly by TUM and FZK. Feedback on errors and problems will immediately be given back to the tool developers in WP2. After successful testing, the tools will be released and distributed to all other testbed sites by FZK and will then be used by the application programmers in WP1. Feedback, lessons learned, necessary software quality improvements, and additional requirements will be collected, evaluated and summarized into a report by TUM. This report will form the basis for the development and implementation of the following prototype versions, which will be handled in the same way.

Workflow and Interfaces to other Workpackages

M 1-3: Requirements Definition Phase

Deliverable for all partners is D2.1 (report) at the end of M 3.

For all partners: Define the technical infrastructure (hardware and software) necessary for the development of these tools in close collaboration with the CrossGrid Architecture Team. Deliver the information to WP4 for initial infrastructure setup. Define common data sources and needs from the applications in WP1. Review existing technologies on the tools to be developed in this workpackage. Specify details of the overall programming environment architecture (refinement of Fig. 4). Specify interfaces between the different components developed in this workpackage and to Task 3.3.

For Task 2.2 (MPI): Codes to be analysed are: (1) already existing, basic and Grid-enabled (NxN) matrix algorithms formulated with the MPI paradigm and provided by CSIC in Task 2.2, (2) HEP application provided by Task 1.3, (3) Weather forecast & air pollution application provided by Task 1.4, (4) Flooding application provided by Task 1.2, (5) benchmarks provided by Task 2.3 after M 12.

Collect a list of MPI calls used in these codes. They will be implemented with high priority to provide a useful functionality of the tool as soon as possible.

For Task 2.3 (Metrics and benchmarks) Analyse basic blocks of all applications in WP1, concerning e.g. dataflow for input and output, CPU-intensive cores, parallel tasks/threads, communication, as well as basic structures of the CrossGrid. Take into account flow charts, diagrams, basic blocks from the applications and get support by the respective application developers. Basic structures of the CrossGrid configurations may be analysed in close collaboration with the CrossGrid Architecture team and with representatives of the testbed sites of WP4.

For Task 2.4: Performance evaluation Analyse basic blocks of all applications in WP1, concerning e.g. dataflow for input and output, CPU-intensive cores, parallel tasks/threads, communication, as well as basic structures of the CrossGrid. Take into account flow charts, diagrams, basic blocks from the applications and get support by the respective application developers. Basic structures of the CrossGrid configurations may be analysed in close collaboration with the CrossGrid Architecture team and with representatives of the testbed sites of WP4.

Define high level performance properties to be extracted from the applications in WP1 in close collaboration with the respective application developers. Select the most important performance properties to be extracted from the applications point of view. These will be implemented on local cluster infrastructures with high priority to provide a useful functionality of the first prototype. Define interfaces to raw performance data provided by the monitoring system of Task 3.3 in close collaboration with the respective developers. Define an interface to the output of benchmark results from Task 2.3. Review the interface to the performance visualization module.

M 4-12: 1st development phase

Deliverables for all partners are D2.2 (report) at the end of M 6 and D2.3 (1st prototypes and reports) at the end of M 12.

M 4-6:

Task 2.2 (MPI) Design architecture of the MPI code verification tool and its implementation into the programming environment architecture. Deliverable is D2.2.

Task 2.3 (Metrics and benchmarks) Propose a set of performance metrics to describe the performance capacity of CrossGrid configurations and application performance. Design benchmarks representative for the applications of WP1 and their implementation into the programming environment architecture. Deliverable is D2.2.

Task 2.4: Performance evaluation Design the architecture of the tools, based on their functional description. Design the hierarchy and naming policy of objects to be monitored. Define types of measurement, those general-purpose and specific to the field of uses. Design the tool/monitor interface, based on the expressing of measurement requests in terms of monitoring specification standard services. Develop the filtering and grouping policy for the tools. Define functions for handling the measurement requests and the modes of their operation. Design the granularity of measurement representation and visualization modes. The design for the first prototype will concentrate on performance measurement and evaluation of applications on the level of local clusters.

For the selected kernels, a first approach of the analytical model will be done, in which the most relevant features, such as the runtime and communications costs, will be included. This initial model will be improved in following development phases, to include new performance features and kernels. Deliverable is D2.2.

M 7 demands:

Grid infrastructure must be available on the FZK testbed site.

M 7-12:

Tasks 2.2, 2.3, 2.4: Write first prototypes on a local cluster at the test site. Perform first simple tests during development e.g. with (NxN) algorithms and HEP application. Fix any errors. Write documentation of design, implementation and interfaces. Design test scenarios especially for tests with applications. Deliverable is D2.3.

M 10-12:

Task 2.5: 1st Integration, testing and refinement Provide all the tools for installation on the FZK test site. Integrate all tools into the programming environment. Test them with simple test cases and with one of the applications (HEP, Task 1.3). Fix any errors. Provide a software package for distribution to other testbed sites. Support installation of the tools on all other testbed clusters. Fix any portability problems. Test the tools with simple test cases on different architectures. Fix any errors. Deliverable is D2.3.

M 13-24: 2nd development phase

Deliverables for all partners are D2.4 (report) at the end of M 15 and D2.5 (2nd prototypes and reports) at the end of M 24.

M 13 demands:

Grid Infrastructure (local Testbeds of WP4) must ready to integrate the tools developed in this WP. Applications of WP1 must be ready to integrate and test the tools of this WP. The extraction of raw performance data (Task 3.3) of applications must be available for local cluster infrastructures.

M 13-14:

Task 2.5: 2nd Integration, testing and refinement Testing of the tools by – and/or in close collaboration with - application developers from WP1. Compare benchmarks, and performance predictions with applications. Collect feedback on the single tools. Feedback will be taken into account in the next prototype version. Collect feedback on a possible user interface to the programming environment, i.e. a common 'look and feel' (e.g. GUIs) for the tools developed in this Workpackage.

Refine the programming environment architecture (if still necessary). Define the user interface to the programming environment according to the demands. Deliverable is D2.4.

M 15-18:

Tasks 2.2, 2.3, 2.4: Refine the architecture design of the tools and their implementation into the programming environment architecture. Design the user interfaces according to the user feedback. Define additional performance data to be measured, evaluated and visualized on a Grid level and design respective visualization modules. Propose strategies for a redesign where benchmarks and performance predictions did not match application performance. Fix bugs found during the 2nd integration, testing and refinement phase. Deliverable is D2.5.

M 18 demands:

Full Grid infrastructure must be available on all testbed sites at the end of M 18. The extraction of raw performance data (Task 3.3) of applications must be available on a Grid infrastructure level.

M 19-24:

Task 2.2, 2.3, 2.4: Write second, fully Grid-enabled prototypes, at the test site. Perform simple Grid tests during development e.g. with (NxN) algorithms and HEP applications. Fix any errors. Write documentation of design, implementation and interfaces. Propose test scenarios on the whole Grid. Deliverable is D2.5.

M 22-24:

Task 2.5: 3rd Integration, testing and refinement Provide all the tools for installation on the FZK test site. Integrate all tools into the programming environment. Test them with simple test cases and with one of the applications (HEP, Task 1.3). Fix any errors. Provide a software package for distribution to other testbed sites. Support installation of the tools on other testbed clusters. Fix any portability problems. Test the tools with simple test cases on different architectures. Fix any errors. Deliverable is D2.5.

M 25-33: 3rd development phase

Deliverables for all partners are D2.6 (report) at the end of M 27 and D2.7 (demo and report) at M 36. The final version of all codes is not a separate deliverable but a milestone at the end of M 33.

M 25 demands:

Application developers must be available to test the tools developed in this WP on the whole Grid.

M 25-26:

Task 2.5: 4th Integration, testing and refinement Testing of the tools by – and/or in close collaboration with - application developers from WP1 on the Grid. Collect feedback on the single tools, on the user interfaces and on the integration of the whole package. Feedback will be taken into account in the final version. Deliverable is D2.6.

Fix the programming environment architecture (if still necessary). It should not be changed after these tests any more. Refine the user interfaces to the programming environment according to the requirements.

M 27-30:

Task 2.2, 2.3, 2.4: Fix the architecture design of all tools and their implementation into the programming environment architecture. Refine the user interfaces according to the demands. Fix bugs found during the 4th integration, testing and refinement phase. Propose a strategy to implement the full MPI standard. Define strategies for a final design where benchmarks and predictions did not match application performance.

M 30-33:

For all: Write final code versions, perform final tests, prepare final integration of the programming environment into Globus. The final version of the programming environment is milestone M2.4 at the end of M 33.

M 33-36: Finishing phase

Demonstration and documentation on the final versions and on the testing of all WP2 software components. Deliverable for all partners is D2.7 (demo and report) at the end of M 36.

Technology required from other CrossGrid WPs

Task 2.1 (requirements definition) needs a strong interaction mainly with WP1, WP3 and with the CrossGrid Architecture Team. The results of this stage will be also a set of requirements to WP3. Tasks 2.2 and 2.4 will use all applications of WP1 for testing and refinement of tools. All tools developed in WP2 need the CrossGrid technical infrastructure (WP4) during the integration, testing and refinement phase of Task 2.5.

Technology required from other EU projects

Technology required from DataGrid will be available through WP3 and WP4. Task 2.4 needs results from the APART working group (design ideas and APART specification language) and then this Task will be in contact with APART2 (started July 2001).

Resources

Total and (funded) Person Months (PM).

Task	Task PM	CYFRONET	FZK	USTUTT	TUM	UCY	CSIC	U.S.C.
2.0	9 (9)		9 (9)					
2.1	32 (19)	6 (3)		3 (3)	8 (4)	6 (3)	3 (3)	6 (3)
2.2	32 (32)			24 (24)			8 (8)	
2.3	28 (18)				6 (3)	22 (15)		
2.4	81 (46)	24 (12)			36 (18)		6 (6)	15 (10)
2.5	51 (32)	8 (4)	4 (4)	4 (4)	18 (9)	8 (4)	3 (3)	6 (4)
Total PM	233	38	13	31	68	36	20	27
Funded PM	156	19	13	31	34	22	20	17

WP3 New Grid Services and Tools

Workpackage manager: Norbert Meyer, PSNC (CR11)

Objectives

The main objective of this workpackage is to develop Grid services and a software infrastructure required to support the Grid users, applications and tools as defined in the workpackages WP1 and WP2. This workpackage includes a set of tools and services, which will define (also including the results of WP2) the middleware layer of the CrossGrid project. The development in WP3 will use and take into account the current and future state of the art. We will also follow the GRID standards which are the subject of the Global Grid Forum meetings [GGF,GGFU,GGFD]. We are aware of other Grid projects, like DataGrid (concerning especially workpackages developing the Grid middleware) or EuroGrid under the subject access points to applications. Therefore, our goal is also to assure the compatibility with the middleware developed within the aforementioned and other projects.

The subject of Task 3.1 is to establish a user friendly Grid environment by portal access to the Grid independently from the user location [Pitoura, 1997; Bethel, 2000]. This task includes two major subtasks. The first subtask will allow one to access applications via portals, which will support the users directly by simplifying the handling of applications. The second subtask will create a mechanism allowing one to enter the Grid environment at any place and from any hardware (independently of the type of the operating system) with the same previously used user's environment. This is a new approach, currently not developed in any other project. Task 3.2 addresses the construction of new resource management techniques, namely the design and evaluation of self-adaptive scheduling agents that will be responsible for scheduling parallel applications submitted to a Grid with the goal to achieve a reasonable trade-off between the efficiency of resource usage and application speedup according to the user's optimisation preferences [Berman, 1998; Casanova, 1999; Heymann, 2001]. In Task 3.3 we propose to develop a prototype infrastructure for the needs of monitoring-related activities for automatic extraction of high level performance properties and for tool support of performance analysis (as described in the workpackage WP2) [Bubak, 2000]. We also propose to build a tool for monitoring network infrastructure. Another aspect of the Grid performance will be solved in Task 3.4: in order to help users in their interaction with large data manipulation, we propose a kind of expert system which will operate as an advisor for the selection of migration/replication policy in a particular environment. Within this task a common middleware for fast tape file access [Slota, 2000] and a system for data access time estimation will also be developed [Nikolow, 2001].

The major relations between WP3 and other WPs have been presented at Fig. 5. Task 3.5 will deliver all information about possible errors detected while the system tests and required corrections which should be done, including e.g. changes in the architecture or the API interfaces between the WP3 tools and other modules (WPs). Task 3.5 will prepare an official prototype version of the whole set of tools of the WP3 – middleware. This middleware software will be used by other WPs, especially the WP1 and WP2. The first prototype of WP3 software for local grids will be available in M 12, and a prototype for a global grid infrastructure will be available in M 24. Because of the needs of WP2.4, which requires grid implementation of Task 3.3 in M 18, a prototype of WP3.3 will be available in M 18. WP3 will need feedbacks in M 15 and M 27 from WP1 (Tasks 3.1, 3.2 and 3.4) and from WP2 (Task 3.3). The testbed supported by WP4 will be used in Task 3.5 as a testing environment, allowing testing separate tools (Tasks: 3.1 ... 3.4) as well as the whole prototype of WP3 (middleware). Task 3.5 requires that the local grid testbed be available in M 6 (with a Certification Authority, Registration Authority, monitoring system and access to data storage systems) and the full grid be available in M 12. Task 2.4 will use the Grid Monitoring tool (3.3). The tools developed in Tasks 3.1 and 3.4 will be used directly by the applications (WP1) under the scope of optimised data access (input data and results of computations) and simplified user access to each of the applications deployed in WP1. The roaming access tool will directly support the end user by making movement easier in the Grid environment.

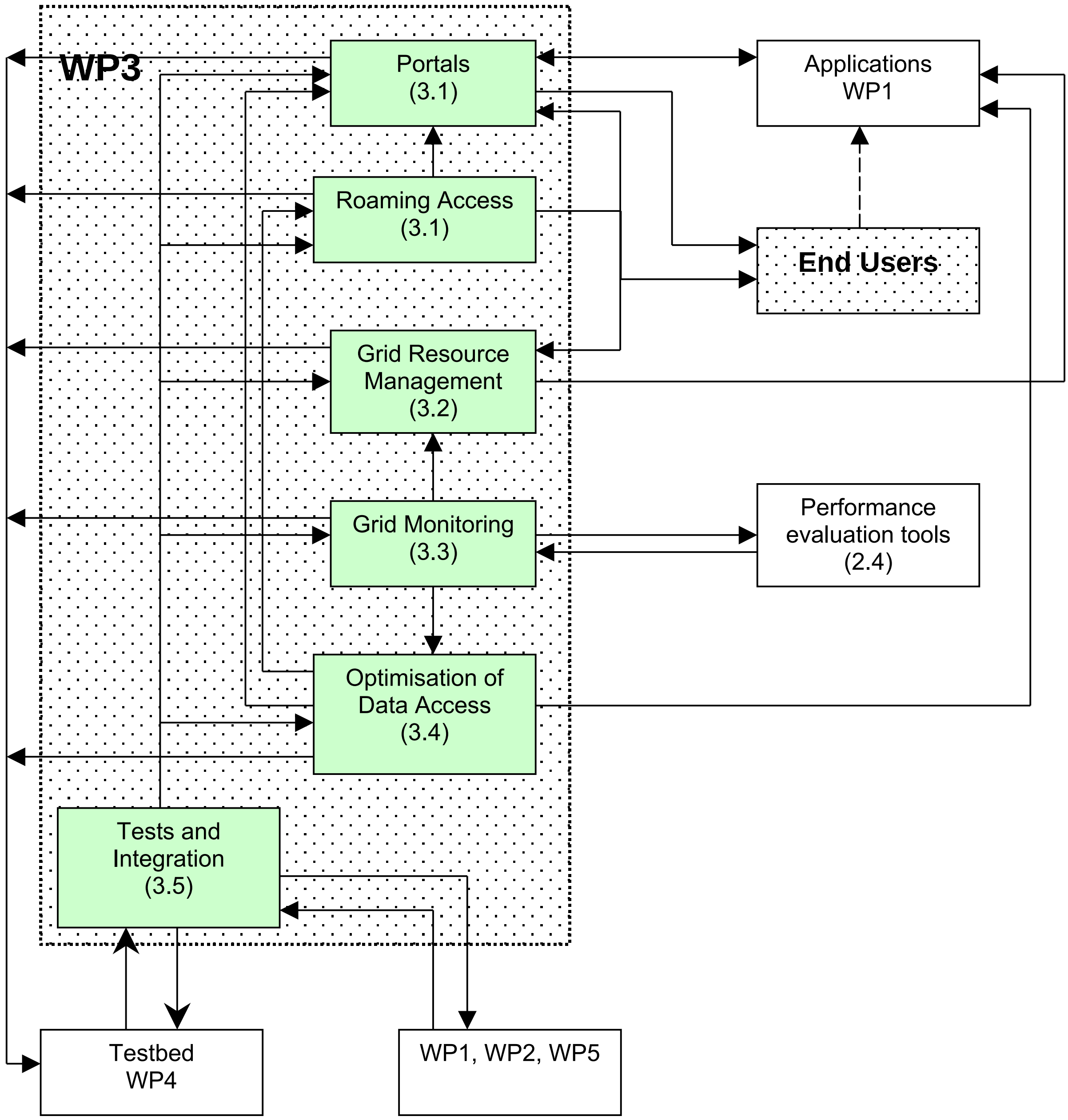


Fig. 5. Relation between WP3 and other WPs of the CrossGrid.

Task descriptions

Task 3.0 Co-ordination and management (Months 1 – 36)

Task leader: Norbert Meyer, PSNC (CR11)

The goal of this task is to control the progress of each task in this workpackage, to establish and synchronise the co-operation between the workpackages and tasks within this workpackage. One of the major tasks will also be to work out a general software architecture within WP2 and WP3 as well as in WP1. The management part of these tasks will also investigate and avoid any overlaps and duplications between these tasks and components which have been developed in other Grid-related projects. The task leaders will be requested to submit reports on task status to the WP manager twice a year, and WP co-ordination meetings will be held with the same frequency. The management duties of WP3 will be done by PSNC (CR11).

Task 3.1 Portals and roaming access (Months 1 – 36)

Task leader: Miroslaw Kupczyk, PSNC (CR11)

This task will develop user portals customised to particular applications. The portals will give a unified and consistent window to the CrossGrid environment. As soon as the Grid tools are developed, they will be integrated into the portal providing easier access and use of the Grid by applications. Reusable components are necessary to build such portals for various classes of applications, allowing a uniform desktop access to remote resources. Such components will include tools and services enabling easy collaboration, job submission, job monitoring, component discovery, and persistent object storage (Tasks 3.2, ... 3.4). The CrossGrid portals will securely authenticate and authorise users to remote resources and help them make better decisions for scheduling jobs by allowing them to view pertinent resource information obtained and stored in a remote database (Task 3.2, Task 3.3). Profiles will be created and stored for portal users allowing them to monitor jobs submitted and view the results. The CrossGrid portal capabilities will include:

- secure user authentication and authorisation to the Grid resources,
- a possibility of running interactive or batch applications,
- remote data access,
- enhanced features for mobile users,
- access to the CrossGrid applications,
- access to the Grid VPN management tools.

One of the subtasks we propose to implement is roaming access i.e. a mobile personalised environment to allow users to access the Grid from different locations and different platforms including mobile terminals. The main objective of this subtask is to build a middleware for users' access to their personal settings; independently of the hardware and software client platform. Personal settings stand for features like the configuration definitions that characterise e.g. links to the user data files, links to applications, access to portals and HPC infrastructure, as well as windows settings.

In a complex corporate network or, furthermore, in a network which implements the computational GRID idea, a user needs something more, namely transparent resource access. By the resource term we understand e.g. databases, tele-immersion laboratory hardware, digital libraries and computational machines as well. A set of privileges that the user owns is associated with every user regarding every resource. In connection with it, complex user authorisation mechanisms are always required, on any resource demand procedure.

The migrating desktop concept, as the main objective, should produce a transparent user work environment, automatic access to resources: network file systems, the same file paths and applications independently of the system version and hardware. Any such solution is nevertheless limited by applications, which are not always binary compatible with the user's current architecture. The term *mobility* in roaming access means not only the geographical localisation of the user but also the type of the hardware and operating system. It means that the user will have his/her working environment accessible independently of the software, hardware and localisation.

Some additional characteristics are:

- uniform data organisation, across different platforms and locations, including user workplace,
- workspace migration,
- integration of monitoring tools and personalised user environments.

The participants working on this task are aware of such systems and tools as: GPDK, Globus, GSI, GRAM, GridFTP, LDAP, Myproxy, OpenSSL, Mozilla Directory API Apache, JSP, EJB, XSIL, CoG toolkit, IAIK SSL API, Java Servlets, and Gateway. Globus allows using common and uniform API for access to the resources, but unfortunately it does not support migration of the users' personal settings.

Summing up, this task will develop new functionalities for the Grid environment. The first (roaming access), a new feature, currently not solved and not even defined in other Grid-related projects, will allow one to use a personal user environment (defined as the whole desktop), independently of the localisation and independently of the terminal type. This tool can be used with the portal, enhancing its functionality, or as an independent Grid middleware tool. The second task (the portal) will connect together all tools developed in WP3 as well as in WP2 and will provide easy access to the applications deployed by WP1. An additional but very important issue is to simplify the user access and the handling of applications and tools by providing a specialised, advanced graphical user interface.

A common software architecture will be worked out by the whole team, based on the set of requirements (feedback from WP1 and WP2, current state of the art.). The design and development phases of the roaming access service will be done mainly by PSNC (CR11) and UCY (AC12). The DATAMAT (AC13) and ALGO (CR23) partners will focus on accessing applications by the portal approach. The test phase will be mainly realized in Task 3.5.

Task 3.2 Grid resource management (Month 1 – 36)

Task leader: Miquel A. Senar, UAB (AC16)

In Task 3.2, we will develop a Grid resource management system based on self-adaptive scheduling agents for scheduling a particular parallel application submitted to a Grid, with the goal of achieving a reasonable trade-off between resource usage efficiency and application speedup, in accordance with the user's optimisation preferences. The system will be supported by the monitoring subsystem that will provide monitoring facilities for the Grid management. It will be adapted to any common parallel job description interface agreed both in this consortium and other related projects such us Datagrid.

Scheduling agents for parallel applications submitted to a Grid will take into account, on the one hand, dynamic information about target resources such as performance heterogeneity, load and availability, location of remote data files and availability of replicas, and network delays. On the other hand, they will also take into account application characteristics in order to adapt their execution to the available resources. The application characteristics will include both static and dynamic information. The static information will be described through a job description language and would include, for instance, information about the architecture and operating system required to run the application, as well as other requirements such as memory, libraries, etc. The dynamic information will be gathered at run-time through services provided by the Grid monitoring facilities, and will be used in conjunction with information from the Grid resources to predict near future requirements and to guide scheduling actions. In this sense, scheduling agents will try to take advantage of the high degree of predictability that many parallel applications exhibit, as they are usually composed of several tasks that tend to behave similarly over time for a given set of initial data. This fact has been observed, for instance, in the context of job batch scheduling.

Agents will be implemented in a parameterised way in order to be easily adapted and extended to different kinds of parallel programming paradigms (which include, for instance, embarrassingly parallel models such as parameter sweep applications or master-worker applications). Their design will also take into account all the scalability and reliability issues needed by user applications. Early experiences with scheduling agents manually written on a case-by-case basis have been presented in the AppLeS (Application-Level Scheduling) system, which uses the services offered by the NWS (Network Weather Service) to monitor the varying performance of available resources. Similar experiences are presented in NetSolve, a client-agent-server system, which enables the user to solve complex scientific problems remotely, and has also been used to develop automatic scheduling agents for simple farming applications [Casanova, 1999]. Finally, early experiences with the use of automatic self-adjusting scheduling agents for generic task farming applications have been shown in [Heymann, 2001].

From the user perspective, scheduling decisions are important as they usually have a direct impact on the overall performance and efficiency achieved by an application submitted to a Grid. Moreover, different users may be interested in running their applications using different performance metrics as targets to be optimised. Management policies defined by system users are submitted to the management system and then automatically deployed in the local management modules. Certain events (e.g., a user login) would then trigger policy-defined local actions (e.g., checking user credentials). Events that are not handled locally would travel up the event hierarchy to be resolved in management nodes with the appropriate polices installed. The management system should be able to control events related to users, processes, resources, i.e., any type of action that might generate an event.

Scheduling agents will make use of resource management mechanisms that are currently available or that are under development in other related projects such as Globus and DataGrid. Such mechanisms are used to find and allocate resources for an application and provide or will provide services for code migration, data migration and remote data access. Our scheduling agents will extend these basic services by providing agents that will take advantage of them to guide the execution of applications in a fully automatic manner. Additional mechanisms, such as task duplication, will also be investigated in this task, in order to compensate for the potential loss or contention of the Grid resources.

All partners will work on the specification of requirements of scheduling agents. The architecture and interface design of scheduling agents will be mainly developed by UAB (AC16), which is also leading this task and will implement the basic elements. DATAMAT (AC13) and CSIC (CR15) will collaborate on the design and implementation of particular agents for certain parallel programming paradigms.

Task 3.3. Grid monitoring (Month 1 – 36)

Task leader: Brian Coghlan, TCD (AC14)

Grid monitoring facilities within Task 3.3 are aimed at underpinning the functionality of various tools, by delivering the low-level data intended for automatic extraction of high-level performance properties and on-line tool support, oriented towards application-bound activities, as well as the data for the needs of network-oriented activities on the Grid. System scalability, flexibility and ease of configuration will be the guidelines during the design and implementation phases.

The tools from within WP2 will get indispensable performance data from the monitoring system developed in Task 3.3, which will offer a uniform interface to feasible monitoring services available on the Grid. Interfacing with the DataGrid monitoring tools will be part of the task.

The application-level monitoring environment will comprise an autonomous monitoring system. For communication with tools the monitoring system will comply to a standardised interface. The *efficiency* of application monitoring will be provided by specialised application monitors, embedded in the application address space, and by efficient performance data storage. The issue of *scalability* will be addressed by distributing the monitoring system into *service managers* (intermediate layer between tools and local monitors), *local monitors* (direct control of local application processes), and *application monitors* (local handling of performance-critical actions in co-operation with local monitors). The OMIS specification will be used to build the communication layers: *tools/service managers* and *service managers/local monitors*. The system will be extensible to provide new, additional functionality.

