



DataGrid

CONTRIBUTIONS TO INTERNATIONAL STANDARDS

WP11: Information Dissemination and Exploitation

Document identifier: **DataGrid-11-D11.9-1081-1_2**
Date: **10/11/2003**
Work package: **WP11:** Information Dissemination and Exploitation

Partner(s): **CNR, CS-SI**
Lead Partner: **CNR**
Document status: **draft**

Deliverable identifier: **DataGrid-D11.9**

Abstract: This document summarizes the contributions to international standards provided by the EDG project during its life-time.

Delivery Slip

	Name	Partner	Date	Signature
From	R.Puccinelli	CNR	10/11/2003	
Reviewed by	German Cancio Mela Erwin Laure Peter Clarke David Kelsey	CERN CERN RAL RAL		
Approved by				

Document Log

Software Products	User files
Word	DataGrid-11-D11.9-1081-1_2.DOC

TABLE OF CONTENTS

• INTRODUCTION.....	5
1. OBJECTIVES OF THIS DOCUMENT	5
2. STRUCTURE OF THE DOCUMENT.....	5
3. APPLICABLE DOCUMENTS AND REFERENCE DOCUMENTS	5
4. DOCUMENT AMENDMENT PROCEDURE.....	6
5. GLOSSARY	6
• EXECUTIVE SUMMARY.....	8
1. CONTRIBUTIONS TO GGF ACTIVITIES	8
2. INTER-PROJECT COLLABORATIONS.....	9
3. PAPERS	10
• CONTRIBUTIONS BY WP.....	11
1. WORK PACKAGE 1 (WORK LOAD MANAGEMENT)	11
1. WORK PACKAGE DESCRIPTION	11
2. CONTRIBUTIONS TO GGF ACTIVITIES	11
3. INTER-PROJECT COLLABORATIONS.....	11
4. PAPERS	12
2. WORK PACKAGE 2 (GRID DATA MANAGEMENT).....	12
1. WORK PACKAGE DESCRIPTION.....	12
2. CONTRIBUTIONS TO GGF ACTIVITIES	13
3. INTER-PROJECT COLLABORATIONS.....	14
4. PAPERS	14
3. WORK PACKAGE 3 (GRID MONITORING)	17
1. WORK PACKAGE DESCRIPTION	17
2. CONTRIBUTIONS TO GGF ACTIVITIES	17
3. INTER-PROJECTS COLLABORATIONS.....	18
4. PAPERS	18
4. WORK PACKAGE 4 (FABRIC MANAGEMENT)	19
1. WORK PACKAGE DESCRIPTION	19
2. CONTRIBUTIONS TO GGF ACTIVITIES	19
3. INTER-PROJECT COLLABORATIONS.....	20
4. PAPERS	20
5. WORK PACKAGE 5	21
1. WORK PACKAGE DESCRIPTION.....	21
2. CONTRIBUTIONS TO GGF ACTIVITIES	21
3. INTER-PROJECT COLLABORATIONS.....	21
Papers	22



6. WORK PACKAGE 7 (NETWORK)	22
1. WORK PACKAGE DESCRIPTION.....	22
2. CONTRIBUTIONS TO GGF ACTIVITIES	23
3. INTER-PROJECT COLLABORATIONS.....	23
4. PAPERS	24
7. SECURITY CO-ORDINATION GROUP	24
1. GROUP DESCRIPTION.....	24
2. CONTRIBUTION TO GGF ACTIVITIES	24
3. INTER-PROJECT COLLABORATIONS.....	25
4. PAPERS	25
• CONCLUSIONS	27

•INTRODUCTION

1.OBJECTIVES OF THIS DOCUMENT

The objective of this document is to draw an exhaustive report of all the contributions provided by the DataGrid project to the establishment of international standards in the field of Grid Computing. Several activities and results of the DataGrid project contributed directly or indirectly to the specification of open standards which are now under discussion or already adopted inside international bodies such as the Global Grid Forum. Contributions have been provided in the following forms:

- participation of DataGrid members in GGF activities;
- collaborations with other Grid projects aimed at designing common protocols, interfaces or database schemas in order to improve interoperability between the respective Grid infrastructures;
- publishing of papers which present the results of the project and stimulate discussions in the community of Grid Computing experts.

Given the amount of results produced by the project, only in the next years we will be able to draw a clear picture of the contributions provided by DataGrid to the world of Grid computing.

2.STRUCTURE OF THE DOCUMENT

We have chosen to start with a general overview and then devote a section of this document to each work package because the project structure provides a consistent and effective classification of the activities and of the results. For each work package we will briefly describe the main tasks and the contributions to international standards, categorized according to the list presented in the previous paragraph.

Obviously the main contributions came from the middleware and infrastructure work packages, hence not all of the work packages are mentioned in this document.

3.APPLICABLE DOCUMENTS AND REFERENCE DOCUMENTS

This document is mainly based on the information provided by the work package managers of the DataGrid project. Additional information have been taken from the following sources:

[1] *DataGrid Work Package 12 – “Response to the 1st EU review report” - Document Identifier: DataGrid-12-NOT-0342-2-0-PY2_followup-to-first-EUreviewResponse*

[2] *DataGrid Work Package 12 – “Response to the 2nd EU review report” – Document Identifier: DataGrid-12-TED-376326*

[3] **“DataGrid Technical Annex v. 6.1”**

[4] **Global Grid Forum website** (<http://www.gridforum.org>)

[5] **DataGrid Dissemination website** (<http://web.datagrid.cnr.it>)

[6] **Gridstart Project website** (<http://www.gridstart.org>)

[7] **DataTAG Project website** (<http://www.datatag.org>)

[8] **EGEE Project website** (<http://www.eu-egee.org>)

[9] Massimo Sgaravatto – “DataGrid deliverable 1.7” – Document Identifier: DataGrid-01-D1.7-0146-1_0

[10] High Energy and Nuclear Physics Intergrid Joint Technical Board website (<http://www.hicb.org/>)

[11] Liberty Alliance Project website (<http://www.projectliberty.org/>).

4.DOCUMENT AMENDMENT PROCEDURE

This report describes the contributions provided by the DataGrid project to the establishment of international standards. It provides a snapshot of the situation at the last month of the project, hence no amendment procedure is foreseen for this document.

5.GLOSSARY

Authentication Authorization and Accounting Research Group (ACCT_RG): this GGF research group was formed to study accounting models for Grid environments.

Authorization Frameworks and Mechanisms Working Group (AuthZ-WG): this GGF group aims to define a conceptual grid authorization framework for grid developers with the goal to provide a basis for the design of grid authorization systems.

Computing Element (CE): a Computing Element is one of the resources available on a Grid based on the EDG middleware. It provides an “abstraction” of the computing power.

