

Next Generation Grid(s)
European Grid Research 2005 - 2010
Expert Group Report
Monday, 16th June 2003

Commission Disclaimer

This document contains information provided by a group of independent experts convened by the European Commission with the objective to identify potential European Research priorities for Next Generation Grid(s) 2005 – 2010. The document does not necessarily reflect the view of the European Commission.

Authors' Institutional Disclaimer

The views represented here are those of the individuals forming the group and do not necessarily represent those of the organizations to which the individuals belong.

Table of contents

Executive Summary	3
Rationale	4
Objectives and goals	4
Deliverables.....	4
Timing.....	4
Scoping	5
Grids.....	5
Next Generation Grid Properties	5
Grid Evolution.....	5
Research Agenda and Topics	6
Visions	6
End-User Vision	6
Architectural Vision.....	6
Software Vision	7
Research Priorities	8
Properties.....	9
Facilities	11
Models.....	12
Conclusion.....	14
Annex: Membership	15

Executive Summary

The expert group convened in two meetings at Brussels to provide an expert's view on the 5 to 7 year research priorities for the European context. This group's opinion was that current Grid implementations lack many essential capabilities, which would enable the vision of complete resource virtualization.

Three perspectives are defined to envision Next Generation Grid(s) (NGG): the end-user vision in which the simplicity of the Grid is exemplified; the architectural vision of a large-scale, self-organizing network; and the software vision of a programmable system. Each of these visions is described in detail to support the assignment of priorities to the research agenda.

In order to realize those visions, a number of research priorities have been identified. They fall into three categories, the first one considering the properties of a Next Generation Grid, the second one specifying the facilities provided by an NGG, and the third one comprising the models that are needed for orchestrating Grid services. For clarity, a number of research topics have explicitly been excluded as falling outside the scope of this paper.

Rationale

This document results from a set of workshops with an authoritative panel of experts convened by the EC. The purpose of the meetings and discussions was to identify potential research priorities for NGG. In order to ensure a comprehensive and consistent outcome experts were invited from across the spectrum of relevant fields.

Grid research is a key Strategic Objective of the IST (Information Society Technology) Programme, Priority 2 of Framework Programme 6 (FP6). A workshop held in Brussels on the 29th and 30th January 2003 on "Grids for Complex Problem Solving" attracted about 200 participants from academia, business and industry. The main objective of the workshop was to offer a forum to the constituency for consolidating and structuring emerging ideas in line with the set objectives of the IST/FP6 work programme 2003-2004. Stimulated by presentations of key players from all levels of the value chain, the participants discussed in parallel working groups the needs and requirements for Grid research from different perspectives.

The working group on the development of the **Next Generation Grid Architectures** suggested that a small group of visionary experts should be set up to identify the research issues and their potential priorities for the next 5-7 years (i.e. beyond the present work programme) and propose suitable means for implementation. Another recommendation, linked to the previous one, was to compile a comprehensive inventory of the research topics already addressed in ongoing Grid projects (EU, US, world-wide) to avoid duplication of efforts. The IST GridStart Project is undertaking this task.

The document can be downloaded from: <http://www.gridstart.org/GRIDSTART-IR-D2.2.1.2-V0.5.doc>

Objectives and goals

The present IST WP2003-2004 Work-programme specifies:

- The research objective:
 - *Overcome present architectural and design limitations hampering the use and wider deployment of computing and knowledge Grids and to enrich its capabilities by including new functionalities required for complex problem solving. This should help the larger uptake of Grid type architectures and extend the concept from computation Grids to knowledge Grids, eventually leading to a "semantic Grid".*
- With the following research focus:
 - *Architecture, design and development of the next generation Grid beyond extensions of existing technologies, leading to open standards, and including security built-in at all levels, programming environments, resource management; economic and business models for new services, customisable middleware, interoperability with existing Grid and Web services. An integrated and comprehensive approach including stakeholders from all relevant levels is required.*

The recent workshop clearly demonstrated that there is a need to identify more specific priorities in the EU research agenda - taking into account national activities - for the next 5-7 years, which will be addressed in upcoming calls. This is in order to assure critical mass and durable impact while avoiding fragmentation and duplication of efforts. With a view to achieving this, an expert group was set up (see membership list in the Annex).