We will develop specific relational producer/consumers of non-invasive performance monitoring information, as well as associated SQL-query-based performance analysis tool support. This is especially desirable for validation of invasive toolsets and simulation models. An OMIS-compliant interface will enable interaction with the above invasive toolset.

Within the Task we will also focus on the Jiro technology, which is a Java language-based implementation of the Federated Management Architecture. Being designed for enterprise-sized applications, and supporting network management protocols such as SNMP, this platform seems to be an appropriate approach to the network-related issues within Grid monitoring. Jiro components, such as the Logging Service, are going to be used for infrastructure monitoring, i.e. collecting information about user and node activity, network and node loads and network configuration tracing. Therefore, the field of uses for this approach will complement that of the OMIS-based one. Moreover, an important feature of the approach is that the management components could communicate with managed objects using SNMP, WBEM or some proprietary protocols. This enables one to create a solution that can leverage the existing "legacy" systems without need to convert them to Java, so the proposed solution is flexible and extensible.

All partners of this task will work out the design specification of the monitoring tools. CYFRONET (CO1) will define the functionality of the monitoring environment to support tools' needs in compliance with a standardised interface as well as develop the corresponding modules of the infrastructure. ICM (AC2) and TCD (AC14) will ensure the co-operation of the SQL-query-based performance analysis tools with other components of the monitoring system. TCD will specify and develop the relational non-invasive producer/consumer and associated SQL-query-based tool support, and will take care of standardized interfaces to the Grid programming environment (WP2) and the application level (WP1). Additionally, the tests and integration will be realized within Task 3.5 and other partners not directly involved in the development phase, which will make the tests more general, as they will be carried out by independent third-party sites.

Task 3.4 Optimisation of data access (Month 1 – 36)

Task leader: Jacek Kitowski, CYFRONET (CO1)

Scientific applications require efficient access to the data. In order to help users in their interaction with data store, a kind of expert system is proposed. It will operate as an advisor for selection of migration/replication policy in defined user states. Its main goal will be to suggest whether and when a chosen data file should be replicated into the local environment of the remote user in order to shorten the waiting time for data availability. The expert system for data management will be used by the Grid resource management system (Task 3.2 of CrossGrid).

In the proposed expert system the number of possible migration strategies will be small, therefore we propose to implement a heuristic reinforcement learning scheme. Based on the current user state, one heuristic migration strategy is to be chosen from the set of possible strategies; its effectiveness will be verified according to the chosen criterion and its future choice probability for the current state then modified. So the system will gather experience of which strategy is the most likely to be chosen in each of the user states. Clustering for similar user states will be applied. Sample heuristics are: greedy, look-back and forward tuning. Prefetching will be based on the Lempel-Ziv scheme or Zipf's distribution. This activity will benefit from Task 2.4 of the DataGrid Project, in which the LDAP protocol is proposed, which provides a structure for organising, replicating and distributing catalogue information. Another goal of this task is to design and implement a mass-storage-system-independent middleware to provide faster access to large tape resident files. This will be done by using a sub-filing strategy based on file partitioning into smaller fragments, which could be retrieved separately according to the requests. Such a strategy is useful for multimedia files, therefore this approach extends WP2 of the DataGrid Project, in which the granularity of access is only on the file level and on the dataset level. Benefits of this subtask will be evident to applications in WP1 of CrossGrid, where the large amount of data cannot reside entirely on the disks. In Task 1.1 that deals with interactive simulation and visualisation of a biomedical system it will serve as a tertiary storage for multimedia medical data, offering the possibility for additional analysis. In Tasks 1.2 and 1.4 it will co-operate with databases that reside on disks so they may use data which is not available on-line. This will make it possible to perform independent analysis of off-line or near-line data, thus extending the possible scope of the analysis.

Finally, within this task we will design and implement a system for data access time estimation. The estimation will depend on many factors like MSMS load, queue length, number of drives, drive throughput, file size, etc. An appropriate function to communicate with the tool should be added to a common API and could be used in the Grid work scheduling. This system will be useful for the users to estimate their waiting time for data availability and also as a supporting module for the prefetching and migration strategies. It will be used in Task 3.2 of CrossGrid.

Since visualisation is an essential tool in computational science and engineering, and on the Grid a scientific application is composed of (possibly) massively parallel computing tasks which process unprecedented amounts of data stored remotely, the use of tertiary storage is an important task. Thus Task 1.5 will directly benefit from the access time estimation subsystem and migration strategies.

All systems from this task will co-operate with the portal described in Task 3.1 of CrossGrid.

Since a similar area is covered by WP2 of the DataGrid Project and also by some current US projects (GryPhyN, PPDG) we will be trying as far as possible to avoid unnecessary duplication of major middleware features by keeping aware of their work and collaborating as fully as possible.

The development of this service tool is mainly devolved to CYFRONET (CO1), an institution having a wide experience in the field of data management and archiver systems. The test phase will be realised within Task 3.5 by other partners, i.e. PSNC (CR11) and UAB (AC16).

Task 3.5 Tests and integration (Month 10 – 36)

Task leader: Santiago Gonzalez, CSIC (CR15)

The main goals of the tests and integration are to deliver a release version of WP3 before each prototype. Such a release will be a base for further development in other workpackages, especially for WP1 and WP2. Task 3.5 is very strongly correlated with the prototypes released in WP4. The four partners of this task: CYFRONET (CO1), PSNC (CR11), CSIC (CR15) and UAB (AC16) will test all tools and services of WP3 to be able to deliver a middleware layer in the CrossGrid testbed, which will be used by the applications (WP1) and the Grid Application Programming Environment (WP2). The integration phase is strictly dependent and correlated with WP4 tasks and its time scheduling. A feedback from other workpackages (WP1, WP2, WP4 and WP5) is highly appreciated and will be taken into account before releasing each prototype version.

Workflow and Interfaces to other Workpackages

M 1-3 Definition of requirements

Define the technical infrastructure (hardware and software) necessary for the development of these tools in close collaboration with the CrossGrid Architecture Team. Deliver the information to WP4 for initial infrastructure setup. Define requirements from the applications in WP1 and WP2. Review existing technologies on the tools to be developed in this workpackage. Specify interfaces between the different components developed in this workpackage. Define the monitoring environment functionality and interfaces to Task 2.4. Define specification of GRM agents. In the definition phase, various tasks will closely collaborate with the CrossGrid Architecture team and the initial tasks in WP1, WP2, and WP4.

Deliverable for all partners is D3.1 (report) at the end of M 3.

M 4-12 Initial implementation and development phase

M 4-6 Designing of architecture, interfaces and security issues.

Design the architecture of the new grid services and tools. Design interfaces between tools. Specify the security issues: design the secure communication between tools, specify ways of secure authentication and authorisation, propose standards of implementing security features.

Deliverable for all partners is D3.2 (report) at the end of M 6.

M 6

Requires: First testbed set-up on selected sites.
Deliverable D3.2

M 6-12 Implementation of 1st prototype (running on local grid)

Create initial Grid-enabled implementations based on existing Grid software and the first versions of software. Deploy this implementation in testbed set-up on selected sites. Perform first simple testing. Write documentation of design , implementation and specification of interfaces.

Task 3.1 batch application running

Task 3.2 introduces self-adaptive scheduling agents

Task 3.3 comprises separate components: monitors, managers

Task 3.4 analytical issue of expert system for migration strategies

Produces: First prototype in M 12 running in local grid infrastructure

Deliverables is D3.3 (prototype and report) in M 12

M 13-24 Second development phase

Co-operation with WP1 and WP2 in testing and improving implemented tools. Correcting bugs. Continue development of software. Implement the 2nd prototype running on the international CrossGrid testbed. Extend the functionality of applications. Write the second, fully Grid-enabled prototypes. Perform simple Grid tests during development. Write the documentation of design, implementation and interfaces. Propose test scenarios on the whole Grid.

Requires: full testbed prototype in the beginning of M 13

M 13-15 Testing, feedback collecting

Testing of the tools by – and/or in close collaboration with - application developers from WP1 and WP2. Compare benchmarks, and performance predictions with applications. Collect feedback on the single tools. Feedback will be taken into account in the next prototype version. Collect feedback on a possible user interface to the programming environment, for the tools developed in this Workpackage.

Deliverable is 3.4.

Produces: The extraction of raw performance data (Task 3.3) of applications available for local cluster infrastructures in M 13. For WP2.4 purposes.

M 16-18 Refinement, implementation of WP3.3

Task 3.3 complies with standards of interfaces, integrates components: monitors, managers as a middleware

Produces: In M 18 -The extraction of raw performance data (Task 3.3) of applications available for Grid infrastructure level. For WP2.4 purposes.

Deliverable is D3.5.

M 18-24 Refinement, implementation of WP3.3.

Integration of Globus Resource Management with Task 2.4. Collect feedback from WP2.4.

M 16-24 Refinement, implementation of other tasks within WP3.

Task 3.1 application runs on distributed grid

Task 3.2 implements interfaces and integrates scheduling agents

Task 3.4 co-operates with portal described in Task 3.1

Produces: First prototype running in distributed grid infrastructure

Deliverable is D3.6 (prototype and report).

M 25-33 3rd development phase

Integrate tasks. Write final code versions, perform final tests, prepare final integration of the programming environment into Globus. The final version of the programming environment is milestone M3.5 at the end of M 33.

M 25-27 Testing, feedback collecting

Co-operation with WP1 and WP2 in testing and improving implemented tools. Correcting bugs. Continue development of software. Testing of the tools by – and/or in close collaboration with - application developers from WP1 and WP2 on the Grid. Collect feedback on the single tools, on the user interfaces and on the integration of the whole package. Feedback will be taken into account in the final version. Deliverable is D3.7.

M 28-30 Refinement

Fix the architecture design of all tools and their implementation into the programming environment architecture. Refine the user interfaces according to the demands. Fix bugs found during the integration, testing and refinement phase.

M 31-33 Implementation, full integration.

Final version of the 2nd prototype with all errors found in tests phase corrected. Implemented some suggestion received as a feedback from applications.

Final deployment in CrossGrid, tasks are integrated.

Task 3.1 supports user environment on the heterogeneous Grid

Task 3.2 scheduling agents for parallel jobs, tested with interaction with end users in WP1.

Task 3.3 integrates monitoring environment

Task 3.4 implements expert system for accessing data; implements mass-storage-system-independent middleware

Task 3.5 tests of the solution

Deliverable is D3.8.

M34 - 36 Final phase

Demonstration and documentation on the final versions and on the testing of all WP3 software components. Deliverable for all partners is D3.9 (demo and report) at the end of M 36.

Technology required from/by other CrossGrid WPs

The implementation and deployment of portals (Task 3.1) will allow the users to have simple and unified access to applications developed in WP1. The portal will simultaneously support access to all other resources and tools of the CrossGrid project. Additionally, Task 3.1 (roaming access service) will give the end user a possibility of a smooth moving within the Grid environment (support for mobile users) by saving the previously used working environment or allowing the choice of a different configuration. Thereby, Task 3.1 will support the end user and is strictly correlated with WP1, giving a value added feature at the application level. Task 3.1 should use an efficient data access mechanism. Therefore we will also use the results and mechanisms developed in Task 3.4 of WP3.

Task 2.4 within WP2 will use raw performance data based on the monitoring infrastructure developed in WP3.

Task 3.2 will make extensive use of basic technologies that will be shared by most workpackages in the project. In particular, the services provided by this task will be interfaced to the common job description and resource specification languages used in the project. Such languages may also be common to the ones used in other related projects such us DataGrid.

Other sources of technology will come from the monitoring task, that will be responsible for providing information about the Grid resources, and from Task 2.4, which will provide valuable information about automatic extraction of high level performance properties.

It is also planned to work with other tasks from workpackage 1 in order to use their applications as practical examples to test the scheduling agents. This means that this task will also take advantage of the infrastructure deployed by the Testbed workpackage to test early prototypes of the scheduling agents.

Other results from the DataGrid project, such us services for code and data migration and access to data replicas, will be used as far as possible.

Task 3.2 will contribute to the design and deployment of scheduling agents that will control the execution of parallel applications running on a Grid. Agents will be implemented in a parameterised way in order to be easily adapted to different kinds of parallel programming paradigms. These agents will be delivered and tested in a strong interaction with end users in WP1.

Technology required from other projects

As we are trying to develop a part of the Grid middleware layer, it is very important to follow the standards and propositions of standards worked out by international organisations and fora, e.g. Global Grid Forum. The CrossGrid project will be based on the Globus system. Therefore, all tasks of WP3 will take into account the tools of the aforementioned system. Additionally, we will be very closely co-operating with the DataGrid project. There will be high focus given to use the tools developed in this project, such as results concerning data access or monitoring systems, e.g. within Tasks 3.3 and 3.4, the monitoring technology developed in DataGrid will be integrated into the functionality. Task 3.2 describes also the results which will be used during the development.

The project is planned for the full 36 months and divided into several prototype phases. Each of the phases will finish with a release version of the applications, middleware and testbed infrastructure. The next phase in WP3 will begin from the one side with the analysis of feedback given from other WPs and from the other side the current state of the art will also be taken into account.

Resources

Total and (funded) Person Months (PM).

Task	Task PM	CYFRONET	ICM	PSNC	UCY	DATAMAT	TCD	CSIC	UAB	ALGO
3.0	18 (1)			18 (1)						
3.1	151 (94)			100 (50)	14 (7)	16 (16)				21 (21)
3.2	58 (58)					18 (18)		10 (10)	30 (30)	
3.3	98 (49)	38 (19)	28 (14)				32 (16)			
3.4	40 (20)	40 (20)								
3.5	56 (36)	20 (10)		20 (10)				6 (6)	10 (10)	
Total PM	421	98	28	138	14	34	32	16	40	21
Funded PM	258	49	14	61	7	34	16	16	40	21

WP4 International Testbed Organisation

Workpackage manager: Jesus Marco, CSIC (CR15)

Objectives

The international testbed is a key component of the CrossGrid Project, as it will provide the framework to run the applications in a realistic GRID environment. In particular, organisational issues and performance and security aspects, including amongst others the network support, may only be evaluated with the help of a testbed relying on a high-performance network (which will be provided as the result of the Géant project), thereby assuring the participation of an adequate number of computing and data resources distributed across Europe. The main objectives of this workpackage are to:

- provide a distributed resource facility where the different WP developments can be tested on a Grid,
- support the construction of testbed sites across Europe, integrating national facilities provided by the involved partners into the CrossGrid framework,
- monitor the network support required in the testbed set-up, and establish the required links with the corresponding network providers,
- integrate the basic middleware software developed or required in the different WP tasks,
- assure the required level of interoperability with other Grid testbeds, firstly with the DataGrid testbed,
- coordinate in practice the software releases that should provide the appropriate documentation and support for the installation across the distributed sites. In particular, assure that GRID applications from WP1 will run in the corresponding setup.

CrossGrid testbed sites will be placed in 16 different institutions distributed across 9 different European countries, expanding the Grid community to these countries.

Planning for WP4

In the first period a common effort will be necessary to select the initial testbed architecture. At this stage, the basic idea is to set up an extension to the DataGrid testbed satisfying the needs arising from the CrossGrid project workpackages. A specific task will deal with the infrastructure support, providing an early output that will help in the installation of clusters of computers in each testbed site. A help desk will share information and experience, avoiding redundant work. This first installation will be done at selected testbed centres, at least one per country and at centers developing applications. The corresponding results and experience will be exported to all other testbed centres to get the first full testbed release.

An incremental evolution will follow in order to be flexible enough to satisfy the needs of other workpackages and to maintain close coordination with other Grid projects, especially with DataGrid. The CrossGrid testbed will be fully coordinated with the DataGrid testbed. It will aim to be a catalyst for the unification of several Grid initiatives, both in Europe and USA, which could be integrated into a single world wide Grid. With this aim, CrossGrid will participate in the GridStart cluster, and is already active in the Global Grid Forum.

Task descriptions

Task 4.0 Testbed coordination and management (Month 1 - 36)

Task leader: Jesus Marco, CSIC (CR15)

The aim of this task is to provide coordination between all the tasks inside this WP and a main interface with other workpackages. This task will be done in coordination with the management workpackage, WP5, and using collaborative tools, like dedicated web-sites, web cast tools, videoconferencing through IP, etc.

An **integration team** will be setup under this coordination, with technical representation from WP1 responsibles for application deployment on the grid, WP2 and WP3 contacts, and responsibles from core testbed sites, with the objective of managing effectively the testbed releases.

Task 4.1 Testbed set-up and incremental evolution (Month 1-36)

Task leader: Rafael Marco, CSIC (CR15)

The initial CrossGrid testbed will be based on a small number of core technologies that are already available. Since the needs from other WPs will be changing dynamically and the coordination with DataGrid will be very important, an incremental approach will be adopted for the testbed evolution. The aim is to obtain an infrastructure as flexible as possible that will provide testbeds for other WPs to develop and test their work. It is likely that several testbeds will have to coexist offering different services and levels of stability. The main subtasks are to:

- assess the hardware and network components available for the testbed,
- define CrossGrid testbed requirements, including use cases from other WPs.
- check the integration (software and hardware) issues with the DataGrid testbed,
- define middleware software and infrastructures to be employed; inputs will come from other WPs (in particular from the architecture team in WP5), the DataGrid project and other middleware projects,
- provide the necessary basic infrastructures for testbed installation (Grid Security Infrastructure, including definition of CA and RA, Grid Information Service, etc.),
- define CrossGrid testbed installations compatible with DataGrid,
- coordinate (together with WP5) the testbed architecture and its evolution,
- deploy testbed releases, making them available to other WPs.
- trace security problems in CrossGrid, and provide solutions,
- provide the final testbed release, demonstrate the success of the project with the deployment and evaluation of final user applications developed in WP1.

A basic assignment of resources will guarantee the setup installation and evolution at each testbed site, depending on its previous experience and complexity:

- Cyfronet (Cracow) (responsible: Andrzej Ozieblo)
- ICM (Warsaw) (responsible: Wojtek Wislicki)
- INS (Warsaw) (responsible: Krzysztof Nawrocki)
- UvA (Amsterdam) (responsible: G.Dick van Albada)
- FZK (Karlsruhe) (responsible: Marcel Kunze)
- II SAS (Bratislava) (responsible: Jan Astalos)
- PSNC (Poznan) (responsible: Pawel Wolniewicz)
- UCY (Cyprus) (responsible: Marios Dikaiakos)
- TCD (Dublin) (responsible: Brian Coghlan)
- CSIC (Santander & Valencia) (responsible: Santiago Gonzalez)
- UAB (Barcelona) (responsible: G. Merino)
- U.S.C. (Santiago) (responsible: Andreas Gómez Tato)
- Demokritos (Athenas) (responsible: Christos Markou)
- A.U.Th. (Thessaloniki) (responsible: Dimitrios Sampsonidis)
- LIP (Lisbon) (responsible: J.P. Martins)

Task 4.2 Integration with DataGrid (Month 1 - 36)

Task leader: Marcel Kunze, FZK (CR8)

The coordination with the DataGrid project will be crucial for the success of this project. The goal of this task is to enhance the cooperation and coordination with DataGrid. The main subtasks are to:

- study and push the compatibility of CrossGrid and DataGrid testbeds:
- coordination of authentication
- common resource access policies
- running environment compatibility
- libraries

- operating systems
- database management systems
- access to temporary mass storage
- command interpreters
- compilers and any required development tools to generate
- the executables
- keep in close contact with DataGrid in order to help in the coordinated evolution of both projects:
- coordinate possible demonstrations and Grid wide testing of selected application
- exchange knowledge and components
- coordinate the development process in order to prevent overlapping between both projects.

Success of this task will allow a common testbed for applications, in particular HEP applications, including DataGrid and CrossGrid testbed sites.

FZK will provide the main contact with DataGrid, while the other institutions will cover specific issues: TCD will cover performance/monitoring issues, adressed in DataGrid WP3, CSIC will be in contact for Data Management (DataGrid WP2) and Fabric Management (DataGrid WP4), UAB for resources and scheduling (DataGrid WP1), and LIP for Network Monitoring (DataGrid WP7).

Task 4.3 Infrastructure support (Month 1 – 36))

Task leader: Josep Salt, CSIC (CR15)

The aim is to provide testbed sites with easy and flexible fabric management tools and network support for the deployment of local grid infrastructures and their interconnection. The results from this task will be critical in order to help in the fabric set-up. This task comprises the following subtasks:

- definition of requirements and coordination with other WPs and DataGrid,
- study of available tools for fabric management for Grids,
- develop or adapt, test and provide flexible tools for fabric management in CrossGrid; simplify the coexistence of different software releases,
- create the Help Desk for assistance in local Grid infrastructure set-up process.

As the result of this task, a user friendly Grid installation kit will be provided for use in every testbed site.

CSIC will be responsible for this testbed kit release, and for maintaining the Help Desk. FZK will address mainly fabric management issues, in collaboration with U.S.C. UCY and A.U.Th will work on the impact of the changes related to new releases.

The success of the CrossGrid testbed will depend on the underlying network. CrossGrid will use the European (Géant) and national research network infrastructures to connect the computational and data resources. Since Géant is expected to provide services by mid 2001, we will use its infrastructure from the beginning of the Project. Network requirements will be reviewed in coordination with the corresponding national network providers, several of them being directly involved in this project. A corresponding promotion and development of the necessary network infrastructure for CrossGrid will follow. Note that explicit collaboration with other Grid projects, in particular with DataGrid WP7 (Network Services), carried in Task 4.2, will avoid the repetition of work already being done and complement other work, such as traffic monitoring and modelling (including performance issues addressed in WP2), and also security aspects. LIP will be the responsible for this point in this Task 4.3.

Task 4.4 Verification and quality control (Month 1 - 36)

Task leader: Jorge Gomes, LIP (AC20)

Verification and evaluation of results as intermediate goals will provide valuable feedback within the project. Creation of quality control mechanisms will help the detection of stacked issues and will provide tools to guarantee an effective use of the workpackage resources. This will be very important to avoid redundant or stopped work since this workpackage will have to combine efforts from many centres.

This task will contribute to the stability of the testbed by assuring the reliability of the middleware and network infrastructures required to develop and test applications. The stability and conformance of the testbed are essential to the smooth operation of the infrastructure and success of the application development and deployment activities, which require a near production facility.

Simultaneously applications will be reviewed and verified against the specifications, the testbed capabilities and middleware functionalities. This will allow feedback on the application implementation while preventing major testbed operational problems caused by incompatibilities between the applications and the middleware.

In particular, verification of middleware and applications conformance according with specifications and design rules will include:

- verification of testbed components before production release,
- verification of middleware and application interoperability,
- verification of policies and practices,
- independent review of applications,
- architecture review of major testbed releases.

LIP will coordinate this verification and quality control, where also Cyfronet, CSIC and Demokritos University will contribute, addressing specific issues: CSIC will review and provide feedback about the testbed setup, Cyfronet will take care of issues related to WP2 and WP3, while Demokritos will establish an independent review on the applications.

Workflow and Interfaces to other Workpackages

M 1-3: Detailed Testbed Planning Phase

Deliverable is D4.1 (report) at the end of M 3.

For all partners: Review the infrastructure (hardware and software), manpower and network resources (including contact with national Geant partners) available at each site. Define site expertise and special hardware infrastructure available. Prepare detailed planning for testbed installation.

For Task 4.0 (Testbed coordination and management): Coordinate report elaboration. Define the Integration Team. Propose collaborative tools. Confirm selected sites for first testbed setup.

For Task 4.1 (Testbed set-up and incremental evolution): Receive the requirements defined by the Architecture Team and by WP1/2/3. Define Security and Administrative policy. Prepare the basic infrastructures for testbed installation (Grid Security Infrastructure, Grid Information Service). Elaborate the testbed incremental project release procedure.

For Task 4.2 (Integration with DataGrid): Review DataGrid testbed situation. Establish participation in DataGrid WP1/WP5 meetings.

For Task 4.3 (Infrastructure Support) Provide expertise on hardware and network configurations. Define User support and Helpdesk requirements.

For Task 4.4: (Verification and Quality Control) Define verification and quality control procedures aiming to establish a validation process.

M 4-6: Setup of first testbed on selected sites

Deliverable is D4.2 (prototype) at the end of M 6.

For selected sites: Testbed installation: hardware installation, manpower allocation, network monitoring, implementation of authentication and authorisation procedures (CA & RA).

For Task 4.0 (Testbed coordination and management): Coordination of first testbed release with intensive Integration Team support.

For Task 4.1 (Testbed set-up and incremental evolution): First testbed project release: extend initial setup at CSIC (two nodes, Santander and Valencia) to FZK, CYFRONET, LIP, PSNC, UvA.

For Task 4.2 (Integration with DataGrid): Comparison with DataGrid testbed situation, and study the possibility of test of a simple application.

For Task 4.3 (Infrastructure Support): Prepare Installation Kit and Help Desk setup.

For Task 4.4: (Verification and Quality Control) Apply first verification procedures.

M 6-10: Experience with first testbed on selected sites and 1st prototype release

Deliverable is D4.3 (report) at the end of M 9 and D4.4 (prototype) in M 10.

For selected sites: Experience with testbed installation, and selective deployment of WP2 middleware. Report on experience with first testbed setup.

For all other sites: Prepare, under supervision of a selected site, for 1st testbed release.

For Task 4.0 (Testbed coordination and management): Coordination of middleware deployment from WP2.

For Task 4.1 (Testbed set-up and incremental evolution): Incremental evolution of testbed setup to follow needs of WP2 middleware. Report on policy procedures.

For Task 4.2 (Integration with DataGrid): Trace interoperability with DataGrid testbed. Report on previous simple application tests.

For Task 4.3 (Infrastructure Support): Update Installation Kit and extend Help Desk support. Report on experience.

For Task 4.4: (Verification and Quality Control) Extend verification procedures to deployed middleware.