Database Access and Integration Services Working Group (DAIS-WG): this GGF group seeks to promote standards for the development of grid database services, focusing principally on providing consistent access to existing, autonomously managed databases.

Discovery and Monitoring Event Description Working Group (DAMED-WG): this GGF working group aims to define a basic set of monitoring event descriptions. These descriptions, or schemas, will describe the information (attributes) associated with a particular data element and will describe conventions for the representation of the value associated with it.

Global Grid Forum (GGF): The Global Grid Forum (GGF) is a community-initiated forum of more than 5000 individual researchers and practitioners working on distributed computing, or "Grid" technologies. GGF's primary objective is to promote and support the development, deployment, and implementation of Grid technologies and applications via the creation and documentation of "best practices" - technical specifications, user experiences, and implementation guidelines.

Grid Check Pointing Recovery Working Group (GridCPR-WG): this GGF working group aims to define a user-level API and associated layer of services that will permit checkpointed jobs to be recovered and continued on the same or on remote Grid resources.

Grid High Performance Network Research Group (GHPN-RG): this GGF research group focuses on the relationship between network research and Grid application and infrastructure development. The objective of GHPN-RG is to bridge the gap between the networking and grid research communities.

Local Centre Authorization Service (LCAS): handles authorization requests to the local computing fabric

Network Measurements Working Group (NM-WG): this GGF working group aims to define characteristics and measurement tools for network performance.

OGSA Data Replication Services Working Group (OREP-WG): this GGF working group is intended to create, review and refine grid service specifications for data replication services. These specifications will conform to the Grid Services Specification being developed by the OGSI Grid Service Infrastructure Working Group.

Open Grid Services Infrastructure Working Group (OGSI-WG): this GGF working group is to review and refine the Grid Service Specification and other documents that derive from this specification, including OGSA-infrastructure-related technical specifications and supporting informational documents.

Persistent Archives Research Group (PA-RG): this GGF research group of the Grid Forum promotes the development of an architecture for the construction of persistent archives. Persistent archives are conceptually equivalent to virtual data grids. The persistent archive research group address issues related to how persistent archives can be built from virtual data grids.

Relational Grid Information Services Research Group (RGIS-RG): this GGF research group serves researchers who are exploring approaches based on data models such as the relational data model for defining, communicating, recording, and querying static and dynamic Grid information.

Storage Element (SE): a Storage Element is one of the resources available on a Grid based on the EDG middleware. It provides an abstraction of the storage space.

WP: work package

•EXECUTIVE SUMMARY

The DataGrid project has been active in discussing, developing and promoting new standards in order to allow interoperability between the various Grid infrastructures. In this section we provide an overview of the contributions to international standards provided by the project.

1.Contributions to GGF activities

The Global Grid Forum (GGF) is a community-initiated forum of more than 5000 individual researchers and practitioners working on distributed computing, or "Grid" technologies. GGF's primary objective is to promote and support the development, deployment, and implementation of Grid technologies and applications via the creation and documentation of "best practices" - technical specifications, user experiences, and implementation guidelines.

GGF is organized into thematic areas. Each area contains several working and research groups. A working group is generally focused on a very specific technology or issue with the intention to develop one or more specific documents aimed generally at providing specifications, guidelines or recommendations. A research group is often longer-term focused, intending to explore an area where it may be premature to develop specifications. For more information see [4].

DataGrid representatives are providing contributions to different GGF Working and Research groups, holding in some cases key positions. It is worth mentioning that Fabrizio Gagliardi, the project manager, is a member of the GFAC – the GGF External Advisory Committee - and Peter Clarke from work package 7 is a member of GFSG – the GGF Steering Group – and “Data Area” director. Besides, several DataGrid members chair or co-chair some of the GGF research and working groups.

Here below is the list (in alphabetical order) of the GGF Working and Research Groups where EDG members provided contributions:

- Accounting Research Group (ACCT_RG)
- Authorization Frameworks and Mechanisms Working Group (AuthZ-WG)
- Certificate Authority Operations Working Group (CAOPs-WG)
- Database Access and Integration Services Working Group (DAIS-WG)
- Discovery and Monitoring Event Description Working Group (DAMED-WG)
- Grid Check Point Recovery Working Group (GridCPR-WG)
- Grid High Performance Network Research Group
- Network Measurements Working Group
- OGSA Data Replication Services Working Group (OREP-WG)
- Open Grid Service Architecture Authorization (OGSA AUTHZ-WG)
- Open Grid Services Infrastructure Working Group (OGSI-WG)
- PERF group
- Persistent Archives Research Group
- Relational Database Information Services Research Group (RDIS-RG)
- Relational Grid Information Services Research Group (RGIS-RG)
- Replication Research Group (R-RG)

Contributions have been provided by participating in discussions, working on GGF documents or proposing some project's results for the adoption as standards. Here below we draw a short summary of the major contributions provided by each work package:

- WP1 proposed:
 - the “Grid check-pointing framework” to the “Grid Check Pointing Research and Working Group” (GridCPR-WG),
 - the “EDG Job Description Language” to the “Job Submission Description Language Working Group” (JSDL-WG), and
 - an economy-based accounting model to the “Accounting Research Group” (ACCT_RG);
- WP2 participated in the preparation of the “Grid Database Service Specification”;
- WP3 provided significant contributions to the definition of the Grid monitoring Architecture and developed a relational implementation (R-GMA), proposed for the adoption within GGF;
- WP4 developed the Local Center Authorization Service (LCAS), which handles authorization requests to the local computing fabric, and proposed it to the OGSA AuthZ Working Group;
- WP7 contributed to the preparation of the “Grid Network Service” outline document within the “Grid High-Performance Networking Research Group”. They also participated to the preparation of some documents within the “Network Monitoring Working Group”.

In the field of security, the experience gained building the DataGrid PKI and inter-Grid authentication with the other projects has been valuable input to the GGF.

2. Inter-project collaborations

DataGrid participated and continues to participate to various inter-project initiatives outside of the GGF.

DataGrid is member and major actor in the EU GridSTART accompanying measure which has in its charter the establishment of de facto standards in all the EU supported Grid projects. GridSTART is an initiative sponsored by the European Commission with the specific objective of consolidating technical advances in Europe, encouraging interaction amongst similar activities both in Europe and the rest of the world and stimulating the early take-up by industry and research of Grid-enabled applications. The initiative brings together technologists, scientists and industry in multi-disciplinary approach to developing the Grid infrastructure. The clear goal is to develop sustainable, effective and universal solutions addressing the needs of science, industry and the public.

