Project no. 034567

Grid4All

Specific Targeted Research Project (STREP)

Thematic Priority 2: Information Society Technologies

D4.1: Specification of scenarios, user requirements, and infrastructure requirements

Due date of deliverable: 1st June 2007

Actual submission date: 28th June 2007

Start date of project: 1 June 2006 Duration: 30 months

Organisation name of lead contractor for this deliverable: UPC

Contributors: Gabriel Belvedere, Antares; Sergio Borgoñoz, Antares; Miguel Bote, UPC; Alícia Bou, Antares; Charalambos Constantinou, UPRC; Jorge García, UPC; Eduardo Gómez, UPC; Joan Manuel Marquès, UPC; Leandro Navarro, UPC; Symeon Retalis, UPRC; Xavier Vilajosana, UPC

Revision R.1

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Table of Contents

Table of figures **3**
Grid4All list of participants **4**
1 Executive Summary **5**
2 Introduction **6**
3 Use scenarios **8**
3.1 Collaborative learning 8
3.1.1 State of the art 9
3.1.2 Scenario: Using Asynchronous Multimedia Collaborative Tools and Grid resources to support Collaborative Learning. 10
3.1.3 Scenario: Collaborative simulation of computer networks 12
3.1.4 Scenario: Distributed Collaborative Software Project Development. 14
3.1.5 Scenario: Live tutorized online sessions 17
3.2 Domestic users 18
3.2.1 State of the art 18
3.2.2 Scenario: Multimedia and entertainment spaces 19
3.2.3 Scenario: Using grid resources and digital home health services to fight various diseases 21
3.3 Small and Medium Enterprises 24
3.3.1 State of the art 24
3.3.2 Scenario: Live sessions using collaborative tools. 24
3.3.3 Scenario: Collaborative software development 26
3.3.4 Scenario: "Earthquake prediction" 28
3.4 For-profit and non-profit communities 29
3.4.1 Scenario: Communities 30
3.4.2 Scenario: Emergency handling 31
3.4.3 Scenario: "Shared virtual libraries" 33
4 User requirements **35**
4.1 Collaborative learning 35
4.1.1 Scenario: Using Asynchronous Multimedia Collaborative Tools and Grid resources to support Collaborative Learning 35
4.1.2 Scenario: Collaborative simulation of computer networks 35
4.1.3 Scenario: Distributed Collaborative Software Project Development. 35
4.1.4 Scenario: Live tutorized online sessions 36
4.2 Domestic users 36
4.2.1 Scenario: Multimedia and entertainment spaces 36
4.2.2 Scenario: Using grid resources and digital home health services to fight various diseases 37
4.3 Small and Medium Enterprises 37
4.3.1 Scenario: Live sessions using collaborative tools. 37
4.3.2 Scenario: Collaborative software development 38
4.3.3 Scenario: Earthquake prediction 38
4.4 For-profit and non-profit communities 39
4.4.1 Scenario: Emergency handling 39
4.4.2 Scenario: Shared virtual libraries 40
4.5 Summary and discussion 40
5 Infrastructure requirements **42**
5.1 Collaborative learning 42

5.1.1 Scenario: Using Asynchronous Multimedia Collaborative Tools and Grid resources to support Collaborative Learning 42
5.1.2 Scenario: Collaborative simulation of computer networks 42
5.1.3 Scenario: Distributed Collaborative Software Project Development 43
5.1.4 Scenario: Live tutorized online sessions 43
5.2 Domestic users 44
5.2.1 Scenario: Multimedia and entertainment spaces 44
5.2.2 Scenario: Using grid resources and digital home health services to fight various diseases 44
5.3 Small and Medium Enterprises 45
5.3.1 Scenario: Live sessions using collaborative tools 45
5.3.2 Scenario: Collaborative software development 45
5.3.3 Scenario: Earthquake prediction 45
5.4 For-profit and non-profit communities 46
5.4.1 Scenario: Communities 46
5.4.2 Scenario: Emergency handling 46
5.4.3 Scenario: Shared virtual libraries 47
5.5 Summary and discussion 48
6 Conclusions 49
7 References 50

Table of figures

Figure 1: Use case diagram for both "Learner" and "Educator" 12
Figure 2: Use cases diagram for "Educator" and "Learner". 14
Figure 3: Use case diagram for "Chief Instructor". 15
Figure 4: Use cases diagram for "Student". 15
Figure 5: Use cases diagram for "Subsystem Responsible", "Project Coordinator" and "Supervisor". 16