M 11-21: Testbed evolution supporting application prototypes 0 and middleware development

Deliverables are D4.5 and D4.6(reports) at the end of M 15 and M 21

For all sites: Incremental testbed installation, and selective deployment of WP2 and WP3 middleware. Report on experience of testbed setup. Selective deployment of WP1 prototype 0 applications.

For Task 4.0 (Testbed coordination and management): Report on experience of testbed evolution to satisfy WP1/2/3 requirements.

For Task 4.1 (Testbed set-up and incremental evolution): Incremental evolution of testbed setup to follow needs of WP2 & WP3 middleware and WP1 applications.

For Task 4.3 (Infrastructure Support): Update Installation Kit. Consideration of special hardware requirements from application side.

For Task 4.4: (Verification and Quality Control) Apply verification procedures to deployed middleware and first application prototypes.

M 22-30: Testbed evolution to support application prototypes 1

Deliverable is D4.7 (report) at the end of M 30

For all sites: Incremental testbed installation to support deployment of WP1 prototype 1 applications, aiming to realistic computing and data management loads.

For Task 4.0 (Testbed coordination and management): Report on experience of testbed evolution regarding application prototype 1.

For Task 4.1 (Testbed set-up and incremental evolution): Incremental evolution of testbed setup to suit WP1 application prototype 1.

For Task 4.3 (Infrastructure Support): Update Installation Kit. Consideration of realistic requirements from application side.

For Task 4.4: (Verification and Quality Control) Apply verification procedures to deployed application prototype.

M 31-36: Final testbed with integration of applications.

Demonstration and detailed testbed reports including the final version of applications. Deliverable for all partners is D2.8 (prototype) at the end of M 33, followed by D4.9 (demo and report) in the final month M 36.

Technology required from other EU projects

As stated above, the CrossGrid testbed aims for full interoperability with the DataGrid testbed. Strong links with the DataGrid testbed exist, as several partners are participating in unfunded mode in this testbed, and will be reinforced to guarantee that the technology employed in their testbed setup is available to the CrossGrid testbed. In particular a very similar scheme will be adopted regarding operational policy issues.

Interaction with other CrossGrid WPs

The CrossGrid testbed will support not only final applications deployment on the GRID, but also the application and middleware development in Workpackages WP1, WP2 and WP3. The following table gives overview of all the partners' contributions to the Project software development and to the testbed support. The figures in the cells show the total PM assigned to each task.

Task	CYFRONET	ICM	INP	INS	UvA	II SAS	Univ. Linz	FZK	USTUTT	TUM	PSNC	UCY	DATAMAT	TCD	CS IC	UAB	U.S.C.	Demo	A.U.Th.	LIP	ALGO
1.0		3			11	3									3						
1.1	12				117		46														
1.2						73															
1.3			46	26				6							49	15					
1.4		65									12				22		28				
2.0								9													
2.1	6								3	8		6			3		6				
2.2									24						8						
2.3										6		22									
2.4	24									36					6		15				
2.5	8							4	4	18		8			3		6				
3.0											18										
3.1											100	14	16								21
3.2													18		10	30					
3.3	38	28												32							
3.4	40																				
3.5	20										20				6	10					
4.0															20						
4.1	30	24		20	12	12		25			18	24		12	16	12	21	27	40	12	
4.2								24						12	8	4				10	
4.3								24				6			16	3	20		34	9	
4.4	10														12			32		18	
5.1	84																				
5.2	44																				
5.3	4												3								33

Resources

Total and (funded) Person Months (PM).

Task	Task PM	CYFRONET	ICM	INS	UvA	II SAS	FZK	PSNC	UCY	TCD	CSIC	UAB	U.S.C.	Demo	A.U.Th	LIP
4.0	20 (20)										20 (20)					
4.1	305 (218)	30 (20)	24 (12)	20 (10)	12 (6)	12 (12)	25 (25)	18 (18)	24 (12)	12 (6)	16 (16)	12 (12)	21 (14)	27 (18)	40 (25)	12 (12)
4.2	58 (52)						24 (24)			12 (6)	8 (8)	4 (4)				10 (10)
4.3	112 (85)						24 (24)		6 (3)		16 (16)	3 (3)	20 (12)		34 (18)	9 (9)
4.4	72 (60)	10 (8)									12 (12)			32 (22)		18 (18)
Total PM	567	40	24	20	12	12	73	18	30	24	72	19	41	59	74	49
Funded PM	435	28	12	10	6	12	73	18	15	12	72	19	26	40	43	49

A minimum hardware required for the deployment of a local CrossGrid testbed site is the following:

- GIIS (Grid Index Information Service) machine,
- Dedicated 100 Mbps switch,
- Three machines (standard type, double processor Intel compatible >1GHz, 512Mb RAM, HD40GB),
- One CA (Certification Authority) machine per country (we plan to set up a CA policy similar to that of the DataGrid).
- Registration Authority web server and LDAP server (1 per country)
- Dedicated gatekeeper (1 per site)
- Dedicated network monitoring system (1 per site)

The total estimated cost of this hardware is approximately 140 kEuro, and it is reflected in the A4 form (cost summary). It is worth noticing that each local computer centre participating in the Project is providing additional funding to complete a reasonable local testbed site.

WP5 Project Management

Workpackage manager: Michal Turala, INP (AC3)

Objectives

In order to establish efficient management procedures it is necessary to take into account the consequences of:

- the relatively large number of Project partners (21) and high number of countries (11) participating in the Project,
- the strong correlation and mutual dependence of the Project workpackages and tasks, as their products must be integrated and implemented for appropriate operation of the international testbed,
- the dependence of the CrossGrid development on results obtained in the framework of other Grid projects in Europe, USA, and Japan.

The CrossGrid Project is large taking into account the number of participating institutions and the number of European countries where the Grid will be spread out. An important positive factor is that all CrossGrid Project participants have very good experience of large international research collaborations, so we are sure the Project will have a considerable impact on the research community in Europe. At the same time we are aware of the possible managerial problems.

The Project is divided into four main workpackages (WP1 – WP4) which deal with the research and technical aspects, and one workpackage dealing with the management and the dissemination. The technical and managerial responsibilities for the workpackages WP1 – WP4 lie primarily with their leaders (Work Package Managers – WPM), who are the top experts in their fields.

The Project Management (PM) will consist of:

- a Project Coordinator (PC) accompanied by the Project Secretary (PS) responsible for representation of the Project at the EU Brussels office, efficient long-term and daily administration of the Project, contacts with partners, conflict solving, and reporting,
- a CrossGrid Architecture Team (AT), responsible for technical co-ordination.

The technical co-ordination will focus on:

- definition of the overall CrossGrid architecture,
- establishing the rules for component transfer from other Grids and integration into CrossGrid,
- selection of appropriate software engineering development methodologies,
- definition of measures necessary for interoperability with other Grids,
- definition of quality assurance criteria and procedures,
- reviewing of the technical progress of the Project.

The above-mentioned actions will be elaborated by the Architecture Team (AT) and then accepted by the Project Steering Group (PSG).

The PSG, consisting of the Project Coordinator (PC), the managers of the workpackages (WPM), the leader of the CrossGrid Architecture Team (ATL) and a few additional experts, will be responsible for the overall project workplan, for the quality assurance, and for the arbitration of local conflicts (if arising). The Steering Group will meet regularly (approximately every two months), either in person or via phone/video-conference.

The Collaboration Board (CB), consisting of the representatives of each partner and chaired by a person elected at the beginning of the Project, will supervise the global progress of the project and will deal with strategic and contractual issues concerning the project. This body will meet yearly.

For efficient internal operations, and for interactions with external communities (public, potential users, professionals), the project partners and the project management will need a help of a small team, which could provide:

- organise and maintain CrossGrid Web pages, containing all essential collaboration information and materials (Intranet), and links to WWW pages of all CrossGrid partners, other Grid projects, and Grid technology working groups,

- collect and disseminate information on forthcoming Grid events, and to stimulate active participation of the CrossGrid partners,
- collect information on ongoing CrossGrid developments and to make them available in an appropriate form (including high quality public relation materials) for internal seminars, teaching) and/or public use (media),
- identify CrossGrid products, which could be of broader interests, and to look actively for potential customers
- help in organisation of yearly collaboration workshops.

To satisfy these needs the project management workpackage will contain also a dissemination-exploitation task, under the responsibility of industrial partners, Algosystems and DATAMAT.

Potential Risks

The CrossGrid Project addresses very advanced information technologies, and it concerns large international communities, and these may give rise to potential risks to its successful completion.

Technical Risks

The creation of a large integrated computing infrastructure, with a common middleware, open and available to every eligible person, is very appealing idea, but its realisation may encounter many conflicting requirements and interests. As there are several strong Grid communities and several Grid projects, possible divergences of their efforts could create a serious problem for the concept of a unified Grid. To minimise this kind of the risk we plan to work hand-to-hand with the other Grid projects (DataGrid, EuroGrid, GriPhyN, PPDG, and others), and to actively participate in the Global Grid Forum and the DataGrid Industry & Research Forum.

As Grid computing offers interesting opportunities for large number of commercial applications, it is very possible that industrial labs are already working on several ideas of our Project. We are sure such competition will bring the best end products to society.

The CrossGrid Project depends on results of the DataGrid Project, so a very close collaboration between them is of key importance. In daily operations the CrossGrid Architecture Team will be in close contact with the DataGrid Project team to watch their progress, gain experience, and obtain any necessary public domain results at the earliest possible time. In the case of changes to the DataGrid schedule/deliverables, the management of the CrossGrid Project will analyse the situation and take the necessary steps to overcome difficulties.

Managerial Risks

The CrossGrid Project is large, with many partners from many institutions and countries. The local centres have local plans and duties which they have to satisfy in a timely way, and therefore the new research activities may compete with the daily demands. We are sure that in this respect additional manpower provided by the CrossGrid Project will help to avoid potential conflict of interests.

The CrossGrid project addresses several, quite different applications. We are sure that the EC funding will play an essential role in allowing for recruitment of new personnel with the right profile, to complement the staff of the institutions involved, who will mainly concentrate on the applications.

Task descriptions

Task 5.1 Project coordination and administration (Months 1 – 36)

Task leader: Michal Turala, INP (AC3)

A Project Coordinator (PC) is appointed by Cyfronet, Cracow. The Project Coordinator will establish regular contacts with all partners of the Project and with the EU Project Office (EU PO). He/ she will follow the daily operations of the Projects, and define the administrative tools and procedures to be followed. He/she will set up the project Steering Group (SG) and the Collaboration Board (CB), and define the timetable and the agendas of the SG meetings. The Project Coordinator will be assisted by the Project Secretary (PS). The Project Office (POC), where all documents will be stored, will be established at Cyfronet, Cracow. The collaborative framework will be based on the use of Internet tools. A Project portal will be set up, simplifying access to the collaborative environment. The complete Project information will be updated on a daily basis, and WP web pages, including discussion fora and e-mail lists, will be created by Algosystems. It will include access to a videoconference virtual room (VRVS).

Task 5.2 CrossGrid Architecture Team (Months 1- 36)

Task leader: Marian Bubak, CYFRONET (CO1)

The definition and improvement of the CrossGrid architecture, which deployed on the final Grid framework, is of great importance. It will start with the definition of the application requirements and a selection of the necessary Grid middleware in the programming environment and the testbeds, and finally will validate it in the testbed environment. The quality assurance criteria and procedures will be defined, and a critical review of the technical progress of the Project will be periodically done and reported. The project architecture should ensure the overall coherence across the different WPs, identify parallel efforts, provide technical guidance, and contact with the adequate committees, fora, external consultants, for assessment of technical issues. The detailed comparison with the DataGrid architecture will also be discussed. The evolution of this architecture, taking into account the development of any new standards, is an obvious requirement. An Architecture Task Leader (ATL) is appointed at Cyfronet Cracow and he will set-up an Architecture Team, which will consist of local experts and representatives of the CrossGrid Work Packages.

The activity of the Architecture Team will be concentrated on

- Specification of the CrossGrid architecture,
- Merging of requirements from WP1, WP2, and WP3,
- Establishing standard operational procedures,
- Specification of the structure of deliverables.

Architecture of the CrossGrid

The most important factor which determines the success of the CrossGrid Project is the definition of its architecture. We are going to build it using results from other Grid projects, experience with component approach and HLA.

We start with a detailed survey of existing technology for integration of the software. First, we plan to review existing Globus and DataGrid architecture.

As the second, we plan to analyse the recent development of the concept of component approach, namely the Common Component Architecture (CCA). The important feature of this approach is that it addresses the needs of high performance community for distributed and parallel scientific computing. This activity is organised by the CCA Forum [CCAFor], and the last implementation called CCA Toolkit was designed for Globus based computational grids. When using such a component oriented programming environment, components can be connected, disconnected, and hot-wired into an application running across globally distributed resources. It seems that component model is natural for Grid computing.

The next step will be devoted to analysis of the applicability for CrossGrid of the High Level Architecture (HLA) [HLA] which was designed to promote standardisation in modelling and simulation and to facilitate the reuse of components. HLA is defined by federation rules, the HLA interface, and the Object Model Template. HLA allows data distribution across heterogeneous computing platforms.

Requirements merge

Since WP1, WP2 and WP3 have planned to complete their applications descriptions and requirement definition after 3rd month of work, Architecture Team will set up requirements merge and refinement activity starting point to this time point. The requirements review with participation of WP leaders should take place in order to align mutual expectations and to verify requirements testability and completeness. As a result, 1st requirements baseline will be established. For each defined requirement, the appropriate acceptance criteria (test case) will be defined to facilitate later validation of a given functionality. Each such test case set will be also baselined as initial reference point for correctness proof. Once requirements will be specified, WP4 may prepare testbed infrastructure proposal with both hardware and software components.

Management and coordination will be concentrated on requirements changes and refinement, integration and testing schedule. Integration and testing schedule will be developed based on schedules of WPs. In general, WPs will realise the unit testing of their software and WP4 is expected to coordinate integration and system tests.

Detailed workplan of the Architecture Team will be determined by initial requirements baseline date (then first architecture sketch may be prepared), and then it will depend of number and kind of problems raised (see below for PR explanation).

Establishing details of the Standard Operational Procedures (SOPs)

SOPs will be established according to common practice of software development [IEEE828, IEEE1058] and will consist of:

1. **Project Repository Access procedure**

   The central repository server to keep baselined software versions should be designated. All project artifacts (technical documents, design documents, code, tests) should be kept under version control. Appropriate version control software shall be selected. Evaluation of existing tools (CVS, Rational ClearCase, Microsoft SourceSafe) should be conducted before the final decision will be taken. Rules and constraints concerning repository access as well as release plan will be described in project Software Configuration Management Plan (SCMP).

2. **Handling Problem Reporting (PR) mechanism**

   It is obvious that all project components will be continuously changed. Each project element change has to be documented. Before each change will be committed to project repository (and become visible for other WPs) the formal PR has to be raised in order to analyse impact of proposed modification.

3. **Handling Change Requests (CR) mechanism**

   After a PR is accepted the CR has to be generated. The repository modifications can be made only based on designated CR. It will eliminate uncontrolled changes in project components. Each CR should be reviewed before implementing it.

4. **Release Preparation procedure**

   Each formal release should be described in SCMP. Release plan should be prepared to fit particular WPs planned application prototypes build dates. For a given release a description of implemented features should be prepared as well as test suite for each of them.

Structure of deliverables

Each work package that is going to produce deliverables containing API to another work packages or using other APIs should will be obliged to:

- Deliver also detailed specifications for any public package they provide or require. The term “public package” refers to any API, protocol, service, etc. intended to be used by other work package software or by applications; it will help to resolve potential conflicts concerning object names, available ports, etc.
- Conform to defined coding standards. Such standard will be proposed and accepted from every programming language used in the framework of the Project.
- Deliver detailed specification of any public resources that they use (e.g. port numbers) which may conflict with other software.
- Specify any required operating system, architecture, compiler, etc.

Task 5.3 Central dissemination (Months 1- 36)

Task leader: Yannis Perros, ALGO (AC21)

The main objective of task 5.3 is to assist the CrossGrid partners, and the project management, in their dissemination duties. This concerns internal CrossGrid communication and operations, as well as interactions with other communities.

The following sub-tasks are defined:

5.3.1 Organisation of CrossGrid Web pages, which will contain all essential information and materials concerning the CrossGrid project: participants, structure, mailing lists, documents, meetings, discussion fora, timetables, minutes, etc. CrossGrid Web pages should provide sufficient information on the project and its partners – corresponding links to be installed. Some of the information, related to internal operations, discussion fora, drafts, etc. will have restricted access (Intranet).

The other function of the CrossGrid Web pages will be to facilitate access to the information of the other Grid projects – corresponding links will be installed. It is planned to inform CrossGrid partners on newly installed links and information.

An area dedicated to users, the CrossGrid Vortal, which will provide specific Grid (and in particular CrossGrid) information and software for accessing CrossGrid services, will be created.

The CrossGrid Web pages need to be maintained for the life of the project. An easy and stable access has to be granted (including a server back up and mirroring). It is planned to follow the experience of the other Grid projects, in particular DataGrid, which Web pages (http://www.eu-datagrid.org/) seem to be satisfying all practical needs. Algosystems will be responsible for this subtask; a prototype design of CrossGrid WWW Pages already exists (http://www.crossgrid.org/) and it is in use by the consortium.

5.3.2 Provide prompt information on ongoing and forthcoming events related to Grid technologies, and to stimulate CrossGrid partners in an active participation. This applies to all kind of activities: new projects, workshops, conferences, working groups, etc. The responsible partner will collect such information via dedicated e-mail box, and in addition it will perform an "active search", to find relevant news. It is expected, that good information channel will be provided by an accompanying action project, GridStart, which involves nine coordinators of European Grid projects, including Cyfronet, the CrossGrid coordinating partner. The relevant information will be posted on the CrossGrid Web pages, but in addition the CrossGrid management, including workpackage and task leaders, will be informed on events of particular interest. This subtask will be performed by Algosystems.

5.3.3 Collect results of the CrossGrid project, which are of wider/general interests, and to prepare them in a form appropriate for external presentations. It is planned to publicise the concept of the Grid – especially in new countries, and the CrossGrid applications – for this purpose high quality PR materials will be prepared (a CrossGrid leaflet). There will be also a need to prepare more elaborated materials for the communities of potentially interested: students, researchers, public administration, etc. – such materials could be assembled as a set of presentations, which could be used by every partner of the consortium, when at need. This sub-task will be done by Algosystems.

5.3.4 Identify CrossGrid products, which could be a broader/ general interest, and to promote them. We believe that CrossGrid special software and tools will be attractive for some professionals, and that the CrossGrid application will be of interest of governmental agencies, hospitals, research communities - an active promotion of our developments will help to transfer this information to the right people. Participation in the Industry&Research DataGrid Forum will provide an opportunity. We also plan that CrossGrid industrial presentations will accompany some international conferences and fairs. CrossGrid industrial partner, DATAMAT, already participating in the DataGrid project, will be responsible for such actions.

5.3.5 The CrossGrid consortium will organise yearly collaboration meetings, with a main purpose to overview the progress of the project. In addition, in the first year we plan to have a "kick-off" meeting. These meetings will be defined by the management, however they will require some preparatory (logistic) and post-meeting work (collection of materials, summaries, etc.). The necessary support will be provided by Algosystems.

Number	Description	Responsible partner	Subtasks	Delivery date	Nature	Dissemination level
D5.3.1	CrossGrid Web pages	ALGO	5.3.1	M3	Prototype	PU
D5.3.2	Dissemination and exploitation plan	ALGO DATAMAT	5.3.1 – 5.3.5	M6	Report	PP
D5.3.3 – D5.3.7	Public (industrial) presentations	ALGO DATAMAT	5.3.2	M6, M12, M18, M24, M30	Event	PU
D5.3.8	CrossGrid Website	ALGO	5.3.1	M9	Prototype	PU
D5.3.9 – D5.3.11	Project leaflet/brochure	ALGO	5.3.3	M11, M23, M35	Other	PU
D5.3.12 – D5.3.14	Dissemination and exploitation report	ALGO DATAMAT	5.3.1 – 5.3.5	M11, M23, M35	Report	PP
D5.3.15 – D5.3.17	Yearly collboration meeting	ALGO	5.3.5	M11, M23, M35	Event	PU
D5.3.18	CD-Rom	ALGO	5.3.3	M30	Demo	PU

The central dissemination task complements the activities of all CrossGrid partners in this domain. The work will be done by the CrossGrid industrial partners, Algosystems and DATAMAT, together with CYFRONET, the project coordinator.

Workflow and Interfaces to other Workpackages

The project coordination and administration must follow closely the progress of the project. The pace will be defined by regular Steering Group meetings (monthly), EU Quarterly Monitoring Reports (QMR) and financial statements, and EU Yearly Reviews and Status Reports (YST).

The Architecture Team has a key role in timing of the project, checking its milestones and deliverables. Detailed workplan of the Architecture Team will be determined by the initial requirements definition (then first architecture sketch may be prepared), and it will depend of number and kind of problems raised (see below for PR explanation).

M 1-3: Requirements definition and merging phase

Deliverable for all partners is D5.16 (report) at the end of M 3. This phase merges all requirements (from WP1, WP2, and WP3) and specifies the structure of the CrossGrid software. This is done in close collaboration with managers and persons responsible for software integration.

M 4-12: 1st development phase

Deliverables for all partners are D5.17 (report) at the end of M 6, and D5.19 (report) at M 12. As the result, first full specification of the CrossGrid architecture as well as operational procedures and the structure of deliverables will be available to all partners.

M 13-24: 2nd development phase

In this phase there will be incremental improvement of specifications of the previous phase basing on the feedback from WPs. Deliverable for all partners are D5.19 (report) at the end of M 15 and D5.20 (report) at the end of M 24, so the complete specification of the CrossGrid architecture will be defined.

M 25-33: 3rd development phase

Deliverable for all partners is D5.21 (report) at the end of M 33.

M 33-36: Finishing phase

Documentation of CrossGrid architecture in the final version.

The dissemination task is associated with the project management and coordination. This position provides continuous and close contacts with all CrossGrid workpackages, their respective WP and task leaders, and guarantees the best access to all relevant information. The workplan of the central dissemination team will follow, on a local scale, the developments of the CrossGrid project, and all important Grid events, on a global scale.

Collaboration with other CrossGrid WPs

Architecture Task requires a strong interaction mainly with WP1, WP2, WP3 and WP4 during the all phases of the Project realisation.

The internal and external dissemination of the CrossGrid developments, progress and results can only work if close relation with the developers are maintained. It is planned that the dissemination task leader will regularly participate in the CrossGrid steering Group.

Technology required from other Grid projects

Close, almost daily collaboration with the DataGrid Architecture Task Force is necessary to reach full integration with the DataGrid testbeds. The Architecture Team will be in contact with other Grid projects in order to introduce common standards in the Grid technology. Active participation in the Grid Forum is a must.

The CrossGrid partners responsible for central dissemination must follow the developments of the other Grid projects. It is supported by Cyfronet being a partner of the GridStart consortium, and DATAMAT being a member of the DataGrid project.

Resources

Total and (funded) Person Months (PM).

Task	Task PM	CYFRONET	DATAMAT	ALGO
5.1	84 (42)	84 (42)		
5.2	44 (22)	44 (22)		
5.3	40 (38)	4 (2)	3 (3)	33 (33)
Total PM	168	132	3	33
Funded PM	102	66	3	33

9.6 Short Descriptions of Workpackages

Each of the workpackages is shortly described below. The effort required to complete each of the tasks is indicated in a table at the end of each description. Both funded and not funded efforts are given (the EU funded effort in brackets). Throughout this section the acronym *PM* means *Person Months*.

Workpackage description - CrossGrid Application Development

Workpackage number:	1	**Start date or starting event:**									Project start
Participant:	CYFRONET	ICM	INP	INS	UvA	II SAS	Univ. Linz	FZK	PSNC	CSIC	UAB	U.S.C.
PM per participant (funded)	12 (6)	68 (34)	46 (23)	26 (13)	128 (72)	76 (76)	46 (23)	6 (6)	12 (6)	74 (74)	15 (15)	28 (17)

Objectives

Data gathering, processing and interpretation in health and environment is driven by modern technology. The geographical distribution of data generators (Medical scanners, Patient Databases, Environment-Actuators and Monitors) requires and integrated approach to access and process these data. With the advent of Grid-based technology many new possibilities for data presentation and fast decision making come within reach. It is the goal of the CrossGrid project to explore, incorporate adapt and validate this technology for those application areas that require interactive access to resources, be it Databases, Supercomputers, Visualisation Engines, Medical Scanners, or environmental data input devices.

Decision making processes stemming from many health and environmental problems require integration of distributed data and near real-time interaction. The objectives of this workpackage are to provide a representative collection of sample applications from various fields that will drive and exploit the specific (interactive) functionalities to be developed in the CrossGrid project. Firstly, through their need for specific services and their performance characteristics, they will provide a driving force for the technology-oriented workpackages. Secondly, they will serve as a benchmark for the performance of the those workpackages, and finally, they will serve as demonstrators, for their respective fields, of the added value provided by the Grid in general, and the technology developed by the CrossGrid project in particular. Each task in this workpackage will focus on the development of one of these applications. Together, they cover a wide range of final user communities, from health care to environment management, and basic research.

