Project no. FP6-004265

CoreGRID

European Research Network on Foundations, Software Infrastructures and Applications for large scale distributed, GRID and Peer-to-Peer Technologies

Network of Excellence

GRID-based Systems for solving complex problems

D.IM.01 – Roadmap version 1 on Grid Information and Monitoring Services

Due date of deliverable: February 28, 2005
Actual submission date: October 2005 (Updated ver. 1.4)

Start date of project: 1 September 2004 Duration: 48 months

Masaryk University Brno (MU Brno)
Poznań Supercomputing and Networking Center (PSNC)

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	PU

Keyword List:
Grid information systems, Grid monitoring, Infrastructure monitoring, Checkpointing, Workflow, Virtual Accounts, Virtual Machines, Accounting

Version: 1.4

Table of content

Paivi Palosaari 8/11/05 12:23
Supprimé: 14

1. Executive Summary

The CoreGRID – Network of Excellence – aims at building a virtual European-wide Research Laboratory that will achieve scientific and technological excellence in the domain of large scale distributed, Grid and Peer-to-Peer (P2P) technologies. The primary objective of the CoreGRID Network of Excellence is to build solid foundations for Grid and Peer-to-Peer both on a methodological basis and a technological basis and to stay at the forefront of Excellence. This will be achieved by structuring research activities, leading to integrated research among experts from the relevant fields and, more specifically, distributed systems and middleware, programming models, algorithms, tools and environments.
The three main project objectives of the CoreGRID Network of Excellence are 1) excellence, 2) integration & sustainability and 3) dissemination.
The Virtual Institute (VI) on **Grid Information and Monitoring Services** (work package WP5) was created to integrate a group of experts in the following GRID areas:

- Information and Monitoring Services
- Checkpointing Services
- Workflow Services
- Accounting and User Management Services.

This report is the first version of the WP5 roadmap, giving an overall description of the work to be done in the first 18 project months. The work package WP5 integrates 13 institutions divided in four groups of experts investigating the Grid areas mentioned above. Each of the GRID area forms a WP5 task (5.1 .. 5.4).

The main objectives of the work package are the following:

- Providing multi-grain and dynamic monitoring for GRID resources and services
- Enabling online monitoring of status and performance for a large range of resources
- Providing GRID resource management framework based on the monitoring infrastructure
- Providing monitoring of the job workflows progress
- Support for extraction and representation of job workflows from programming models
- Realizing middleware support for complex job workflow execution
- Investigating accounting services.

2. Introduction

Context

The purpose of the CoreGRID Network of Excellence is to promote first class joint research on Grids and Peer to Peer systems. This broad research topic has been structured into six complementary and mutually interdependent research areas, each characterized by an important topic. These areas are structured by six CoreGRID Virtual Institutes, parts of the virtual CoreGRID Research Laboratory:

- WP2: Knowledge & Data Management (KDM)
- WP3: Programming Model (PM)
- WP4: System Architecture (SA)
- WP5: Grid Information and Monitoring Services (IM)
- WP6: Resource Management and Scheduling (RMS)
- WP7: Systems, Tools, and Environments (STE).

As mentioned, the primary objective of the CoreGRID Network of Excellence is to build solid foundations for Grid and Peer-to-Peer both on a methodological basis and a technological basis, and to stay at the forefront of Excellence. This will be achieved by structuring research activities, leading to integrated research among experts from the relevant fields and, more specifically, distributed systems and middleware, programming models, algorithms, tools and environments. The three main project objectives of the CoreGRID Network of Excellence are 1) excellence, 2) integration & sustainability and 3) dissemination.

The roadmap for one of these Virtual Institutes—the Grid Information and Monitoring Services (CoreGRID Work Package 5)—is presented in this document to help coordinate research and promote scientific expertise between its members. The Grid Information and Monitoring services gather, transport and provide vital information about the Grid, its individual components and their combination, provides the workflow mechanism and information, supports the end user and administrator by defining multi-level checkpointing, and user account management. Finally, it delivers accounting information about the used resources.

These services are used by other Grid components and also by Grid users to make qualified decisions about the management and use of the Grid and make it as efficient as possible. While often hidden and only indirectly accessed, the information and monitoring services provide an information infrastructure that is indispensable for all the other Grid components and users. Also, this infrastructure creates a specific Grid, which must be built and maintained in a similar way as the Grid directly accessible by end users.
Besides the integration of research already done by individual partners within this Virtual Institute, these services (information and monitoring, checkpointing, workflow, accounting and user management) are also essential to other Virtual Institutes as they provide
data for evaluation of the efficiency of systems and tools resulting from their research and support core functionality necessary for production grid environments. Strong collaboration is thus essential and we foresee mutual benefits resulting from the common work within the CoreGRID NoE. This synergy is to be especially expected between this Virtual Institute and WP2, WP4, WP6 and WP7, as all the developed tools and environments must become part of the monitored infrastructure and will provide data to be evaluated and fed into the scheduling systems. Also, the CoreGRID testbed could not be built without strong interaction with this Virtual Institute.
The work will be mostly organized into Research Groups, units of two or more CoreGRID partners collaborating closely together on common goals.

The CoreGrid NoE will integrate the research results in each Virtual Institute. The work package WP5 defined several research groups in each of the four tasks. The added value of WP5 is a common architecture in terms of monitoring and information infrastructure, checkpointing, workflow services, user account management and accounting. The WP5 partners are engaged in national and international projects, the major R&D results will be taken from the following projects:

- SZTAKI: “Hungarian Supercomputing GRID” (No.: IKTA4-075), “Cluster Programming Technology and Its Application in Meteorology” (No.: IKTA3-029), “Chemistry GRID and its Application for Air Pollution Forecast” sponsored by the Ministry of Education (No.: IKTA5-137, “Hungarian SuperCluster project” (No.:IKTA-00064/2003)
- PSNC: Pionier national programme (www.pionier.gov.pl), Progress (“Polish Research on Grid Environment for SUN Servers” project no. 6.T11.069.2001C/5688), SGIGrid (“High Performance Computing and Visualization with the SGI Grid for Virtual Laboratory Applications”, no. 6 T11 0052 2002 C/05836), Clusterix (National Cluster of Linux Systems, 2003-2006)
- FHG: K-Wf Grid ("Knowledge-based Workflow System for Grid Applications", STREP project), Fraunhofer Resource Grid, D-Grid (German Grid Initiative)
- FZJ: UNICORE, OpenMolGRID, NextGRID
- MU: MediGRID (national project), GridLab (IST-2001-32133, www.gridlab.org)
- UoW: UK EPSRC funded OGSA testbed project.

It is worth mentioning that many of WP5 partners are active in the research and working groups of Global Grid Forum.

Cbitoune 18/10/05 18:45
Mis en forme: Retrait : Gauche : 0 cm, Avec puces + Niveau : 1 + Alignement : 0,63 cm + Tabulation après : 1,27 cm + Retrait : 1,27 cm, Tabulations : 0,63 cm, Tabulation de liste + Pas à 1,27 cm

Problem(s)

A large-scale heterogeneous GRID is subject to frequent changes and disruptions in the service of a huge number of components it is built of. To manage such a dynamic system and its resources, online monitoring of the resources to determine their availability is required. Without this information, the challenging goals set up in the report “Next Generation Grid(s), European Grid Research 2005-2010” (also known as the NGG report) [35,36] can not be achieved. While the monitoring is not extensively mentioned in the NGG report, it is an essential part of any Grid infrastructure, helping to hide its complexity. Current Grid information and monitoring frameworks have identifiable drawbacks, as they are either too focused on specific aspects or do not scale enough. The same information is often used in many places, at a different aggregation level, and allowing each component needing such information to collect it independently will impose on high overhead in the whole Grid. The centralized approach used currently for the monitoring and management of large scale Grids, does not really scale beyond tens or at most hundreds of nodes/resources and is not sufficient when dealing with problems spanning several administrative domains. Despite the recent progress in this area, the level of architectures and concepts (like the

Grid Monitoring Architecture [37])is still either too high or there are too specific implementations, like MDS [39], R-GMA [37], Mercury [40], or iGrid [38], which are not yet able to provide a reliable information and monitoring system for a general heterogeneous large scale Grid. The most important design issues are shared with the WP4 Virtual Institute—namely scalability, reliability and robustness of such a system. Additional problems lie in a blur line between information and monitoring systems (Is it possible to merge the two into one infrastructure?) and also in further processing of monitoring information it is not possible to collect all the data, so powerful filters, aggregation functions and distributed storage for most important logs must be developed. Also, in a web service oriented world, what are the "natural" interfaces to the monitoring information and to which extent an active monitoring system based on probes and sensors could be complemented with a passive system based on the service and resource instrumentation. The throughput of network lines between Grid nodes adds another complexity layer, as there may not even be enough processing power to work in the real time. In the network layer of the grid, we would like to investigate services like accurate traffic monitoring and classification, scalability properties, performance aspects, provision of fault detection, dependability and security services for Grids.

The Virtual Institute will focus its research on better understanding of the reasons and finding models and frameworks to overcome these limitations. In close collaboration with the Virtual Institute on system architectures, appropriate architectures for scalable and dependable information and monitoring infrastructures will be investigated and deployed. A possibility to converge the currently distinct approaches to information services and monitoring services will also be taken into account, with the aim to propose a unified framework.

Monitoring is not used only to detect failures or other types of problems. The collected data could be further processed to give an understanding of the overall Grid performance. We plan to develop new models and approaches that would provide metrics for Grid performance evaluation. These will be used, among others, to compare the influence of different Grid components on the overall behaviour of the Grid.

While monitoring is able to provide information about the failed node, the checkpointing is needed to minimize effects of such failures on user jobs and tasks. The contemporary operating systems were not designed to directly support checkpointing even in a single processor on network environments. As a result the higher level models that harness the checkpointing mechanism are not well developed either. However, it is a very important research activity, as checkpointing not only increases efficiency, but in some cases it is the only way to get results in a prescribed timeframe even in the presence of failures. Despite numerous technological issues the new checkpointing implementations for different platforms still emerge, but their functionality is diverse and is not directly usable in the heterogeneous Grid environment. Additionally, due to technological and semantic issues, some limitations according to applications that can use checkpointing are imposed. To obtain the ability to utilize the existing and future low level checkpointing implementations, there is a need to define the alignment of that technology in the context of Grid environment. Additionally, the interfaces to the surrounding Grid environment and low level checkpointing mechanisms must be defined, including possibly direct interfaces with the monitoring infrastructure.

Complex job workflows represent another challenge, as the monitoring information must be almost synchronously gathered from many different sources and appropriately processed to provide a coherent view (state information) of the whole workflow and its components. The job workflow itself must be extracted from programming models, the monitoring and information services must be tightly coupled with job checkpointing and migration support to provide an environment where even complex job workflows could be easily deployed, executed, and monitored. Models and methods to provide a virtualized end user account system are a specific part of combined job flow support and information services.

The problem of managing user accounts becomes a non-trivial one in a distributed environment (especially in a grid environment) that includes many independent sites and virtual organizations with hundreds or even thousands of user accounts. The complexity rises from the point of view of time required for administration tasks and automatization of these tasks. The accounting issue becomes impossible in the distributed environment, but still possible on a single computer system.