The Expert Group's Terms of reference were to:

- Identify research priorities for the next 5-7 years and propose pragmatic steps to be taken;
- Propose a roadmap for the implementation of these steps (in view of upcoming calls for proposals);
- Align technology priorities and means of implementation with policy objectives, e.g. European Research Agenda, European Centres of Excellence, etc.;
- Network/liaise and discuss findings with the Grid research community;
- Propose actions to increase efficiency in international collaboration.

Deliverables

The Expert Group was to produce a report in line with the objectives defined above.

Timing

Two plenary face-to-face meetings took place in the offices of the European Commission in Brussels. The first took place on the 2nd-3rd April, the second on the 10th-11th June 2003. The final version of this report was presented to the Commission on 16th June 2003.

Scoping

Grids

A Grid provides an abstraction for resource sharing and collaboration across multiple administrative domains. The term resource covers a wide range of concepts including physical resources (computation, communication, storage), informational resources (databases, archives, instruments), individuals (people and the expertise they represent), capabilities (software packages, brokering and scheduling services) and frameworks for access and control of these resources (OGSA - Open Grid Services Architecture, The Semantic Web). At present multiple different Grid technologies co-exist, which stimulates creativity in the research community. Ultimately, however, we envision one Grid based on agreed interfaces and protocols just like the Web. Within that environment virtual organisations can co-exist, evolve and interact with each other in a secure way. This should avoid a proliferation of non-interoperable Grids, which would hamper the wide acceptance of Grid technology.

Next Generation Grid Properties

Next Generation Grid(s) should possess the following properties:

- Transparent and reliable
- Open to wide user and provider communities
- Pervasive and ubiquitous
- Secure and provide trust across multiple administrative domains
- Easy to use and to program
- Persistent
- Based on standards for software and protocols
- Person-centric
- Scalable
- Easy to configure and manage

The current Grid implementations clearly do not individually possess all of these properties. Future Grids not possessing these properties are unlikely to be of significant use and, therefore, inadequate from both research and commercial perspectives.

Grid Evolution

We observe several distinct phases in the evolution of Grids. The Grid started out as a means for resource sharing; initially supercomputers were shared by means of meta-computing. As an extension, the sharing of data in general was added and special devices like scientific instruments, telescopes and medical equipment were included. The marriage of the web technology with the first generation Grid technology led to new and generic Grid services.

The focus evolved to knowledge sharing to enable collaborations between different virtual organizations while respecting their individual security. The knowledge Grid facilitates data-mining across the Internet. This requires techniques for abstracting heterogeneous data, creating meta-data, publishing, discovering and describing data in the Grid (e.g., using meta-data descriptions).

An ongoing process in this evolution is a complete virtualization of Grid resources. Virtualization is a way to expose only high-level functionalities in a standard and interoperable way hiding the implementation details. Virtualization of the constituents of the Grid will enable an abstract and scalable exploitation of all kinds of resources.

Research Agenda and Topics

Visions

The following sections present the visions, as seen by the authors, from three perspectives:

We start out with the **end user perspective**. Only with an indication of how the Grid might be deployed in the everyday life of individuals and businesses can the importance of Grid technology be addressed.

The **architectural perspective** envisions the Grid as a structural entity with a collection of capabilities and properties. The perspective is critical in providing an indication of the scale expected from NGG, in term of numbers, geography and administrative domains.

The **software perspective**, finally, focuses on how programming on a Grid will look, which algorithmic problems have to be solved in order to build an infrastructure as described from the architectural perspective, and which constraints have to be observed when developing a NGG.

End-User Vision

To start with, obvious as it may sound, it is important to stress that for the end users, the Next Generation Grid must enable to solve their problems in a simple way, without them having to be aware of all the complexity hidden behind. The Grid can only be claimed a pervasive success when there is no longer any need to define it.

Persistent life-support for (business) data and processes:

Businesses and organisations are carrying out an ever-increasing proportion of their core functions and communications via digital media. NGG will need to support and enable those functions. It will be an operating environment where complex, high-value data can be manipulated, moved and stored for long periods (over 50 years for some industries and government institutions and perpetually for museums and some libraries). A similar requirement exists for knowledge resulting from business processes. There is a requirement for these to be supported and adapted in the face of changes in the underlying software and hardware technologies with minimal human intervention required. Eventually it should be possible to define business data and processes in high-level terms alone and have them persist independently of changes in hardware and the underlying service environment.