Description of work

Task 1.0 Co-ordination and management

Task 1.1 Interactive simulation and visualisation of a biomedical system

Task 1.2 Flooding crisis team support

Task 1.3 Distributed data analysis in High Energy Physics

Task 1.4 Weather forecast and air pollution modelling

Deliverables

Number	Description	Responsible partner	Sub-task	Delivery month	Nature	Dissemination level
D1.0.1	Joint requirements definition document	UvA	1.0.2	3	Report	CO
D1.1.1	Application description, including use cases for task 1.1	UvA	1.1.1, 1.1.2, 1.1.3	3	Report	PU
D1.2.1	Application description, including use cases for task 1.2	II SAS	1.2.1, 1.2.2, 1.2.3	3	Report	PU

D1.3.1	Application description, including use cases for task 1.3	CSIC	1.3.1, 1.3.2, 1.3.4	3	Report	PU
D1.4.1	Detailed planning for air/sea pollution application including use cases	ICM CSIC U.S.C	1.4.1 – 1.4.4	3	Report	PU
D1.0.2	Semi-annual managerial report for WP1	UvA	1.0.1	6	Report	CO
D1.0.3	Extended yearly managerial report for WP1	UvA	1.0.1	12	Report	CO
D1.1.2a D1.1.2b	Internal progress report First internal software release	UvA	1.1.1 - 1.1.4	12	Report SW	CO
D1.2.2a D1.2.2b	Internal progress report First internal software release	II SAS	1.2.1 - 1.2.4	12	Report SW	CO
D1.3.2a D1.3.2b	Internal progress report First internal software release	CSIC	1.3.1 1.3.4	12	Report SW	CO
D1.4.2	Results of migration of data mining algorithms to GRID structure	ICM CSIC	1.4.2	12	Report	CO
D1.0.4	Intermediate report on coherency	UvA	1.0.3	18	Report	CO
D1.0.5	Semi-annual managerial report for WP1	UvA	1.0.1	18	Report	CO
D1.1.3	Report on application prototype 0 run	UvA	1.1.4	18	Report	CO
D1.2.3	Report on application prototype 0 run	II SAS	1.2.4	18	Report	CO
D1.3.3	Report on application prototype 0 run	CSIC	1.3.3	18	Report	CO
D1.4.3	Integration of parallel codes for air quality models into GRID structure Prototype A phase 0	U.S.C.	1.4.4 1.4.5	18	Report + Proto-type	CO
D1.0.6	Extended yearly managerial report for WP1	UvA	1.0.1	24	Report	CO
D1.1.4a D1.1.4b	Internal progress report Second internal SW release	UvA	1.1.1 - 1.1.4	24	Report SW	CO
D1.2.4a D1.2.4b	Internal progress report Second internal SW release	II SAS	1.2.1 - 1.2.4	24	Report SW	CO
D1.3.4a D1.3.4b	Internal progress report Second internal SW release	CSIC	1.3.1 - 1.3.4	24	Report SW	CO
D1.4.4	Integration of distributed atmosphere and wave models into GRID structure). Prototype B phase	ICM	1.4.3 1.4.5	24	Report + Proto-type	CO
D1.0.7	Semi-annual managerial report for WP1	UvA	1.0.1	30	Report	CO
D1.1.5	Report on application prototype 1 run	UvA	1.1.4	30	Report	CO

D1.2.5	Report on application prototype 1 run	II SAS	1.2.4	30	Report	CO
D1.3.5	Report on application prototype 1 run	CSIC	1.3.3	30	Report	CO
D1.4.5	Prototype A runs (phase 1) (integration of parallel codes for air quality models into GRID structure).	U.S.C	1.4.4 1.4.5	30	Report + Proto-type	CO
D1.4.6	Prototype B runs (phase 1) (integration of distributed atmosphere and wave models into GRID structure).	ICM	1.4.3 1.4.5	30	Report + Proto-type	CO
D1.0.8	Final report on coherency	UvA	1.0.3	36	Report	CO
D1.0.9	Final managerial report for WP1	UvA	1.0.1	36	Report	CO
D1.1.6	Application final demonstration and report	UvA	1.1.4, 1.1.5	36	Report, Demo	PU
D1.2.6	Application final demonstration and report	II SAS	1.2.4	36	Report, Demo	PU
D1.2.7	Final report on dissemination and exploitation	II SAS	1.2.5	36	Report, Demo	Confidential
D1.3.6	Application final demonstration and report	CSIC	1.3.3, 1.3.4	36	Report, Demo	PU
D1.4.7	Application final demonstration and report	ICM CSIC U.S.C	1.4.5	36	Report, Demo	PU

Milestones[9] and expected results		
Milestone	Month	Results
M1 - 1	3	Application descriptions and use cases for all applications Requirements to other workpackages Identification of principal common tasks
M1 - 2	12	First software releases
M1 - 3	18	First prototypes up and running on local testbed
M1 - 4	24	Second software releases
M1 – 5	30	Second prototypes up and running on global testbed
M1 - 6	33	Final prototypes ready for testing on distributed testbeds

[9] Milestones are control points at which decisions are needed; for example concerning which of several technologies will be adopted as the basis for the next phase of the project.

Workpackage description – Grid Application Programming Environment

Workpackage number:	2	Start date or starting event:			Project start	
Participant:	CYFRONET	FZK	USTUTT	TUM	UCY	CSIC	U.S.C.
PM per participant:	38 (19)	13 (13)	31 (31)	68 (34)	36 (22)	20 (20)	27 (17)

Objectives

Specify, develop, integrate and test the tools that facilitate the development and tuning of parallel distributed, compute and data intensive, interactive applications on Grid.

Description of work

Task 2.0 Co-ordination and management. Ensure the seamless integration of the interfaces and software components

Task 2.1 Tools requirements definition. Define the needs of all classes of users, review existing grid monitoring services and tools

Task 2.2 MPI code debugging and verification. Develop a tool that verifies the correctness of parallel, distributed Grid applications using the MPI paradigm

Task 2.3 Metrics and benchmarks. Develop and implement benchmarks representative for Project applications to estimate values of performance metrics

Task 2.4 Interactive and semiautomatic performance evaluation tools. Develop on-line tools to measure, evaluate and visualise the performance of Grid applications

Task 2.5 Integration, testing and refinement. Implement, test and refine all software components produced by this WP

Deliverables

D2.1 (Report) Month 3: General requirements and detailed planning for programming environment

a) requirements on the hardware and software infrastructure for tools development and testing,

b) requirements of end user applications,

c) interfaces to Grid monitoring services and definition of a performance data model,

d) specification of performance metrics and benchmarks,

e) review of state-of-the-art and current related techniques,

f) definition of the architecture of the application programming environment

D2.2 (Report) Month 6: WP2 internal progress report

a) design of interfaces between tools,

b) design of MPI code verification tool,

c) design of metrics and benchmarks,

d) design of performance analysis tool

D2.3 (Prototypes and reports) Month 12: Demonstration and report on WP2 1st prototypes

a) test scenarios, evaluation suite,

b) documentation of design, implementation, and interfaces

D2.4 (Report) Month 15: Internal progress report on WP2 software evaluation and testing

a) feedback from applications, new requirements,

b) detailed integration plan

D2.5 (Prototypes and report) Month 24: Demonstration and report on WP2 2nd prototypes

a) test scenarios, evaluation suite

b) documentation of design, implementation, and interfaces

D2.6 (Report) Month 27: Internal progress report on WP2 software evaluation and testing

a) feedback from applications

b) detailed integration plan

D2.7 (Demo and report) Month 36: Demonstration and documentation on the final versions and testing of all WP2 software components

Milestones[10] and expected result

M2.1 Month 3: Definition of requirements

M2.2 Month 12: 1st prototype – separate tools

a) of Task 2.2 contains the subset of MPI calls required by the end user applications and is running on a local environment;

b) of Task 2.3 contains the benchmarks running on a local environment

c) of Task 2.4 contains the instrumentation software and local monitoring modules; for homogeneous platform, with few examples of specialized properties automatically extracted, selected visualizers

M2.3 Month 24: 2nd prototype – concurrently running tools, no interactions

a) of Task 2.2 is running on the Grid environment

b) of Task 2.3 contains the 2nd prototype of benchmarks running on the Grid

c) of Task 2.4 includes the upper layer of the monitoring system and the Grid-oriented tools; generic Grid implementation of automatic assessment of selected performance properties; visualisation of performance data collected on the Grid level

M2.4 Month 33: Final version – fully integrated, interoperable set of tools; full functionality, including analytical model

a) of Task 2.2 supports the full MPI 1.2 standard and is running on the Grid

b) of Task 2.3 of benchmarks running on the Grid

c) of Task 2.4 of performance monitoring tools for applications running on the Grid

[10] Milestones are control points at which decisions are needed; for example concerning which of several technologies will be adopted as the basis for the next phase of the project.

Workpackage description – New Grid Services and Tools

Workpackage number :	3		**Start date or starting event:**		Project start			
Participant:	CYFRONET	ICM	PSNC	UCY	DATAMAT	TCD	CSIC	UAB	ALGO
Person-months participant:	98 (49)	28 (14)	138 (61)	14 (7)	34 (34)	32 (16)	16 (16)	40 (40)	21 (21)

Objectives

The main objective of this workpackage is to develop, test and deploy the new Grid services and technologies that will enable to develop the Grid-infrastructure and run the Grid enabled applications on it. Emphasis will be put on the user-friendly environment based on the portal access to the Grid independently of the user location. The mechanism allowing one to enter the Grid environment at any place and from any hardware will also be developed. Another task addresses the construction of new resource management techniques, namely the design and evaluation of self-adaptive scheduling agents that will be responsible for scheduling parallel applications submitted to a Grid. There is a task with a proposal of developing a prototype infrastructure mainly for the needs of monitoring-related activities for automatic extraction of high-level performance properties and for tool support of performance analysis. In order to help users in their interaction with large data manipulation, a kind of expert system is proposed, which will operate as an advisor for the selection of migration/replication policy in a particular environment. Within this task a common middleware for fast tape file access and a system for data access time estimation will also be developed.

Description of work

Task 3.1: Development of user portals customised to particular applications and implementation of a roaming access which allow users to access the Grid from different locations and different platforms.

Task 3.2: Development of a Grid resource management system based on self-adaptive scheduling agents.

Subtask 3.2.1: New technologies for the Grid resource management

Subtask 3.2.2: Monitoring for purposes of the Grid resource management

Task 3.3: Development of Grid monitoring services for various tools support and for Grid infrastructure monitoring.

Subtask 3.3.1: Autonomous monitoring system for on-line and automatic performance analysis

Subtask 3.3.2: SQL-query-based tool support and interfaces to Grid application programming environment

Subtask 3.3.3: Grid Jiro-based services for Grid infrastructure monitoring

Task 3.4: Development of an expert system that will operate as an advisor for the selection of migration/replication policy in defined user states in order to optimise data access.

Subtask 3.4.1: Expert system for data management

Subtask 3.4.2: Common middleware for fast tape file access

Subtask 3.4.3: Data access time estimation system

Deliverables

D3.1 (Report) Month 3: Detailed planning for all the tools and services including use cases for WP3

a) requirements of end user,

b) definition of the monitoring environment functionality and of interfaces to Task 2.4

c) specification of GRM agents

d) review of state-of-the-art and current related techniques

D3.2 (Report) Month 6: WP3 internal progress report

a) architecture of the New Grid Services and Tools

b) design of interfaces between tools

c) security issues – proposal of standard

D3.3 (Prototype) Month 12: 1st prototype – first WP3 software release for health and HEP applications build on local grid infrastructure.

a) test scenarios, evaluation suite,

b) documentation of design, implementation, and interfaces

D3.4 (Report) Month 15: Internal progress report on WP3 software evaluation and testing

a) feedback from applications, new requirements

b) detailed integration plan

D3.5 (Prototype, Report) Month 18: Report on the results of the WP3 2nd prototype for Task 3.3

a) homogeneous mobile access

b) local monitors for Grid monitoring

D3.6 (Prototype, Report) Month 24: Report on the results of the WP3 2nd prototype for Tasks 3.1, 3.2, 3.4

a) applications run in distributed grid infrastructure

b) test scenarios, evaluation suite

c) documentation of design, implementation, and interfaces

d) integration of Globus Resource Management with Task 2.4

D3.7 (Report) Month 27: Internal progress report on WP3 software evaluation and testing

a) feedback from applications, new requirements

b) detailed integration plan

D3.8 (Prototype, Report) Month 30 : Report on the results of WP3 prototype 1 final version run

D3.9 (Demonstration, Report) Month 36: WP3 final report

Milestones[11] and expected result

M3.1 Month 3: Definition of requirements

M3.2 Month 12: 1st prototype (running on local grid) for health and HEP applications

a) Task 3.1 batch application running
b) Task 3.2 introduces self-adaptive scheduling agents
c) Task 3.3 comprises separate components: monitors, managers
d) Task 3.4 analytical issue of expert system for migration strategies

M3.3 Month 18: 2nd prototype for WP3.3

Task 3.3 integrates components: monitors, managers as middleware

M3.4 Month 24: 2nd prototype for WP3.1, WP3.2, WP3.4 running on the international CrossGrid testbed

a) Task 3.1 application runs on distributed grid
b) Task 3.2 implements interfaces of scheduling agents
c) Task 3.4 co-operates with portal described in Task 3.1

M3.5 Month 33 : Final deployment in CrossGrid, tasks are integrated:

a) Task 3.1 supports user environment on the heterogeneous Grid,
b) Task 3.2 scheduling agents for parallel jobs, tested with interaction with end users in WP1
c) Task 3.3 integrates monitoring environment
d) Task 3.4 implements an expert system for accessing data; implements mass-storage-system-independent middleware
e) Task 3.5 tests of the solution

[11] Milestones are control points at which decisions are needed; for example concerning which of several technologies will be adopted as the basis for the next phase of the project.

Workpackage description - International Testbed Organisation

Workpackage number :	4	**Start date or starting event:**	Project start					
Participant:	CYFRONET	ICM	INS	UvA	II SAS	FZK	PSNC	UCY
PM assigned (funded)	40 (28)	24 (12)	20 (10)	12 (6)	12 (12)	73 (73)	18 (18)	30 (15)
Participant:	TCD	CSIC	UAB	U.S.C.	Demo	A.U.Th.	LIP	
PM assigned (funded)	24 (12)	72 (72)	19 (19)	41 (26)	59 (40)	74 (43)	49 (49)	

Objectives

Provide a distributed resource facility where the WPs developments can be tested in a grid environment. This package will assure the integration in testbeds of the applications. It will allow one to achieve a realistic proof of the success of the grid concept, showing how the developed applications take advantage of the CrossGrid infrastructure.

Description of work

Task 4.0 Testbed coordination and management

Task 4.1 Testbed set-up and incremental evolution

Task 4.2 Integration with DataGrid

Task 4.3 Infrastructure Support

Task 4.4 Verification and quality control

Deliverables

D4.1 (Report) Month 3: Detailed Planning for Testbed Setup (CSIC)

Infrastructure, manpower and network resources review at each site (ALL)

Requirements defined by Architecture Team and platform support policy (CYFRONET)

Application & Middleware requirements from other CrossGrid WP:

WP1 (UvA), WP2 (FZK), WP3 (PSNC)

Definition of site expertise and special hardware infrastructure (massive storage, specific platform…) (CSIC)

Interoperability and coordination with DataGrid (FZK)

Security and administrative policy (CSIC)

Testbed incremental project releases procedure (CSIC)

Integration process definition and Integration Team setup (CSIC)

Validation process definition (LIP)

User support and Helpdesk requirements (CSIC)

Description of proposed collaborative tools (CSIC)

D4.2 (Prototype) Month 6: First testbed set-up on selected sites

Procedures as detailed in D4.1

Deployable software will include (in due time): WP1 application prototypes, WP2 & WP3 distributed tools & middleware, plus a DataGrid simple application example to test interoperability.

Definition of extra-support at selected sites for deployable software.

Installation kit and documentation web site (CSIC)

Setup of national CA,RA, monitoring system for this first testbed.

Setup of repository software and HelpDesk at CSIC.

Testbed evolution: initial setup at CSIC, one-by-one extension policy reaching all selected sites; list will include at least all CR (CYFRONET, UvA, FZK, PSNC) plus validation site (LIP).

D4.3 (Report) Month 9: WP4 status internal report (CSIC)

General report on the testbed setup status with specific sections on:

First testbed experience review

Integration team experience

User support and HelpDesk

Progress report at each tesbed site

Perspectives of deployment of applications & middleware

Interoperability with DataGrid

Validation procedures

Collaborative tools experience

Preparations for testbed extension:

Installation kit revision

Requirements from other WP and deployable CrossGrid software update

Perspectives at each site

Definition of site expertise and special hardware infrastructure (massive storage, specific platform…)

Network requirements

D4.4 (Prototype) Month 10: 1st Testbed prototype release (ALL)

Revised released procedures

Deployable software will include that available at selected testbed sites

Installation kit and documentation web site update (CSIC)

Testbed evolution: initial setup at selected sites, one-by-one extension policy from these selected sites reaching all testbed sites.

Setup of national CA,RA, LDAP servers and monitoring system for the extra sites

Update of repository software and HelpDesk at CSIC.

D4.5 (Report) Month 15 : WP4 status internal report (CSIC)

General report on the 1st testbed prototype release and evolution with specific sections on:

Testbed experience review

Integration team experience

User support and HelpDesk

Progress report at each tesbed site

Deployment of applications & middleware

Interoperability with DataGrid

Validation procedures

Collaborative tools experience

Preparations for next testbed releases

Installation kit revision

Requirements from other WP and deployable CrossGrid software update

Perspectives at each site

Requirements on expertise and special hardware infrastructure (massive storage, specific platform…)

Network use

D4.6 (Report) Month 21 : WP4 status internal report update(CSIC)

Update of D4.5 with additional emphasis on:

DataGrid interoperability (project ends Dec.2003) and next steps (CERN-GRID)

Experience with first deployed applications and middleware

Status regarding successive incremental testbed releases

D4.7 (Report) Month 30: WP4 status internal report (CSIC)

Update of D4.6 with special sections on:

Planning for the final testbed

Applications detailed requirements and impact on network QoS

Final feedback to applications and middleware development

D4.8 (Prototype) Month 33: Final testbed with all applications integrated (ALL)

Deployed software will include: WP1 final applications, WP2 & WP3 distributed tools & middleware.

Final installation kit and documentation web site, software repository and HelpDesk (CSIC)

D4.9 (Demo and Report) Month 36: WP4 final demo and report (CSIC)

Milestones[12] and expected result

M4.1 Month 3: Testbed setup plan

M4.2 Month 6: 1st testbed on selected sites

M4.3 Month 10: 1st CrossGrid testbed release

M4.4 Month 33: Final testbed

[12] Milestones are control points at which decisions are needed; for example concerning which of several technologies will be adopted as the basis for the next phase of the project.

Workpackage description - Project Management

Workpackage number:	5	**Start date or starting event:**	Project start	
Participant number:		CYFRONET	DATAMAT	ALGO
Person-months per participant:		132 (66)	3 (3)	33 (33)

Objectives

The long-term decisions, risk management, daily Project administration, the technical coordination, and central dissemination activities.

Description of work

Task 5.1 Project coordination and administration (Months 1 - 36) – daily management, periodic reports, conflict resolution.

Task 5.2 CrossGrid architecture team (Months 1 - 36) - specification of the CrossGrid architecture, merging of requirements from WPs, establishing standard operational procedures, specification of the structure of deliverables.

Task 5.3 Assisting the CrossGrid partners and the project management in their dissemination and exploitation duties: internal CrossGrid communication and operations, interactions with other communities.

Deliverables

D5.1.1 – D5.1.12 (Report) Quarterly reports (each 3 months)

D5.2.1 (Report) Month3: Quality assurance plan

D5.2.2 (Report) Month 3: The CrossGrid architecture requirements and the first definition of the architecture

D5.3.1 (Prototype) Month 3: CrossGrid Web pages

D5.2.3 (Report) Month 6: Full description of the CrossGrid standard operational procedures and specification of the structure of deliverables.

D5.3.2 (Report) Month 6: Dissemination and exploitation plan

D5.3.3 – D5.3.7 (Event) Months 6, 12, 18, 24, 30: Public (industrial) presentations

D5.3.8 (Prototype) Month 9: CrossGrid Wesbite

D5.3.9 – D.5.3.11 (Other) Months 11, 23, 35: Project leaflet / brochure

D5.3.12 – D5.3.14 (Report) Months 11, 23, 35: Dissemination and exploitation report

D5.3.15 – D5.3.17 (Event) Months 11, 23, 35: Yearly collboration meeting

D5.1.13 – D5.1.15 (Report) Yearly reports

D5.2.4 (Report) Month 12: Report on requirements for integration and interoperability with DataGrid

D5.2.5 (Report) Month 15: Detailed report on the CrossGrid architecture

D5.2.6 (Report) Month 24: Current CrossGrid architecture assessment report

D5.3.18 (Demo) Month 30: CD-Rom

D5.2.7 (Report) Month 33: Final CrossGrid architecture and interoperability requirements description

Milestones[13] and expected result
M5.1 Month 12: Successful completion of 1st EU review
M5.2 Month 24: Successful completion of 2nd EU review
M5.3 Month 36: Successful completion of 3rd EU review

[13] Milestones are control points at which decisions are needed; for example concerning which of several technologies will be adopted as the basis for the next phase of the project.

9.7 Compiled list of deliverables[14]

Deliverable No.	Deliverable title	Delivery month	Nature	Dissemination level
D1.0.1	Joint requirements definition document	3	R	CO
D1.1.1	Application description, including use cases for Task 1.1	3	R	PU
D1.2.1	Application description, including use cases for Task 1.2	3	R	PU
D1.3.1	Application description, including use cases for Task 1.3	3	R	PU
D1.4.1	Detailed planning for air/sea pollution application including use cases	3	R	PU
D2.1	General requirements and detailed planning for programming environment	3	R	PU
D3.1	Detailed planning for all the tools and services including use cases for WP3	3	R	PU
D4.1	Detailed Planning for Testbed Setup	3	R	PU
D5.1.1 – D5.1.12	Quarterly reports	each 3 months	R	PU
D5.2.1	Quality assurance plan	3	R	PU
D5.2.2	The CrossGrid architecture requirements and the first definition of the architecture	3	R	PU
D5.3.1	CrossGrid Web page	3	P	PU
D1.0.2	Semi-annual managerial report for WP1	6	R	CO
D2.2	WP2 internal progress report	6	R	CO
D3.2	WP3 internal progress report	6	R	CO
D4.2	First testbed set-up on selected sites	6	R	PU
D5.2.3	Full description of the CrossGrid standard operational procedures and specification of the structure of deliverables.	6	R	PU
D5.3.2	Dissemination and exploitation plan	6	R	PP

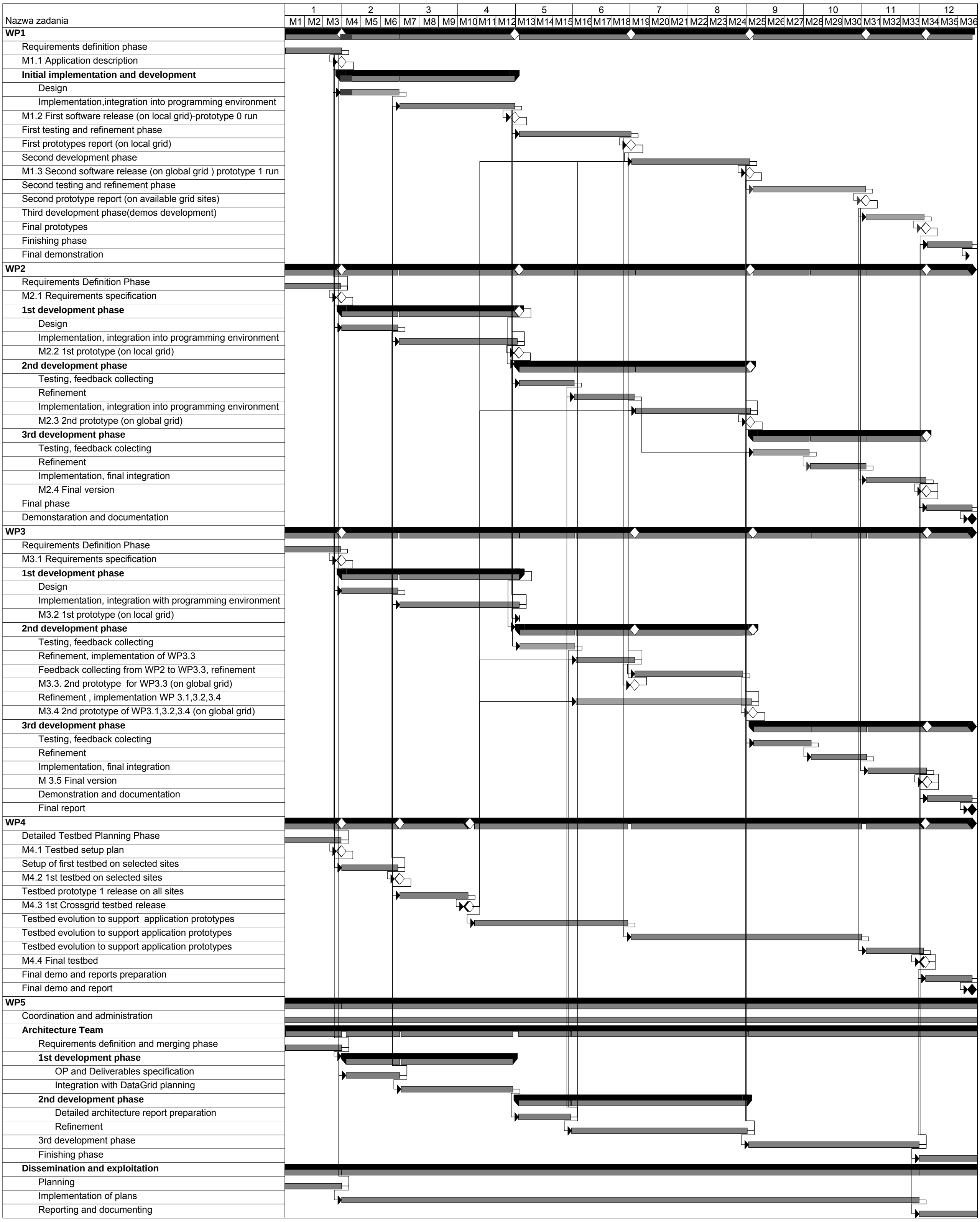
[14] Abbreviations used in tables are explained in Appendix B.