The next important problem is accounting in distributed systems. The solutions existing/available on the market allow the accounting of resources used for only one system or, at most, for local and homogeneous clusters. The problem arises when we make the environment more complex by adding virtual organizations and dynamically assigned accounts. In the existing testbed installations the problem is frequently neglected. However, the demand for a solution will significantly increase in the near future, especially in context of production grids and grid economy. The GGF created the Accounting Models Research Group, whose goal is to work out the rules of data exchange and a general communication interface between sites, allowing information about resources used in grids to be collected. The problem becomes more complex the greater the environment is, i.e. the number of VOs and the number of users accessing the resources. Obviously, the accounting mechanism should provide us with some information independently of the service scale and range (a single system, local cluster or Grid environment). The next important feature is the scalability and low time overheads generated by this mechanism. It is significant in a production environment characterized by highly dynamical changes. To ensure a real practical accounting mechanism for the Grid, it should have a decentralized, scalable and flexible structure. It should interfere with local domain policies as little as possible. The accounting should allow the data for a single user, group or the whole VO to be processed. The remote system tracks all information about resource usage by the user. The information is sent after the completion of the computations to the VO the remote user belongs to. To meet these requirements, distributed resource allocation accounting models, which would properly work at various sites with different administrations and resource management policies, would be necessary. As mentioned above, the possibility of accounting the resources used in grids forms the basis for introducing the Grid economy concept.

Objectives

The main objective of this Virtual Institute is to provide a general framework for large scale distributed information and a monitoring service for Grids. While the main focus will be on the general infrastructure, the work will also be targeted to provide the design interfaces for a distributed passive traffic monitoring infrastructure in order to enable accurate network performance monitoring, and Grid security services on top of this infrastructure. The provided services must be able to adapt to perennial changes in the GRID state, providing the necessary level of robustness and fault tolerance. Scalability of the provided services and anonymity of information exchanged between monitoring resources that span a number of different administrative domains is of paramount importance.

The Virtual Institute also focuses on the definition of the position of checkpointing technology within the Grid environment and work out proposals of checkpointing-related services and interfaces. Contemporary Grid environments are featured by an increasingly growing virtualization and distribution of resources. Such situations impose greater demands on load-balancing and fault-tolerant capabilities. The checkpoint-restart functionality is one of the components which can fulfil the specific requirements.
The main objective of the task is to define a high-level checkpoint-restart Grid Service and to locate it among other Grid Services. We aim to define both the abstract model of that service and the lower layer interface that will allow the service to cooperate with diverse existing and future checkpoint-restart tools. The model will be described under the Grid Checkpointing Architecture.
The specific objectives include :

- analysis of the existing approaches
- analysis of user demands
- proposition of the Grid Checkpointing Architecture, including high and low level checkpointing services (kernel, user and application levels)
- presentation of the first version of integrated functionality: kernel level and MPI.

Another aspect of the research provided by this Institute is the study and development of services able to coordinate the reliable execution of vastly complex compound GRID jobs and realize middleware support for complex job workflow execution. This will also include an adequate description and modelling of workflows, mapping abstract onto concrete workflows and providing services for the monitoring of workflows, thus adding support for dynamic workflows on non-reliable GRID resources. Future Grids must support the management of complex jobs and service-level agreements. Those jobs have workflows with co-allocation constraints and dependencies that must be considered for a wide range of diverse resources. Present systems have architectural and design limitations that make them usable in a productive manner only for simple workflows. This is not sufficient for highly complex GRID applications to be expected in important application domains such as industrial design, engineering, drug design and bioinformatics. The current limitations must be overcome by a common set of job management and execution services based on a powerful model.
The main objectives of the Workflow Services Task of WP5 are the research and collaboration on the following topics:

- services able to coordinate the reliable execution of vastly complex compound GRID jobs and realizing middleware support for complex job workflow execution
- user interfaces and Web portals for GRID workflow management
- adequate description and modelling of workflows
- services for the monitoring of workflows
- dynamic workflows on non-reliable GRID resources

- mapping abstract onto concrete workflows
- relation between Business and Grid workflows.

Within its last task, the work will focus on the development of new user account management (and accounting) services suitable for the Grid environment, giving adequate encapsulation, scalability, and robustness.
The main objective of the user management system is controlled, secure access to grid resources. The considerations we try to introduce must take into account the fact that we are dealing with a production GRID environment, which has the ability to dynamically change its configuration. Security requires authentication of the user and authorization based on combined security policy from the resource provider and virtual organization of the user. The second important issue is a possibility of logging user activities for accounting and security reasons and then gathering these data both by the resource provider and virtual organization of the user. From the user's point of view, an important feature is single sign-on.

The problem of user management is a non-trivial one in an environment that includes a bulk number of computing resources, data, and hundreds or even thousands of users participating in numerous virtual organizations. The complexity rises from the point of view of time required for administration tasks and automation of these tasks. There are many solutions that attempt to fulfil these basic requirements and solve the mentioned problem, but none of them, according to our best knowledge, solve the problem in a complex and satisfactory way.
The specific objectives are the following:

- defining and understanding the general requirements coming from the end user and service provider
- collecting and analyzing best practices of systems developed by the research groups
- working out a model of full user accounts virtualization in terms of user account management and Grid accounting
- defining a preliminary version of services able to be deployed in production Grid-based environments.

The overall objective of this Virtual Institute is to bring together the fragmented research of individual institutes in the area of GRID information and monitoring systems and to provide a critical mass of researchers to achieve excellent research results. This will also include continuous identification of weak areas and gaps in knowledge, where new research will be pointed out. .

Tasks

The Virtual Institute - Grid Information and Monitoring Services defined the following tasks in the JPA1:

- **Information and Monitoring Services (5.1)**
- **Checkpointing Services (5.2)**
- **Workflow Services (5.3)**
- **Accounting and User Management Services (5.4).**

The Information and Monitoring Services (task 5.1) will concentrate on the following specific subtasks:

- Establishing a definition for 'GRID performance'. Investigating the relationship of 'application performance' and 'infrastructure performance'
- Elaborating GRID performance metrics. Current performance metrics are not characteristic, comparable and interpretable in a multi-site, diverse and heterogeneous environment.
- Enable instrumentation through provision of notification, specification of events and logging of these events.
- Designing a framework for performance steering and management. It involves analyzing the monitored data, defining appropriate levels for making (possibly symbolic) decisions based on the performance and giving feedback both to the infrastructure and the application. The essence of this design is establishing an appropriate transformation between multidimensional local and GRID metrics and vice versa.
- Exploring the statistical properties and establishing stochastic models for 'application performance', 'infrastructure performance' and 'GRID performance' for supporting resource brokering together with an advanced information system.
- Scalability to a large number of resources and a multitude of events produced by these resources
- Provision of different levels of granularity to ensure a consistent view on the GRID through the monitoring and to enable prompt reactions on sudden changes
- Exploitation of monitoring data for fault and security management.

The primary task in Checkpointing Services (task 5.2) according to the aforementioned problems is to locate the checkpointing technology within the Grid environment. The joint outcome of the established research group is to work out a general model of the Grid Checkpointing Architecture. A general understanding of the Grid layers and physiology is required to accomplish this task. The following subtasks have been defined for the first 18 months of the project :

- Analysis of the overall Grid architecture to gain knowledge allowing to locate checkpointing within the Grid context
- Utilizing the gained knowledge to place the checkpointing technology within the Grid. It is important to make a precise definition of the lower and higher layers and services
- Defining API interfaces to layers. At present we can assume that interfaces to higher and equivalent layers will base on OGSA services while the interfaces to lower layers will use some general techniques allowing to reference and access the existing and future, checkpointing systems
- Definition of the Grid Checkpointing Architecture, integrated with the Grid architecture.
- Providing a preliminary version of reference implementation, including kernel and MPI level checkpointing services. The working model of the checkpointing service will utilize one of the checkpoint restart packages developed by PSNC and PGRADE development framework developed by MTA SZTAKI.

The Workflow Services (task 5.3) will focus on the following specific issues:

- Analysis of workflow approaches in Grid computing
- Knowledge interchange in workflow orchestration
- Workflow description languages and using coloured Petri nets for Grid workflows
- Collaborative workflow-oriented portals related to collaborative workflow management.
- Fault-tolerant workflow manager service taking into consideration the CoreGRID requirements and the current trends of service-oriented Grid development
- Investigating the problem of workflow management over different Grid middlewares
- Identifying the appearance of transactions in Grid environments, the roles of transaction managers in Grids
- Proposed solutions to execute different parts of a workflow in different security domains or in Grids that use different middlewares.

Parts of the roadmap of task 5.3 are strongly related to the other tasks of this work package, e.g.:

- Information and monitoring services (Task 5.1) are required for mapping abstract workflows and for monitoring workflows.
- Checkpointing services (Task 5.2) can be used in order to achieve workflow checkpointing for workflow rollback (transactional safety) or migration of sub workflows.
- Accounting and User management Services (Task 5.4) are also important on the workflow level.

The activities of task 5.3 will have a strong interaction with Task 6.3, which covers scheduling strategies for GRID workflows.

The Accounting and User Management Services task (5.4) will focus on the following specific subtasks:

- Analysis of requirements (basing on the current state of development)
- Design of a mechanism that will allow management of user accounts adjusted to the CoreGrid testbed, local security policies and VO resource management policies
- Definition of an interface for exchanging information between different user account management systems from different grids
- Attempt to standardize the above interface
- Analysis of requirements for the accounting gathering system
- Definition of a minimum set of resources that the accounting system in the production grid must gather and process
- Definition of a database structure that will allow to store other and non-standard accounting data
- Definition of a protocol for exchanging the accounting data
- Design of a model architecture for the accounting system that will comply with different resources from different grids and will take into account local and VOs policies
- Implementation of a model system.

Section 4 is positioning all the specific subtasks according to partners and research groups defined in the Virtual Institute - WP5.

Drivers

The driving forces of the planned WP5 work is to give the best service which would allow to release a production Grid environment. The vision described by a group of experts in NGG and NGG2 reports gives [35,36] key research priorities in terms of properties, facility and models, cit.:

The motivating examples indicate that a Grids environment (particularly the Grids Service Middleware) should offer (in addition to those characterised in NGG1 Report) the following:
(a) ***flexible, dynamic, reconfigurable resources available on de****mand and to the level required for the application and / or end-user;*
(b) accurate, ***relevant information presented*** *in the optimal way for the end-user which implies reconciliation across heterogeneous distributed information systems;*
(c) context awareness, task awareness and service negotiation capability driven from the user interface;
(d) pre-emptive and proactive services as seen by the end-user;

The main goal of R&D to be done in the future is to have a dynamic and reconfigurable environment, which would meet the requirements of industry and scientific communities.

Therefore, it is very important to work out services defined by tasks 5.1 .. 5.4 of the work package WP5.

3. Positioning

State of the art

This section outlines the existing approaches and their problems, tradeoffs, and limitations. It is structured according to the topics covered by the work package .

Monitoring and Information Services

Over the last few years an increasing number of passive monitoring infrastructures have recently started to appear, focusing mostly on the network monitoring. For example, NLANR, the National Laboratory for Advanced Network Research in the United States has installed a large number of passive Internet monitors operating at speeds between 155 Mbps and 2.5 Gbit/s. In passive monitoring systems, network sensors capture all packets, including both headers and payload that pass through their monitored network. Based on the headers and payloads of captured packets, passive monitors are able to produce a wealth of information including high-level performance metrics, as well as detection of attacks.