A Grid where people are nodes:

The integration of the person into the system in a seamless and comfortable way is paramount to obtain maximal benefit. People are commonly data providers, decision-makers, and facilitators and thus become a process within the workflow to achieve an objective in the Grid environment.

Affordable knowledge discovery service:

Affordable techniques aimed at exploiting the Grid are of great interest in the near future. An emerging and important issue should be how to discover and extract useful knowledge from huge amounts of data. Knowledge discovery services providing consistent, efficient and pervasive access to Grid resources are required.

Architectural Vision

The next generation Grid will virtualise the notion of distribution in computation, storage and communication over unlimited resources. A Grid will be a virtual, pervasive organisation with specific computational semantics. It performs a computation, solves a problem, or provides service to one single client or to millions of clients. Grids will pervade into everyday life, sometimes in the form of ambient intelligence. Grid may consist of millions of interconnected nodes.

A Grid node is an atomic unit forming an abstraction over resources, entailing what is hidden by the interfaces it provides.

Nodes may provide new services, functions, or even new concepts that are unknown to clients. The semantics of such services, functions or concepts are defined by semantic description languages or ontologies.

Nodes can organise, on the fly, into a group in order to provide functionality and behaviour that none of its individual members has. The new “sociology” for a specific service can be transient, persistent, or a combination of both. Any new organisation of nodes can be made available to every client of the Grid system. The self-organising capabilities of nodes aim at establishing higher robustness and lower costs for systems management. These capabilities are provided through a small, common set of facilities, such as highly scalable protocols for communication and membership management. A new ecology of computers, data entities and communication links will interconnect, interact, interoperate, interfunction syntactically and semantically in a societal way. The dynamics of a Grid (dynamic behaviour, dynamic configuration management...) will allow more freedom for new potential services.

A key aspect of Grids is the capability to negotiate with agents to provide a collection of services and facilities that satisfies the end-user requirement for a price within a certain time. This requires a common understanding of the services and facilities, which necessitate semantic interoperability and a mechanism of assurance to ensure delivery of service (trust, provenance, accountability, auditability, traceability).

The next generation Grid will have some mandatory architectural properties:

- Simplicity to allow for easy life cycle management and smooth evolution;
- Subsidiarity of control and management, and scalability of services;
- Resilience, through redundant, self-organising components to minimize points of failure;
- Transparency to allow many virtual organisations to run over it;
- Straightforward administration and trouble-free configuration management.

Software Vision

From a software perspective, i.e. from the point of view of what it means to build something that qualifies as an NGG, the crucial thing is to liberate one’s thinking from the current “web model” of the Grid. It is important to consider more general models of networking, higher levels of semantic model, people on one hand and very simple devices on the other hand as nodes in such networks, highly individual and dynamic security schemes, services with huge amounts of state, and so on.

The following list of important goals of research into NGGs from a software perspective is by no means intended to be complete. It rather is intended to give a flavour of what type of problems need to be addressed – but there are many more.

Interoperability as a basic means for problem solving:

In many situations, an end-user or a business actor will have a problem that leads to a requirement that needs to be translated into terms understandable in a Grid environment. However, there is no single Grid environment because the data sources, software sources, compute and storage resources, network resources and detection equipment resources are heterogeneous and distributed. Thus Grids can be seen as multiple overlapping n-dimensional constructs, which have to interoperate to provide solutions to dynamically changing requirements. The interoperation will be through metadata describing the sources and resources (and users) represented by agents and mediated through brokers, together with any constraints, policies, billing schemes, etc. that are of relevance.

Programming Grids through abstraction:

When the first computers were built, the programmers needed to have detailed knowledge of all machine instructions to program them. Then programming languages and operating systems were developed, which shielded the programmers and users more and more from the hardware details. Nowadays, a Grid infrastructure is fully exposed to the programmers, meaning that they have to mastermind the usage of different kinds of resources (or nodes), which are geographically dispersed.

Despite several research and engineering efforts in the design of first generation Grid middleware, the Grid infrastructure is still visible.