D5.3.3 – D5.3.7	Public (industrial) presentations	6, 12, 18, 24, 30	E	PU
D4.3	WP4 status internal report	9	R	CO
D5.3.8	CrossGrid Website	9	P	PU
D4.4	1st Testbed prototype release	10	P	PU
D5.3.9 – D5.3.11	Project leaflet / brochure	11, 23, 35	O	PU
D5.3.12 – D5.3.14	Dissemination and exploitation report	11, 23, 35	R	PP
D5.3.15 – D5.3.17	Yearly collaboration meeting	11, 23, 35	O	PP
D1.0.3	Extended yearly managerial report for WP1	12	R	CO
D1.1.2a D1.1.2b	Internal progress report First internal software release	12	R SW	CO
D1.2.2a D1.2.2b	Internal progress report First internal software release	12	R SW	CO
D1.3.2a D1.3.2b	Internal progress report First internal software release	12	R SW	CO
D1.4.2	Results of migration of data mining algorithms to GRID structure	12	R	CO
D2.3	Demonstration and report on WP2 1st prototypes	12	R, P	PU
D3.3	1st prototype – first WP3 software release for ealth and HEP applications build on local grid infrastructure.	12	P	PU
D5.1.13 – D5.1.15	Yearly reports	each 12 months	R	PU
D5.2.4	Report on requirements for integration and interoperability with DataGrid	12	R	PU
D2.4	Internal progress report on WP2 software evaluation and testing	15	R	CO
D3.4	Internal progress report on WP2 software evaluation and testing	15	R	CO
D4.5	WP4 status internal report	15	R	CO
D5.2.5	Detailed report on the CrossGrid architecture	15	R	PU
D1.0.4	Intermediate report on coherency	18	R	CO

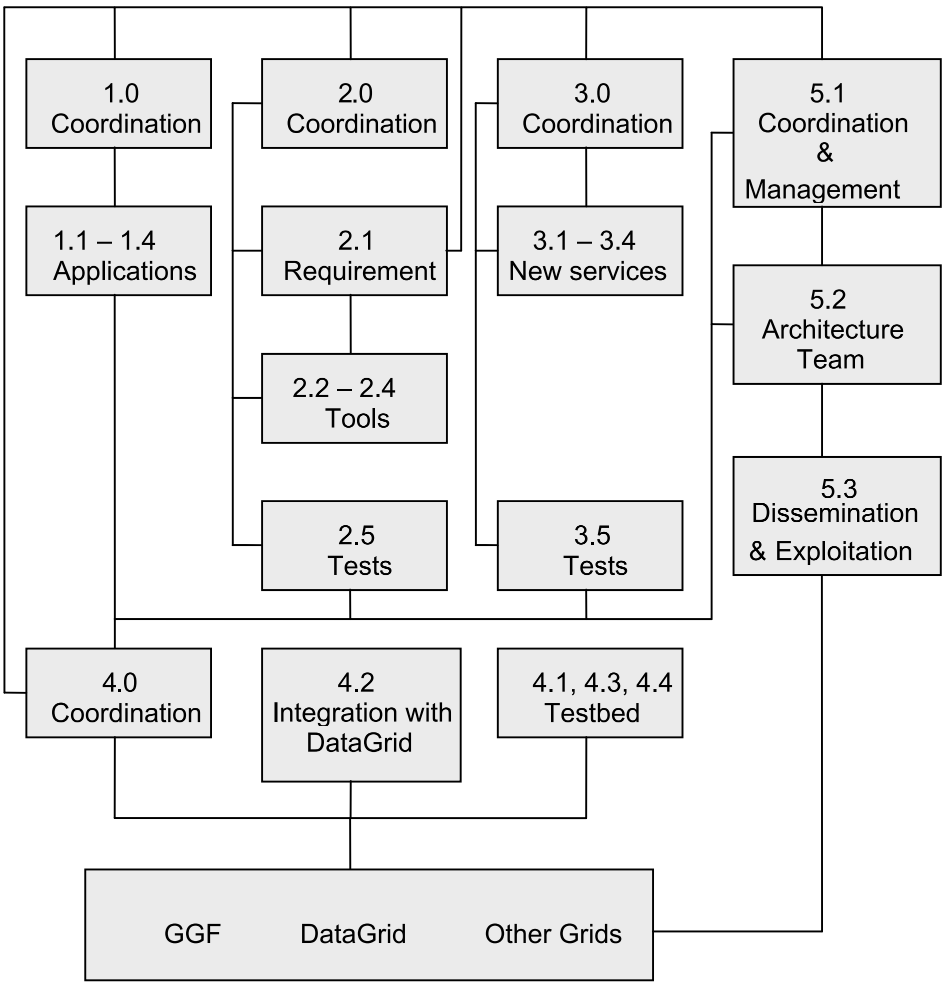
D1.0.5	Semi-annual managerial report for WP1	18	R	CO
D1.1.3	Report on application prototype 0 run	18	R	CO
D1.2.3	Report on application prototype 0 run	18	R	CO
D1.3.3	Report on application prototype 0 run	18	R	CO
D1.4.3	Integration of parallel codes for air quality models into GRID structure Prototype A phase 0	18	R, P	CO
D3.5	Report on the results of the WP3 2nd prototype for Task 3.3	18	R, P	PU
D4.6	WP4 status internal report update	21	R	CO
D1.0.6	Extended yearly managerial report for WP1	24	R	CO
D1.1.4a D1.1.4b	Internal progress report Second internal SW release	24	R SW	CO
D1.2.4a D1.2.4b	Internal progress report Second internal SW release	24	R SW	CO
D1.3.4a D1.3.4b	Internal progress report Second internal SW release	24	R, SW	CO
D1.4.4	Integration of distributed atmosphere and wave models into GRID structure). Prototype B phase	24	R, P	CO
D2.5	Demonstration and report on WP2 2nd prototypes	24	R, P	PU
D3.6	Report on the results of the WP3 2nd prototype for Tasks 3.1, 3.2, 3.4	24	R, P	PU
D5.2.6	Current CrossGrid architecture assessment report	24	R	PU
D2.6	Internal progress report on WP2 software evaluation and testing	27	R	CO
D3.7	Internal progress report on WP3 software evaluation and testing	27	R	CO
D1.0.7	Semi-annual managerial report for WP1	30	R	CO
D1.1.5	Report on application prototype 1 run	30	R	CO
D1.2.5	Report on application prototype 1 run	30	R	CO

D1.3.5	Report on application prototype 1 run	30	R	CO
D1.4.5	Prototype A runs (phase 1) (integration of parallel codes for air quality models into GRID structure).	30	R, P	CO
D1.4.6	Prototype B runs (phase 1) (integration of distributed atmosphere and wave models into GRID structure).	30	R, P	CO
D3.8	Report on the results of WP3 prototype 1 final version run	30	R, P	PU
D4.7	WP4 status internal report	30	R	CO
D5.3.18	CD-Rom	30	D	PU
D4.8	Final testbed with all applications integrated	33	P	PU
D5.2.7	Final CrossGrid architecture and interoperability requirements description	33	R	PU
D1.0.9	Final managerial report for WP1	36	R	CO
D1.1.6	Application final demonstration and report	36	R, D	PU
D1.2.6	Application final demonstration and report	36	R, D	PU
D1.2.7	Final report on dissemination and exploitation	36	R, D	CO
D1.3.6	Application final demonstration and report	36	R, D	PU
D1.4.7	Application final demonstration and report	36	R, D	PU
D2.7	Demonstration and documentation on the final versions and testing of all WP2 software components	36	R, D	PU
D3.9	WP3 final report	36	R, D	PU
D4.9	WP4 final demo and report	36	R, D	PU

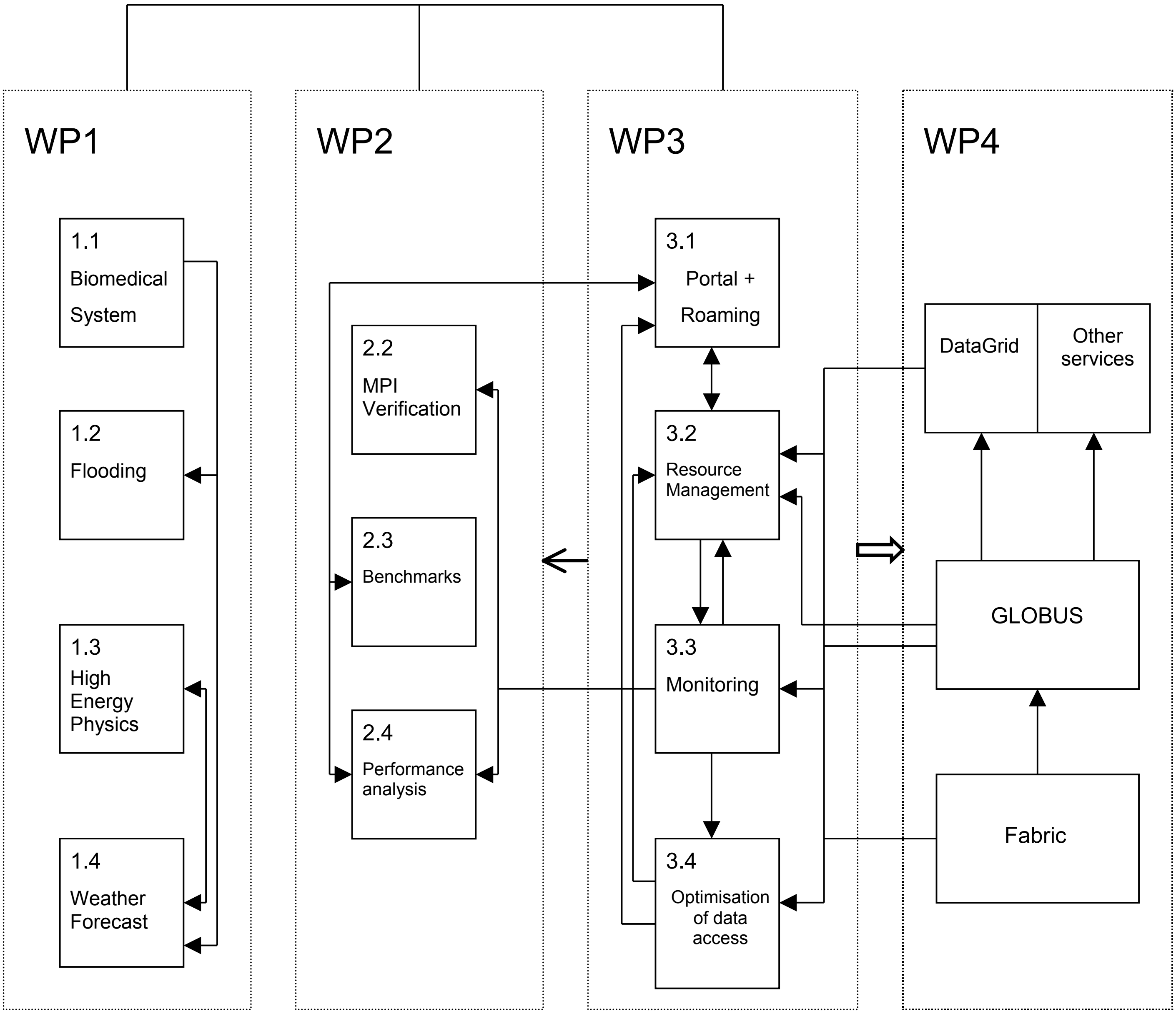
9.8 Project planning and timetable



9.9 Graphical presentation of project components



9.10 CrossGrid architecture with regard to WPs and Tasks



10 REFERENCES

[Belleman2000]

R.G. Belleman and P.M.A. Sloot: The Design of Dynamic Exploration Environments for Computational Steering Simulations, in M. Bubak; J. Moscinski and M. Noga, editors, Proceedings of the SGI Users' Conference 2000, pp. 57-74, Academic Computer Centre CYFRONET AGH, Krakow, Poland, October 2000, ISBN 83-902363-9-7.

[Berge1997]

Berge, E. Transboundary Air Pollution in Europe. MSC-W Status Report 1997. Part 1 and 2 EMEP/MSC-W Report 1/97. The Norwegian Meteorological Institute, Oslo, Norway, 1997. http://www.emep.int.

[Berman, 1998]

F. Berman, R. Wolski, S. Figueira, Jennifer Schopf and G. Shao, "Application-Level Scheduling on Distributed Heterogeneous Networks", Proc. of Supercomputing'96.

[Bethel, 2000]

W.Bethel, B.Tierney,J.Lee, Using High-Speed WANs and Network Data Caches to Enable Remote and Distributed Visualization, SC2000 conference materials, Dallas 2000

[Bubak, 2000]

Bubak, M., Funika, W., Żbik, D., van Albada, D., Iskra, K., Sloot, P., Wismueller, R., and Sowa-Pieklo, K.: Performance Measurement, Debugging and Load Balancing for Metacomputing, The 1st European GRID Forum Workshop, Poznan, April 12-13, 2000, Proc. of ISThmus 2000 -- Research and Development for the Information Society, pp. 409-418, Poznan, 2000.

[Bubak, 2001]

Bubak, M., Funika, W., Balis, B., and Wismueller, R.: On-line OCM-based Tool Support for Parallel Applications. In: Yuen Chung Kwong (ed.): Annual Review of Scalable Computing, Vol.3, Chapter 2, pp. 32-62, World Scientific Publishing, Singapore, 2001.

[Carmichael, 1986]

Carmichael, G.R., Peters, L.K., Kitada, T. A second generation model for regional-scale transport/chemistry/deposition. Atmospheric Environment, 20, pp. 173-188, 1986.

[Casanova, 1999]

H. Casanova, M. Kim, J. S. Plank and J. Dongarra, "Adaptive scheduling for task farming with Grid middleware", International Journal of Supercomputer Applications and High-Performance Computing, pp. 231-240, Volume 13, Number 3, Fall 1999.

[CCAFor]

CCA Forum: http://www.acl.lanl.gov/cca/

[Dikaiakos, 1998]

Dikaiakos, M., Rogers, A., and Steiglitz, K.: Performance Modeling through Functional Algorithm Simulation. In: Zobrist, G., Bagchi, K., and Trivedi, K. (eds): Advanced Computer System Design, Chapter 3, pp. 43-62, Gordon & Breach Science Publishers, 1998.

[Fahringer, 2000]

Fahringer, T., Gerndt, M., Riley, G., and Treff, J.L. : Specification of Performance Problems in MPI-Programs with ASL. International Conference on Parallel Processing (ICPP'00), pp 51-58, 2000.

[Gancio2001]

The DataGrid Architecture Version 1, G.Cancio et al., presented in the European DataGrid Conference, Amsterdam, 7 March 2001.

[GGF]

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

[GGFD]

Remote Data Access Working Group http://www.sdsc.edu/GridForum/RemoteData/

[GGFU]

Grid Forum User Services Working Group http://dast.nlanr.net/GridForum/UserServ-WG/

[Gropp1994]

William Gropp and Ewing Lusk and Anthony Skjellum. USING MPI Portable Parallel Programming with the Message-Passing Interface. The MIT Press. Cambridge Massachusetts and London England. 1994.

[Heymann, 2001]

E. Heymann, M. A. Senar, E. Luque and M. Livny, "Self-Adjusting Scheduling of Master-Worker Applications on Distributed Clusters", Proc. of EuroPar Conference, , Manchester, UK, 2001.

[HLA]

HLA: http://hla.dmso.mil/

[HLA1999]

Defense Modeling and Simulation Office, Dept. of Defense (USA), High Level Architecture Run Time Infrastructure, 1999 http://hla.dmso.mil/.

{Hluchy2001a]

L. Hluchy, V.D. Tran, J. Astalos, M. Dobrucky: Parallel numerical solution for flood modeling systems. Parallel Processing and Applied Mathematics PPAM'2001, Naleczow, Poland, September 2001. To appear.

[Hluchy2001b]

L. Hluchy, G. T. Nguyen, L. Halada, V. D. Tran: Parallel Flood Modeling, Parallel Computing PARCO'2001, Naples, Italy, September 2001. To appear.

[Hluchy2001c]

L. Hluchy, V. D. Tran, G. T. Nguyen: Cluster Computation for Flood Simulations. High-Performance Computing and Networking HPCN'2001, pp. pp. 425-434, Amsterdam, The Netherlands, June 2001.Lecture Notes on Computer Sciences 2110, Springer Verlag. ISBN 3-540-42293-5. ISSN 0302-9743.

[IEEE828]

IEEE Standard for Software Configuration Management Plans. IEEE Std 828-1998

[IEEE1058]

IEEE Standard for Software Project Management Plans. IEEE Std 828-1998.

[Iskra2000]

K.A. Iskra; F. van der Linden; Z.W. Hendrikse; B.J. Overeinder; G.D. van Albada and P.M.A. Sloot: The implementation of Dynamite - an environment for migrating PVM tasks, Operating Systems Review, vol. 34, nr 3 pp. 40-55. Association for Computing Machinery, Special Interest Group on Operating Systems, July 2000.

[Kalas1998]

Kalas M., K. Szefler (1998) The test of application of the WAM model for the wind wave forecasting of the Baltic Sea. Bull. Maritime Inst. 25 No 2, 73-82.

[Kalas1999]

Kalas M., B. Jakubiak (1999) Assessment of numerical wind forecasts for the Gdansk Bay area during the POLRODEX'97 experiment. Bull. Maritime Inst. 26 No 1, 13-25.

[LHC2001]

Report of the Steering Group of the LHC Computing Review, CERN/LHCC/2001-004

[Mellor1998]

Mellor, G. L., 1998: User's guide for a three-dimensional, primitive equation, numerical ocean model, 41 pp., Prog. in Atmos. and Ocean. Sci, Princeton University.

[Mouriño2001a]

J. C. Mouriño, D. E. Singh, M. J. Martín, F. F. Rivera, R. Doallo and J. D. Bruguera. The STEM-II Air Quality Model on a Distributed Memory System. Proc. Workshop on High Performance Scientific and Engineering computing with Applications (HPSECA-2001) in conjunction with the International Conference on Parallel Processing (ICPP-2001) . Valencia, Spain. (2001)

[Mouriño2001b]

J.C. Mouriño, D.E. Singh, M.J. Martin, J.M. Eiroa, F.F. Rivera, R. Doallo, J.D. Bruguera, Parallelization of the STEM-II air quality model, pp.543-550 in L.O. Hertzberger; A.G. Hoekstra and R.D. Williams: Proceedings High-Performance Computing and Networking (HPCN Europe 2001), Amsterdam, The Netherlands, in series Lecture Notes in Computer Science, vol. 2110, Springer Verlag, 2001. ISBN 3-540-42293-5.

[Niezgodka1998]

Niezgodka M., B. Jakubiak, 1998: Numerical weather prediction system: Scientific and operational aspects. Proceedings of the III Symposium on Military meteorology, 191-197.

[Nikolow, 2001]

Nikolow, D., Slota, R., Kitowski, J., Benchmarking Tertiary Storage Systems with File Fragmentation, PPAM'2001, Naleczow, Sept. 10-12, 2001.

[Paplinska1999]

Paplinska B., Wave analysis at Lubiatowo and in the Pomeranian Bay based on measurements from 1997/1998 – comparision with modelled data (WAM4 model). Oceanologia, 41(2), 1999. pp. 241-254

[Pitoura, 1997]

E. Pitoura, G. Samaras, Data Management for Mobile Computing, Kluwer International Series on Advances in Database Systems, 10, 1997

[Preussner1981]

Preussner, P.R., Brand, K.P. Application of a semi-implicit Euler method to mass action kinetics. Chem. Engng. Sci, 10, pp. 1633-1641, 1981.

[Singh2000]

D.E. Singh, M.Arenaz, F.F. Rivera, J.D. Bruguera, J. Touriño, R. Doallo, M.R. Mendez, J.A. Souto and J. Casares. Some proposals about the vector and parallel implementations of STEM-II. Proc. 8th Int. Conf. Development and Application of Computer Techniques to Environmental Studies. ENVIROSOFT 2000.Bilbao, Spain. pp. 57-66. (2000).

[Slota, 2000]

Slota, R., Kosch, H., Nikolow, D., Pogoda, M., Breidler, K., Podlipnig, S., MMSRS - Multimedia Storage and Retrieval System for a Distributed Mediacal Information System, in: Bubak, M., Afsarmanesh, H., Williams, R., Hertberger, B., (Eds.), Proc. Int. Conf. High Performance Computing and Networking, Amsterdam, May 8-10, 2000, Lecture Notes in Computer Science 1823, pp. 517-524, Springer, 2000.

[WAMDIGroup1988]

WAMDI Group, 1988, A Third Generation Ocean Wave Prediction Model. J. Phys. Oceanogr. 18, 1775.

[Wismueller, 2000]

Wismueller, R., Ludwig, T.: Interoperable Run-Time Tools for Distributed Systems – A Case Study. The Journal of Supercomputing, 17 (3) pp. 277-289, 2000.

11 SHORT DESCRIPTION OF THE CONSORTIUM

1. CYFRONET Academic Computer Centre of the University of Mining and Metallurgy, Krakow, Poland, is the coordinating partner (leader of the WP5), and contributes to the programming environment (WP2), grid services and tools (WP3) and testbeds (WP4).

2. The Interdisciplinary Centre for Computational and Mathematical Modelling, Warsaw, Poland, coordinates the air pollution applications (WP1) and contributes to the testbed developments (WP4).

3. The Institute of Nuclear Physics, Krakow, Poland, contributes to the HEP applications (WP1) and the testbed developments (WP4).

4. The Institute for Nuclear Studies, Warsaw, Poland, contributes to HEP applications (WP1) and the testbed developments (WP4).

5. The Universiteit van Amsterdam, The Netherlands, is coordinating the workpackage on applications (WP1). It is responsible for the medical applications (WP1).

6. The Institute of Informatics, Bratislava, Slovakia, is responsible for the development of flood prediction applications (WP1).

7. The University of Linz, Austria, is responsible for the development of visualisation software (WP1).

8. Forschungzentrum Karlsruhe, Germany, coordinates the workpackage on the development of programming environment (WP2), and contributes to the testbed developments (WP4) as well as to the implementation of the HEP application (WP1).

9. HLRS Stuttgart, Germany, contributes to the development of programming environment (WP2).

10. Technische Universität München, Germany, contributes to the development of programming environment (WP2).

11. Poznan Supercomputing and Networking Centre, coordinates and contributes to the workpackage on grid services and tools (WP3), and contributes to the testbed developments (WP4).

12. The University of Cyprus, contributes to the development of programming environment (WP2) and the testbed developments (WP4).

13. DATAMAT, Rome, Italy, contributes to the development of new services and tools (WP3) and to central dissemination (WP5).

14. Trinity College Dublin, Ireland, contributes to the development of new grid services and tools (WP3) and testbed developments (WP4).

15. CSIC Santander, Valencia and Madrid, Spain, coordinates the workpackage on testbeds (WP4). It is responsible for the HEP applications (WP1), and contributes to the programming environment (WP2).

16. The University Autonoma of Barcelona, Spain, contributes to the development of HEP applications (WP1), grid services and tools (WP3), and the testbed development (WP4).

17. The University Santiago de Compostela, Spain, contributes to the air pollution applications (WP1), programming environment (WP2) and testbed developments (WP4).

18. NCRS “DEMOKRITOS”, Athens, Greece, contributes to the testbed developments (WP4).

19. The Aristotle University of Tessaloniki, Greece, contributes to the testbed developments (WP4) and the dissemination (WP5).

20. LIP, Lisbon, Portugal, contributes to the testbed developments (WP4).

21. Algosystems S.A., Athens, Greece. It works on dissemination (WP5), and contributes to the development of grid services and tools (WP3)

12 PROJECT MANAGEMENT

The CrossGrid project is large, as the GRID idea calls for broad dimensions and integration, and it requires close cooperation of many institutions on the international scale. Such large projects are not unusual these days, and the experience with multi-national collaborations in research is very positive. We believe that the large size of the project will generate a bigger impact on the research community in Europe, and a larger weight in the international fora, however we are also aware of the managerial problems in large international collaborations.

Managing of workpackages

The CrossGrid project is divided into four workpackages which deal with the technical aspects of the project, and one workpackage dealing with management, dissemination, and exploitation. All the technical tasks are coupled and they form one logical structure, with applications at one end and the infrastructure at the other, and with intermediate programming environment and software tools in-between. The whole project is factorised into a number of individual tasks, each of which has one or a few institutions attached. The technical and managerial responsibilities for the four workpackages lie primarily with their leaders, who are the top experts in their fields, and who originate from the principal contractors of the Consortium. It is up to these managers to prepare the work plan for these workpackages and to report periodically on the progress. All these tasks have to be synchronised by the central management.

Central management and coordination

Over these workpackages will preside a central managerial structure, comprising a Project Coordinator (PC) and a Project Secretary (PS), who will form a Project Office (POC) established at Cyfronet, and a CrossGrid Architecture Technical Team (ATT) formed also at Cyfronet and lead by an Architecture Team Leader (ATL).

The Project Coordinator, accompanied by the Project Secretary, will be responsible for daily and long-term administaration of the project, contacts with the IST Brussels office, organisation and running of the Steering Group meetings, contacts with partners, conflict solving, and reporting. Project Office will deal with all the administrative and financial issues. It will be small (a single person), but it will be necessary, as the CrossGrid Consortium is large, and to follow all correspondence, prepare quarterly and yearly reports, and to keep track of all these and many other documents will require quite an effort.

The role of the Architecture Team is crucial for the project. It will:

- define and supervise the interlinks between different components (tasks) of the CrossGrid project including the timing (overall progress will be conditioned by timely delivery of some components),

- follow very closely the development of the other Grid projects, to profit from their knowledge and experience, and to use some of their results, and to allow for integration on a global scale.

The coordinating institution, Cyfronet, has a long record of providing good computational services to the research communities, and is one of the strongest partners in the Consortium. On a daily basis a scientist with good managerial experience on the international scale will coordinate the project.

Steering Group and Collaboration Board

The technical and managerial matters will be handled by the *CrossGrid Steering Group (SG)*, consisting of the Project Coordinator (PC), the leaders of the workpackages (WPL), the leader of the Architecture Team (ATL), plus few experts. This group will also be responsible for the overall project work plan, for quality assurance, and for arbitration of local conflicts (if arising). The Steering Group could organise internal technical review panels whenever necessary. The group will meet regularly, either in person or via phone/video-conference; meetings will occur approximately every second month.