Besides the US-based NLANR infrastructure, there are also European passive network monitoring projects. SCAMPI (A Scaleable Monitoring Platform for the Internet), for example, is an IST-funded project that builds a hardware monitor along with the necessary system and application software to facilitate passive network monitoring at speeds as fast as 10 Gbit/sec. It is planned underway to deploy SCAMPI to several places in Europe through LOBSTER (Large-Scale Monitoring for Broadband Internet Infrastructures) a new European project partly supported by the European Commission (http://www.ist-lobster.org), and GÉANT2. SCAMPI, LOBSTER, and GN2 may significantly enhance our understanding of Internet traffic including Internet-based attacks. Being based on passive monitors, and having access to all network traffic including all packet's headers and payloads, these projects may have significant amounts of information that can be used to identify attacks. This information may even be used to automatically generate signatures for new, not previously seen, attacks. Therefore, the mentioned projects but also passive monitoring systems in general, may significantly help us towards improving our Internet security.

Flow-level Statistics: Network traffic monitoring has been around for quite some time now since the days of the ARPANET and the NSFNET. Currently, the most common network traffic monitoring systems are based on flow-level statistics, such as those provided by Cisco's Netflow, IPFIX, FlowScan, and similar approaches. Such systems provide high level statistics, such as volume of traffic, timestamps, number of packets, etc. on a per-flow basis. A flow is loosely defined as a set of all packets that share a common source IP address, destination IP address, source port, destination port, and protocol number. Since flow-based statistics do not collect any information about the packets' payload, they provide inadequate support for monitoring applications that depend on information found on the payload.

Summarizing, although flow-level-statistics-based approaches have been widely used for network monitoring, they provide limited monitoring support against applications that use dynamically-generated ports (such as peer-to-peer systems), and slowly-moving or carefully-hidden cyber attacks.

Active Monitoring: Having realized the limitations of flow-level statistics, several network monitors resort to active monitoring. Active monitoring is a broad term that collectively describes a family of monitoring methods based on sending probing packets from a sender towards a (usually) cooperating receiver. Based on the packets' arrival, the sender is able to find several performance aspects of the network, including latency, bandwidth, and error rate, etc. Active monitoring is currently being widely used in several countries. For example, RIPE NCC has installed more than 50 test boxes, which periodically probe each other to find out the status and performance of their Internet connection. At the other side of the Atlantic, several organizations, including NLANR, CAIDA, and SLAC have been working in active monitoring. Recently, Internet2 has started installing an active monitoring infrastructure known as E2E PIPES (End-to-End Performance Improvement Performance Environment System). Based on tools such as iperf, and traceroute, PIPES measures latency, bandwidth, and connectivity among various hosts connected in the Internet.

To sum up, although active monitoring is being widely used to identify performance characteristics of the Internet, it provides limited support for monitoring cyber attacks (especially the stealth ones), and provides no support for accurate traffic characterization.

Passive Monitoring: In addition to active monitoring infrastructures, passive monitoring infrastructures have recently started to appear as well. For example, NLANR, the National Laboratory for Advanced Network Research in the United States has installed a large number of passive Internet monitors. Although most of them operate at speeds between 155 Mbps and 2.5 Gbps, there are plans to scale the infrastructure to speeds up to 10Gbps. The counterpart projects to NLANR in Europe are SCAMPI and LOBSTER, which aim to provide passive monitoring infrastructures (single node and multimode, respectively), able to monitor traffic at very high speeds (above Gbps).

Of all the related approaches described so far, passive monitoring is closest to the monitoring needs we have identified for CoreGrid. Indeed, CoreGrid (or more precisely, the Grid community) needs infrastructures that use passive monitoring as their core monitoring mechanism. CoreGrid monitoring requirements focus on security applications: it uses passive monitoring to provide statistics and early warnings about the spread of cyber attacks on the Grid.

Also, CoreGrid focuses on real-time processing of the gathered data in order to support collective monitoring experiments. In addition, CoreGrid will pay particular attention to the anonymization of the monitoring data that are going to be exchanged between Grid nodes that belong to different administrative domains. Thus, although administrators have access to the data needed, at the same time the privacy of the users is fully protected.

Checkpointing Services

Low-level Checkpointing Mechanism

There are three methods of implementing the checkpoint mechanism: the kernel level (operating system), the user level, and the application level. The mechanism included in the operating system takes care of storing the job image. In this case checkpointing is transparent for the applications, which means that the user does not need to modify or extend his/her applications. Using checkpointing on the user level is less convenient. The user must recompile his/her application and link it with additional checkpointing libraries. Most of the libraries periodically store the job's state. On this level the checkpointing is transparent for the applications but it is onerous for the user because of the necessity to link with additional libraries. The third method (the application level checkpointing) is even more complicated. The user must implement the checkpoint and the restart mechanism in his/her application by himself/herself.

Due to historical reasons the checkpointing mechanism is not widely available. When modern operating systems were designed, the checkpointing facility was not considered. Consequently it is hard to provide these operating systems with the checkpointing mechanism. Only few operating systems have been built in this checkpointing (IRIX, UNICOS MK). Additionally there are a few projects that provide checkpointing packages for other platforms. The links to these projects are collected on the http://www.checkpointing.org site. In the aftermath of the aforementioned historical reasons, all available checkpointing mechanisms impose some limitations on programs that are to be checkpointed. However, the new checkpointing packages for new platforms with improved functionality are still developed. For instance, PSNC has developed three checkpointing packages. These packages can be downloaded from the http://psnc.checkpointing.pl page.

Checkpointing in the context of resource management software and clusters

Some of clustering or resource management systems are able to utilize the low level checkpointing mechanisms. For example, the Sun Grid Engine (SGE) has an interface that allows using the checkpointing facility on nodes which posses this functionality. Further, the Condor project is shipped with a proprietary user-level checkpointing library and allows to submit jobs that are to be checkpointable.

An additional challenge is to checkpoint distributed applications based on the PVM/MPI communication model. The package that addresses that area is the PGRADE development framework developed by MTA SZTAKI which allows to checkpoint the PVM and MPI application based on the low level checkpointing mechanism like e.g. Condor.

Checkpointing in the context of Grid environment

Currently the Grid virtually lacks in well-developed checkpointing-related services and interfaces. As this problem has been noticed by the Global Grid Forum, they established the Grid Checkpoint Recovery (GridCPR) Working Group. The group aims to define the user-level API and associated layer of services. The home page of that group is http://gridcpr.psc.edu/GGF/.

Workflow Services

The main objectives of the Workflow Services Task of WP5 are the research and collaboration on the following topics:

- Services able to coordinate the reliable execution of vastly complex compound GRID jobs and realizing middleware support for complex job workflow execution
- User interfaces and Web portals for GRID workflow management
- Adequate description and modelling of workflows
- Services for monitoring of workflows
- Dynamic workflows on non-reliable GRID resources
- Mapping abstract onto concrete workflows
- Relation between Business and Grid workflows

There is no complete environment covering all the above-mentioned topics. Instead, there are a lot of projects related to the specific grid infrastructure or resulting from specific local needs. WP5.3 partners are involved in a number of projects dealing with workflows.

K-Wf Grid is a STREP project funded by EC that started in 09/2004 under the leadership of Fraunhofer FIRST with the objective to develop a "Knowledge-based Workflow System for Grid Applications".
The K-Wf Grid system will assist its users in composing powerful Grid workflows by means of a rule-based expert system. All interactions with the Grid environment are monitored and evaluated. The knowledge about the Grid itself is mined and reused in the process of workflow construction, service selection and Grid behaviour prediction. Workflows are dynamic and fault-tolerant beyond the current state of the art.

The OPENMOLGRID project will provide a unified and extensible information-rich environment for solving molecular design and engineering tasks. The project contains the high-level work flow description & processing in UNICORE

The goal of the NextGRID project is to develop next-generation architectural components enabling a more economically viable use of the current Grid infrastructure for business and research as well as the general public. This project will aim to work out a universal model to integrate resources in work flow processing.

The GGF Grid Scheduling Architecture Research Group was set up to standardize workflow efforts. The goal of this research group is to define a scheduling architecture that supports cooperation between different scheduling instances for arbitrary Grid resources. The considered resources include network, software, data, storage and processing units. The research group will particularly address the interaction between resource management and data management. Co-allocation and the reservation of resources are key aspects of the new scheduling architecture which will also include the integration of user or provider defined scheduling policies.

Another workflow-related project is the Ganga grid user interface which will be deployed on the LHC Computing Grid to manage analysis jobs for high energy physics. It is a front-end for the configuration, submission, monitoring, bookkeeping, output collection, and reporting of computing jobs run on a local batch system or on the grid. In particular, Ganga handles jobs that use applications written for the gaudi software framework shared by the Atlas and LHCb experiments. Ganga exploits the commonality of gaudi-based computing jobs, while insulating against grid-, batch- and framework-specific technicalities, to maximize end-user productivity in defining, configuring, and executing jobs.

UOW implemented an environment where users can construct, manage and execute visual workflows from a Grid portal in order to support the execution of Grid-enabled Legacy Codes. The middleware called GEMLCA (www.cpc.wmin.ac.uk/gemlca) has been developed by the University of Westminster during the UK EPSRC funded OGSA testbed project. GEMLCA uses Globus Toolkit 3 and is being migrated to GT4 (WSRF). GEMLCA has been demonstrated at IST 2004 in Hagues and there is an article about it in the 'Next Generation Grids' publication document.

User account management

The problem of the single sign is very important and two kinds of solutions can be considered. Information connected with user accounts can be replicated to other hosts or user accounts can be assigned dynamically. The first solution is more popular, because it has been more suitable to the local clusters and local networks of computers.

A very popular solution was to use NIS (Network Information System), which allows different machines to access the same database with user accounts. More secure solutions use Kerberos authentication to access remote hosts. The Distributed Common Environment uses Kerberos to authenticate users across the DCE cell. Also, some distributed file systems like DFS or AFS employ Kerberos authentication to access files at a remote location. All these solutions were designed for local networks and do not solve the problem of grids with hundreds of thousands of users.

The most common system to implement a Grid environment nowadays is Globus. In the Globus environment users are authenticated with X509 certificates. Logging into the Globus environment only involves a creation of an X509 proxy, and this proxy works on behalf of the user without any need for further authorization. However, in order to use any Globus resources, the users must be included in the grid-mapfile on remote hosts, where the resources are located. The grid-mapfile is checked by the Globus-gatekeeper to find the local Unix account which should be used to run processes on behalf of the Globus user ID. Users can have individual accounts or many users can be mapped to the same Unix account. Some enhancements for Globus authorization were made by the DataGrid project. DataGrid is based on the idea of virtual organizations. Each virtual organization (VO) runs a dedicated server that maintains certificates of all people that are working on the same experiments. Each computing machine downloads a list of authorized users from all approved virtual organization servers. Then these users are added to the grid-mapfile and are mapped to the pool of accounts dedicated to this VO. The accounts from this pool are assigned to consecutive users when needed, but are not recycled or reused. Hence, all users have their own accounts on all the machines they use, but they do not need to apply for them. All the users need to do is enter any approved VO.