The DataGrid project has also cooperated with the DataTAG project providing substantial contributions to the GLUE initiative mainly in the definition of the GLUE schema. The goal of the DataTAG project is to create a large-scale intercontinental testbed for data-intensive Grids. The project is focused on high speed networks and grid interoperability. GLUE is an initiative born within DataTAG and stands for Grid Laboratory Universal Environment. It is operating under the recommendation of the High Energy and Nuclear Physics Intergrid Joint Technical Board, whose mandate is to design, develop and deploy a consistent open source standards-based global Grid infrastructure. GLUE aims to sponsor and enable interoperability between the EU physics grid project efforts (EDG, DataTAG, etc.) and the US physics grid project efforts (iVDGL, PPDG, GriPhyN). The Glue-schema effort is a sub-group of GLUE. The schema provides a uniform way of representing information about Grid resources. It has been separated into Computing Element, Storage Element and

Network Element schemas, which have been extensively tested by the DataTAG project. The GLUE schema is now adopted by the information systems of the DataGrid infrastructure.

WP2 and WP5 participated to the Storage Resource Management (SRM) interface definition. This is an effort led by Arie Shoshani from Lawrence Berkeley National Lab. Other participants are Jefferson Lab and Fermi National Accelerator Lab. A GGF Working Group should be instantiated which will take over the results of this initiative. The EDG Storage Element Control Interface is an implementation of the Storage Resource Manager protocol.

Version 1 of the SRM protocol is the only common standard for mass storage system access between the American labs mentioned above and it is now gaining wider support with CERN and RAL providing implementations. Other communities are looking at SRM interfaces as well. In particular, LCG have decided to use the SRM protocol to access mass storage systems.

In the field of networking, collaborations have been established by WP7 with the following projects and initiatives:

- “Internet2 End-to-End Performance Initiative (E2Epi) Performance Environment System”, a project designed to allow users and administrators to locate and diagnose network problems along the complete end-to-end path of a connection;
- “Dante Inter-Domain Performance Monitoring group”, concerned with enabling measurements to be made between domains within the Géant project;
- “UK GridMon”, a UK e-Science project concerned with enabling network monitoring across the UK. It uses tools based on earlier work by WP7: PingER, IperfER and UDPMon. The development of web services for accessing the monitoring data is also part of this project.

In the field of security, the DataGrid WP6 “Certificate Authorities Group” has built a large-scale Public Key Infrastructure (PKI) to meet the needs not only of Authentication within DataGrid, but also to facilitate cross-authentication between Grid projects. This collaboration has involved active participation of CA managers from several other Grid projects across many different countries resulting in a current total of some 20 approved CAs.

Besides the initiatives mentioned above, application portals are developed in activities closely related to DataGrid to make application porting easier and less dependent on the underlying middleware development.

3.Papers

EDG members submitted many papers to Grid related conferences and journals. Those papers describe the major achievements of the project and stimulate discussions within the Grid community. A list of the project’s published papers is available on the dissemination portal at the URL below:

http://web.datagrid.cnr.it/pls/portal30/GRID.RPT_DATAGRIDPAPERS.show

In the following sections we will provide for each work package the list of the published papers.

•CONTRIBUTIONS BY WP

1.WORK PACKAGE 1 (WORK LOAD MANAGEMENT)

1.Work Package Description

The goals of this Work Package were to define and implement an architecture for distributed scheduling and resource management in a Grid environment. Because of the general-purpose nature of the Grid, this infrastructure implies an unpredictable, chaotic workload generated by relatively large numbers of independent users.

Many challenging issues have been considered:

- optimal choice of execution location based on the availability of data, computation and network resources;
- optimal co-allocation and advance reservation of CPU, data and network resources;
- definition and implementation of uniform interfaces to different local resource management systems;
- definition of resource usage policies.

The workload management system developed by WP1 has been designed to be used in real production environments. This means that reliability, fault tolerance and scalability have been taken as key requirements.

2.Contributions to GGF activities

WP1 have been constantly participating to the activities of the GGF Scheduling and Resource Management areas.

WP1 members have participated to the following research and working groups:

- **Grid Check Pointing Research and Working Group (GridCPR-WG)**: WP1 was following since the beginning the activities of this WG, presenting the Grid checkpointing framework implemented in the context of the DataGrid project. WP1 is now active in the definition of a Grid Checkpointing Architecture and in the standardization of the Grid Checkpointing APIs.
- **Job Submission Description Language (JSDL-WG)**: WP1 documents concerning the EDG Job Description Language have been submitted to this WG for evaluation.
- **Accounting Research Group (ACCT RG)**: WP1 proposed an economy-based accounting model.

3.Inter-project collaborations

When planning the first Workload Management System (and the first release of the EDG software), it became quite clear that the default schema provided by the Globus MDS Information Services was not suitable to describe the characteristics and status of the Grid resources (in particular Computing Elements and Storage Elements¹), since many important (at least from the EDG point of view) information could not be represented. WP1 triggered the discussion, whose outcome was the definition of a schema for the EDG Information Service, which was used for the first release of the EDG software. This effort has then been carried on within the DataTAG 'GLUE' activity.

¹ In the rest of the document we will often use the acronyms CE for Computing Element and SE for Storage Element. See the glossary for their definitions.

4.Papers

C. Anglano et al. - **“Integrating Grid tools to build a Computing Resource Broker: Activities of DataGrid WP1”** - *Proceedings of the Conference on Computing in High Energy Physics 2001 (CHEP01), Beijing, September 2001.*

G. Avellino et al. - **“The First deployment of Workload Management Services on the EU DataGrid testbed: feedback on design and implementation”** - *2003 Conference for Computing in High-Energy and Nuclear Physics (CHEP 03), La Jolla, California, 24-28 Mar 2003*

G. Avellino et al. - **“The EU DataGrid Workload Management System: towards the second major release”** - *2003 Conference for Computing in High-Energy and Nuclear Physics (CHEP 03), La Jolla, California, 24-28 Mar 2003.*

L. Matyska, A. Krenek, M. Ruda, M. Vocu, Z. Salvat, J. Sitera, J. Pospisil and D. Kouril - **“The Grid Job Monitoring Service”** - *Terena Networking Conference, 2002*

A. Werbrouck, R. Piro, A. Guarise - **“An Economy-based Accounting Infrastructure for the DataGrid”** - *4th International Workshop on Grid Computing (Grid2003), Phoenix, Arizona, 17 November 2003*

G. Avellino, S. Beco, F. Pacini, A. Maraschini, A. Terracina - **“Job Description Language and User Interface in a Grid context: the EU DataGrid experience”** - *MiniSymposium on Grid Computing, Parallel Computing 2003, September 2-5, 2003, Dresda.*

WP1 had also the opportunity to present its activities at the “New network technologies, Grids and portals” conference, organized on the occasion of 10th anniversary of Poznań Supercomputing and Networking Center (<http://www.man.poznan.pl/10years/main.html>), as invited talk. The corresponding paper (“The learning experience of finding resources on the EU DataGrid testbed”), to be submitted by January 2004, is under preparation.