The *Collaboration Board (CB)*, consisting of the representatives of each partner and chaired by a person elected at the beginning of the project, will supervise the global progress of the project and will deal with strategic and contractual issues concerning the project. This body will also be in charge of resolving global problems or conflicts in the Collaboration. It is expected to meet yearly, but it should also take action whenever necessary.

In a large and broadly distributed collaboration the access to information and the communication is essential. Collaboration Web pages will be set up and maintained by one of the partners over the life of the project. A repository will be created to keep all the documentation and publications. E-mails will be the main communication tool for daily operations; operational meetings will be conducted via phone and/or video-conference (VRVS). Annual workshops will be held, at which the main results will be presented, giving a basis for the annual reports.

Cooperation with the other GRID projects

As already stated, close links with other Grid projects are essential to the whole Grid concept. The CrossGrid project relies strongly on the developments of the others, in particular of the DataGrid, and the interoperability of the most essential components of these two projects is essential.

The connections of the CrossGrid partners to the other Grid projects are very good. Tha DataGrid project is a particular case, as the staff members of INP Cracow, INS Warsaw, Univ. of Cyprus University, CSIC Santander, Univ. of Barcelona, Univ. of Santiago de Compostela, University of Madrid, Demokritos, Univ. of Thessaloniki and LIP Lisbon worked closely with the staff of all principal contractors of DataGrid (CERN, CRNS, INFN, NIKHEF and PPARC) over many years in the area of physics and research computing. Already now several CrossGrid partners are involved in the DataGrid projet; and this concerns TC Dublin and several Spanish institutions. The management of both projects plans to establish closer contacts on technical and managerial levels as soon as the CrossGrid project gets lunched. In several cases CrossGrid partners participate also in the other Grid projects, which is the case for ICM Warsaw, who is a partner in Eurogrid and Grip projects, and PSNC Poznan is a partner in Gridlab project.

Recently proposed accompany measure GridStart project would be very desirable, as it would offer the appropriate framewok for collaboration between all European Grid projects, and it would motivate all teams to work closely together, and support the common objective of building one Grid.

Potential risks

One should not hide that there are technical and managerial risks involved. The CrossGrid project addresses advanced information technologies and concerns large international communities – both these facts engender some risk.

Technical Risks

The idea of the Grid is very appealing, however it is also very difficult to implement. The creation of a large integrated computing infrastructure, with a common middleware, open and available to every eligible person, is a formidable task, as there are many conflicting requirements and interests. Therefore the realisation of this great idea, and the success of all Grid projects, cannot be guaranteed. Nonetheless, the idea is very definitely worth the effort, as it promises to open new perspectives to the whole of mankind.

There are several strong Grid communities and several Grid projects. Possible divergences of their efforts could create a serious degradation of the whole idea of a unified global Grid. The risk exists and we have to face it. The only way to minimise the risk is to work together, hand-to-hand with other Grid projects (DataGrid, EuroGrid, GridLab, Grip, GriPhyN, PPDG, and others), and to participate actively in the Grid coordination bodies, such as the Global Grid Forum (some members of the CrossGrid collaboration are already involved) and the DataGrid Industry & Research Forum (again, some partners are already members).

Computing is a hot commercial topic and every new idea, carrying potential for new business, is rapidly explored. As distributed computing offers interesting opportunities for a number of applications, it is very possible that industrial labs are already working on several ideas presented by us in this proposal. However, we believe that on a global scale such competition is healthy, as it brings the best end products to society.

The CrossGrid project will rely very much on the model and results of the DataGrid project, its testbeds and the middleware. Close collaboration between the two projects is of key importance, and a GridStart accompany measure project is under preparation. In daily operations the CrossGrid architecture team will interface with the DataGrid project, to watch their progress, gain experience, and gain access to necessary results at the earliest opportunity. This team, being part of the CrossGrid management, will closely follow the DataGrid project, and will inform the whole collaboration on progress and delays, if relevant. In the case of a change to the DataGrid schedule/deliverables, the management of the CrossGrid project will analyse the situation and take necessary steps to overcome difficulties, if they arise. As stressed above, the Architecture Team fulfils an essential role in guaranteeing full interoperability between the two projects.

Managerial Risks

The CrossGrid project is large, with many partners from many institutions and countries, who will integrate a number of local computing sites into one Grid, and thereby allow the solution of problems that are difficult to address in any other way. These local centres have local plans and duties, which they have to satisfy in a timely way. Therefore it is clear that the new CrossGrid R&D activities will compete with their daily demands. We are sure that in this respect additional resources (manpower) provided by the CrossGrid project will play a key role, and it will help to avoid potential conflict of interests.

The CrossGrid project addresses several, quite different, applications. However there is a common denominator, which unifies them all. It is the need for a large computing infrastructure, which can be obtained by joining the individual centres into the Grid, and which will allow interactive execution of complex jobs, requiring high throughputs and handling large distributed data volumes. We believe that the EU funding will play an essential role, as it will allow for recruitment of new personnel with the right profile, to complement the staff of the institutions involved, who will mainly concentrate on their applications.

The number of CrossGrid participants creates some problems. However, such large multinational collaborations, with a strong voluntary ethos, are not unusual these days. Good examples come from the field of physics, where international collaborations of hundreds of participants are reasonably common. We are convinced that the CrossGrid managerial structure, which factorises the project into workpackages and tasks, with integrating managerial bodies to give guidance, will be able to face all the technical and managerial problems.

Dissemination

Dessemination is a duty of every CrossGrid partner, and specific plans for each of them are presented in Chapter 9.3 of this Annex1. It is expected that representatives of every institution will spread the information concerning the Grid via lecturing, seminars and contacts with local media, promote local developments on a broad scale on conferences and workshops, and writing scientific notes and publications, and play a visible role in the global development of the Grid by active participation in appriopriate working groups.

However the project will also requires a centralised effort to:

- design and maintain CrossGrid Web pages, which will include a lot of practical information (address and mailing lists, conferences) for the whole life of the project,
- collect information on ongoing internal developments, and to make them available in an appriopriate form for internal and/or public use,
- identify the CrossGrid products which could be of a broader interest, and to look actively for potential customers (*active promotion*),
- prepare CrossGrid products, including the documentation, in a form which could be accepted by customers, help in technology transfer, provide necessary advicing and support,
- prepare high quality PR materials.

Therefore the managerial structure of the CrossGrid is complemented with the dissemination workpackage, under the responsibility of Algosystems, who is an industrial partner, with an experience in industrialisation, *active promotion* and public relations.

13 APPENDIX A – CONSORTIUM MEMBERS

1 Academic Computer Centre CYFRONET AGH, Krakow (CYFRONET, CO1)

Description

CYFRONET Academic Computer Centre, a separate organisational and financial unit of the University of Mining and Metallurgy in Krakow (since 1999), is one of the five Polish supercomputer and networking centres. The Centre, an independent, non-profit organisation under the authority of the Ministry of Education, was established in 1973 as a result of a joint initiative of the authorities of Krakow universities and research institutes. The Centre comprises the Division of Supercomputer Operations, Programming Division, Division of Computer Network Operations and Development, and Autonomous Section of Network Research and Development. CYFRONET staff numbers about 60 employees (three professors, four PhD's).

CYFRONET has been appointed by the State Committee for Scientific Research (KBN) as the leading unit in the field of developing and running the Metropolitan Area Network in Krakow, and as the leading unit in the field of high-performance computers. KBN finances the purchase, operation, and maintenance of computer equipment and software, as well as network development. The Centre's mission is to:

- offer access to its computational facilities and network services to universities and research institutes;
- create, maintain and develop its computer and network infrastructure;
- perform tasks related to the national policy in the field of seeking, evaluating, promoting and utilising new computer and network techniques in science, education, management, and economy;
- perform research activities in the area of utilisation high-performance computers, computer networks, information and tele-communications services;
- perform consulting, expertise, training and educational activities in the field of computer science, networks, high-performance computers and information services.

CYFRONET, authorised by the Ministry of Telecommunications for the network activities, functions as the main network node in southern Poland. The Krakow MAN provides and supports access to interurban and international connections over the POL-34 network (155 Mb/s) with direct connection to the European TEN-155 (155 Mb/s) network, as well as over two slower networks: Polpak-T and NASK. CYFRONET has designed its own fiber-optic infrastructure (several dozen kilometres in length) within the city of Krakow. Moreover there are lot of leased lines in usage. Modern telecommunications equipment (mainly from Cisco Systems), two HP network servers, and some workstations have been installed for network purposes. CYFRONET operates a variety of “classical” Internet services, develops virtual networks, and maintains regional domains. The Introduction and evaluation of multimedia services has become very important to CYFRONET. The Krakow MAN is a base for the VTLS integrated system of library services. The Centre is also involved in establishing, developing, and maintaining computer networks in southern Poland.

CYFRONET hosts three high-performance parallel computers: a SGI 2800 (128 processors), a HP Exemplar S2000 (16 processors), and a HP Exemplar 1600/XA (32 processors). Archiving facilities consist of two robotic tape libraries from ATL Products (over 17 TB of archiving capacity) and a HP Optical Jukebox 660ex (660 GB). The Centre's software resources comprise the newest versions of system software and a wide range of application software (including Oracle database services).

CYFRONET has contributed greatly to progress in computer science, offering a variety of consulting services, training, and seminars. The Centre is involved in research projects (such as the *Establishment of a Centre of Excellence in Poland - Krakow Centre for Telemedicine and Preventive Medicine* Phare SCI-TECH II Project, and the *HPC/HPV Grid Based on the SGI Cluster for Virtual Laboratory and Tele-Immersion* and BIOPORTAL projects under the *PIONIER – Polish Optical Internet* Program). A high priority is given to international collaboration. The Centre's staff members participate in boards of computer organisations (e.g. C20 IUPAP, HP2EUP, CPG EPS, Advisory/Executive Committee of IEEE Task Force on Cluster Computing) as well as in programme committees of scientific conferences (HPCN, HiPer, EuroPVM/MPI, PPAM, ICA3PP). Between 1994 and 2000, CYFRONET organised six international conferences on computer science: ECUC'94, SupEur'96, PC'96, HiPer'97, PVM/MPI'97, and SGI'2000.

CVs of key personnel

Marian Bubak

MSc degree in Technical Physics and PhD in Computer Science from the University of Mining and Metallurgy (AGH), Krakow, Poland, in 1975 and 1986, respectively. Since graduation he has been working at the Institute of Physics and Nuclear Techniques AGH, and in 1982 he moved to the newly established Institute of Computer Science AGH. At present he is an assistant professor (adjunct) at the Institute of Computer Science AGH and holds a staff position at the CYFRONET Academic Computer Center AGH, Krakow, Poland. His current research interests include parallel and distributed computing as well as tool support for distributed applications and systems. He is author and co-author of about 120 papers and co-editor of 8 proceedings of international conferences. He has served as a program committee member and organiser of several international conferences in both Computational and Computer Science (e.g. Board of the Computational Physics Group of EPS, C20 IUPAP, HPCN, EuroPVM/MPI, SupEur, HiPer, Physics Computing). He was the leader of several Polish and international projects in the high-performance computing area.

Wlodzimierz Funika

PhD in Computer Science. He has 22-years of working experience in software design. He has led and been involved in many system and application projects spanning from distributed system facilities for academic research to office automation. He received his MSc degree at the University of Kishinev where he majored in Mathematics. In 1993 he joined the University of Mining and Metallurgy (AGH) and is currently involved in designing tools for parallel and distributed programming on heterogeneous platforms. He has experience in database systems, parallel and distributed programming, operating systems, parallel tools construction, all within national and international projects. He will be responsible for the "Performance evaluation tools" Task within WP2.

Jacek Kitowski

Graduated in 1973 in Electrical Department of University of Mining and Metallurgy in Krakow (Poland). He obtained a PhD in 1978 and a DSc in 1991, both in Computer Science at the same University. His main fields of interests are: programming models and simulation algorithms for parallel and distributed computing with reference to particle methods and finite element methods, computer architectures and high availability systems, storage, visualization and tele-education systems.

Krzysztof Zielinski

MSc in Control Engineering and PhD and DSc in Computer Science, at the University of Mining and Metallurgy (UMM) in Kraków, Poland in 1976, 1979 and 1985, respectively. In the years 1988-1990 he worked at the Olivetti Research Ltd in Cambridge, UK and the University of Cambridge Computer Laboratory. Since 1991 he has been a Professor and the Head of the Distributed Systems Research Group at the Dept. of Computer Science of UMM and the Chief Computer Networking Advisor at CYFRONET. His interests focus on networking, distributed computing, object-oriented distributed systems engineering and networked multimedia. He has been Technical Manager and Workpackage manager in a number of national and international research projects, many of which were EU-funded. He is now also the Director of Telemedicine and Preventive Medicine Centre of Excellence in Kraków and a Workpackage manager in IST project 6WINIT (IST-2000-25153).

2 ICM, University of Warsaw (ICM, AC2 – CO1)

Description

Founded in 1816, Warsaw University is the largest university in Poland, with about 2,900 academic staff among its 5,300 employees, and almost 51,000 students. Offering courses of studies in 32 fields of the arts and sciences, its 18 faculties include natural sciences, the social sciences, the humanities and 30 extra-departmental and inter-faculty centres and programmes. The Interdisciplinary Centre for Mathematical and Computational Modelling was established in 1993 by the Senate of the University of Warsaw. The research of the Centre focusses on the area of mathematical, natural and computational sciences, networking and information technology.

The major research projects include theoretical and large-scale computer simulations in natural sciences, in particular on studies of biomolecular systems, spatial structure formation in complex systems, global minimization and new areas in applied non-linear analysis. ICM has been developing new visual modelling methods and parallelisation of computing processes and was honoured with nomination to the Computerworld Smithsonian Award in 1995. ICM plays the role of national laboratory in computational sciences, with over 120 related grants currently running. The IT infrastructure of ICM represents a balanced mixture of various classes of parallel systems (CRAY T3E, SUN E10000, SGI Onyx) with vector-parallel systems (CRAY J916, NEC SX4/B) and advanced ATM networking.

The (project-related) experience of the ICM group in computational modelling includes:

- Development of the quantum-classical methods for investigation of dynamics of large molecular systems,
- Development of Approximated Valence Bond method and its application to the study of enzymatic reactions,
- Development of applications of finite element methods to the study of electrostatic properties of large, biologically active molecules,
- Development of novel methods for visual analysis and modelling in medicine.

The experience of ICM team in the subject of wide area data replication and caching covers:

- A very active role from the beginning of the European Internet caching task force,
- Close collaboration on caching product development with Cisco and Network Appliance,
- Creation of one of the first and largest European national web caching hierarchies (since 1995).

CVs of key personnel

Bogumil Jakubiak

Studied at the Technical University in Warsaw. Received PhD in 1981 at IMWM for work on data assimilation. In the years 1969-1976 he was a computer programmer, then from 1977-1986 a researcher, and from 1986-1995 the head of Long-range Weather Forecast Department at the Institute of Meteorology and Water Management in Warsaw. Coordinator of research projects at IMWM during the years 1990-1995. Main area of activity: development of the meteorological information processing system, elaboration of NWP models, and research on climate change. Since March 1997 he has been working as a scientific researcher at ICM – Warsaw University. He participated in several research projects working on implementation and development of the mesoscale version of the UKMO Unified Model, development of a telematics project of 4th Framework Program: PRESTO *Precision Weather Forecast System for Multimodal Transport*. His main scientific interest is a development of the variational methods of the assimilation of new remote sensing meteorological data. He is a member of a working group on data assimilation in hydrological and NWP models in EU project COST 717.

3 The Henryk Niewodniczanski Institute of Nuclear Physics (INP, AC3 – CO1)

Description

The Henryk Niewodniczanski Institute of Nuclear Physics (INP), Krakow, Poland, has about 500 employees (28 professors, 25 associate professors and 110 research assistants) working on high-energy and elementary particle physics, physics of the structure of the nucleus and of nuclear reaction mechanisms, studies of the structure, interactions and properties of condensed matter, and applications of nuclear methods in geophysics, radio-chemistry, medicine, biology, environmental physics and materials engineering. About 250-300 scientific papers are published annually in world-class scientific journals. Each year the Institute organises international schools and conferences. The Institute closely cooperates with other institutions in Poland, such as the Jagiellonian University, University of Mining and Metallurgy, Institute of Nuclear Problems, Warsaw University, the Centre of Oncology in Krakow and many others. The Institute is a member of many international collaborations; in particular experimental groups of the Institute will take part in the ALICE, ATLAS and LHCb experiments at the Large Hadron Collider at CERN. The development of a Grid infrastructure in Poland is a necessity for full participation in these experiments. In the past the Institute played a leading role in developing computing facilities for the academic community in Krakow. The Institute was the main founder of the CYFRONET Academic Computer Centre which provides computing power and data storage facilities, and which has created and maintains the metropolitan research network in Krakow. Recently INP physicists have stimulated an effort towards the development of a distributed computing infrastructure for research in Poland, and the creation of the *Polish DataGrid Consortium*.

CVs of key personnel

Michal Turala

Professor of physics at the Institute of Nuclear Physics, Krakow, Poland, since 1989. MSc in electric engineering, PhD in technical physics at JINR Dubna, habilitation in physics at INP Krakow. Main interests: high-energy physics (HEP) instrumentation. He is participating in many HEP experiments at JINR Dubna, IHEP Protvino, CERN Geneva, and SLAC Stanford. He has published more than 200 reports and articles, and has had a number of long stays at JINR Dubna, MPI Munich, CERN, and UC Santa Cruz/SLAC. Since 1968 he has managerial responsibilities at JINR Dubna and INP Krakow (a section leader). From 1981-85 he was deputy head of the HEP department and during 1992-1994 became deputy director for research at the INP Krakow (600 staff). He has been the organiser of several international schools and conferences. During 1990-94 he was a member and the chairman of the CERN Detector Research and Development Committee. From 1995-97 he was leader of the CERN Division for Electronics and Computing for Physics (about 300 people, yearly budget >10 MCHF). From 1999-2000 he was technical coordinator of the ATLAS silicon tracker project.

Piotr Malecki

Professor of physics at the Institute of Nuclear Physics, Krakow, Poland. MSc in electric engineering, PhD and habilitation in physics. He works in experimental particle physics on data analyses, Monte Carlo simulations of physics processes and detector performance and on real time systems for the detector electronics control. From 1970 – 1998 he was the head of the Data Processing and Computer Group. From 1993 - 99 he was deputy director of the INP Krakow on technical matters. Since 1996 he has been a member of the Restricted European Committee for Future Accelerators. Currently he is the head of the Krakow ATLAS group.

4 The Andrzej Soltan Institute for Nuclear Studies (INS, AC4 – CO1)

Description

The Soltan Institute for Nuclear Studies Institute carries out pure and applied research on subatomic physics, i.e. elementary particle and nuclear physics, hot plasma physics and related fields. It also produces specialised equipment for various applications (notably for medicine and environmental protection). Its staff includes a number of eminent scientists and technology experts. The Scientific Council of the Institute grants PhD and habilitatio degrees. The Institute conducts advanced training for various inside and outside groups. The staff of the institute consists of over 300 employees. The main focus of the institute is on nuclear physics, particle physics, accelerators, plasma physics and medical physics.

The HEP experimental groups at the Institute take active part in the collaboration with CERN, the European Laboratory for Particle Physics. Experimenters study the structure of the nucleon with muon beams, look in electron-positron collisions for the fundamental Higgs particle that is presumably generating all other particle masses, and actively pursue preparations for next generation experiments at the Large Hadron Collider. Another group of physicists studies cosmic radiation, analysing spectra and content of particles bombarding Earth and coming from outer space.

CVs of key personnel

Wojciech Wislicki

PhD in physics in 1987, high-energy particle and nuclear physics, Soltan Institute for Nuclear Studies, reward for outstanding thesis. Staff member of the Institute since 1987. Habilitation in physics in 1998. 18 years of experience in experimental particle and nuclear physics in experiments at JINR Dubna, GSI Darmstadt and CERN Geneva, in experiments on high-energy medium and heavy-ion interactions, deep inelastic lepton-nucleon scattering and spin structure of the nucleon, discrete symmetries violation in weak decays. Some 4.5 years spent at CERN, GSI, and Dubna and shorter visits at many universities and national laboratories worldwide. He has published over 60 papers in scientific journals, and also teaches at the University of Warsaw since 1999.

5 Universiteit van Amsterdam (UvA, CR5)

Description

The Universiteit van Amsterdam will participate through its Computational Science section in its Informatics Institute, which has a good record in related projects. The Computational Science section is headed by Prof. Dr. P.M.A. Sloot, a leading expert in the field of modelling and simulation. The section started in 1988 as a nucleus of parallel computing activity at the Universiteit van Amsterdam, and as such has participated in many HPC projects, at the national and international level. For example, the section participated in many European projects, such as CAMAS, HPCnet, Cauchy^{par}, and Dynamite. The results of the research is published in almost 400 peer reviewed and technical papers, and has resulted in seven PhD theses in the field of computational science. At present the section consists of one full professor, three assistant professors, three technical programmers, and seven PhD students. The section hosts several (international) guests.

Currently, the research activities of the Computational Science section are aimed at modelling and simulation, and evolve around four major themes:

- theory of modelling and simulation,
- massively parallel particle models,
- natural solvers,
- performance research.

Within these themes items such as virtual reality, interactive simulation, distributed computing in grid based heterogeneous environments, and particle models for fluid flow, are being investigated. The section participates in the Virtual Lab (VL) project, which is a very large nationally funded project with many partners, all at the campus of the University. The VL infrastructure, with a whole range of virtual reality environments, will be available to this project. We also want to establish a link between CrossGrid and the VL project.

CVs of key personnel

Peter M.A. Sloot

MSc in Chemical Physics and PhD in Computer Science. He was appointed full professor of Computational Science at the Universiteit van Amsterdam in 1997. He performs a variety of official national and international functions. He has well over 150 peer-reviewed research papers to his name. He has participated in thirteen EU-funded projects.

A.G. Hoekstra

MSc in Physics and PhD in Computer Science. He was appointed assistant professor in 1996. His research interest lies in particle models and interactive simulation. He has published almost 100 peer reviewed and technical papers.

G.D. van Albada

MSc and PhD in Astronomy, but has worked in Computer Science since 1984. He was appointed assistant professor at the Department of Computer Science of the UvA in 1988. He teaches operating systems. He has participated in various EU-funded projects, including Dynamite, often as local project manager.

6 Ustav Informatiky, Slovenska Akademia Vied (II SAS, AC6 – CR5)

Description

The Institute of Informatics was established on 1st January 1999 after gradual transformation from the Laboratory of Mechanics and Automation (1956), to the Institute of Technical Cybernetics of the Slovak Academy of Sciences (1966) and Institute of Computer Systems (1991). The aim of the Institute of Informatics is to conduct research and scientific work in informatics and information technologies, oriented in two directions:

- Distributed information processing
- Design of information system elements and of their technological implementations.

Within the program *Distributed information processing* the following basic research problems are studied:

- Program and architecture models for distributed systems
- Strategies, methods, algorithms, taxonomy and implementation for optimisation of distributed programs.
- Integrated tools for program development environment for distributed systems.
- Algorithms and methods for reliable and efficient transfer in parallel/distributed systems
- HPCN implementation of large-scale applications
- Intelligent agents
- Metacomputing, distributed computing, Internet/Intranet computing

Within the program *Design of information systems and of their technological implementations* further problems of basic research are studied:

- Submicro- and nanometer microtechnology
- Microelectronic sensor structures
- Design of specific ASIC's and FPGA's.
- Thin organic ordered films for optical recording and sensor technique

The Institute is taking part in numerous national and international projects (COPERNICUS, ESPRIT, NATO, EUREKA, IST and other). The Institute of Informatics of the Slovak Academy of Sciences is also the Organiser of international events (conferences and workshops), the most significant of them being the international conference *Artificial Intelligence and Information-Control Systems of Robots*.

CVs of key personnel

Ladislav Hluchy

Director of the Institute of Informatics SAS since 1992. He is Head of the Department of Parallel and Distributed Computing. He was Guest Professor at the Institute of Robotics and Informatics, Technical University of Braunschweig, Germany (1992); the University of Westminster, Centre for Parallel Computing, London, United Kingdom (1993); and the University of Manchester, Department of Computer Science (1993). He is experienced in distributed processing, metacomputing, intelligent agents, Internet/Intranet computing, scheduling, mapping, load balancing and monitoring for distributed computing. He is the author and co-author of more than 100 published papers and two patents. He has extensive experience in leading European projects, including *High Performance Computing Tools for Industry* (HPCTI), Copernicus Programme CP:5383 (1994-1996); *Software Engineering for Parallel Processing* (SEPP), CIPA-CT93-0251

(1994-1998); *Stimulation of European Industry Through High Performance Computing* (SEIHPC), Copernicus Programme, COP-94-00774 (1995-1998); INCO-ESPRIT: 977100 (KIT) *Keep in touch*: Parallel Processing Tools: Integration and Result Dissemination (1998-1999); *datA fusioN for Flood Analysis and decision Support* (ANFAS), IST-1999-11676 (2000-2002).

Viet Dinh Tran

Member of the Department of Parallel and Distributed Computing at the Institute of Informatics SAS. He is experienced in parallelisation of large-scale applications (ground and surface water flow modelling), distributed processing, tools for distributed computing, metacomputing, Internet/Intranet programming. He is the author and co-author of more than 30 published papers. He has worked on several European projects, including *Software Engineering for Parallel Processing* (SEPP), CIPA-CT93-0251 (1994-1998); *Stimulation of European Industry Through High Performance Computing* (SEIHPC), Copernicus Programme, COP-94-00774 (1995-1998); INCO-ESPRIT: 977100 (KIT) *Keep in touch*: Parallel Processing Tools: Integration and Result Dissemination (1998-1999); *datA fusioN for Flood Analysis and decision Support* (ANFAS), IST-1999-11676 (2000-2002).