Several Grid environments employ the idea of virtual accounts that are temporarily mapped to a pool of physical accounts when needed. Such accounts are called virtual, scratch, generic, template or shadow accounts. In PUNCH users have their own logical user-accounts and the system manages its own physical accounts on remote resources and dynamically recycles them among users as necessary. Generic accounts can also be used in Condor and Legion, but in both systems it is recommended that users have their own accounts on every machine. Condor uses a nobody UID to run jobs for users that do not have an account in a Condor flock. Legion also manages the pool of generic accounts that are assigned for Legion use. To the best of our knowledge, none of the systems using a pool of virtual accounts allows for full accounting of resources used. Some work is underway to define the mechanisms of distributed accounting on the Grid.

Accounting

There are four research groups within GGF currently working on accounting-related standards. The Grid Resource Allocation Agreement Protocol (GRAAP) working group is currently specifying a Service Level Agreement (SLA)-based protocol called Agreement-based Service Management (WS-Agreement). The Resource Usage Service (RUS) is a service that can be used to publish and query resource usage data. It heavily relies on the Usage Record (UR) format, which is a standard XML document composed of the various usage properties. The Grid Economic Services Architecture (GESA) essentially extends the OGSI Grid service model into an economic service model, where you can charge consumers for service usage. In order to achieve this, GESA defines an architecture based on OGSI-Agreement contracts, and Resource Usage services. This combines all accounting-related efforts within GGF into a common model.

The Datagrid and EGEE projects implemented the DGAS component for scheduling jobs so that resources are fully utilized to the lowest possible price. Each user has an account in a local bank called the Home Location Registry (HLR). When a job is submitted by the user, the resource broker receiving the request contacts a pricing authority at various resources and the local bank to check whether there are sufficient funds to run the job. If the bank grants the transaction, the request is passed to the job controller which sends it to an available resource that matched the user requirements. The resource monitoring service then tracks the job status and the resource usage and sends periodic reports to the HLR. When the job completes, the total cost is calculated and possible holds on amounts in the HLR are unlocked and the credits spent are withdrawn from the user HLR and deposited into an account in the resource HLR.

SweGrid (Swedish Grid) is preparing the system (SGAS) for allocating resources to project and account used resources. The system bases on and extends GGF standards.

Another approach is SNUPI, the System, Network, Usage and Performance Interface which provides an interface for resource utilization reporting for heterogeneous computer systems, including Linux clusters. SNUPI provides data collection tools, recommended RDBMS schema design, and Perl-DBI scripts suitable for portal services to deliver reports at the system, user, and job for heterogeneous systems across the enterprise, including Linux clusters.

4. Vision, Strategy and Roadmap

Vision and Scenarios

The CoreGRID aims at building a virtual European-wide Research Laboratory that will achieve scientific and technological excellence in the domain of large scale distributed, Grid and Peer-to-Peer (P2P) technologies. The three main project objectives of the CoreGRID Network of Excellence are

1) excellence
2) integration & sustainability and
3) dissemination.

The Virtual Institute (VI) on **Grid Information and Monitoring Services** is one of the 6 packages which will focus on working out the knowledge and excellence on various disciplines of Grids and P2P technologies.

> ”The Grid concept demands novel approaches to system design and management – and thus to operating system behaviour, middleware requirements and services offered to applications. *A three-level conceptual architecture is emerging. The application requirements provide a specification for the required services in the Grids Service Middleware layer, and this in turn drives requirements of the Operating System layer including the Grids Foundation Middleware required to elevate the interface of each Operating system to that required for the Grids Service Middleware.*
>
> *Furthermore, current operating systems (dominantly LINUX in work to date but increasingly Symbian and others used in embedded systems) do not provide the necessary services for effective operation of the Grids Service Middleware layer. Again, the model of the operating system controlling a single node and managing its resources exclusively (security, scheduling) is at variance with the philosophy of Grids. This leads to the concept of augmenting the existing operating systems with Grids Foundations Middleware to provide the required functionality.”*
> – Next Generation Grid report [35, 36]

The IST group of experts identified in the NGG and NGG2 reports some crucial features which will allow a migration from the ‘scientific’ Grid to the production Grid, allowing to introduce computing on demand and make the Grid suitable for industry challenges.
Some of the features are mentioned below:

”In addition to the properties mentioned in the NGG1 Report, a Next Generation Grids
environment should have the following properties in order to satisfy the requirements of the scenarios considered:
(a) ***pervasive****, with mobility as the cornerstone enhanced with more advanced pervasive computing facilities;*

(b) ***self-managing*** *with the ability to handle highly dynamic and unpredictable configuration of demanders and suppliers;*
(c) ***resilient*** *with the ability to handle highly dynamic and unpredictable configuration of the network connecting the computing nodes;*
(d) ***flexible*** *to handle various types of computing nodes and highly dynamic distribution of computation tasks among involved resources;*
(e) ***resilient*** *with the ability to handle intermittent connectivity and associated synchronisation of information sources;*
(f) ***easy to program*** *with a high-level, functional programming interface reusing the existing software modules;*
(g) ***flexible in trust*** *to allow business operations to work effectively and efficiently as virtual organisation and distributed collaborations;*
(h) ***secure*** *to assure confidence in its use for business purposes."*

The vision of the WP5 Virtual Institute is coherent to the main outcome of the NGG document, i.e. provides the Grid middleware layer with services allowing a stable migration to industry-based Grids with reliable behaviour, secure, pervasive and scalable, being able to make self-reconfiguration.

All services must be designed to establish a fault-tolerant and flexible behaviour in a large-scale heterogeneous environment. It is impossible to achieve that without relevant information about the state of services, properly collected, merged and filtered if necessary. Current models do not scale to the GRID level or are focused on specific aspects. The primary objective of the proposed Virtual Institute is to study and provide general information and monitoring service for the underlying GRID management required by the Next Generation GRID. The GRID management services considered here include GRID middleware (core) services and components as well as higher-level services and components on top of the GRID middleware.
Therefore, WP5 scope of tasks will focus on the following objectives:

- Providing multi-grain and dynamic monitoring for GRID resources and services
- Enabling online monitoring of status and performance for a large range of resources
- Providing GRID resource management framework based on the monitoring infrastructure
- Providing monitoring of the job workflows progress
- Support for extraction and representation of job workflows from programming models
- Realizing middleware support for complex job workflow execution
- Investigating accounting services.

This Virtual Institute will not be isolated. This synergy is to be especially expected between WP5 and the rest of WPs, especially WP2, WP4, WP6 and WP7, as all the tools and environments developed must become part of the monitored infrastructure and will provide data to be evaluated and fed into the scheduling systems.

Strategy

The main objective of the Virtual Institute on Monitoring and Information Services is to deliver experts' knowledge and integrated solutions in terms of the following services: monitoring and information, checkpointing, user management, accounting and workflow. Detailed objectives were mentioned in the vision of WP5.
The strategy on achieving the mentioned objectives includes the following actions:

- Definition of short-term research groups within each task.
  The research groups will focus on few/several? goals, defining the subjects of research and short term milestones with clearly defined outcomes. Several research groups are foreseen during the project life time.
- Clearly defined milestones and deliverables of WP5:
  - The milestones and deliverables take into account the integration work of research groups
  - After providing common solutions within tasks, co-operation between tasks has been foreseen to integrate a joint architecture of this Virtual Institute
  - WP5 will finally provide the architecture which is planned to be used by other WPs. The WP5 work package will use the results of other CoreGrid Virtual Institutes as well, e.g. WP2 or WP4.

Two WP5 meetings were organized:

- The CoreGRID Kick-off Meeting, 13–14 September, 2004, Charleroi (BE)

During the first CoreGrid meeting we organised the first WP5 meeting in order to present all WP5 tasks and establish first steps towards the creation of research groups. All partners presented their possible contribution to WP5. This was also an opportunity to meet other partners and define possible collaboration

- The First CoreGrid WP5 Meeting, January 19, 2005, Crete (Greece)
  At this meeting the WP5 partners presented their research activities carried out. This led to the first approach towards the creation of research groups defined after the kick-off meeting. A plan of short visits between partners was also prepared during this meeting.

The strategy of the first project time period includes the following general activities:

- Presentation of the results and work done by each partner in R&D projects in the scope of WP5 activities
- Working out research groups for each tasks (short term activities) and the vision of the final expected results (long term activities)
- Defining a common understanding and naming of each service (tasks 5.1 .. 5.4 defines a set of services – high and low level)
- Integrating the R&D results
- Working out a common architecture of the services
- Integrating services and interface into a consistent WP5 architecture.

The mentioned activities will be implemented by WP5 meetings, short visits between partners, longer visits (fellow programs), joint CoreGrid reports, publications and thematic workshops. Common demonstrators were planned as a proof of the concept.

The Virtual Institute on Grid Information and Monitoring Services is focusing on expertise in the following activities:
Task 5.1: Information and Monitoring Services
Task 5.2: Checkpointing Services
Task 5.3: Workflow Services
Task 5.4: Accounting and User Management Services.

Roadmap and phases

The Virtual Institute WP5 defined 4 deliverables in the JPA1:

- D.IM.01: Roadmap version 1 on Grid Information and Monitoring Services
- D.IM.02: Proposal of architecture for scalable GRID monitoring infrastructure
- D.IM.03: Integrated Framework for Monitoring Infrastructure and Job Workflows
- D.IM.04: Roadmap version 2 on Grid Information and Monitoring Services

The list of WP5 deliverables of the first 18 months identified the time schedule and key issues ordered in the following major phases:

Month 6:

Information exchange between partners about the knowledge, experience, research interests and Development made in R&D projects correlated with CoreGrid
Collection of the existing ideas circulating among the partners, possibly coming from the different "local"
projects the partners are involved in.

Month 6-8:

Identified research groups within each task
Common understanding of the problems being the subject of research and integration
Definition of common terms and naming used for future models

Month 12:

Consolidation of common perception of the features to be included in the model.
Definition of models and architecture
Integration work
Joint reports and publications

Month 12-20:

Proof of concepts

The following sections include detailed information about contributions, common interests and integration work, which defines the roadmap of work to be done in each task within the established research groups.