To harness the power of an NGG without being overwhelmed by its complexity, abstraction mechanisms must be provided. Those mechanisms will keep all the intricacies of resource allocation and scheduling, data movement, synchronisation, error handling, load balancing, etc. transparent to the user and developer. A Grid will possess abstraction layers, presented at the user interface by an agent controlled by metadata, itself interacting with other agents (and metadata) representing the other entities in other Grids via brokers.

Future generations of Grids should be programmed through generic and problem-specific abstractions, supported by an appropriate programming environment. In order to achieve the ambitious goal of making all the technical and structural aspects of a Grid transparent, one needs to study and adapt existing programming models to the Grid context – which may require the definition of new programming models, combining parallel and distributed programming practices in a coherent way.

A key strategic result of this will be the emergence of standards for metadata representation and agent - broker communication based on some messaging paradigm and with bridges to existing similar paradigms such as Web services or OGSA.

A global information model for common things:

Grid and Web services, as they are defined now, provide a way of publishing information by specifying an invocation model and standards for naming, authentication, authorization, etc. However, there are no standards for units, accuracy, precision (errors), terminology (measured quantities), representation, and so on. What is needed is a semantic object model (more than XML - eXtensible Markup Language) for methods and instance data that allows us to specify cross-domain terminology (metric systems) and schema building tools that allow us to (automatically) integrate local schemas from different domains into one global information schema¹. The challenge here is to come up with a small number of high level ontologies with wide applicability, say: one for the sciences, and one for business, rather than add to the already long list of narrow, domain-specific XML–extensions, because those, without additional mechanisms, do not facilitate inter-operability beyond their own domain.

Research Priorities

Based on the visions described above, a collection of research priorities emerges. In fact, the process is somewhat iterative as important research areas also inspire further visions of what is possible in Next Generation Grids. Only with research in the areas sketched below, at both a practical level and a deeper more theoretical level, can the visions above be realized. These areas were then characterised and grouped into categories: Properties, Facilities, Models; the reasoning behind this is that facilities (useful constructs for the system architect) exhibit and depend upon properties (characteristics necessary for the implementation of complex systems) to implement models (for the User, Business, Society...) which form part of the vision of the knowledge society.

In order to focus discussion on the fundamental issues of next generation Grids, the following research topics have been explicitly excluded from consideration:

- a) Safety critical systems;
- b) Mission critical systems;
- c) Grid applications;
- d) Problem solving environments;
- e) High performance computing;
- f) Parallelisation of software code.

This does not qualify these topics as less interesting or important. It just reflects the fact that they go beyond the scope of this paper – and, apart from that, some are covered elsewhere in the FP6 work programme.

¹ This is based on an excerpt from slides used by Jim Gray in his presentation at the Grid Open Issues Panel (Globus World, January 2003).

The Next Generation Grid must provide an infrastructure for business as well as academic research. Research is required on the constraints that business practices and international laws may place on Grid technologies, and on lessons learned from existing international and inter-organisation distributed systems, which might be relevant to the development of the Grid. Examples include constraints on security and access to data, data protection and the movement of data across national boundaries, the legal framework for the use of encryption, differing standards for audit and accountability between industries, and the issues in sharing resources between organisations with different legal and financial constraints.

Properties

Reliability:

As users come to rely on the Next Generation Grid(s) to support their everyday business, recreational and social needs, reliability becomes paramount. To paraphrase Leslie Lamport, "You know that you are dealing with a distributed system when you are prevented from getting your work done because a node you never heard of has crashed." The traditional approach to solving this problem has been the so-called "single system image", i.e. the attempt to hide all aspects of a distributed system from the user (application programmer, etc.) and make it look like a single system. In case of the Grid this may not be possible (latency or the effects of low bandwidth can't be "hidden", for example), and in some cases it may not even be useful. Nevertheless, the complexity of the Grid infrastructure must not translate into more complex failure modes from the perspective of the higher-level clients. Applications should be able to get back to a well-defined state in case something goes wrong that exceeds the specifications for fully automatic recovery, such as the loss of a resource for which no equivalent backup is available, long-lasting disconnection of parts of the network, etc.

Key research topics:

- a) New models for fault tolerance and dependability;
- b) Models for quality and accessibility of data sources, incl. versioning, recoverability, etc.;
- c) Execution and workflow models;
- d) Transactional and reconciliation models.