Two other papers (under preparation at the time of writing) will be submitted to the Terena Network Conference 2004: one on recent advances in the LB service, and the other on the connection with R-GMA.

2.WORK PACKAGE 2 (GRID DATA MANAGEMENT)

1.Work Package description

The goals of this Work Package were to specify, develop, integrate and test tools and middleware components designed and developed to coherently manage and share petabyte-scale information volumes in high-throughput production quality Grid environments. The work package developed general purpose information sharing solutions which aim at providing unprecedented automation, ease of use, scalability, uniformity and transparency. Those solutions allow to securely access massive amounts of data in a universal global name space, to move and replicate data at high speed from one

geographical site to another. WP2 also developed solutions for automated data-caching on wide-area networks.

2. Contributions to GGF activities

Most of the WP2 contributions to international standards have been provided in the context of the Global Grid Forum. WP2 actively participated in the following groups:

- **Replication Research Group**: Peter Kunszt co-chairs this RG;
- **Database Access and Integration Services WG**: The co-chair of this WG used to be from WP2;
- **Relational Database Information Services Research Group**;
- **Authorization Working Group**;
- **Security Working Group**.

WP2 members also participated to the discussions of the following GGF Research and Working Groups:

- **Open Grid Services Infrastructure Working group**;
- **Persistent Archives Research Group**;
- **OGSA Replication Services Working Group**.

WP2 members participated in the preparation of the following GGF documents:

M. Antonioletti, M. Atkinson, A. Krause, S. Malaika, G. McCance, J. Magowan, N. W. Paton, G. Riccardi - “Grid Database Service Specification” - Data Access and Integration Services Group Working Draft, Global Grid Forum 8, Seattle, USA, June 2003.

N. P. Chue Hong, A. Krause, S. Malaika, G. McCance, S. Laws, J. Magowan, N. W. Paton, G. Riccardi - “Grid Database Service Specification” - Data Access and Integration Services Group Working Draft, Global Grid Forum 7, Tokyo, Japan, March 2003.

A. Krause, S. Malaika, G. McCance, J. Magowan, N. W. Paton, G. Riccardi - “Grid Database Service Specification” - Data Access and Integration Services Group Working Draft, Global Grid Forum 6, Chicago, USA, October 2002.

The following documents have been presented during GGF conferences:

Leanne Guy, Peter Kunszt, Erwin Laure, Heinz Stockinger, Kurt Stockinger – “Replica Management in Data Grids” - Technical Report, Global Grid Forum - GGF5, Edinburgh, Scotland, July 2002.

William H. Bell, Diana Bosio, Wolfgang Hoschek, Peter Kunszt, Gavin McCance, Mika Silander. Project Spitfire - Towards Grid Web Service Databases. Technical Report, Global Grid Forum - GGF5, Edinburgh, Scotland, July 2002

Leanne Guy, Peter Kunszt, Erwin Laure, Heinz Stockinger, Kurt Stockinger. – **“Replica Management in Data Grids”** - *Informational Document Global Grid Forum 5 Edinburgh, Scotland, July 21-24*

William H. Bell, Diana Bosio, Wolfgang Hoschek, Peter Kunszt, Gavin McCance, and Mika Silander – **“Project Spitfire - Towards Grid Web Service Databases”** - *Informational Document Global Grid Forum 5 Edinburgh, Scotland, July 21-24*

3. Inter-project collaborations

Standardization efforts outside GGF include :

- Standardization of a Storage Resource Management (SRM) interface.
- Involvement in the Liberty Alliance project²

4. Papers

Wolfgang Hoschek, Javier Jaen-Martinez, Asad Samar, Heinz Stockinger, Kurt Stockinger - **“Data Management in an International Data Grid Project”** - (*“distinguished paper” award*) *1st IEEE/ACM International Workshop on Grid Computing (Grid'2000), Bangalore, India, December 17-20, 2000.*

Luciano Serafini, Heinz Stockinger, Kurt Stockinger, Floriano Zini - **“Agent-Based Query Optimisation in a Grid Environment”** - *IASTED International Conference on Applied Informatics (AI2001), Innsbruck, Austria, February 19 -22, 2001.*

Heinz Stockinger - **“Distributed Database Management Systems and the Data Grid”** - *18th IEEE Symposium and 9th NASA Goddard Conference on Mass Storage Systems and Technologies (MSS2001), San Diego, April 17-20, 2001.*

Heinz Stockinger, Omer F. Rana, Reagan Moore, Andre Merzky - **“Data Management in a Grid Environment”** - *European High Performance Computing & Networking Conference (HPCN2001), Amsterdam, June 25-27, 2001.*

Heinz Stockinger, Asad Samar, Bill Allcock, Ian Foster, Koen Holtman & Brian Tierney - **“File and Object Replication in Data Grids”** - *Proceedings of the 10th IEEE Symposium on High Performance and Distributed Computing (HPDC-10), San Francisco.*

² The Liberty Alliance Project was formed in September 2001 to develop open standards for federated network identity management and identity-based services. Its goals are to ensure interoperability, support privacy, and promote adoption of its specifications, guidelines and best practices. The Liberty Alliance Project develops open specifications and does not deliver specific products or services. For further information see [11].

Dirk Duellman, Wolfgang Hoschek, Javier Jaen-Martinez, Asad Samar, Heinz Stockinger, and Kurt Stockinger - **“Models for Replica Synchronisation and Consistency in a Data Grid”** - *Proceedings of the 10th IEEE Symposium on High Performance and Distributed Computing (HPDC-10), San Francisco (to appear).*

Paolo Busetta, Mark Carman, Luciano Serafini, Kurt Stockinger & Floriano Zini -**“Grid Query Optimisation in the Data Grid”** - *IRST Technical Report 0109-01, Istituto Trentino di Cultura, September 2001.*

Heinz Stockinger - **“Database Replication in World-wide Distributed Data Grids”** - *PhD Thesis.*

Kurt Stockinger - **“Multi-Dimensional Bitmap Indices for Optimising Data Access within Object Oriented Databases at CERN”** - *PhD Thesis.*

Mark Carman, Floriano Zini, Luciano Serafini, Kurt Stockinger – **“Towards an Economy-Based Optimisation of File Access and Replication on a Data Grid”**- *International Workshop on Agent based Cluster and Grid Computing at International Symposium on Cluster Computing and the Grid (CCGrid'2002), Berlin, Germany, May 2002, IEEE Computer Society Press*

Wolfgang Hoschek – **“A Database for Dynamic Distributed Content and its Application for Service and Resource Discovery”**. *In Int'l. IEEE Symposium on Parallel and Distributed Computing (ISPDC 2002), Iasi, Romania, July 2002.*

Kurt Stockinger. – **“Bitmap indices for speeding up high-dimensional data analysis”** *In International Conference on Database and Expert Systems Applications, Aix-en-Provence, France, September 2002. (to appear) Springer-Verlag.*