Grid Information and Monitoring Services

Identified research groups:

- FORTH, INFN: integration of Passive and Active Monitoring for GRID infrastructure (transport) monitoring, distributed configuration of GRID passive and active monitoring sensors, passive traffic monitoring for Grids
- FORTH, MU: GRID security - exploring the use of Network Intrusion Detections/Prevention Systems and passive monitoring for securing the GRID infrastructure

Partner	Contributions/Interests and R&D background projects	Collaborations
MTA SZTAKI	Gathering information about the state and performance of a large scale distributed system is essential for system management, failure detection, accounting, auditing, performance tuning and intelligent steering. Providing such information is not easy, however. There were several monitoring systems created in the past (some are described in section 2) that cover more or less aspects of grid monitoring but there is still no generally accepted architecture for unifying the existing approaches. The Grid Monitoring Architecture (GMA) is an abstraction of the essential characteristics needed for scalable high performance monitoring on a large distributed computational Grid. The GMA proposal [61], published by the GGF Performance Working Group, defines the components producer, consumer and directory service. Besides the component definition, a proposal for interaction between components and required operations for each components is also defined. Basic requirements on monitoring architecture are defined as low latency, capable of high data rate, minimal overhead, security and scalability. The Mercury Monitoring System developed by SZTAKI in the GridLab project is a generic GMA-compliant grid monitoring system. It is designed to support both resource and application monitoring. Mercury also supports actuators. Mercury has a modular infrastructure where individual sensors and actuators are implemented as loadable modules, thus providing easy extensibility. Mercury uses a compact binary protocol to reduce network traffic and communication overhead. Mercury provides an instrumentation library that applications can use to publish their own internal data like performance counters or application events. Applications can also register controls that can be invoked through the monitoring system therefore providing remote steering capability. The research tasks are the following: ○ working on a commonly accepted protocol for interacting with a monitoring system. ○ solving problems in Resource monitoring, Application and job monitoring, service monitoring ○ interoperability between monitoring systems.	MU
FORTH	Integration of Passive and Active Monitoring for GRID infrastructure (transport) monitoring, distributed configuration of GRID passive and active monitoring sensors, passive traffic monitoring for Grids	INFN
FORTH	Exploring the use of Network Intrusion Detections/Prevention Systems and passive monitoring for securing the GRID infrastructure	MU

Checkpointing Services

Identified research groups:

- PSNC-SZTAKI co-operation in terms of integrating kernel and application level checkpointing systems. The first goal of the research group is to release a Grid Checkpointing Architecture

- During the WP4 and WP5 meetings in Barcelona (July 2005)PSNC, SZTAKI and UCO (Univ. Coimbra) initiated a horizontal activity between WP4.5 (Dependability in Grids) and WP5.2 (Checkpointing Services).
  Joint activity on the following issues:
  - Design of Grid Services for Checkpointing Support - WP5.2 and WP4.5
  - Checkpointing schemes for Desktop Grids (Boinc/XtremWeb) - WP4.5
  - Checkpointing schemes for Parallel Applications (MPI) - WP5.2
  - Experimental evaluation - WP4.5 and WP5.

Partner	Contributions/Interests and R&D background projects	Collaborations
PSNC	The main objective is to define the position of checkpointing technology within the Grid environment and work out a concept of checkpointing-related services and interfaces. Owing to the aforementioned problems, the primary task is to locate the checkpointing technology within the Grid environment. To accomplish this task, an understanding of the Grid layers and physiology is required. The first task is to study the overall Grid architecture to gain the knowledge allowing to locate the checkpointing within the Grid architecture. The second task is to utilize the gained knowledge to place the checkpointing technology within the Grid. It is important to work out a precise definition of the higher and lower layers and services. Another task is to define the interfaces to layers defined with the preceding task. At present we can assume that interfaces to higher and equivalent layers will base on the OGSA services while the interfaces to lower layers will use general techniques allowing to harness the different underlying, existing and future, checkpointing mechanisms. It is apparent that appropriate meta data models that will describe the available facilities and support information management and exchange must be worked out as well. As the checkpoint mechanism will be placed within the Grid environment and the appropriate services and interfaces will be defined, we will try to provide reference implementation of them. Assuming that the reference implementation of OGSA services will succeed, the next task will include a reference integration of the developed services with an underlying checkpointing mechanism. PSNC has been working on the following checkpointing issues: • User level checkpointing for Solaris OS • Kernel level checkpointing for Solaris OS under the SPARC platform • Kernel level checkpointing for Intel IA64 platforms. The SPARC implementation was done in the R&D project PROGRESS ("Polish Research on Grid Environment for SUN Servers" project no. 6.T11.069.2001C/5688), which was co-funded by the Polish Ministry of Scientific Research and Information Technologies and SUN Microsystems. Within the confines of the PROGRESS project we developed both the kernel-level and user-level checkpointing mechanisms. The checpointing modules were integrated with the Sun Grid Engine (SGE). The checkpoint restart functionality for Intel IA64 platforms was developed in the SGIgrid project ("High Performance Computing and Visualization with the SGI Grid for Virtual Laboratory Applications" – project no. 6 T11 0052 2002	SZTAKI

	C/05836), co-funded by the Polish Ministry of Scientific Research and Information Technologies and Silicon Graphics.	
MTA SZTAKI	In the recent years MTA SZTAKI aimed at making parallel – GRAPNEL, PVM, MPI – applications capable of being checkpointed and restarted on clusters. GRAPNEL is a graphical language to define parallel applications in the PGRADE development framework. Until now MTA SZTAKI has been working on the following checkpoint issues: • User-level, automatic, binary checkpoint for GRAPNEL applications • Automatic checkpoint/restart mechanism for GRAPNEL applications under Condor scheduler (without any modifications in the scheduler) • User-level, automatic, binary checkpoint for native PVM applications • User-level, automatic, binary checkpoint for MPI applications (just started recently) All the issues mentioned above are mainly targeted in a way that the underlying single process checkpointer is taken off-the-self, which means we do not deal with research on single process checkpointer, we are focusing on handling parallelism and the software architecture that services the mechanism. MTA SZTAKI had several local projects in the last 3-4 years. They were about to sponsor the research and development regarding the above mentioned areas. These are the "Hungarian Supercomputing GRID" sponsored by the Ministry of Education (IKTA4-075), "Cluster Programming Technology and Its Application in Meteorology" sponsored by the Ministry of Education (IKTA3-029), "Chemistry GRID and its Application for Air Pollution Forecast" sponsored by the Ministry of Education (IKTA5-137) and the currently active project is the "Hungarian SuperCluster project" sponsored by the National Office for Research and Technology (IKTA-00064/2003) which aims at giving parallel checkpoint support for the nationwide Hungarian Clustergrid infrastructure which is the largest cluster-based grid in Hungary.	PSNC

Workflow Services

Identified research groups:

- SZTAKI, UoW: introducing collaborative workflows in legacy code environments
- UMUE, FHG: workflow description languages using high-level Petri nets for Grid workflows
- UMUE, FHG: it is planned to collaboratively develop advanced performance prediction and scheduling mechanisms on top of workflow definition language and to study mapping schemes between different workflow formalisms and languages in relation with WP6
- SZTAKI, MU - extending the SEAGRIN semantic "overlay grid" infrastructure (a potential candidate for content-based monitoring infrastructure) with the collaborative workflow management support of the P-GRADE portal
- FHG, UMUE, INFN: c*ompatibility and conversion of different Grid workflow description languages*. The objective of this research group will be a joint survey about compatibility and conversion (mapping) issues between commonly available workflow description languages.

Partner	Contributions/Interests and R&D background projects	Collaborations
FHG	**K-Wf Grid:** K-Wf Grid is a STREP project funded by EC that started in 09/2004 under the leadership of Fraunhofer FIRST with the objective to develop a "Knowledge-based Workflow System for Grid Applications". The K-Wf Grid system will assist its users in composing powerful Grid workflows by means of a rule-based expert system. All interactions with the Grid environment are monitored and evaluated. The knowledge about the Grid itself is mined and reused in the process of workflow construction, service selection and Grid behaviour prediction. Workflows are dynamic and fault-tolerant beyond the current state of the art. At the beginning of the project we will analyze the existing technologies in the domain of workflow composition and execution. After this we will specify a graph-based workflow model that overcomes the disadvantages of the commonly used Directed Acyclic Graphs. Next, an XML-based Grid job description language will be specified, contributing to the standardization efforts of the Global Grid Forum. The workflow description itself will include semantically-rich metadata on an abstract level that is independent of the Grid infrastructure. Furthermore, we will develop a "Grid Workflow Execution Service", acting as a software layer between the user-interactive Grid Application Building layer and the service-oriented lower-level Grid middleware. Other developments will include user interfaces for execution control and monitoring of Grid jobs as well as administrative tools. **Fraunhofer Resource Grid:** Within the operation of the Fraunhofer Resource Grid and proposed contributions to D-Grid (German Grid Initiative), further research and development on our Petri-Net-based workflow approach will be done. The scientific goals related to Task 5.3 are: - Improvement of the Petri-Net-based Grid job description language (GJobDL) - Service for the execution and monitoring of Grid jobs - Research on interoperability with gLite and WSRF - Bringing further applications to the Fraunhofer Resource Grid - Higher level checkpointing of Grid workflows	CYFRONET, UMUE

FZJ	- Current work:
	  o OpenMolGRID: High-level work flow description & processing in UNICORE
	  o NextGRID: Definition of a universal model to integrate resources in work flow processing
	  o Evaluation of workflow & job description languages
	  o Contributing to GGF's Grid Scheduling Architecture RG
	- Also active in CoreGRID task 6.3 "Workflow Grid Scheduling Strategies"
	- Integration of workflow management into Grid architectures
	- Job models
	- Workflow management vs. job management vs. service management	

INFN	**Matchmaking of a set of activities involved in a workflow schedule:** Matchmaking sets of activities has been addressed with the Gangamatching approach that does not provide a high degree of orthogonality between coordination and computation. We are basing on some standard workflow patterns that have been formalized with Petri nets but miss the description of important patterns as long running transactions. Negotiation protocols have been investigated in the area of Distributed Artificial Intelligence. A classic protocol is the Contract Net Protocol (CNP),	FHG, UMUE

Cbitoune 18/10/05 18:45
Mis en forme: Retrait : Gauche : 0,06 cm, Retrait négatif : 0,25 cm, Avec puces + Niveau : 1 + Alignement : 0,63 cm + Tabulation après : 1,27 cm + Retrait : 1,27 cm, Tabulations : 0,31 cm, Tabulation de liste + Pas à 0,63 cm +