Security and trust across multiple administrative domains:

The basic concepts of security and trust are well understood and defined in terms familiar to Grid users and developers. However, these concepts need to be codified in an agreed (and eventually standardized) framework that would allow automatic or semi-automatic assessment of the levels of trust and security offered and accepted by Grid services of the Next Generation Grid. In some cases, solutions will require an assessment of the legal and legislative implications.

Key research topics:

- a) Exploration of languages and models to assist in reasoning about security and trust properties, rather than simply the implementation of static policies;
- b) Research that goes beyond technical mechanisms for the five security fundamentals (authentication, authorization, integrity, confidentiality, non-refutation).

Persistence:

Grids must support persistent behaviour in data, computing facilities, networks, detectors and their use. There is a need for research in local and personal persistence as well as global persistence. However, simple persistence is insufficient. As well as recording state and preserving it, it is also necessary to keep track of history and be able to recreate previous states. This is provenance management.

The goal of transparent, dynamic resource allocation that adapts to load patterns, availability, pricing schemes, etc. is one of the hallmarks of Grid computing. On the other hand, in many application areas there is the need to be able to reproduce exactly a computation. Research is needed into mechanisms that will help in resolving this apparent contradiction. Finding solutions will be critical for the success of Grid computing, because in a number of fields law requires strict reproducibility, and this will probably extend into other areas in the future.

Even where no requirement exists for exactly reproducible results, an audit trail that extends beyond basic accounting may be needed. For example, a complex distributed query may access many,

different data sources (data bases, experimental devices, etc). These sources may be transitory (experiments) or rapidly evolving (live data bases). Strategies are needed for capturing and retaining the state of these sources for comparison with later queries that may only be different in the time that they were performed, i.e. only the state of the sources has changed.

Key research topics:

- a) Local and personal persistence as well as global persistence;
- b) Strict reproducibility;
- c) Audit trails over transient state.

Scalability:

The Next Generation Grids will incorporate significantly larger numbers of 'nodes' (the end points, services, resources, users, applications, etc.) than current Grid systems. The complexities of maintaining consistent access, providing naming infrastructures, routing mechanisms, support for fail over, discovery services, the ability to distinguish many very similar entities, etc, need to be addressed. The scalability of these systems needs to grow to the significant numbers envisioned without losing the other properties required.

Key research topic:

Research into scalable solutions to naming, accessing, discovering, defining, and maintaining Grid entities.

Open to Wide User Communities:

Grids, to realise their potential, must reach as far as (and beyond) the current Web users. This can only be achieved by integrative design across many properties, integrated into facilities such that a model constructed for the end-user can operate effectively and efficiently.

Key research topic:

Architectural congruence (self similarity) from end-user to base hardware through all intermediate levels.

Pervasive and ubiquitous:

For Grids to be effective they must be seen as the anytime-anyhow-anywhere paradigm. Grid technology must penetrate into every aspect of society from business to learning, from culture to social appointments. This implies a significant intersection with ambient computing.

Key research topics:

- a) Appropriate communications protocols;
- b) Synchronisation of information;
- c) Dealing with interrupted connections.

Transparent, easy to use and program:

Users have a low effort threshold to utilise a technology. Much of the success of Web may be attributed to its low threshold to get started. For both end-users (using facilities provided) and programmers (building those facilities) Grids must be easy to handle and transparent, so the system complexity is invisible to both the end-user and developer.

Key research topics:

- a) For the end-user: intelligent user interfaces;
- b) Representative intelligent agents and brokers;
- c) For the programmer both expressive, representative languages and labour-saving systems development environments.

Person-centric:

The "Personal Grid" concept requires that the interaction between a person and the Grid is based on the person's knowledge (probably stored for that person in metadata) of Grid, and the Grid's knowledge of the person are used to best effect to achieve what is required. This concept widely intersects with previous properties.

Key research topics:

- a) Representation of personal preferences;
- b) Transactions to maintain this view;
- c) Protection of personal information.

Based on standards for software and protocols:

Standards emerge from R&D activities. All actions in the workprogramme are expected to conform to and develop further proposals for standardisation and take them through the appropriate standardisation processes.