Wolfgang Hoschek – **“Dynamic Timeouts and Neighbor Selection Queries in Peer-to-Peer Networks”** - *In Int'l. Conf. on Networks, Parallel and Distributed Processing and Applications (NPDP 2002) (to appear), Tsukuba, Japan, October 2002.*

Wolfgang Hoschek. – **“The Web Service Discovery Architecture”** - *In Proc. of the Int'l. IEEE Supercomputing Conference (SC 2002) (accepted), Baltimore, USA, November 2002.*

Ann Chervenak, Ewa Deelman, Ian Foster, Wolfgang Hoschek, Adriana Iamnitchi, Carl Kesselman, Peter Kunszt, Matei Ripeanu, Heinz Stockinger, Kurt Stockinger, and Brian Tierney – **“Giggle: A Framework for Constructing Scalable Replica Location Services”** - *In Proc. of the Int'l. IEEE Supercomputing Conference (SC 2002) (accepted), Baltimore, USA, November 2002*

Wolfgang Hoschek - **“Web Service Discovery Processing Steps”** - *In Proc. of the Int'l. WWW/Internet 2002 Conference (to appear), Lisbon, Portugal, November 2002.*

Kurt Stockinger. – **“Bitmap Indices for Speeding Up High-Dimensional Data Analysis”** - *International Conference on Database and Expert Systems Applications, Aix-en-Provence, France, September 2002, Springer-Verlag.*

Kurt Stockinger, Kesheng Wu, and Arie Shoshani. **“Strategies for Processing ad-hoc Queries on Large Data Warehouses”** *ACM Fifth International Workshop on Data Warehousing and OLAP (DOLAP 2002), McLean, VA, USA, November 2002. ACM Press.*

William H. Bell, David G. Cameron, Luigi Capozza, A. Paul Millar, Kurt Stockinger, Floriano Zini - **“Simulation of Dynamic Grid Replication Strategies in OptorSim”** - *Proc. of the ACM/IEEE Workshop on Grid Computing (Grid 2002), Baltimore, USA, November 2002. Springer-Verlag.*

Ann L. Chervenak, Ewa Deelman, Ian Foster, Adriana Iamnitchi, Carl Kesselman, Wolfgang Hoschek, Peter Kunszt, Matei Ripeanu, Bob Schwartzkopf, Heinz Stockinger, Kurt Stockinger and Brian Tierney, - **“Giggle: A Framework for Constructing Scalable Replica Location Services”** - *Proc. of the IEEE Supercomputing Conference (SC 2002), Baltimore, USA, November 2002. IEEE Computer Society Press*

Kurt Stockinger, Kesheng Wu, and Arie Shoshani. - **“Strategies for Processing ad-hoc Queries on Large Data Warehouses”** - *ACM Fifth International Workshop on Data Warehousing and OLAP (DOLAP 2002), McLean, VA, USA, November 2002. ACM Press.*

William H. Bell, David G. Cameron, Luigi Capozza, A. Paul Millar, Kurt Stockinger, Floriano Zini, - **“Simulation of Dynamic Grid Replication Strategies in OptorSim”** - *Proc. of the ACM/IEEE Workshop on Grid Computing (Grid 2002), Baltimore, USA, November 2002. Springer-Verlag.*

Ann L. Chervenak, Ewa Deelman, Ian Foster, Adriana Iamnitchi, Carl Kesselman, Wolfgang Hoschek, Peter Kunszt, Matei Ripeanu, Bob Schwartzkopf, Heinz Stockinger, Kurt Stockinger and Brian Tierney, **“Giggle: A Framework for Constructing Scalable Replica Location Services”** *Proc. of the IEEE Supercomputing Conference (SC 2002), Baltimore, USA, November 2002. IEEE Computer Society Press.*

“Evaluation of an Economy-Based File Replication Strategy for a Data Grid”. In *International Workshop on Agent based Cluster and Grid Computing at CCGrid 2003, Tokyo, Japan, May 2003. IEEE Computer Society Press.*

William H. Bell, David G. Cameron, Luigi Capozza, A. Paul Millar, Kurt Stockinger, Floriano Zini, **“Simulation of Dynamic Grid Replication Strategies in OptorSim”**, *Proc. of the ACM/IEEE Workshop on Grid Computing (Grid 2002), Baltimore, USA, November 2002. Springer-Verlag.* got ranked among the topic three papers of this workshop. A longer version will appear as journal paper:

William H. Bell, David G. Cameron, Luigi Capozza, A. Paul Millar, Kurt Stockinger, Floriano Zini, **“Simulation of Dynamic Grid Replication Strategies in OptorSim”**, *To appear in International Journal of High Performance Computing Applications, Vol.17 (4), 2003.*

Tapio Niemi, Marko Niinimäki, Jyrki Nummenmaa, Peter Thanisch: “**Applying Grid Technologies to XML Based OLAP Cube Construction**”, *CERN Open Preprint series ref. no. CERN-OPEN-2003-004*.

William H. Bell, David G. Cameron, Ruben Carvajal-Schiaffino, A. Paul Millar, Kurt Stockinger, and Floriano Zini. **Evaluation of an Economy-Based File Replication Strategy for a Data Grid**. In *International Workshop on Agent based Cluster and Grid Computing at CCGrid 2003, Tokyo, Japan, May 2003*. IEEE Computer Society Press.

D. G. Cameron, R. Carvajal-Schiaffino, A. P. Millar, C. Nicholson, K. Stockinger, F. Zini. **UK Grid Simulation with OptorSim**. In *2nd e-Science All Hands Meeting, Nottingham, UK, September 2003*.

William H. Bell, David G. Cameron, Luigi Capozza, A. Paul Millar, Kurt Stockinger, and Floriano Zini. **OptorSim - A Grid Simulator for Studying Dynamic Data Replication Strategies**. *International Journal of High Performance Computing Applications*, 17(4), 2003.

David G. Cameron, Ruben Carvajal-Schiaffino, A. Paul Millar, Caitriana Nicholson, Kurt Stockinger, Floriano Zini, **Evaluating Scheduling and Replica Optimisation Strategies in OptorSim**, to appear in *International Workshop on Grid Computing (Grid2003), Phoenix, Arizona, November 2003*, IEEE Computer Society Press.

Kurt Stockinger, Heinz Stockinger, Lukasz Dutka, Renata Slota, Darin Nikolow, Jacek Kitowski, **Access Cost Estimation for Unified Grid Storage Systems**, to appear in *International Workshop on Grid Computing (Grid2003), Phoenix, Arizona, November 2003*, IEEE Computer Society Press.

Peter Kunszt, Erwin Laure, Heinz Stockinger, and Kurt Stockinger. **Advanced Replica Management with Reptor**, *International Conference on Parallel Processing and Applied Mathematics, Czestochowa, Poland, September, 2003*, Springer-Verlag.