Cbitoune 18/10/05 18:45
Mis en forme: Retrait : Gauche : 0,56 cm, Retrait négatif : 0,5 cm, Avec puces + Niveau : 2 + Alignement : 1,27 cm + Tabulation après : 1,9 cm + Retrait : 1,9 cm, Tabulations :Pas à 1,9 cm

	proposed in the scenario of distributed problem solving. CNP has then been adapted to different types of multi agent systems. INFN studies the formal definition of negotiation protocols, describing similarities and differences with respect to well known distributed transaction protocols as BTP that have already been formalized with process algebras. Some efforts are addressing the definition of standards for negotiation in the Grid/Web Service scenario (e.g. WS-Agreement by GGF), mostly in an agnostic way with respect to the underlying protocol. WS-Negotiation considers negotiation in all its facets from a Web Service perspective but it lacks a formal description of the negotiation protocol. **Semantic resource discovery in Grid Environments (in cooperation with WP2)**: With the aim of enabling automation of services matchmaking, in the context of WP2 we are developing an extension of a service description language (possibly OWL-S) that takes into account non-functional parameters and negotiable parameters. The goal of this project is to investigate the usage of the service descriptions expressed in the proposed OWL-S extensions. Within this project we aim to investigate the problems of workflow matchmaking and the negotiation issues. We aim to clarify which workflow patterns to consider in out formal framework basing on the existing work on Petri nets for standard workflow patterns and mainly on pi calculus concerning transactional behaviour. We intend to define a set of service features for each pattern that should be considered in the matchmaking of a composition. Or goal is to consider the use of inferential Grid information system, a definition of the requirements for a negotiation protocol in a Grid scenario, investigation on the relationship between negotiation and Web-distributed transactions, the formal definition of a suitable negotiation protocol.	
MU	**MediGRID – the workflow part** Development of concepts and prototype implementation of workflow systems combined with the collaborative environment. Finding ways to create, manage and inspect workflows within the collaborative environment, where different partners have different knowledge of and interest in the complex workflow parts (e.g., processing of laboratory data and processing of some medical images), while sharing a common goal expressed by the workflow (e.g., to converge on a diagnosis). The secondary goal is to study the use of ontologies to glue different workflow components into a consistent whole. Scientific Goals: To provide workflow within the collaborative environment. To provide a framework for semantic consistency checking.	SZTAKI, UoW
SZTAKI	**Workflow Management in Grid Environments:** Creation of a prototype implementation of a workflow management infrastructure which enables the automated execution and monitoring of workflows consisting of parallel or sequential applications. Building onto the transparent job-level checkpointing and migration support provided by the components SZTAKI will develop as part of its 5.2 contribution, the manager will be able to execute job-workflows in a fault-tolerant way in grids. Research on collaborative workflow execution where the workflow components can be executed on different grids based on different certificates owned by different members of a collaborative group. The workflow manager will be built onto the core scheduler functionality of Condor DAGMan. Beside the previously discussed fault-tolerant behaviour, DAGMan will be extended with the multi-grid support mentioned above.	MU, FHG, UoW, UMUE, INFN, UNICAL
UNICAL	The project is concerned with the adoption of WF as a model for defining	FHG, UWE

	complex service-oriented Grid applications involving a number of jobs interacting with each other in various ways. Scientific goals: Defining a specification mechanism simple enough to be used in several application domains and powerful enough to exploit the Grid infrastructure.	
UOW	**GEMLCA** We have an environment where users can construct, manage and execute visual workflows from a Grid portal in order to support execution of Grid-enabled Legacy Codes. The middleware called GEMLCA (www.cpc.wmin.ac.uk/gemlca) was developed by the University of Westminster during the UK EPSRC funded OGSA testbed project. GEMLCA uses Globus Toolkit 3 and is being migrated to GT4 (WSRF). GEMLCA was demonstrated at IST 2004 in Hagues and there is an article about it in the 'Next Generation Grids' publication document. The current workflow system includes support for GT2 and service-oriented GT3 Grid using Gemlca jobs. Gemlca supports the execution of Grid-enabled legacy codes. The current workflow execution manager is based on Condor DAGman. Scientific goals: Research tasks include the following: - Workflows o Specifying workflow (language to take into account Grid QoS, flow conditions, service aggregation, etc) o Scheduling, brokering, and support for different Grids (interoperability issues) - Support for execution of legacy codes in service-oriented Grids - Dynamic scheduling of workflow jobs onto the Grid o Assigning tasks of a workflow application onto Grid resources with a view to minimize workflow execution time. o Novelty of this approach is based on computational economy together with quality of service requirements. - Grid portals	UNICAL, SZTAKI
UMUE	**Using Coloured Petri Nets for Grid Workflows:** The usage of Coloured Petri Nets for describing complex Grid Workflows is investigated. A prototype system will be implemented, which takes local scheduling policies of resources into account. In a theoretical part, the task of modelling different aspects (e.g. cost information) of Grid services or components using Coloured Petri Nets are studied in order to analyse complex Grid workflows and allow for efficient execution. The system should make use of available Grid middleware but not be bound to a certain one. Therefore, a common interface to different underlying middlewares has to be developed. The scientific goals are to provide a system environment for users to describe complex grid workflows intuitively. Upon execution, the scheduling policy of used resources should be taken into account when distributing subtasks of the workflow over the Grid. Deliverables: The project already started (still in its early stage), first concepts have been sketched to Nov/Dec 2005: prototype system information to mid 2006: Enhancement of the system, modelling procedures, extensive tests	FhG, UNIDO

Cbitoune 18/10/05 18:45
Mis en forme: Retrait : Gauche : 0,56 cm, Retrait négatif : 0,5 cm, Avec puces + Niveau : 2 + Alignement : 1,27 cm + Tabulation après : 1,9 cm + Retrait : 1,9 cm, Tabulations :Pas à 1,9 cm

Cbitoune 18/10/05 18:45
Mis en forme: Retrait : Gauche : 0,56 cm, Retrait négatif : 0,5 cm, Avec puces + Niveau : 2 + Alignement : 1,27 cm + Tabulation après : 1,9 cm + Retrait : 1,9 cm, Tabulations :Pas à 1,9 cm

Accounting and User Management Services

Identified research groups:
- PSNC, MU: state of the art, defining common base of understanding, naming; working out a joint model of user account management

- PSNC, FGH FIRST: accounting and user management in grid environments.

Partner	Contributions/Interests and R&D background projects	Collaborations
PSNC	Poznan Supercomputing and Networking Center has performed research on user accounts management and the distributed accounting information system for several years. The project name is the Virtual Management System (VUS). The idea of VUS is universal and can be implemented with a different job management system. VUS is just an extension of the system that runs users' jobs (e.g. queuing system, Globus Gatekeeper, etc.) and allows to run jobs without having a user account on a node. This allows to minimize overhead related to creating and maintaining additional user accounts. On the contrary to other solutions, VUS assures an accurate security level achieved by user authorization and possibility of charging the user with costs of resource usage. Additionally, it respects the local policy of sites and makes it possible for the local administrator to differentiate between local and remote users. The system is transparent for the users and to some extent also for administrators. Users just run their jobs and do not care about account assignment. Administrators just configure access to the machine for VO and do not need to create accounts and manage grid-mapfile. The system also does not interfere with local access policy. The history of VUS started in 1998 [54,55]. The first implementation of VUS was an extension to queuing systems (e.g. LSF) and it was successfully exploited 3 years ago in the Polish national cluster which connected several HPC centres in Poland [56,57,58,59]. Later we focused on Globus. The first version of VUS for Globus 2.4 was using a modified Gatekeeper. The current implementation is GRAM “callout”, a mechanism introduced in GT 3.2. The major advantage of callouts is that there is no need to modify Globus Toolkit (GT) codes to install VUS (the older versions of GT needed slight modifications in the Gatekeeper and GridFTP). The research on account management and distributed accounting processing is performed in several projects. - National Computing Grid – set up by PSNC as an attempt to connect supercomputers from different Polish supercomputing centers with the LSF queuing system. - SGIgrid – "High Performance Computing and Visualization with the SGI Grid for Virtual Laboratory Applications". The goal is to design and implement an environment for remote access to unique and therefore expensive laboratory equipments. The environment will support the end user by delivering enough computational power for pre- and post-processing computations and data intensive visualization. One of the project goals is to provide a backup data center for the national Institute of Meteorology and Water Management. The project started in December 2002 and will finish in November 2005. - Clusterix – “National CLUSTER of LInuX Systems”. The goal of the project is to build a new generation distributed PC cluster by connecting 64bit PC from 12 Polish computing centers. Middleware implemented in this project will allow test application to run in a dynamic environment with a changing number and configuration of computing and network infrastructure. Some cooperation has also been done within the 5FP EU GridLab Project.	MU, FHG
MU Brno	The Perun Project [60] started as a user management tool for the Czech national Grid encompassing heterogeneous resources from supercomputing	PSNC

Norbert Meyer 18/10/05 14:23
Supprimé: Saut de page

	centers across the Czech Republic. The system itself is not a symmetric component of the authorization service. On the contrary, it provides a repository of complex authorization data, as well as tools to manage the data. The data are used to generate configuration of the authorization services themselves (starting from UNIX user accounts through grid-mapfiles to the VOMS database). In turn, these services are used to enforce authorization policies. Hence the centralized Perun architecture does not introduce a single point of failure of the whole Grid authorization infrastructure. Perun makes use of central configuration repository which models an *ideal world*, i. e. what the resources should look like. In this central repository, all the necessary (and possibly very complex) integrity constraints are relatively easy to be enforced. The repository is complemented with a change propagation mechanism which detects the changes, generates consistent configuration snapshots of atomic pieces of managed systems, and tries to deliver them to their final destinations, dealing with resource or network failures appropriately. In this way, the *real world* is forced to follow the ideal one as closely as possible. The core of the system is completely independent of the structure and semantics of the configuration data, hence the system is easily extensible. In the current deployment the managed configuration items include user accounts on UNIX machines, Kerberos realms, and user home directories on both UNIX filesystems and AFS. **Research tasks performed in Perun, GridLab and EGEE projects:** o extending the functionality to cover additional requirements to support for binding X509 identities to physical users, o maintenance of a set of trusted certification authorities, o generating *grid-mapfile*s (mapping of X509 certificate subjects to local user names) directly, o integrating different authorization services: LCMAPS and VOMS.	

Mechanisms

Partner Meetings/Workshops
The all hands gathering are considered to be the most important tool for collaboration and integration promotion. They will be held every 4 to 6 months, either as independent meetings (probably related to similar meetings of other Virtual Institutes, esp. WP4) or associated to some public workshop.
The purpose of these meeting is manifold:

- To serve as a place for the presentation of results, with emphasis on work done in collaboration between partners
- To serve as a discussion venue for future work and direction of research (both operational and strategic decisions will be prepared)
- To provide a forum for the preparation of deliverables and other CoreGRID related documents
- To serve as a research/development environment (extended meetings for several days, focused on a particular research topic to be approached from different perspectives by several partners in the collaboration).

The first meeting of this Virtual Institute was held in Heraklion (at the FORTH premises, Cyprus, Greece) on January 19th, 2005. This meeting followed a one-day WP4 organized workshop and a one-day WP4 meeting, thus strengthening also the inter Institute collaboration. The major purpose of this meeting was the identification of individual partners' contribution to the Roadmap and therefore to the actual scientific work of this Virtual Institute.

E-meetings and tele-conference meetings
Most of the information exchange between partners will occur through e-mail exchange (either individual or through the Virtual Institute e-mail list) and through the document sharing on the CoreGRID portal. The work package and task leaders will have semi-regular tele-conference meetings (usually with bi-weekly periodicity).

Partners will try to setup an e-collaborative environment, using either publicly available systems like VRVS, through H.323 videoconferencing tools where available or using simpler tools like MBone tools with some reflector. These potential e-meeting will be mostly focused on specific research topics, they will not make a complete substitute for the face to face partners meetings.

Researcher/Student Exchanges
On top of short visits funded directly by the CoreGRID, partners plan to promote joint PhD programs in relevant areas. The activity will start with short visits, students' exchange and is expected to culminate in shared PhD supervision.
If knowledge/expertise gaps are discovered, partners plan to invite also researchers external to the CoreGRID or to visit institutes with relevant scientific knowledge.

A common experimental testbed
All the services developed within this Virtual Institute will be deployed on the CoreGRID experimental testbed. While some of these deployments will probably be used only by the research teams who invented and implemented them, we will focus on providing (together with the WP7 and the testbed group) an experimental but still relatively stable environment where the services will be directly used by all the CoreGRID partners. Some Virtual Institute partners already demonstrated the capability to manage and run textbeds with production level quality of service even within a highly dynamic and evolving environment (e.g. within the EU GridLab project). The early wide deployment of this Virtual Institute services will serve both as a feedback to its research activities and also as a valuable tool for other Virtual Institutes.