7 GUP, Johannes Kepler Universität Linz (Univ. Linz, AC7 – CR5)

Description

The research activities at GUP, University Linz focus on parallel processing and computer graphics and the connections between these areas. In the parallel computing domain, investigations are centred on parallel algorithms and parallel software engineering, with special attention to parallel programming tools for error detection and performance tuning as well as graphical data distribution. The research in the domain of computer graphics covers parallel graphics algorithms for realistic image generation (raytracing, radiosity, volume-rendering), and parallel and distributed virtual reality. For all these research topics, GUP, University Linz has established a significant degree of expertise, as can be seen from the publication record, and has attracted a corresponding body computing hardware. This includes several parallel high-performance computers as well as virtual reality gadgets such as Head-Mounted Displays (HMDs), Shutter-Glasses, Trackers, a Dataglove, and a 3-dimensional CAVE Automatic Virtual Environment.

In terms of co-operation, GUP, University Linz, is engaged in several national efforts, such as the Austrian Centre for Parallel Computing and the upcoming Austrian Grid Consortium, and national and international research projects, e.g. with partners in Hungary, France, Belgium, and Germany.

CVs of key personnel

Jens Volkert

Full Professor at the Johannes Kepler University Linz since 1989, Head of the Department for Graphics and Parallel Processing (GUP), and Head of the Institute of Technical Computer Science and Telematics, and Spokesman of the Austrian Centre for Parallel Computation (ACPC) since 1994. He has worked in the area of high performance computing and computer graphics since 1970, and has participated in several research projects supported by the German Ministry of Defence, the German Ministry of Research, the Austrian Ministry of Science and Research and other institutions. He published several books (as editor), contributed chapters to books, and has published around 100 scientific papers in journals and conference proceedings on computer architecture, parallel and distributed computing, computer graphics, and Virtual Reality.

Dieter Kranzlmüller

Has worked in parallel computing and computer graphics since 1993, with a special focus on parallel programming and debugging. His interests include parallel computer graphics with aspects covering computer graphics software and hardware as well as Virtual Reality. He has participated in several national and international research projects with partners in France, Hungary, Germany, Italy, and Belgium. He has co-authored around 60 scientific papers in journals, and conference proceedings. He has presented his scientific results in lectures and interactive demonstrations.

8 Forschungszentrum Karlsruhe GmbH (FZK, CR8)

Description

Forschungszentrum Karlsruhe (FZK), Technik und Umwelt (Research Center Karlsruhe, Technology and Environment) is a member of the Hermann von Helmholtz Association of National Research Centers in Germany (Hermann von Helmholtz-Gemeinschaft Deutscher Forschungszentren, HGF). With about 3500 employees, FZK is one of the largest non-commercial science and engineering research institutions in Germany. It is guided by the research policy goals of its two partners, the Federal Republic of Germany and the State of Baden-Wuerttemberg. The Centre works on research and development problems of public interest, exclusively for peaceful purposes, in the fields of technology and environment. Its *Research for Environmentally Sustainable High Technologies* program is concentrated on five main areas: energy, environment, structure of matter, health, and key technologies. The application-oriented activities of the Centre comprise all stages of research, from basic findings to pre-product developments. In these fields, the Centre cooperates closely with hundreds of partners in science, especially other HGF-centres, universities and industry.

The FZK Computing Centre (Central Information and Communication Technologies Department; Hauptabteilung für Informations- und Kommunikationstechnik, HIK) will participate in this project. HIK operates several high performance computers of different architecture (VPP300/16, VPP5000/8 parallel vector computers, 64-processor IBM RS/6000 SP *Nighthawk*, 32-processor Intel-based Linux Cluster), two StorageTek data robots of 240TB capacity and a LAN with about 7500 computing devices. HIK provides the users with all kinds of software support, from single PC installations over central software services up to optimisation and parallelisation of individual scientific codes.

In 1996, the Computing Centres of the Research Centre Karlsruhe and of the University of Karlsruhe founded the Virtual Computing Centre Karlsruhe to establish common local grid structures. The very close and trusting cooperation of both institutions found expression in a shared ownership of two high-performance computing systems: an IBM RS/6000 SP and a Fujitsu Siemens VPP 300. Common administrative procedures have been introduced to provide easy migration from one system to another within the Virtual Computing Centre, and both sites have set up a shared DCE cell which extends over both campuses.

Recently, the Research Centre Karlsruhe has been selected to build and operate a German Tier-B computing centre for high-energy physics data from the BaBar experiment at the Stanford Linear Accelerator Centre SLAC. Within these activities, FZK will found a research group for Grid computing.

CVs of key personnel

Holger Marten

Masters degree and PhD in Astrophysics from the University of Kiel. He has 7 years experience in model development, simulations of radiation hydrodynamics, and time dependent ionisation of astrophysical gaseous nebulae, and 5 years experience in numerical modelling of radiation hydrodynamics, shock wave phenomena and equations-of-state of dense plasmas in inertial confinement fusion. He was involved in several international WTZ and INTAS projects and is now working in the high performance group of the FZK Computing Centre to support users in numerical code development, optimisation and parallelisation. His skills and competencies include hardware configurations, programming languages, compilers, programming tools and project management. He is the FZK project manager for the construction of the German Tier-B computing centre for BaBar experiments.

Marcel Kunze

Diploma in Physics in 1985 at Karlsruhe University. In the following years he was delegated to CERN to head a group for the construction of the trigger system and data acquisition for PS 197. After his graduation in 1990 he went to Bochum University where he started to work in the field of neurocomputing, in close collaboration with the institute for neuroinformatics. In 1996 he received his habilitation on the use of artificial neural systems in particle physics. As an associate professor he is now teaching software and neuroinformatics. Besides particle physics activities he has participated in projects for brain modelling and 3D television. He is member of the advisory committee for ICANN and ACAT, and was amongst the first to actively promote the use of object oriented programming in particle physics.

9 Universität Stuttgart (USTUTT, AC9 – CR8)

Description

The German Federal High Performance Computing Centre RUS/HLRS hosts one of the largest European supercomputer resources. The high performance computers acquired by RUS/HLRS together with computers from industry are operated by hww GmbH, a company 50% publicly (University Stuttgart, University Karlsruhe, local state) and 50% industry owned (40%Debis, 10% Porsche). The same resources are thus accessible to industry and academia. RUS/HLRS is a national supercomputing centre providing support to users all over Germany.

The main task of a high-performance computing centre is to provide services and support to its users. Services comprise the efficient usage of the HPC resources through all phases of problem handling. Therefore RUS/HLRS employs a group of specialists that advise users how to efficiently use its hardware and software resources, and a further group of specialists who have expertise in different application domains. In addition, courses about parallel programming, including MPI, are offered throughout Germany.

In order to technically support the dissemination of information and usage of HPC systems, RUS/HLRS is actively pursuing the goal of achieving a distributed working environment for its users that allows them to see and use all resources in a seamless way. The first steps towards this have been made within the German UNICORE project and the European METODIS (*Metacomputing Tools for Distributed Systems*) and DAMIEN (*Distributed Applications and Middleware for Industrial Use of European Networks*) projects.

RUS is a technology and knowledge provider with expertise in such fields as distributed simulation, MPI implementation, visualisation, distributed software environments, collaborative working, virtual reality technologies and networking technologies.

CVs of key personnel

Matthias Müller

Research Scientist at HLRS. He received his Diploma in Physics at the University of Stuttgart. His masters thesis on quasicrystals was carried out at the Department of Theoretical and Applied Physics. He wrote his PhD thesis at the Institute of Computer Applications/Stuttgart on fast algorithms for particle simulations. During this period he started to work in a metacomputing project where he participated with his applications. Since June 1999 he has been working at RUS/HLRS. His scientific interests are in metacomputing with a focus on applications, physics on high-performance computers and object oriented programming for parallel computing. He was the project manager of an international metacomputing effort participating in Supercomputing ’99 (HPC games award for the most challenging application).

Edgar Gabriel

Diploma degree at the University of Stuttgart in 1998 in mechanical engineering. Since June 1998 he is a member of the Parallel Computing Department at the High Performance Computing Centre Stuttgart (HLRS). He is involved in metacomputing activities, mainly the development of PACX-MPI. He gained experience in European projects within the METODIS and DAMIEN projects.

Michael M. Resch

Diploma degree in Technical Mathematics from the Technical University of Graz, Austria, 1990. From 1990 to 1993 he was employed by JOANNEUM RESEARCH, a leading Austrian research company, working on numerical simulation of groundwater flow and groundwater pollution on high performance computing systems. Since 1993 he has been working at the High Performance Computing Centre Stuttgart (HLRS). The focus of his work is on parallel programming models. He is responsible for the development of message-passing software and numerical simulations in national and international metacomputing projects. Since 2000, he is heading the HLRS and is responsible for all projects, research and services for users.

10 Technische Universität München (TUM, AC10 – CR8)

Description

Technische Universität München (TUM) teaches about 20,000 students. TUM's computer science department is among the biggest within Germany, with approximately 25 professors, 200 researchers, and 1,600 students. The chair named *Lehrstuhl für Rechnertechnik und Rechnerorganisation / Parallelrechnerarchitektur* (LRR-TUM) has been working in the field of parallel processing for over ten years. It currently comprises two professorships and about 25 scientists.

LRR-TUM's major research is in interactive run-time tools for parallel and distributed systems. LRR-TUM has developed performance analysers, debuggers, program flow visualisers, and the necessary on-line monitoring techniques, but also application oriented tools, e.g. for computational steering. Two of these tools, the performance analyser PATOP and the debugger DETOP, have been successfully transferred to commercial products. Currently, the main focus in this area is the definition of a standard interface for monitoring parallel and distributed programs (OMIS) and the interoperability of run-time tools.

In addition, LRR-TUM investigates high-performance cluster computing, based on SCI interconnection technology. Research in this area includes the efficient use of hardware-supported distributed shared memory and the design of hardware monitoring to provide detailed information to performance analysis and debugging tools. Finally, LRR-TUM is also working on parallel applications in close contact with industry and end-users.

LRR-TUM has successfully participated in several ESPRIT projects and working groups related to tools and distributed computing (e.g. PUMA, HAMLET, PREPARE, SEEDS, NEPHEW, APART), with a focus on performance analysis. From 1990 to 2000, the head of LRR-TUM, Prof. Arndt Bode was head of the special research grant SFB 342 *Tools and Methods for the Use of Parallel Systems* funded by the German Science Foundation (DFG), with approximately 100 staff.

CVs of key personnel

Roland Wismüller

Head of the research group on programming environments and tools at LRR-TUM. He received a diploma in computer science and a doctorate from TUM in 1988 and 1994, respectively, and he has finished his habilitation on Interoperable Run-time Tools for Parallel and Distributed Systems in August 2001. He has more than 10 years of research experience in the field of tools for parallel and distributed systems, including debugging and performance analysis, monitoring techniques and interoperability issues. He is participating in the APART and Peridot projects aiming at automatic performance analysis. In the CrossGrid project, he will be involved in Workpackage 2.

Michael Gerndt

Graduated in 1985 with a Diploma in Computer Science from the University of Bonn. Starting in 1986, he worked on the topic of automatic parallelization for distributed memory machines in the framework of the SUPRENUM project at the University of Bonn and received a PhD in Computer Science in 1989. In 1990 and 1991, he held a postdoctoral position at the University of Vienna. He joined the Research Centre Jülich in 1992, where he concentrated on programming and implementation issues of shared virtual memory systems. Since July 2000 he is Professor for Architecture of Parallel and Distributed Systems at TUM. He was coordinator of the APART working group on automatic performance analysis and will coordinate the follow-up working group APART2. He is leader of the project Peridot (funded by KONWIHR) which is developing an automatic performance analysis system for the first European Teraflop Computer, the Hitachi SR8000 at Leibniz Computer Centre in Munich. In the CrossGrid project, he will be involved in Workpackage2.

11 Poznan Supercomputing and Networking Centre (PSNC, CR11)

Description

Poznan Supercomputing and Networking Centre (PSNC) was established in 1993 by the State Committee for Scientific Research. PSNC today is organised into 5 departments: Applications Department, Supercomputing Department, Network Services Department, Networking Department and Administration Department. The total number of employees equals 60.

From the very beginning PSNC served as the HPC service provider and the operator of Poznan Metropolitan Area Network (POZMAN). As the most active centre in Poland, PSNC soon became an operator of the national research network POL34/155. Today it is also the operator of the Polish international link to the European TEN-155 network and is the national networking centre for excellence.

Each of the departments has an active computer science research group working on such aspects of the computer science as: middleware, tools and methods for Grid computing, resource management for Grids, large scale Grid applications (Application Department), user accounting on the Grid, Grid security mechanisms and policies, Grid fabric management tools (Supercomputing Department), application portals, Grid portals, multimedia services, mobile user support technologies and services, digital libraries (Network Services Department), tools for network management, optical networks, QoS management (Networking Department). All the departments cooperate strongly with other institutes across Poland, Europe and the Atlantic on Grid-related projects.

PSNC is the main author of the national IST and GRID-related program, called '*PIONIER - Polish Optical Internet, Applications, Technologies and Tools*' for the Polish national information infrastructure, which has been accepted by the Polish Parliament for the years of 2001-2005.

PSNC Resources: As a consequence, PSNC is building a metacomputer. The factors determining the functionality of the computing system are the supercomputing paradigms, scale and importance of the problems being solved, as well as the need for a specific kind of computing. Considering the limitations of uniprocessors, high-performance computations cannot be located on a single computer system. The current architectures for metacomputers consist of systems with shared memory architecture (scalar: SGI Power Challenge, vector: Cray J916, Cray SV1), distributed memory (IBM SP-2), massively parallel machines (Cray T3E-900, ccNUMA SGI ORIGIN3200C), PC clusters, data archiving systems and visualisation laboratories (SGI ONYX2).

PSNC has been involved in building the European Grid Forum (EGrid). The first EGrid workshop has been held in Poznan, in April 2000. Originally EGrid was an initiative started by some of the proposers (ZIB, Lecce and PSNC) but evolved very rapidly to a large forum, with several working groups and over 200 people from Europe involved. From the very beginning PSNC was co-leading the organization while having still better and better co-operation with other European centres and industrial partners.

CVs of key personnel

Jan Weglarz

Professor (Ph.D. 1974, Dr. Habil. 1977), in years 1978-83 Associate Professor and then Professor in the Institute of Computing Science, Poznan University of Technology, member of the Polish Academy of Sciences (PAS), Director of the Institute of Computing Science, Poznan University of Technology and its predecessors since 1987, Director of Poznan Supercomputing and Networking Center, Vice-President and Scientific Secretary of the Poznan Branch of the PAS, Vice-President of the Committee for Computer Science of the PAS, member of the State Committee for Scientific Research, Principal Editor of the Foundations of Computing and Decision Sciences, member of several editorial boards, among others Internat. Trans. Opnl. Res. and European J. Opnl. Res. Representative of Poland in the Board of Representatives of IFORS and in EURO Council (President of EURO in years 1997-98). He is a member of several professional and scientific societies, among others the American Mathematical Society and the Operations Research Society of America. Author and is co-author of 11 monographs, 3 textbooks (3 editions each) and over 200 papers in major professional journals and conference proceedings. Frequent visitor in major research centres in Europe and in the U.S.A. Co-laureate of the State Award (1988) and the EURO Gold Medal (1991), laureate of the Foundation for Polish Science Award (2000).

Maciej Stroinski

Received the P.hD. degree in Computer Science from the Technical University of Gdansk in 1987. Currently he is Technical Director of the Poznan Supercomputing and Networking Center. He is also lecturer in the Institute of Computing Science of the Poznan University of Technology. His research interests concern computer network protocols and management. He is a chairman of the technical committee of the Polish Academic Network POL-34/155. He is author or co-author of over 100 papers in major professional journals and conference proceedings.

Norbert Meyer

Received the M.Sc. degree in Computer Science from the Poznan University of Technology (1993), Ph.D. degree from the same university. Currently he is the head of the Supercomputing Department at Poznan Supercomputing and Networking Center. His research interests concerns resource management and security issues, mainly in the aspects of connecting independent, geographically aware Grid domains. He participated in several national projects concerning HPC technology, leader of the concept of connecting Polish supercomputer centres into Polish national cluster. His research interests concern resource management in GRID environment, GRID accounting (Global Grid Forum), network security, technology of development graphical user interfaces. He is also an author and co-author of several reports and over 20 papers in conference proceedings and professional journals.

Miroslaw Kupczyk

Received the M.Sc. degree in Computer Science from the Poznan University of Technology (1999). Presently he is a PhD student at the same university. Currently he is employed as a HPC specialist position at the Supercomputing Department in Poznan Supercomputing and Networking Center. His research interests concern resource management in the GRID technology and graph algorithms. He is an author and co-author of several reports and papers in conference proceedings and journals.

12 University of Cyprus (UCY, AC12 – CR11)

Description

The Department of Computer Science is one of the four departments of the School of Pure and Applied Sciences of the University of Cyprus, with a full-time staff of 14 faculty members, over 10 full-time researchers and visiting staff, and over 120 undergraduate and graduate students. Its main research focus includes Parallel and Distributed Systems, High-Speed Networks, Internet Technologies, Mobile Computing, Parallel Processing, Intelligent Systems, Medical Informatics and Telemedicine, and Multimedia Systems. The Department runs a General Research Lab with a number of multiprocessor servers running AIX, Solaris and Linux (these are to be connected over a gigabit switch during 2001), various workstations running Windows 2000, Solaris, and Linux, and three Multimedia Stations; a Unix Lab hosting 20 powerful IBM 43P workstations running AIX; an NT Lab with 18 PCs running Windows NT; a Digital Lab with equipment for Digital Design, and the Multimedia Research and Development Lab (MRDL). These Laboratories are connected with the Cyprus Research and Educational Network (CYNET), hosted at the University of Cyprus and participating in the Geant Consortium. Currently, the Department runs a number of research programs related to Health Telematics, Mobile Computing, and Internet Infrastructures, funded by the European Commission, Cypriot Government organizations, and local industry. In the recent past, the Department has participated in a number of European Union projects under the ESPRIT and INCO-DC programmes including AMBULANCE, ET-ASSIST, ILP2, MEDUSE, MEDWATER, CJIS, Compulog-Net, ERUDIT, NeuroNet, and Q-Med.

The High-Energy Physics Laboratory (HEPL) of the University of Cyprus participates actively in international collaborative large-scale experiments, such as the L3 Experiment at LEP and the CMS Experiment at LHC, at the European Laboratory for Particle Physics (CERN). HEPL is currently a member of an international Consortium for the construction of the Barrel Yoke and Vacuum Tank of the CMS superconducting magnet and a member of the Electromagnetic Calorimeter group of CMS.

CVs of key personnel

Marios D. Dikaiakos

PhD, Assistant Professor, holds an Electrical Engineering Diploma from the National Technical University of Athens, Greece (summa cum laude, 1988), a MA and a PhD in Computer Science from Princeton University, USA (1991 and 1994 respectively). Since 1998, he has been an Assistant Professor of Computer Science at the University of Cyprus. Before that, he worked at Princeton University (Research Assistant, 1988-1993), the Paris Research Laboratory of Digital Equipment Corporation (Fall 1990), the University of Washington in Seattle (Research Associate, 1994, 1995), and the University of Cyprus (Visiting Assistant Professor, 1996). He has over 10 years of experience in research projects funded by NASA's Earth & Space Sciences program, the National Science Foundation of the US, the U.S. Army Research Office, and the Advanced Research Project Agency (ARPA) of the USA. More recently, he has participated in research projects funded by the European Commission, the Research Promotion Foundation of Cyprus and the Research Committee of the University of Cyprus. Currently, he is principal investigator in two projects funded by the Research Promotion Foundation of Cyprus, the Planning Bureau of the Republic of Cyprus and the General Secretariat of Science and Technology, Greece. Since February 2000, he is serving as the Representative of Cyprus to the IST Committee, the high-level body overseeing the IST Programme of the European Union. He has also served as an External Evaluator for the IST Program (KAII), an Information Technology Node co-ordinator for Web and Multimedia in Cyprus (a DGIII initiative), and a consultant on Internet Security for the Public Power Corporation of Greece. His research interests include Parallel and Distributed systems and Internet Computing.

Panos A. Razis

PhD, Associate Professor, he holds a Physics Diploma (summa cum laude, 1980) from the University of Athens, Greece, and a MSc (1982), M.Ph. (1982) and PhD (1986) in Physics from Yale University, USA. Since 1991 he is an Associate Professor of Physics at the University of Cyprus. He previously served as a Postdoctoral Fellow at Yale University, Wissenschaftlicher Mitarbeiter at the Swiss Federal Institute of Technology (ETH Zurich) and as a Research Associate at the European Laboratory for Particle Physics (CERN) for the University of Alabama. He currently participates in the CMS Experiment at the Large Hadron Collider (LHC) and the L3 Experiment at the Large Electron Positron collider (LEP), both at CERN. In the framework of the above experiments he served as a Coordinator for the New Particle Search Group, Representative of the Other Non-Member States of CERN, Cyprus National Coordinator in the Research Review Board of CERN and Member in the Finance, Conference and Collaboration Boards of the experiments. His research interests include: high-energy and elementary-particle physics, with emphasis on electro-weak physics and experimental testing of unifying field theories; vector bosons production and their decays through rare reactions; analysis of the Standard Model and search for new particles and phenomena (Higgs, supersymmetry, multiphoton and rare reactions, lepton number violation, compositeness etc.); development of detectors for elementary particles; alignment, calibration and electronic data acquisition systems for particle detectors; study of neutrino oscillations; cosmology; medical radiation physics.

13 DATAMAT S.p.A. (DATAMAT, AC13 – CR11)

Description

DATAMAT S.p.A., established in 1971, is one of the key players in the Information and Communication Technology industry in Italy, the only one with thirty years of experience in state-of-the-art technology, active in: development and sale of software products; software and system integration projects; internet solutions; on-line services. DATAMAT is organised in four independent divisions: Banking and Finance, Telecommunications and Utilities, Aerospace, Defence and Environment, Public Administration and Health Care.

DATAMAT's goals are to grow, both domestically and internationally, and to create value for its shareholders and clients. Thanks to its technology and reputation, DATAMAT has been way up front to satisfy its clients' needs and, today, to help their transition from the old to the new economy. A consistent portion of resources in the various business units are employed in product development and innovation. Traditionally, DATAMAT is very active in basic research and development particularly within European R & D programs.

DATAMAT is certified according to NATO AQAP 110 and 150 (for military programs) and ISO 9001.

DATAMAT is involved in a number of projects relevant to the CROSSGRID proposal:

- **ENVISAT Payload Data Segment**: DATAMAT has been working with the European Space Agency from 1994-2001 to collect and deliver to users the ENVISAT-1 products they need in a timely manner. This work involves systems and services to handle products from raw to processed data, trough acquisition, processing, archiving, quality assessment and dissemination. The user services which will be offered are clearly similar to those proposed for data intensive GRIDs.
- **Nuovo Sistema Elaborazione Dati (NSED) of CNMCA**: DATAMAT is leading the design and development of the new data processing system for the National Meteorological Service in support of daily forecasting activity. The main purpose of the system is to support the customer in the execution of all the activities concerned with the weather forecast service. At present, more than 400 operators are involved in service provision, 100 of them in the central site nearby Rome. DATAMAT is in charge of the entire programme management, including: design and development of logistic infrastructures, planning, preparation and conduct of training courses for operational personnel, support to operational start-up, and planning, management and conduct of H/W and S/W assistance services for the whole territory in order to ensure service continuity.
- **Multi-mission User Information System (MUIS)**: Within the ESA Earth Observation Programme, there is a need to constitute an ESA open User Information Service to provide a common and easy access to EO data and information collected from different sources. Up to now this has been provided by several organisations and by different means, in general based on ad-hoc, standalone services. The gap will be filled by the Multi-mission User Information System (MUIS), which will to provide internal and external users, connected by different client software applications, with: Multi Mission Guide and Directory Services, Multi Mission Inventory Services, Multi Mission Browse Services, Multi Mission on-line Ordering Services. DATAMAT has been involved in the following subsystems: User Management System, System Monitoring and Control and File Server.
- **EU DataGRID**: DATAMAT is the only Italian industrial partner involved in the European R&D EU DataGRID project, co-funded by the European Commission as part of the Information Society Technologies (IST) activities within the 5th Framework Programme. This project is led by the CERN and involves scientific, institutional and industrial partners from all Europe. DATAMAT is involved as Assistant Contractor of INFN (the Italian Institute of Nuclear Physics) in the job and workload management workpackage, one of the workpackages dedicated to build up the EU DataGRID middleware. DATAMAT also provides technological and system skills to the European Space Agency for the Earth Observation application work package.
- **SpaceGRID**: DATAMAT has awarded this contract from the European Space Agency, leading a consortium composed by European scientific and industrial partners. The objectives of the SpaceGRID study are to analyse the issues and propose a roadmap for the seamless access to and exploitation of distributed data, applications and resources in a network of heterogeneous computing resources and for different Space application domains. The study shall analyse the technical issues related to the management, access and exploitation of large amounts of data, high throughput computing GRIDs, and the Agency's applications adaptation for heterogeneous distributed systems.

CVs of key personnel

Stefano Beco

He will assume the role of Task Manager for DATAMAT. He has a Doctor's Degree in Electronic Engineering from the University of Rome "Tor Vergata". His skills and competencies include project management, interoperability protocols, user services, and software languages and operating systems. Since end of 2000, he is the Programme Manager of GRID activities in DATAMAT, Task Manager in DataGRID WP1 / WP9 (EC 5FP IST Project) and Project Manager in SpaceGRID (ESA Project). Since 1999 has been Project Manager for three major projects: CEO Catalogue interoperability Experiment and User Metadata Tool (EU JRC), CIP/IMS Gateway for CEO Search Release B (Logica) and DATAMAT's MUIS-related development activities (ESA ESRIN). Prior to this he was a study leader and then project leader in User Services definition and development for the ENVISAT PDS project.