3.WORK PACKAGE 3 (GRID MONITORING)

1.Work Package Description

The goals of this Work Package were to specify, develop, integrate and test tools and infrastructures to enable end-user and administrator access to status and error information in a Grid environment and to provide an environment in which application monitoring can be carried out. This allows both for job performance optimisation as well as problem tracing and is crucial to facilitate high performance Grid computing.

2.Contributions to GGF activities

WP3 members actively participated in the following GGF working groups.

- **Relational Grid Information Services Research Group (RGIS-WG)**: Steve Fisher (RAL) co-chairs the RGIS-WG.
- **PERF group**: Steve Fisher has made contributions to the Grid Monitoring Architecture discussions within the PERF group, and has submitted R-GMA as a GMA implementation. Other members of PERF group coming from WP3 are James Magowan and Manfred Oevres.
- **Discovery and Monitoring Event Description Working Group (DAMED-WG)**: James Magowan (IBM-UK) co-chairs the DAMED-WG.
- **Data Access and Information Services Working Group (DAIS-WG)**: James Magowan and Paul Taylor are members of DAIS-WG.
- **Open Grid Services Infrastructure Working Group (OGSI-WG)**: Abdeslem Djaoui is member of OGSI-WG.
- **Open Grid Services Architecture Working Group (OGSA-WG)**: WP3 are preparing submissions about GMA to the OGSA-WG showing use cases and a service overview.

The following documents have been presented during GGF meetings:

S.M.Fisher - "More thoughts on Architecture and Evolution" - Presentation to Grid Protocol Architecture session of Global Grid Forum 3 in Frascati Rome, 7-10th October 2001.

A.Djaoui, L.Field, S.M.Fisher, J.Magowan, G.McCance and M.Oevers - "Protocols in the EU DataGrid R-GMA" - Presentation to the Perf session of Global Grid Forum 3 in Frascati, Rome, 7-10th October 2001.

3. Inter-projects collaborations

WP3 participated to the definition of the GLUE schema.

Efforts have been made to integrate R-GMA into Globus Toolkit 3 preserving at the same time the original APIs.

4. Papers

Zoltán Balaton, Péter Kacsuk, Norbert Podhorszki and Ferenc Vajda – “From Cluster Monitoring to Grid Monitoring Based on GRM and PROVE” - Submitted to EuroPar'2001, Manchester, UK

Steve Fisher – “Relational Model for Information & Monitoring” - Submitted to PERF-Working Group at Global Grid Forum-1, Amsterdam

Brian Coghlan (Trinity College Dublin), Abdeslem Djaoui (RAL), Steve Fisher (RAL), James Magowan (IBM-UK), Manfred Oevers (IBM-UK) – “Time, Information Services and the Grid” - 18th British National Conference on Databases (BNCOD) and to published in CLRC Conference Proceedings RAL-CONF-2001-003

Z. Balaton, P. Kacsuk, N. Podhorszki, F. Vajda - "From Cluster Monitoring to Grid Monitoring Based on GRM" - Proceedings. 7th EuroPar'2001 Parallel Processing, Manchester, UK. pp. 874-881., 2001

Norbert Podhorszki – “Párhuzamos programok monitorozása klaszteren és griden” (“Monitoring of Parallel Applications in Clusters and in Grids”) in Hungarian - *Networkshop 2002, March 2002, Eger, Hungary.*

WP3 - “R-GMA: A Relational Grid Information and Monitoring System” *Invited talk and paper at Cracow '02 Grid Workshop. ISBN: 83-915141-0-2*

WP3 - “Relational Grid Monitoring Architecture” - *A poster at GlobusWorld 2003 in San Diego, 2003*

WP3 - “R-GMA: First results after deployment” - *CHEP03 at La Jolla, California, 2003*

N. Podhorszki, P. Kacsuk - “Presentation and Analysis of Grid Performance Data” - *EuroPar'2003, Klagenfurt, Austria, 2003*

N. Podhorszki - “Pulse: A Tool for Presentation and Analysis of Grid Performance Data” - *Submitted to MIPRO'2003, Opatija, Croatia*

N. Podhorszki, P. Kacsuk - “Monitoring Message Passing Applications in the Grid with GRM and R-GMA” - *EuroPVM/MPI'2003, Venice, Italy, 29 Sept. 2 Oct.*

Andy Cooke, Alisdair Gray, Lisha Ma, Werner Nutt - “R-GMA: An Information Integration System for Grid Monitoring” - *Tenth International Conference on Cooperative Information Systems (CoopIS) 3-7 November 2003, Catania, Sicily, 2003*

WP3 – “Relational Grid Monitoring Architecture (R-GMA)” - *Submitted to UK e-Science All Hands Meeting, September 2-4 2003, Nottingham*

4.WORK PACKAGE 4 (FABRIC MANAGEMENT)

1.Work Package Description

The goals of this Work Package were to develop new automated system management techniques to allow the deployment of very large computing fabrics constructed from mass market components with reduced systems administration and operations costs. Those management techniques should support an evolutionary model allowing the addition and replacement of components, and the introduction of new technologies.

2.Contributions to GGF activities

With the development of Local Centre Authorization Service (LCAS)³ WP4 defined the reference framework for including site-dependent authorization mechanisms in Grid job submission. This is being fed back via the OGSA-AuthZ Working Group (co-chaired by Von Welch, Andrew McNab and

³ Local Centre Authorization Service (LCAS) handles authorization requests to the local computing fabric.

Rebekah Lepro). WP4 members are closely following that WG, as well as some others EDG members in the Security Group (most notably Andrew McNab, but also Linda Cornwall and David Kelsey).

3. Inter-project collaborations

WP4 cooperated with the LCFG project in order to test new mechanisms and procedures for system installation and maintenance which could offer an alternative to existing procedures. In particular LCFG component scripts can effectively replace the Unix System V service management subsystem and the LCFG installer provides a good alternative to system base installation tools such as Anaconda/KickStart for RedHat Linux. Nevertheless the experience has shown that it is better to use standard systems whenever possible.

WP4 was involved in the definition and implementation of the GLUE schema for the Computing Element.

As regards LCAS, besides direct incorporation of the component in other Grid deployment efforts (notably in the Site Authorization Service at the US Fermi National Laboratory), this component triggered the development of the authorization call-out specification within Globus.

4. Papers

Reinefeld, V. Lindenstruth - "How to Build a High-Performance Compute Cluster for the Grid"- In: Proceedings of the ICPP'2001, International Workshop on Metacomputing Systems and Applications MSA'2001, IEEE Computer Society Press.

F. Schintke, J. Simon, A. Reinefeld - "A Cache Simulator for Shared Memory Systems"- International Conference on Computational Science ICCS 2001, San Francisco, CA, Springer LNCS 2074, vol. 2, pp. 569-578.