Dissemination of results
Due to the primary research focus of the CoreGRID, the primary dissemination platforms are scientific journals, conferences, workshops, and similar venues. Partners expect to submit joint publications, which will also be available on the CoreGRID portal (unless the publishing policy does not prohibit such form of dissemination). Apart from joint publications, partners will extensively use the results achieved within the CoreGRID in their own work, extending thus strength and knowledge of the CoreGRID-related research. The monitoring system is a primary candidate to be also used in publication and other work of partners working within other Virtual Institutes.
The CoreGRID public portal pages will also be used to promote the knowledge gained through joint work within this Virtual Institute.

CoreGRID portal
All partners have access to the CoreGRID portal and it is used to exchange documents and other digital material related to the CoreGRID activities. The portal is also extensively used for cross Virtual Institutes collaboration. Both internal and external web pages are maintained by the partners.

Exchange of documents
All documents are stored on the CoreGRID BSCW server, which provides a uniform www-based document and file sharing platform. As all the CoreGRID participants have secure access to the BSCW server, it is used by the Virtual Institute for easy upload and sharing of all textual information (documents, source code, meeting minutes, etc.).

Short visits
Short visits are considered to be one of the most important tools to initiate and later strengthen the collaboration between partners. The individual short visits will always have a clear purpose, usually a discussion, writing joint scientific paper or preparing a deliverable contribution.
Most short visits are around 3 days and will not exceed a week. The actual length of stay may change even during the stay if an interesting research subject is discovered.
The list is not complete, as short visits are expected to be a flexible tool used as necessary (Table 1).

From	To	TODO
PSNC	MTA SZTAKI	Knowledge exchange about PSNC kernel level checkpointing and MTA SZTAKI development framework (P-GRADE) in the cluster/grid environment, agreement on the checkpointing service vision
MTA SZTAKI	PSNC	Architecture of the checkpointing service
INFN	MU	Discuss how content-based networking can provide alternative, content-aware transport protocols for monitored information and

		implement a data transformation and processing layer inside the monitoring system
INFN	MTA SZTAKI	Discussion on scalability and reliability issues in large-scale distributed monitoring systems
INFN	UNICAL	Discussion on QoS aspects in large-scale Grid monitoring systems
INFN	ICS-FORTH	Discuss and evaluate the active vs. passive monitoring approach in Grid scenarios
PSNC	MU	User account service, use of virtual machines
PSNC	FhG	User account service architecture
MU	ICS-FORTH	Joint visit with INFN, passive vs. active monitroing
MU	PSNC	Architecture of virtual account service over VM
MU	MTA SZTAKI	Discuss the integration of passive monitoring with the P-GRADE system, first design
MU	UoW	Imprecise service discovery
UNICAL	FhG FIRST	Insights into the K-Wf Grid project and other workflow activities of Fraunhofer FIRST, presentation of the workflow activities of both partners
FhG FIRST	Uni Manchester	Interchange knowledge onworkflow orchestration and execution and getting insights to the MyGrid project
Uni Munster	FhG FIRST	Interchange knowledge on workflow description languages and using coloured Petri nets for Grid workflows
MTA SZTAKI	MU	Presentation on the ongoing development on a collaborative workflow-oriented portal. Discussing problems related to collaborative workflow management
MTA SZTAKI	FhG FIRST	Presentation on workflow research activities of the past, present and future in SZTAKI and in FhG FIRST. Discussing the concept of a fault-tolerant workflow manager service taking into consideration the CoreGRID requirements and the current trends of service-oriented Grid development
MTA SZTAKI	UoW	Investigating possible ways to integrate the multi-grid workflow manager system of SZTAKI and the GT3 based GEMLCA middleware of UoW
MTA SZTAKI	WWUM	Investigating the problem of workflow management over different Grid middlewares. Introducing SZTAKI's proposed model for multi-grid access
MTA SZTAKI	INFN	Identifying the appearance of transactions in Grid environments, the roles of transaction managers in Grids and estimating possible use-cases for transaction-based grid-resource access
MTA SZTAKI	UNICAL	Discussing problems of multi-environmental workflows. Presenting SZTAKI's proposed solutions to execute different parts of a workflow in different security domains or in grids that use different middlewares

Table 1: A list of short visits between partners planned in the first stage of WP5

Future steps

While the CoreGRID itself only adds on top of the actual, independently funded partners' research activities, the partners' contribution already covers the most important parts of the area of grid information and monitoring services. It is expected that the collaboration will allow more focused activities of individual partners, accompanied with the identification of weak points (knowledge/expertise gaps) where new research activities will be oriented. The partners plan to look for possible sources of joint research funding (for better coordination on the international level the international programs will be given priority, but implicitly coordinated national funding programs will also be considered). The number of proposals and their acceptance rate will be considered as one of the quality measures within the Virtual Institute.

The short visits and the fellowship program will be used not only to strengthen the current and planned research work, but also as a means to actively search for new venues for strengthening the collaboration. The expected joint research proposals will be accompanied with collaboration on the PhD training level, including shared PhD supervision or extensive medium term PhD students' visits (financially supported by funding outside the CoreGRID, using long term programs for student exchange).

5. Links with other CoreGRID scientific workpackages

Access to information about the state of the Grid and its individual components is vital to most of the other Virtual Institutes. On the other hand, the architecture of grid information and monitoring services itself creates a specific grid, with its own problems related to scalability, adaptability, scheduling etc., i.e. issues studied by the research team of other Virtual Institutes. Therefore, we plan to have a strong collaboration with other Virtual Institutes.

The topics of the four major tasks of the work package WP5 are linked to several activities throughout the other scientific workpackages of CoreGRID. In the following, the links for the design of the comprehensive Grid system are described for each task of the work package WP5. In the description, the relations between activities are classified as follows:
1) providing a basis for other activities (technology supplier)
2) using other workpackages' activities (technology user)
3) general exchange of technical knowledge.

Task 5.1 Information and monitoring services
This task acts as a technology supplier for different topics of other workpackages. For a comprehensive Grid system, it is a crucial operation to obtain static and dynamic information about grid resources and networks as well as monitor the execution of jobs and services in the grid. Therefore, the goal of this task's activities is to provide a basis for the following tasks: allocation and scheduling of resources (work package 6), detection of critical conditions (work package 4) and efficient distribution and storage of data in the grid (work package 2).

In order to supply adequate information and monitoring services matching the requirements of the other workpackages, detailed coordination concerning the content and timeliness of delivered information and monitoring data as well as the interface to the provided services is required.

The information and monitoring services have to be highly scalable to operate in a large scale Grid system. Therefore, task 5.1 will exchange knowledge with the task working on scalable architectures in work package 4. Additionally, security monitoring and management is an integral part of the monitoring services. For this security aspect, the work package 5 is linked with all other work packages and will coordinate the work on security topics in the horizontally integrating security workgroup.

Task 5.2 Checkpointing services
The checkpointing services provide a technical basis for fault-tolerant grid programming. Therefore, this task is linked with the work on grid programming models (work package 3) and has to address the requirements of the programming models. Additionally, there has to be an extensive exchange of knowledge with work package 4 regarding this topic, because the issue of dependability of grid system architectures focused on in work package 4 includes the work on general and scalable approaches for checkpointing, job migration and job recovery.

Task 5.3 Workflow services
Workflows provide a high-level approach to compose compound Grid jobs and programs. The execution of such a composed grid program is managed by workflow services which map the static, compile-time job flow description onto the dynamic Grid system at runtime. Therefore, these services have to access static information provided by the program itself and the workflow description.
Therefore, this task is strongly linked with the work on Grid programming models done in work package 3. Additionally, the runtime mapping has to acquire dynamic information from monitoring services and take resource management information and scheduling policies into account. Therefore, this task working on workflow services is linked as a technology user with work package 6, which provides a technical basis for the retrieval of resource management and scheduling information. Additionally, there is a strong connection to the task working on grid architecture adaptability in work package 4. The workflows provided by work package 5 provide input data for the analysis of overall demand for grid resources and therefore provide a decision base for automatic Grid reconfiguration.

Task 5.4 Accounting and user management service
The information of used resources is a crucial issue for resource management systems to take a decision based on the past information about resource usage, application and jobs characteristics (work package 6). Work package 2 will also use the grid accounting service to keep the information about past storage infrastructure activities.
User management service is a functionality positioned below the brokerage functionality (work package 6) and allows to communicate between different domain, keeping the appropriate authorization and authentication features.

6. References

[1] K. Keahey, M. Ripeanu, and K. Doering. *Dynamic Creation and Management of Runtime Environments in the Grid.* Workshop on Designing and Building Grid Services, GGF-9, October 8, 2003. http://www.globus.org/research/papers/dynamiccreation.pdf
[2] K. Keahey, K. Doering, and I. Foster. *From Sandbox to Playground: Dynamic Virtual Environments in the Grid.* 5th International Workshop in Grid Computing (Grid 2004), Pittsburgh, PA, November 2004. http://www-unix.mcs.anl.gov/~keahey/papers/grid_2004_pub.pdf
[3] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebar, I. Pratt, and A. Warfield. *Xen and the Art of Virtualization.* Proceedings of the ACM Symposium on Operating Systems Principles (SOSP), October 2003 http://www.cl.cam.ac.uk/netos/papers/2003-xensosp.pdf
[4] C. Clark, K. Fraser, S. Hand, J. G. Hansen, E. Jul, C. Limpach, I. Pratt, and A. Warfield. *Live Migration of Virtual Machines.* Proceedings of the 2nd Symposium on Networked Systems Design and Implementation (NSDI '05), May 2005, Boston, MA (to appear)
[5] K. Fraser, S. Hand, R. Neugebauer, I. Pratt, A. Warfield, and M. Williams. *Safe Hardware Access with the Xen Virtual Machine Monitor.* Proceedings of the 1st Workshop on Operating System and Architectural Support for the on demand IT InfraStructure (OASIS), October 2004, Boston, MA http://www.cl.cam.ac.uk/Research/SRG/netos/papers/2004-safehw-oasis.pdf
[6] L. Kleinrock and W. Naylor, *On the Measured Behavior of the ARPA Network*, AFIPS Proceedings, 1974 National computer Conference, pp. 767-780, 1974
[7] K. C. Claffy, *Priorities and Challenges in Internet Measurement Simulation and Analysis*, 2003, available at: http://www.caida.org/outreach/presentations/2003/lsn20030610/
[8] E.P. Markatos, *Tracing a large-scale Peer to Peer System: an hour in the life of Gnutella*, Proceedings of the CCGrid 2002: the second IEEE International Symposium on Cluster Computing and the Grid, May 2002
[9] D. Plonka, *FlowScan: a network traffic flow reporting and visualization tool*, Proc. of LISA 2000 Conference, 2000
[10] Oliver Spatscheck, *How to monitor network traffic 5 Gbit/sec at a time*, Workshop on Management and Processing of Data Streams, 2003
[11] GÉANT2: the seventh generation of pan-European research and education network – more information available at: www.geant2.net
[12] A. Tirumala and J. Ferguson, *Iperf 1.2 -- The TCP/UDP Bandwidth Measurement Tool*, 2001, available at: http://dast.nlanr.net/Projects/Iperf/
[13] J. Quittek and T. Zseby and B. Claise and S. Zander, *Requirements for IP Flow Information Export*, February, 2003, Internet Draft, draft-ietf-ipfix-reqs-09.txt (work-in-progress)
[14] Manish Jain and Constantinos Dovrolis, *End-to-end available bandwidth: Measurement methodology, dynamics, and relation with TCP throughput*,
[15] Proceedings of SIGCOMM 2002, August, Pittsburgh, PA, US, 2002, available at: citeseer.nj.nec.com/jain02endtoend.html
[16] LOBSTER: Large-Scale Monitoring for Broadband Internet Infrastructures – more information available at www.ist-lobster.org
[17] Cisco Systems, *Cisco IOS Netflow*, Technical Report, 2003, available at: http://www.cisco.com/warp/public/732/netflow/
[18] NLANR: National Laboratory for Advanced Network Research – more information available at: http://moat.nlanr.net
[19] K. C. Claffy and G. Polyzos and W. Werner Braun, *Traffic Characteristics of the T1 NSFNET Backbone*, Proc. IEEE Infocom 93, pp 885-892, 1993
[20] Fotis Georgatos and Florian Gruber and Daniel Karrenberg and Mark Santcroos and Ana Susanj and Henk Uijterwaal and René Wilhelm, *Providing Active Measurements as a Regular Service for ISP's*, Proceedings