Fabrizio Pacini

He has a Doctor's Degree cum laude in Mathematics awarded by the University of Rome "Tor Vergata". His skills and competencies include object oriented and Yourdon-DeMarco analysis and design, programming in C++, C, Fortran, Pro*C, database skills in PL-SQL RDBMS, Oracle and Ingres, interoperability protocols and experience of Unix and VAX-VMS. He is Task Leader in DataGRID WP1 (EC 5FP IST Project). Most recently he has been project leader of the CIP/IMS Gateway project for CEO Search Release B (with Logica). Prior to this he was responsible as project leader for the design, development and testing activities in the HELIOS I program for the User Services subsystem in the frame of CSU3.0 contract and as team leader for the integration activities at system level in France for Matra Marconi Space during a period of one year.

14 Trinity College Dublin (TCD, AC14 – CR11)

Description

Trinity College is a university in the liberal arts tradition, established in 1592 by Queen Elizabeth I. The Computer Science Department came into being in 1969 and today consists of more than 50 academic staff, 15 support staff and 170 postgraduate students. It runs the most comprehensive range of full-time and evening courses of any university in Ireland covering all aspects of Computer Science and Information Systems. Overall the department teaches over 2000 students in its 8 undergraduate and 4 postgraduate degree courses. Central resources include a Sun cluster, a Linux cluster, etc. There are nearly one thousand networked machines within the department. There are over a dozen teaching laboratories, hardware and software.

Research groups made up of staff and post-graduate students are active in a broad range of projects undertaken in association with national industry and European research consortia. Over the years, research carried out within the department has spawned a number of campus companies that have been highly successful both in Europe and the USA. The primary research groups are: Applied Information Systems, Artificial Intelligence, Computer Architecture, Computer Vision, Distributed Systems, Formal Methods, Human Computer Interaction, Image Synthesis, Knowledge & Data Engineering, Networks and Telecommunications, and Multimedia Systems.

The Computer Architecture Group, which is represented in this project, is interested in combining both hardware and software expertise in the design and construction of innovative computer subsystems. It has spent many years developing non-invasive hardware trace collection, simulation and analysis tools, most recently for the Scalable Coherent Interconnect. It will bring this expertise to the CrossGrid project.

CVs of key personnel

Brian Arthur Coghlan

B.E. (Western Australia) in 1967, PhD in Physics (London) in 1979. He is a Senior Lecturer in Computer Science at Trinity College Dublin. His current research areas include parallel computing, clustering and metacomputing, fault tolerant and real-time systems, high-performance I/O subsystems and distributed shared memory, particularly the Scalable Coherent Interconnect (SCI). He is Joint Leader of the Computer Architecture Group at TCD. He is actively promoting grid computing in Ireland, and is a Director of Grid-Ireland.

15 Consejo Superior de Investigaciones Científicas (CSIC, CR15)

Description

CSIC (Consejo Superior de Investigaciones Científicas) is the largest Spanish research institution. With more than 100 centres distributed across Spain, and 2000 scientists, its annual budget is above 350 MEuros, half funded by the Ministerio de Ciencia y Tecnologia, with research projects in the areas of Biology, Physics, Materials Science, Chemistry, Social Sciences and Environment. It has a long experience in European projects, being involved in more than 500 with funding in excess of 60 MEuros.

IFCA (Instituto de Física de Cantabria) is located in in the University Campus in Santander. It is involved in international research projects in high-energy physics, astrophysics, and statistical physics. Projects in high-energy physics include participation in the CMS experiment for LHC, the DELPHI collaboration at LEP, and the CDF collaboration at Fermilab (Chicago, USA). A more detailed description can be found in http://ifca.unican.es/.

IFIC (Instituto de Física Corpuscular) is a non-profit research institute located on the Campus of the University of Valencia. IFIC is involved in many experimental and theoretical activities. On the experimental side, the projects are related to high-energy physics programs, such as DELPHI and ATLAS at CERN. In ATLAS, it participates in the Forward SCT detector and in the Tile Calorimeter detector. There are also research activities in Neutrino Physics, for example, ANTARES, which is an atmospheric neutrino experiment under the sea. Moreover there are active groups working on Gamma Spectroscopy and in Medical Physics. The theoretical department of the IFIC is very active and includes several research groups (Phenomenology, Astroparticle, Lattice, etc.). An overall view of the institute's research activities can be seen at http://ific.uv.es.

RedIRIS is the institution responsible for the academic and research network in Spain. Founded in 1988, it is managed by CSIC since 1994. RedIRIS is the main tool of the National Programme of Applications and Telematic Services and assumes the responsibility for providing the required network services and current and future support of the infrastructure. About 250 institutions are connected to RedIRIS, mainly universities and R&D Centres. The Services offered by RedIRIS require that the support of the basic infrastructure be technologically adapted to the requirements of the connected centres. These services are provided in collaboration with other academic networks and other international fora: RedIRIS is a partner of DANTE, a Trans-European network operator offering international network services to the European academic networks. RedIRIS also participates in the European R&D networks association TERENA (formerly known as RARE) and in RIPE (the European forum of Internet service providers), as well as in the creation and funding of the European Internet coordination centre. RedIRIS collaborates as well with DG XIII of the European Union, working in advanced networking projects. The staff of the Centro de Comunicaciones CSIC-RedIRIS specialises in communications technology, and is organised in three different areas: Network Services, Application Services Area and Institutional Relations.

CVs of key personnel

Jesus Marco

PhD in 1989 at the University of Cantabria, on high-energy physics. Since 1991 he holds a staff position at IFCA (CSIC). He is member of the DELPHI and CMS collaborations at CERN, with responsibilities in physics analysis and software development. He has worked in the project Fast Simulation for DELPHI, and has been responsible for the Higgs Research Team in this collaboration in recent years. He has also been involved in the development of the OO simulation for CMS (Oscar), and in studies comparing the use of OO and O/R DBMS in HEP. He is co-author of more than 250 papers in international journals on HEP. He will lead WP4.

Celso Martinez-Rivero

Staff member at IFCA, he received his PhD in 1995 at the University of Cantabria, and enjoyed a Marie-Curie fellowship at LAL (Orsay) until 1998. He received a two-year CERN fellowship in 1999, and a position as tenured scientist in CSIC in 2001. He has worked in physics analysis and reconstruction software in the DELPHI and CMS experiments at CERN. He will be involved in WP1, task I.3, distributed physics analysis in HEP.

Jose Salt

Senior Physicist at IFIC (UV-CSIC). He obtained his Doctor degree in 1987 at Universitat de Valencia. He has extensive experience in software development for HEP applications and Physics Analysis (GEANT, PAW, HIGZ, etc), having worked at CERN (1988-1990) in the GEANT Project, in LAL (Orsay) and at LPNHE (Paris). He is involved in the ATLAS experiment, working in the event generation and in the analysis of the SCT (Silicon Tracker detector) data. Jose Salt will be responsible for IFIC participation in Workpackage 1.

Eduardo Ros

Senior Physicist at IFIC (UV-CSIC). He obtained his Doctor degree in 1986 at Universitat Complutense (Madrid) and has extensive experience in high-energy physics, in particular in data analysis and computing. He has worked at post-doctoral level at DESY (1987-90) and at CERN as staff member (1991-1997). He will be involved in Work Package 4.

Antonio Fuentes

System Engineer in RedIRIS, at Madrid. He qualified as a Computing Engineer (University of Murcia) in 1999, and is now preparing his PhD in the Data Mining and Distributed Computing Department. He has previously worked as system manager in Airtel (a Spanish telecommunication company) and at the University of Murcia.

16 Universitat Autonoma de Barcelona (UAB, AC16 – CR15)

Description

Universitat Autonoma de Barcelona (UAB) is a Spanish university, founded in 1968, that carries out academic and research activities. It is has 37,000 undergraduate students, 6,000 graduate students, more than 2,600 academic staff and 1,200 technical and administrative staff. Academic activities are organised around 12 faculties and 2 technical schools. Research activities involve around 46 departments and 32 institutes and research centres that either belong to or are linked to UAB. The Computer Architecture and Operating Systems (C.A.O.S.) Group and the High-Energy Physics Group that participate in this project are associated to the Computer Science Department and the Sciences Faculty, respectively.

The UAB Computer Architecture and Operating Systems group at UAB is led by Prof. Emilio Luque. The group employs 24 people, which includes 20 academic staff members and PhD-students in research teams and 4 technical and administrative staff members. The main research focus of the CAOS group includes parallel, distributed and concurrent programming, programming environments, supercomputing, cluster computing and metacomputing. The group has been involved in several national projects and some EU projects related to the research areas mentioned above, such as the SEPP (No. CP 93: 5383), HPCTI (No. CIPA-C193-0251) and COPERNICUS projects, as well as in two TEMPUS projects (S_JEP-08333-94, S_JEP 12495-97). In those projects, the CAOS group worked on the design and implementation of simulation tools for complex, parallel and distributed systems, and in the design and evaluation of static and dynamic allocation strategies for parallel applications.

The group has also been involved in the application of high performance computing to forest fire simulation as a subcontractor of the EU INFLAME project. Currently, they are working in other EU projects such as APART 2, which is devoted to the design and study of tools for automatic performance evaluation applied to high performance applications. They are also cooperating with the Condor team at the University of Wisconsin Madison within the framework of scheduling policies for parallel applications running on cluster and grid environments.

The UAB high-energy physics group is a research group associated with the Sciences Faculty of the Universitat Autonoma de Barcelona in Bellaterra-Spain, led by Prof. Enrique Fernandez. The group is involved in many experimental activities, which are mainly related to high-energy physics programs such as ALEPH, ATLAS and LHCB, all at CERN, and Hera-B and the Magic Telescope to be located at the Canary Islands. The HEP group has good experience in large software projects, and was responsible, within the ALEPH experiment, for setting up the FALCON computing facility, where the raw data from the experiment was processed in a quasi-online mode, and for developing the corresponding software. This facility operated very successfully in ALEPH from 1989 to 1995. Current software activities of the group are: testing of algorithms for the ATLAS Event Filter, development of Data-Flow and Monitoring Software for the Event Filter computer farm, Monte Carlo simulation of hadronic showers, studies of the performance of the calorimeters for Jet/E_T^{miss} identification, and setting up the computer infrastructure at the group to prepare for all the tasks related to computing in the LHC era.

The group and CNRS-Marseille are both members of the ATLAS experiment and collaborate in developing the software for the third-level trigger (or event filter) stage. The Event Filter is a computing processor farm that receives candidate events, each 1MB in size, at a 1kHz rate, and through filtering algorithms achieves a factor of ten reduction. Simultaneously with rejecting bad events it also tags the events that pass the selection, classifying them into appropriate classes, as needed for further analysis. The system requires a complex and large data flow. Within this project, the HEP Group will contribute to the development of the software needed to adapt all the HEP code to the Grid infrastructure by attempting realistic analysis problems on simulated data, both within the Jet/E_T^{miss} analysis and the Event Filter tasks, and will participate in the "Mock Data Challenges" project in ATLAS.

CVs of key personnel

Miquel A. Senar

Associate Professor at the Computer Science Department of the UAB. He received the Licenciate and PhD degrees in Computer Science from the University Autonoma of Barcelona (UAB), Spain, in 1988 and 1996, respectively. He has been working in the research projects of the CAOS group in topics related to resource management and scheduling in distributed and parallel systems, software support for parallel and distributed computing, modelling of parallel programs and architectural alternatives for high performance computing. He has published more than 25 scientific papers on these topics. He will be responsible for the UAB CAOS group's participation in Workpackage 3.

Andreu Pacheco

Senior Applied Physicist at UAB/IFAE, where he obtained his Doctorate in 1990. Before joining the staff of IFAE he worked at the UAB Computer Department and at CERN, where he was a Technical Fellow from 1995 to 1998. He played a key role in setting up the FALCON computing facility for the ALEPH experiment and is now responsible for all the computing facilities at the IFAE, in addition to work on the Event Filter software for the ATLAS experiment. He will be responsible for the UAB High-Energy Physics group's participation in Workpackage 1.

17 Universidade de Santiago de Compostela (U.S.C., AC17 – CR15)

Description

The Computer Science Institute of the University of Santiago de Compostela is a scientific and technical development centre. The Institute has two main objectives:

- to promote the co-ordination of several research groups of the University of Santiago de Compostela having research interests in the fields of Computer Science and Information Technologies, and
- to transfer the latest scientific advances in the field of Computer Science and Information Technologies to the Galician and Spanish society. In order to attain the later, the Institute collaborates with companies and public administrations in the realisation of research projects and in the organisation of advanced courses and PhD programs.

The University of Santiago will subcontract the services of the Centro de Supercomputacion de Galicia (CESGA) to provide support for testbed deployment and the University of A Coruña to work together in the performance prediction tools, where they have all together a strong experience (see http://www.cesga.es/ca/defaultC.html?Proxectos/PPT.html&2). Research groups from the University of Santiago (see http://www.cesga.es/en/defaultE.html?ProjMedio.html&2) and from the Spanish based multinational energy corporation ENDESA have collaborated in various projects with CESGA over the last 8 years. Such collaborations have led to the development of computing tools for the simulation of pollutant emission, photochemical alteration, transport, dissemination, and inmission for its use at the As Pontes coal fuelled power plant. CESGA's computing resources are used on a daily basis by University of Santiago researchers to run numerical simulations used to generate daily weather forecasts.

Also, the University of A Coruña and CESGA have collaborated in the development of the Superordenador Virtual Gallego (SVG, Galician Virtual Supercomputer). This is a distributed computer model that is in-production since 2000. SVG's computing resources are spread though out three research institutions in Galicia. Two of these institutions are located 90km apart. It was designed to run parallel scientific and commercial programs that do not ask for frequent communications between CPUs. The SVG has been successfully used for rendering the first 3D European animation film "The Living Forest".

Finally, the University of Santiago has a HEP group working in several international collaborations and it is working now in one of the international experiments that will run at LHC. So, they will be one of the stronger users of GRID facilities in the next future.CESGA, the Galician based supercomputing centre, has accrued significant experience in the management of high performance computing and networking resources. CESGA has provided horizontal services to the scientific community of Galicia and to the research centres of the National Higher Research Council (Consejo Superior de Investigaciones Científicas - CSIC) in the areas of computing, data storage and communications since 1993. CESGA manages and provides access to the following resources:

- High-performance computing systems of various architectures (Vector Processors, Vector Parallel Processors, Massively Parallel Processors, Shared Memory Processors and Distributed Clusters Systems).
- ATM broad band communication networks supported on fibre-optic and radio links.
- Massive data storage equipment.
- Large scientific databases.

CESGA's personnel have extensive knowledge and experience in the areas of resource monitoring, administration (focused on obtaining maximum performance) and accounting for all systems mentioned above. CESGA is a Limited Company providing horizontal services to different institutions with varied scientific research and development interests (computational chemistry, meteorology, high-end computing architectures and methods, telecommunications technologies, particle physics, engineering, etc.).

CVs of key personnel

Francisco Fernández Rivera

Associate Professor of Computer Architecture at the University of Santiago de Compostela (Spain). He attained his B.S. degree in 1986, and his PhD in 1990 at the University of Santiago de Compostela. Throughout his career he has conducted research and published extensively in the areas of computer-based applications, parallel processing and computer architecture. His current research interests include compilation of irregular codes for parallel and distributed systems, and its applications, the analysis and prediction of performance on parallel systems and the design of parallel algorithms of sparse matrix algebra.Javier D. Bruguera B.S degree in Physics and PhD from the University of Santiago de Compostela in 1984 and 1989, respectively. Currently he is a Professor with the Department of Electronic and Computer Engineering at the University of Santiago de Compostela. His research interests are in the area of computer arithmetic and processor design, VLSI design for signal and image processing and parallel architectures. He has participated in several research projects related with the development of parallel codes for some important applications and currently he is participating on the development of a parallel version of a known and widely used pollution model. He is author or co-author of a large number of papers in scientific journals and symposiums.

Jose Angel Hernando

Assistant Professor at the University of Santiago de Compostela, Spain. He worked initially on the development of the microvertex detectors on the DELPHI experiment at LEP. He was part of the NOMAD collaboration during 1994-1997, a neutrino oscillation experiment at CERN. Later, he worked on the GLAST experiment, a satellite telescope that will orbit Earth in 2005 to study gamma rays, in particular on the design of the telescope. GLAST was approved by NASA in competition with other proposals in 1999. He has worked at CERN, Switzerland, for 4 years (1993-1995 and 1997), as a student of the University of Valencia, Spain. He spent a year (1995-1996) at Harvard University, Massachusetts, as a visiting scholar of the University of Massachusetts, Amherst. He has also been a postgraduate researcher for three years (1997-2000) at the University of California, Santa Cruz. Currently, he is working on the inner tracker of the LHCb experiment, using microstrips silicon detectors, as a member of the Santiago group.

Ramon Doallo Biempica

Ramón Doallo Biempica is a Tenured Professor in the Electronics and Systems Department at the University of A Coruña. He obtained his Ph.D.in Physics from the University of Santiago in 1987. He has extensive research experience in the area of massively parallel computing and he has participated in several research projects related with the development of parallel codes. Among other significant research projects, it is worth mentioning his contributions to the parallelisation and optimisation of the STEM-II model that will be used in this project. He has published 19 scientific papers related to high performance

18 National Centre for Scientific Research ‘DEMOKRITOS’ (Demo, AC18 – CR15)

Description

The Institute of Nuclear Physics (INP) of NCSR Demokritos is one of the eight Institutes of the National Centre for Scientific Research ‘Demokritos’. The High-Energy Physics Group of the Institute has 6 senior physicists, 3 post-doctoral researchers and several PhD students. Members of the Institute participate in the ALEPH and DELPHI experiments on the LEP collider of CERN, Geneva, as well as in the CMS experiment on the LHC. The HEP group of INP participates in the CMS Trigger/DAQ and the electromagnetic calorimeter. It has significant expertise in large-scale computations both in the context of high-energy physics as well as in the area of image and speech processing. The group has a local cluster of PC’s which is currently used for Monte Carlo simulations of physics processes. The Institute has a long tradition of collaboration with other HEP Institutions in Greece, such as the University of Athens, the National Technical University of Athens and the University of Ioannina.

CVs of key personnel

Christos Markou

Research Scientist at NCSR, he received his PhD in Physics from Imperial College in 1989. He worked as a post-doctoral researcher at Imperial College, then joined the PPE division at CERN in 1994 as a fellow. He is group leader of the ALEPH group in NCSR Demokritos, Director of the Artificial Neural Networks Lab, and a member of the trigger/DAQ group for CMS. His main research areas include high-energy physics, statistical techniques, neural networks, and image and speech recognition.

19 Aristotle University of Thessaloniki (A.U.Th., AC19 – CR15)

Description

The Aristotle University of Thessaloniki, the biggest in the Balkans, has a wide range of undergraduate programs and rich research programs in various fields. Currently it spends in excess of 60 MEuro on research per year. The University offers education to 80,000 graduate and postgraduate students and, at the same time, gives the opportunity to 4,000 scientists and researchers to work in their fields, to cooperate with other universities and public and private bodies, and to present the results of their research activity to the international scientific community.

The Physics Department has about 100 faculty members and several projects in applied and basic research. The Division of Nuclear and Particle Physics has a Theory group in nuclear and particle physics and an Experimental group with activities in nuclear and particle physics. The Nuclear and Particle Physics Laboratory is part of the Physics Department. The main research activities of the Laboratory are: elementary particle physics, positron physics, nuclear reactions, environmental radioactivity and termoluminescence. It consists of 11 faculty members, three post-doctoral researchers and seven PhD students.

The High-Energy Physics group is participating in the ATLAS experiment The group is heavily involved in the construction and test of the MDT chambers of the ATLAS Muon Spectrometer.

Within the CrossGrid project, the HEP group will contribute to the development of Grid applications, in particular to distributed data analysis for HEP and to the organization of testbed activities, in particular will contribute in the integration of local testbeds and test applications for HEP.

CVs of key personnel

Chara Petridou

Associate Professor in Physics at the University of Thessaloniki. She has 25 years of experience in particle physics, working at Brookhaven National Laboratory in USA in proton-antiproton experiments, and at CERN and INFN (in Pisa and Trieste) in the UA2 and DELPHI experiments. She was involved in online-data acquisition in HEP, experimental data reconstruction and analysis, and in detector development. Since 1995 she is a Professor in Physics at the University of Thessaloniki and the team leader of the ATLAS group. She organised the construction of the MDT chambers for the ATLAS Muon Spectrometer in Thessaloniki and developed a laboratory for detector construction and testing. She is a member of the Greek CERN scientific committee and a member of the governing body of the Greek Physical Society for the Study of High-Energy Physics. She will be the coordinator of the Thessaloniki University group in CrossGrid.

20 Laboratorio de Instrumentacao e Fisica Experimental de Particulas (LIP, AC20 – CR15)

Description

LIP is the Portuguese High-Energy Physics laboratory. LIP is a non-profit technical and scientific association for research in the field of experimental high-energy physics and associated instrumentation. LIP was created in 1986 to consolidate the Portuguese high-energy physics community after Portugal joined CERN. LIP has two main centres in Lisbon and Coimbra with competence in Data Acquisition Systems, instrumentation, networking, computing and development of particle detectors. The Lisbon centre has a branch in the University of Algarve in Faro.

LIP is involved in several experiments at CERN, namely ATLAS, CMS, DELPHI and NA50. LIP's activities also include the AMS experiment, which will be installed in the International Space Station, and the development of detectors and computational methods for medical applications based on high-energy physics technologies.

LIP is already involved in grid activities. A project for the deployment of a Grid testbed site in Lisbon was submitted in September 1999 to a Portuguese funding agency, and started in January of 2001. A computing farm is being implemented, composed of commodity systems running Linux with the Globus grid toolkit. A certification authority was also implemented in the context of the grid activities to provide certificates for system and user authentication.

Contacts have been made with FCCN, the institution in charge of the Portuguese academic and research network, which will be a key player in providing the network services required for the grid connectivity both at national and international level. FCCN is the Portuguese partner in the GÉANT project and will be responsible for the management of the virtual private network and quality of service mechanisms required for proper connection between the Lisbon testbed site and other Grid sites.

CVs of key personnel

Gaspar Barreira

Research Co-ordinator and Director of LIP. He is a research scientist with extensive experience in the fields of data acquisition systems, trigger technologies and data transmission technologies. He is also the Portuguese representative at the High-Energy Physics Computing Co-ordination Committee (HEPCCC) and has been involved in several European and international projects.

Jorge Gomes

LIP Senior Research Scientist in the field of computing, with experience in telecommunications, networking, security and system administration. He has been involved in several national and international projects on data acquisition systems and networking for High Energy Physics. Presently he is head of the LIP computer centre in Lisbon. He is also working on the deployment of the Portuguese LHC computing infrastructure.

Joao Varela

Research Physicist at LIP, Associate Professor at the Technical University of Lisbon, and project associate at CERN. He was director of LIP from 1986 to 1993. Coordinator of the CERN CMS calorimeter trigger since 1994 and resource manager of the CMS trigger and data acquisition project since 1995. He is also coordinator of the Portuguese participation in the CMS experiment.

21 Algosystems S.A. (ALGO, AC21 – CO1)

Description

Algosystems S.A. was incorporated in Athens, Greece in 1986. Its primary activities are in the IT, telecommunications and networking system integration business. It is also involved in automation and control systems and systems integration. The company is actively involved in applied research projects specializing in multi-service networks, wide-area environmental management and surveillance systems and the application of robotics in metal manufacturing processes. It employs about 90 people, 55 of whom are software specialists and engineers.

Algosystems main activity consists of designing and implementing customer- oriented turnkey solutions, and providing project management services for the whole process from inception through to commissioning. It also provides advanced customer training and on going after sales hardware maintenance and software support services.

In industrial automation and wide-area surveillance, Algosystems designs, and implements, integrated multi vendor applications in building management systems (BMS), shipping, transportation, petroleum processing and industry in general as well as in environmental surveillance. An extension of this activity is the provisioning of metrology services. These services include the calibration of industrial test equipment used in the manufacturing and process industries under strict quality control conditions. To deliver these services, Algosystems has developed state-of-the-art. laboratories, two of which are mobile, and has been qualified to ISO-9002 and accredited to EN-45000 standards.

The Communications & IT Division provides total end-to-end solutions and applications in the area of telecommunications and networks. Algosystems has the know-how to offer complete solutions through a line of services ranging from consulting, design, pre-sales support, installation, maintenance and training and after-sales support through to remote help desks. In order to cover the increasing needs for custom made solutions in its Application Software Section, Algosystems has an *Application Software Team*, which is an inseparable part of its turnkey solutions.

The company has invested, along with CISCO SYSTEMS Corp., in the creation of a new IP-ATM technology laboratory in Greece. This lab will be sponsored by the E.U., and will allow Algosystems to demonstrate new IP-based technologies and solutions in Greece.

CVs of key personnel

Yannis Perros

Electronic and Computer Engineer specialising in computer networking and Internet services. Ten years of experience in IT applications developments, systems integration, environmental management applications, network and service management, mobile telephony management, advanced services, and service engineering. He was leader of Algosystems contribution in the DEDICS project (DG XIII), and he is currently responsible for the development of Internet services and applications with the R&D Division of Algosystems S.A.

Catherine Makri

Electrical and Computer Network research engineer. Responsible for the design, analysis, implementation and testing of network testbeds, and alpha testing of communication software. She participated in European research projects related to telecommunications and computer networking (TEQUILA, NABUKO, REFORM).

Peter Sevdagian

Electrical and Network research engineer. Responsible for the design and testing of integrated networking systems (LAN & WAN) and call centres.

Constantine Chassapis

Physicist, PhD in computer simulation of physicochemical systems. Specialised in multidimensional data analysis and time-series prediction issues. He is a HTML awarded programmer and has experience with XML applications development.

14 APPENDIX B – ABBREVIATIONS USED IN TABLES

CO – Confidential, internal
PP – Partially public
PU – Public
D – Demonstration
E – Event
P – Prototype
O – Other
R – Report
SW – Software