Keller, A. Reinefeld - "Anatomy of a Resource Management System for HPC-Clusters" - In: Yuen Chung Kwong (ed.), Annual Review of Scalable Computing, Vol. 3, Singapore University Press, ISBN 981-02-4579-3, 2001, pp. 1-31.

Lionel Cons, Piotr Poznanski - "PAN: A High-Level Configuration Language"- Invited paper at 2002 LISA XVI - November 3 - 8, 2002 - Philadelphia, PA.

Alexander Reinefeld, Florian Schintke - "Concepts and Technologies for a Worldwide Grid Infrastructure" - In Euro-Par 2002 Parallel Processing, volume 2400 of Lecture Notes in Computer Science, pages 62--71, Springer 2002. (c) Springer-Verlag

Olof Barring -, "Towards automation of computing fabrics using tools from the fabric management workpackage of the EU DataGrid project" - CHEP '03 conference

D. L. Groep, W. Som de Cerff, M. Steenbakkers, G. Venekamp - “Managing dynamic user communities in a Grid of autonomous resources” – CHEP '03 conference. WP4 provided contributions about LCMAPS and LCAS

Rafael Garcia Leiva – “Automating network administration tasks under Linux” - Submitted to the 6th Hispalinux Congress.

5.WORK PACKAGE 5

1.Work Package description

This work package worked to implement solutions involving Mass Storage Management Systems (MSMS) to make data more accessible for testbed users as well as providing an interface to the higher level data access architectures as they develop.

The goals of this Work Package were:

- Recognising the use of different existing MSMSs by the user community, provide extra functionality through common user and data export/import interfaces to all different *existing* local mass storage systems used by the project partners.
- Ease integration of local mass storage system with the Grid data management system by using these interfaces and through relevant information publication.

The goals listed above have been achieved by:

- defining and implementing a common API to all MSMSs of interest;
- defining and implementing an interchange mechanism for physical media between heterogeneous MSMSs together with exchange of the relevant meta data;
- publication of information about the data held (meta data) and about the MSMS itself.

2.Contributions to GGF activities

John Gordon has tracked the work of and participated in various research and working groups in the Data Area. Among them is worth mentioning:

- *GridFTP*,
- *Data Transport*,
- *Data Replication*
- *Open Grid Services Architecture - Database Access and Integration Services Working Group*;

3.Inter-project collaborations

WP5 have actively participated in the SRM (Storage Resource Manager) collaboration building an SRM interface to Castor, and are currently integrating a more generalised SRM interface to the Storage Element.

Furthermore, WP5 have actively participated in the development of the latest version of the SRM protocol, version 2.1. This version has just been finalised and no implementation is available at the moment.

Most of the version 1 SRMs do not provide secure interfaces, but both of the SRM implementations provided by WP5 improve on the standard by implementing secure interfaces using Globus (GSI) security. Indeed, as of version 2.1, this is now a requirement.

WP5 actively participated to the design of the GLUE schema in the following ways:

- participation in the GLUE schema conference discussions during 2002;
- definition of the following GLUE schema components:
 - Storage Element schema, versions 1.0 and 1.1 (final),
 - CE-SE Bind , versions 1.0 and 1.1 (final);
- implementation of the Glue Schema within the EDG Storage Element software and more precisely:
 - Implementation of Glue Schema for MDS 2.x (LDAP schema) used by the ‘Classical SE’ configuration within EDG,
 - Implementation using MDS 2.x with the Glue Schema – this generates information for the Storage Service and Storage Space with basic support for the Storage Library objects. This configuration is used by the ‘WP5 –SE’ Storage Element.

Papers

“**DataGrid Overview**” - *IAEA TCM2001 Workshop, Padova, Italy 16-18 July 2001*

Jules Wolfrat, Pieter de Boer, Walter de Jong, Ron Trompert - "A mass storage solution for Grid environments" - proceedings of the EUNIS 2003 conference, Amsterdam July 2-4 2003, p 354.

J Gordon, J Jensen, O Syngé - "Enabling Access to Mass Storage" - proceedings of the UK e-Science "All-Hands" meeting, 2-4 September 2003.

6.WORK PACKAGE 7 (NETWORK)

1.Work Package description

The goals of this Work Package were:

- Defining the network service requirements of DataGrid and make detailed plans in collaboration with the European and national actors involved;
- studying, deploying and managing advanced network and transport services;
- monitoring the traffic on and performance of the network, and develop models and provide tools and data for the planning of future networks, especially concentrating on the requirements of grids handling significant volumes of data.
- dealing with the distributed security aspects of DataGrid.

WP7 broadly used current emerging standards in order to achieve maximum interoperability with other Grid components. Web and Grid services standards in particular seemed to be mature enough to be used.

2. Contributions to GGF activities

WP7 is leading or involved in several network groups in GGF, including:

- **Grid High-Performance Networking Research Group**, Richard Hughes-Jones has contributed to the first deliverable document, and Peter Clarke is contributing to the “Grid Network Service” outline document.
- **Data Transport Research Group**: Pascale Primet (chair)
- **Network Monitoring Working Group (NMWG)** Richard Hughes-Jones (Chair), Paul Mealor working on XML and OGSA publication
- **Data Area**, Peter Clarke is Director within the GGF, with responsibility for all network oriented groups.

3. Inter-project collaborations

WP7 actively participates in the definition of the “Network Element” within the GLUE schema in order to achieve a standard network information model to store all related metrics. The GLUE “Network monitoring schema group” provides simple APIs for accessing the monitoring data and the relationships between the components of the model. The monitoring data itself is presented with a similar structure to that used within WP7 at present – that is the information content is the same as that used within WP7.

Much of the work in building an implementation of the GLUE schema has been carried out at INFN. The implementation of the GLUE Schema APIs is based on a combination of MySQL and R-GMA for the storage system, and Perl as the programming language. A prototype is being tested at INFN.

The WP7 team has collaborated in the definition and implementation of the schema, especially in providing R-GMA info providers suitable for the GLUE schema.

WP7 cooperated with the following initiatives or projects:

- “Internet2 End-to-End Performance Initiative (E2Epi) Performance Environment System”, a project designed to allow users and administrators to locate and diagnose network problems along the complete end-to-end path of a connection;
- “Dante Inter-Domain Performance Monitoring group”, concerned with enabling measurements to be made between domains within the Geant project;
- “UK GridMon”, a UK e-Science project concerned with enabling network monitoring across the UK. It uses tools based on earlier work by WP7: PingER, IperfER and UDPMon. The development of web services for accessing the monitoring data is also part of this project.