Facilities**Information representation of elements of the Grid:**

The Next Generation Grids will provide and use data from a wide variety of sources and technical domains. Research is required into techniques to allow interoperability between different data sources. Schema conversion technologies via common meta-models and ontologies are required to allow data to be moved between storage resources and shared between tools with different data format requirements. The construction of Grid resources to automate this process and build on-line libraries of conversion and mapping tools will be of use in a variety of scientific, social and business domains. Research into semantic mapping using ontologies and automated inference should also be applied to Grid technologies where appropriate, to enrich and annotate the data made available through the data Grid and to describe the Grid itself.

Key research topic:

Definition of a semantically rich meta-data model that supports the integration of data from heterogeneous sources and their conversion into the user's target format without loss of data quality, accuracy, etc. The descriptive dimensions should include the standard ones like format, but must also include units, accuracy, conversion rules, validity ranges, prices, access policies, constraints of all sorts, etc. Mappings into the standard data and access models must be provided.

Co-ordination and orchestration of Grid elements:

A Next Generation Grid will have to handle a large number of heterogeneous resources (computing, storage, networks, sensors, services, applications,...). One of the key challenges in the use of such a computing infrastructure will be the co-ordination and orchestration of resources to solve a particular problem or perform a business process. The current generation of Grid architecture relies heavily on the program designer or the user to express their requirements in terms of resource usage. Such requirements are usually expressed in a hardwired way in a program using low-level primitives. A Next Generation Grid will handle resources in a more dynamic way making a static description of resource usage impossible. Grid applications will require the co-ordination and orchestration of Grid elements in a dynamic way during its execution.

Key research topic:

Defining dynamic and reactive co-ordination and orchestration mechanisms or paradigms using languages (with sufficient power of expression to denote and manipulate real world objects easily), models (object, component, service) or frameworks (spanning input/update, query, transformation, data analysis, data mining, visualisation, deduction, induction).

Systems Management:

Systems management for Next Generation Grids encompasses several important issues, including maintenance and deployment. Scalability is crucial; solutions should be able to support Grids consisting of billions of nodes. This requirement immediately stresses the importance of finding an appropriate trade-off between systems control as exerted by users and administrators, and the automatic control that is realised by the system itself.

For fault-tolerance, the systems management and control must be completely autonomous. This includes redundant resources (computation, data base, network), and also an autonomous self-regulatory model that ensures the proper working of the systems management itself.

How to establish automatic control in such large-scale distributed systems is seen as an important research area within systems management. Such control encompasses mechanisms for automatically configuring and organising (groups of) nodes for providing services, for automatically detecting and recovering from failures, maintaining sufficient quality of service and performance and for automatically adapting the system to changing external conditions.

In other words, NGGs should be largely closed systems with internal feedback control loops that can scale to vast numbers of nodes. Only when strictly necessary, users should be able to exert control on parts of the system.

Key research topics:

- a) The balance between user-based system control and automatic system control;
- b) System management tools to instrument, manage and reconfigure Grids.

Virtual organizations:

The next generation Grid will allow the dynamic creation of virtual organisations, either as a goal in itself or a prerequisite to support some wider activity. Research is required into techniques for creating such organisations, defining membership, allocating resources, and managing the termination of such resources. Examples could include the creation of a project between several organisations with shared databases, compute resources, and mailing lists: the Next Generation Grid should provide a standard technology for this so that any individual or group with access to a networked computer could become a member, with the initial barrier to entry as low as possible in terms of cost and availability of access software. The target would be the same level (or better) of accessibility as the Web.

Research is needed in the legal aspects and business practice aspects and their mapping into an electronic form. A key element is the trust model associated with each component of a virtual organisation. This trust model must be supported in its description, acquisition, distribution, maintenance, and dissolution.

Key research topics:

- a) Techniques for creating such organisations, defining membership, allocating resources, and managing the termination of such resources;
- b) Legal aspects and business practice aspects and their mapping into an electronic form;
- c) Trust models associated with each component of a virtual organisation.

Models

Business models:

As long as a service provider retains total control of all aspects of the service provided (systems, licenses, personnel, etc.), the basic ASP (Application Service Provider) model works effectively. However, when the complexity of Grid-based services extends beyond a single provider, this simple approach breaks down. This gives rise to a number of new research problems.