of the Passive and Active Measurements Workshop (PAM2001), 2001, available at: citeseer.nj.nec.com/georgatos01providing.html
[21] SCAMPI: A Scaleable Monitoring Platform for the Internet – more information www.ist-scampi.org
[22] http://www-iepm.slac.stanford.edu/
[23] S. Savage, *Sting: a TCP-based network measurement tool*, Proc. USENIX Symposium on Internet Technologies and Systems (USITS '99), 1999
[24] V. Jacobson, "Traceroute", 1989, available at: ftp://ftp.ee.lbl.gov/traceroute.tar.Z
[25] W. Dymaczewski, N. Meyer, M. Stroiński, P. Wolniewicz, *Virtual Users Account System for Distributed Batch Processing*, in: P. Sloot, M. Bubak, A. Hoekstra, B. Hertzberger (Eds.): HPCN 1999, LNCS 1593, Springer-Verlag, 1999, 1231-1234.
[26] M.Kupczyk, M. Lawenda, N. Meyer, M.Stroinski, P.Wolniewicz: *Experiences with Virtual Users' Accounts System in Polish National Cluster*, CUG, Indian Wells, May 2001
[27] M. Kupczyk, M. Lawenda, N. Meyer, P.Wolniewicz: *Using Virtual User Account System for Managing Users Account in Polish National Cluster*, HPCN,Amsterdam, June 2001, 587-590.
[28] N. Meyer, P. Wolniewicz, *Virtual User Account System, Conference* – 10th Anniversary of Poznań Supercomputing and Networking Center – New Network Technologies, Grids and Portals, Poznań, October 2003.
[29] SGAS http://www.sgas.se/
[30] SNUPI http://snupi.sdsc.edu/snupidoc.html
[31] P. Kacsuk, G. Dzsa, J. Kovcs, R. Lovas, N. Podhorszki, Z. Balaton, and G. Gombs. Pgrade: a grid programming environment. Journal of Grid Computing, 1(2):171–197, January 2003.
[32] P.Stelling, I.Foster, C.Kesselman, C.Lee, G.Laszewski. *A Fault Detection Service for Wide-Area Distributed Computations*, Proc. 7th IEEE Symposium on High-Performance Distributed Computing, 1998, pp. 268-278
[33] M.Baker, G.Smith. *GridRM: A Resource Monitoring Architecture for the Grid*, Technical Report University of Portsmouth UK, June 2002
[34] GMA-WG: Grid Monitoring Architecture Working Group, http://www-didc.lbl.gov/GGF-PERF/GMA-WG/
[35] Expert Group. *Next Generation Grids 2 (NGG2)—Requirements and Options for European Grids Research 2005–2010 and Beyond*, EC Report, ftp://ftp.cordis.lu/pub/ist/docs/ngg2_eg_final.pdf, 2004
[36] Expert Group. *Next Generation Grids—Requirements and Options for European Grids Research 2005–2010 and Beyond*, EC Report, http://www.cordis.lu/ist/grids/pub-report.htm, 2004
[37] S. Fisher : *Relational Model for Information and Monitoring.* Technical Report, GWD-Perf-7-1, GGF, 2001.
[38] G. Aloisio, M. Cafaro, I. Epicoco, S. Fiore., D. Lezzi, M. Mirto, and S. Mocavero : *iGrid, a Novel Grid Information Service*. Proc. EGC 2005, Amsterdam, The Netherlands, 2005 ; revised in LNCS 3470, Springer-Verlag, pp.1-9, 2005.
[39] (MDS 2) Czajkowski, K., Fitzgerald, S., Foster, I., Kesselman, C.: *Grid Information Services for Distributed Resource Sharing.* In Proceedings of the 10th IEEE International Symposium on High-Performance Distributed Computing (HPDC-10), IEEE Press, August 2001.
[40] (Mercury) Balaton, Z., Gombás, G.: *Resource and Job Monitoring.* In the Grid. Proc. of the Euro-Par 2003 International Conference, Klagenfurt, 2003. 404–411
[41] Sitera L., Matyska L., Krenek A., Ruda M., Vocu M., Salvet Z., Mulac M.: *Capability and Attribute Based GRID Monitoring Architecture.* Proc. CGW04, 2004.
[42] Holub, P., Kuba, M., Matyska, L., Ruda, M.: *Grid Infrastructure Monitoring as Reliable Information Service.* Proc. 2nd European AcrossGrids Conference, LNCS 3165, Springer-Verlag, pp.220-229, 2004.
[43] Zanikolas S., Sakellariou R.: *A Taxonomy of Grid Monitoring Systems*. FGCS, 21(1), 163–188, 2005
[44] Ceccanti A. Panzieri F.: *Content-based Monitoring in Grid Environments*. In Proc. 13th IEEE International Workshop in Enabling Technologies, WETICE 2004.
[45] Ciuffoletti A.: *Reliability versus cost: design of a probabilistic broadcast algorithm*. Distributed Computing, 7, 115–127, 1994
[46] Demers, A, Greene, D., Hauser, C., Irish, W., Larson, J., Shenker, S., Sturgis, H., Swinehart, D.. Terry, D.: *Epidemic Algorithm for replicated database maintenance*. Proc. of ACM Symposium on Principles of Distributed Computing, Vancouver, pp. 1–12, 1987.
[47] Jelasity M., Kowalczyk W., van Steen M.: *Newcast Computing*. Technical Report, Vrije Universiteit Amsterdam, IR-CS-006, 2003.
[48] Jelasity M., Babaoglu O.: *T-Man: Fast Gossip-based Construction of Large-Scale Overlay Topologies*. Technical Report, Comp. Sci. Department, University of Bologna, UBLCS-2004-7, 2004.
[49] Andreozzi S., Ciuffoletti A., Ghiselli A., Vistoli C.: *Monitoring and Connectivity of a Grid.* 2nd Workshop on Middleware for Grid Computing, Toronto, pp. 47–51, 2004.

[50] Eidson J., Fischer M., White J.: *IEEE-1588TM Standard for a Precision Clock Synchronization Protocol for NBetworked Measurement and Control System*. 34th PTTI, Reston, VA, pp. 243–255, 2002.
[51] Carzaniga A., Rosenblum D.S., Wolf A.L.: *Design and Evaluation of a Wide-Area Event Notification Service*. ACM Trans on Computer Systems, 19(3), 332–383, 2001.
[52] Chand R., Felber P.A.: A scalable protocol for content-based routing in overlay networks. NCA'03, 2003.
[53] Diao Y., Rizvi S. Franklin M.J.: *Towards on Internet-Scale XML Dissemination Service*. VLDB 2004.
[54] W. Dymaczewski, N. Meyer, M. Stroiński, P. Wolniewicz, *Virtual Users Account System for Distributed Batch Processing*, in: P. Sloot, M. Bubak, A. Hoekstra, B. Hertzberger (Eds.): HPCN 1999, LNCS 1593, Springer-Verlag, 1999, 1231-1234.
[55] M.Kupczyk, M.Lawenda, N.Meyer, M.Stroinski, P.Wolniewicz: *Experiences with Virtual Users' Accounts System in Polish National Cluster*, CUG, Indian Wells, May 2001
[56] M.Kupczyk, M.Lawenda, N.Meyer, P.Wolniewicz: *Using Virtual User Account System for Managing Users Account in Polish National Cluster*, HPCN,Amsterdam, June 2001, 587-590.
[57] N.Meyer, P.Wolniewicz, *Virtual User Account System*, Conference – 10th Anniversary of Poznań Supercomputing and Networking Center – New Network Technologies, Grids and Portals, Poznań, October 2003,
[58] SGAS http://www.sgas.se/
[59] SNUPI http://snupi.sdsc.edu/snupidoc.html
[60] Ales Kˇrenek and Zora Sebestianov´a. Perun – Fault-Tolerant Management of Grid Resources, Cracow '04 Grid Workshop Proceedings, December 2004.
[61] B. Tierney, R. Aydt, D. Gunter,W. Smith, V. Taylor, R.Wolski, and M. Swany. A grid monitoring architecture. Technical Report GWD-PERF-16-2, Global Grid Forum, Jan. 2002.

7. Participants

			Task			
Partner	**No.**	**Researchers and Ph.D. students**	**T1**	**T2**	**T3**	**T4**
FHG	8	Andreas Hoheisel Thilo Ernst	X		X	X
FORTH	11	Spyros Antonatos Michalis Polychronakis Paraskevi Fragopoulou Evangelos Markatos Panos Trimintzios	X			
FZJ	9	Achim Streit Bernd Schuller Philipp Wieder			X	
INFN	13	Andrea Ceccanti Antonia Ghiselli Edgardo Ambrosi Laura Bocchi Augusto Ciuffolotti Tiziana Ferrari	X		X	
INRIA	14	Emmanual Cecchet Jean-Bernard Stefani	X			
MU	16	Daniel Kouril Ondrej Krajicek Ales Krenek Ludek Matyska Miroslav Ruda Lukas Hejtmanek Jiri Denemark	X		X	X
PSNC	17	Gracjan Jankowski Michal Jankowski Norbert Meyer Dominik Stoklosa Radek Januszewski Rafal Mikolajczak Pawel Wolniewicz		X		X
SZTAKI	20	Gabor Terstyanszky Gabor Gombas Gergely Sipos Jozsef Kovacs	X	X	X	
UCAM	24	Mark Hayes Andy Parker Mark Calleja Rizos Sakellari		X	X	
UNICAL	23	Antonio Congiusta Domenico Talia Paolo Trunfio Domenico Talia				
UOW	37	Tamas Kiss Steve C. Winter Tamas Kiss Vladimir Getov Thierry Delaitre Ariel Goyeneche	X		X	
UMUE	22	Sergei Gorlatch J. Mueller Martin Alt			X	
UNI DO	29	Ramin.Yahyapour			X	

Table 1: Participants of the Virtual Institute on Grid Information and Monitoring Services (WP5)