The capability for high throughput transfers between Grid nodes has been largely demonstrated inside WP7 and in other Grid projects, especially in the DataTAG Project. DataTAG is currently writing several *how-to* documents to disseminate these results, and in particular a “*How to tune TCP*”

for gigabit networks” document. Several authors of this *how-to* are also members of EDG WP7, and WP7 will also perform a general review of it.

4.Papers

Ferrari, T.; Giacomini, F. - “Network Monitoring for Grid Performance Optimization” - Submitted for publication on Computer Communications, March 2002.

Harakaly R., Primet P., Bonnassieux F., Gaidioz B. - “Probes Coordination Protocol for Network Performance Measurement in Grid Computing Environment” - submitted to Journal of PDCP - Special issue for Internet computing

Stockinger.K., Stockinger H., Harakaly R., Bonnassieux F., Primet P- “Optimisation of File Replication in a Data Grid Using Network Cost Function” - submitted to Journal of Grid Computing in September 2003

7.SECURITY CO-ORDINATION GROUP

1.Group Description

The DataGrid Security Co-ordination Group (SCG) is formally part of WP7. It was formed in November 2001 with the following mandate.

- To produce the EDG deliverables of WP7 on Security (D7.5, D7.6 and D7.7)
- To help coordinate, where necessary, the various Security activities taking place in Work Packages 1 to 5 and 7.
- To liaise with the WP6 Certificate Authorities and Authorisation groups, national Grid projects and Globus
- To contribute to the various versions of the Architecture of the EU DataGrid via input to ATF.

The majority of the members of the SCG are developers from the middleware Work Packages. There is only limited effort available within SCG for any independent security development. Members of the group have written the 3 Deliverable documents and worked on international standards in GGF and collaborative security activities with other Grid projects.

Two other groups, Certificate Authority and Authorization, are sub-groups of EDG WP6 charged with managing the operational issues of the Authentication and Authorization systems. As they have always worked closely with the SCG, with many of the same participants, the combined security activities are described here under one section.

2.Contribution to GGF activities

Grid Security is an important low-level foundation of all Grid middleware. The EDG SCG has always recognised the importance of international standards in both Authentication and Authorization so has been an active participant in most of the working groups and research groups in the GGF Security Area.

The GGF groups in which SCG members have made substantial contributions are as follows:

- **Certificate Authority Operations Working Group (CAOPs-WG):** The EDG experience with building the largest Grid PKI to date is vital input for the discussions and papers of this group. Policy documents from EDG CAs, for example, were used in the preparation of the GGF Certificate Policy Model document.
- **Authorization Frameworks and Mechanisms Working Group (AuthZ-WG):** This group is co-chaired by Andrew McNab, a member of the SCG. The draft paper from this group contains a description of the EDG Authorization mechanisms as an example of a current implementation.
- **Open Grid Service Architecture Authorization Working Group (OGSA-AUTHZ-WG):** This group is co-chaired by Andrew McNab.

SCG members also participate in discussions and review documents in these GGF groups:

- **Authority Recognition (ARRG-RG)**
- **Site Authentication, Authorization, and Accounting Requirements (SAAA-RG)**

David Kelsey, the leader of SCG, gave a plenary talk at GGF8 in Seattle (June 2003) entitled “EU DataGrid CAs” as the European contribution to a panel session on Global Authentication issues.

3. Inter-project collaborations

One of the major achievements of DataGrid has been the successful implementation of an Authentication system which can be used for applications running across several Grid projects. The DataGrid WP6 Certificate Authorities Group has built a large-scale Public Key Infrastructure in collaboration with the other projects. To achieve this, the group invited active participation of CA managers from several projects across many different countries. Some of the EU CrossGrid CAs were founder members of the group early in 2001, with the US DOE Grids CA joining soon afterwards. Further expansion of the group, driven mainly by the identity requirements of the global LCG project, a customer of EDG, has resulted in the approval of the remainder of the CrossGrid CAs together with Grid Canada and ASGC Taiwan, to a current total now of some 20 approved CAs. Many other national CAs have since joined the EDG CA group and are working towards approval. The current list of new CAs includes Hungary, Israel, Pakistan, Belgium, and Armenia. The experience gained building the DataGrid PKI and inter-Grid authentication with the other projects has been valuable input not only to the GGF as described above, but also to related work to establish worldwide trust via multiple Policy Management Authorities (PMA) under the auspices of the GridPMA body.

Authorization activities have also involved much inter-project collaboration. One of the major components of DataGrid Authorization, the Virtual Organisation Membership Service (VOMS), has been a joint development with the EU DataTAG project. VOMS is now under active use and/or evaluation by several Grid projects, including the US CMS VOX project and the LCG. In both of these cases, active collaboration has been involved.

4. Papers

R. Alfieri,, R. Cecchini, V. Ciaschini,, L. dell’Agnello,, A. Frohner, A. Gianoli, K. Lorentey, F. Spataro – “VOMS, an Authorization System for Virtual Organizations” – Presented at the 1st European Across Grids Conference, Santiago de Compostela, February, 2003

R. Alfieri, F. Bonnassieux, P. Broadfoot, R. Cecchini, V. Ciaschini, L. Cornwall, L. dell’Agnello, A. Frohner, A. Gianoli, D. Groep, J. Jensen, D. Kelsey, D. Kouril, G. Lowe, A. McNab, O. Mulmo, F.

*Spataro - “Managing dynamic user communities in a Grid of autonomous resources”
- Proceedings of the Conference on Computing in High Energy Physics 2003 (CHEP03), La Jolla,
California, March 2003.*

*A. McNab - “Grid-based access control and user management for Unix environments,
Filesystems, Web Sites and Virtual Organisations”- Proceedings of the Conference on Computing
in High Energy Physics 2003 (CHEP03), La Jolla, California, March 2003.*

*L. Cornwall, D Kelsey, A. McNab - “EU DataGrid and GridPP Authorization and Access
Control” - Proceedings of the UK e-Science All Hands Conference, Nottingham, September 2003*

•CONCLUSIONS

In this document we have drawn a report of all the contributions provided by the DataGrid project to the definition of international standards in the field of Grid computing. We have tried to provide a coherent organization of the information by categorizing the contributions according to the following schema:

- contributions to GGF activities;
- inter-project collaborations;
- publishing of papers.

We have started from a general overview of the project's contributions and then illustrated the contributions provided by each single work package. Further information can be found in the final reports of the technical work packages.

We have shown that significant contributions have been provided in all of the fields addressed by the project: workload management, data management, monitoring, fabric management, mass storage management, network management and security.

The work started within the DataGrid project will be continued within other projects, first of all EGEE. This project will integrate current national, regional and thematic Grid efforts to create a seamless European Grid infrastructure for the support of the European Research Area. This European Grid infrastructure will allow to test scalability, fault-tolerance and resilience of the DataGrid middleware. This will hopefully favour the adoption of DataGrid solutions as international standards.