Key research topics:

- a) Techniques, models and languages to develop, deploy, and evaluate alternative business models for exploiting the Next Generation Grid;
- b) Issues such as accountability, Quality of Service (QoS), micro payments, performance guarantees, pricing, and dynamic negotiation of Service Level Agreements (SLA) need to be addressed.

Grid Economies:

Early visions of the Grid stemmed from a metaphor based on the electrical power grid. Whereas the ubiquitous nature of electrical power is still a significant part of this vision, the reality is that Grid services (information, compute, facilities access, service aggregation, etc.) lack any of the uniformity inherent in the power grid metaphor. The only way to achieve ubiquity, in such a heterogeneous environment, is to develop an economy for the Grid. Grid researchers have shied away from a simple approach based on pricing services in terms of "hard cash" for a variety of reasons, but partly because local service provision has been based on a "capital" cost model rather than an "operating" cost model.

Key research topic:

Strategies that, while supporting the accounting and tracking required by a "hard cash" model, go beyond simple monetary approaches. The scope of such research should include free (and controlled) market based models, mechanisms for accounting and pricing, and service exchange brokers, as well as the legal and socio-political implications of a global market in Grid services.

User Interface:

The complexity of Grid facilities facing an end-user requires that the user interface, if it is to be effective, to be intelligently assisted. In essence, the complexity of the services offered must not translate into a complex interface.

The intersection of user interface and virtual organisations covers cooperative working, from simple chat protocols through to formal workflow processes.

In addition, special attention needs to be focused on user modelling and profiling. It is only through improved knowledge about the user (preferences, physical location, temporal properties, relationships, etc) that user interfaces can approach the level of sophistication needed to expose the complex services possible in the next generation Grid. Those data, of course, need to be thoroughly protected.

As the quality of information available to a service increases, so does the need to provide the user with auditable assurances that their information is available only as they have directed. Research into legal issues is clearly required, as are civil rights issues (such as implied by a user's unwitting release of more information than they realise).

Key research topics:

- a) Architectures and standards for metadata and agents representing the user;
- b) The legal rights of an end-user and civil rights issues;
- c) Architectures and models for cooperative working;
- d) Agents and brokers and the metadata to control and parameterise them;
- e) Interaction models for e-learning.

Conclusion

We were asked by the European Commission to address the research issues in Next Generation Grid Architectures for the next five to seven years. We have presented a collection of visions that cannot be realized in full based on existing or near term Grid technologies (although many of the properties and capabilities required exist in isolation or small collections). We have identified areas of research that need to be addressed in order to realize these visions.

The Grids that evolve along these lines will need to exhibit all the properties, most importantly, interoperability. In this way a collection of Grids, each providing specialized facilities, will interoperate to form a single Grid at the interface level.

We believe that addressing these research priorities (with the representative balance between the pure technical and the social/business oriented issues) in the 2005 to 2010 timeframe will enable Europe to take the global lead in Grid technology and its application in the Information Society of the future.

Annex: Membership

Experts with the collective expertise to cover all the relevant topics have been selected to form the expert group. A strong commitment is expected from each member in one or more of the main research topics. The group selected two co-chairs from among its members (marked with a *), and the Commission provided secretarial support.

Expert	Country	Affiliation
H. Bal	NL	Dept. of Computer Science, Vrije Universiteit, Amsterdam
C. de Laat	NL	Faculty of Science, Informatics, University of Amsterdam
S. Haridi	SE	SICS
K. Jeffery	UK	RAL
J. Labarta	ES	Universitat Politecnica Catalunya
D. Laforenza	IT	ISTI-CNR
P. Maccallum	UK	Life Science Informatics Solutions Ltd.
J. Massó	ES	GRIDSYSTEMS
L. Matyska	CZ	Faculty of Informatics, Masaryk University
T. Priol *	FR	IRISA
A. Reinefeld	DE	Konrad-Zuse-Zentrum für Informationstechnik
A. Reuter	DE	European Media Lab GmbH
M. Riguiedel	FR	ENST
D. Snelling *	UK	Fujitsu Laboratories of Europe
M. van Steen	NL	Dept. of Computer Science, Vrije Universiteit, Amsterdam

Commission officer in charge: Roman Tirler,
Information Society-DG, Grids for Complex Problem Solving