Grid4All list of participants

Role	Participant N°	Participant name	Participant short name	Country
CO	1	France Telecom	FT	FR
CR	2	Institut National de Recherche en Informatique en Automatique	INRIA	FR
CR	3	The Royal Institute of technology	KTH	SWE
CR	4	Swedish Institute of Computer Science	SICS	SWE
CR	5	Institute of Communication and Computer Systems	ICCS	GR
CR	6	University of Piraeus Research Center	UPRC	GR
CR	7	Universitat Politècnica de Catalunya	UPC	ES
CR	8	Antares S.L.	Antares	ES

1 Executive Summary

This documents reports work done in task 4.1 regarding the definition of realistic scenarios where Grid4All applies with advantages to other models. The user requirements are elicited based on existing studies, applications and interviews with practitioners. Based on the scenarios and user requirements, the operational and functional requirements for the infrastructure (infrastructure requirements) to support these applications and their usages are specified.

The cases evaluated are: collaborative learning, domestic users, small enterprises, communities with a total of 12 different scenarios presented.

2 Introduction

This deliverable occurs in the context of task 4.1:

T4.1: Usage scenarios, collection of user requirements and translation into infrastructure requirements (M0-M5; UPRC, UPC, FT): Define realistic scenarios where Grid4All applies with advantages to other models; elicit user requirements based on existing studies, applications and interviews with practitioners. Based on the scenarios and user requirements, specify operational and functional requirements for the infrastructure to support these applications and their usages. These scenarios, user and infrastructure requirements, the rationale, and the relations between them will be presented and justified in D4.1. This task will iteratively provide the inputs to WP1-3 thus allowing concurrent progress.

The following scenarios were mentioned in the description of work:

1. **Collaborative learning**: A virtual organization formed by students from different schools. This can be applicable to students at any educational level (primary or secondary school and university level), to students from a single school working from home or at the school, to international collaboration among different schools. An increasing number of activities in schools require resources beyond the typical environment of a single student (with his textbooks, paper, pencil and computer). Students usually need to consult external experts, perform experiments using external resources, launch computer intensive tasks such as simulations, communicate and share information with other students in groups, move to different locations (home, school, library, etc) while they continue working on the same task (same documents). These activities will benefit from an infrastructure such as Grid4All's provides, and particularly, group applications to negotiate and create shared knowledge (e.g. a shared whiteboard) or a shared source of information and group awareness to coordinate groups (group repository with documents, annotations and events denoting changes or awareness information and a shared calendar to articulate work and meetings).
2. **Domestic multimedia processing**: (video conversion, image processing, photos storage, etc.) based on the "on demand" paradigm. People may want to use Grid4All to have at their disposal resources with capabilities beyond their individual PC to share and use external resources to store, share, annotate, transform media content that otherwise (1) would risk to be lost on a failure (destruction or disappearance) of his own PC or the underlying storage media where the files are stored, (2) may require very long delay for processing, (3) may require services or applications not installed or licensed locally, (4) may desire to share with some other friends or colleagues, or (5) may just desire to be able to access as he moves to different locations with different computers without having to carry a copy.
3. **Small enterprises and supply chains**: these organizations could borrow or share the resources of friends, collaborators or clients to provide services (e.g. a digital photo processing service using storage and processing resources of the local community; a local digital TV, video-club or radio station providing service, audio or video content, based on the resources contributed by the community and other resource providers)
4. **Communities of for-profit or non-profit organizations**: the membership of a network of similar enterprises could share resources to help each other during periods where some members have a peak of demand. For instance, a network of content distribution providers may cooperate to sometimes use resources from other members to overcome situations of excessive demand. It can even be possible to have "virtual members" that lack any resource at all and it relies completely on resources from the network.

The above mentioned scenarios are detailed to produce different use-case scenarios to illustrate better and capture the practical requirements from the user point of view, taking into account:

- The targeted user population
- The environment and context within which the middleware will be deployed and hence the functional and non-functional requirements imposed by these.

Based on the detailed use-case scenarios, previous research from Computer-Supported Collaborative work (CSCW) and Computer-Supported Cooperative Learning (CSCL) fields, analysis of existing and current practices in Grid platforms and infrastructure, analysis of existing applications used in these scenarios and interviews with practitioners in the relevant communities, user requirements are defined.

These user requirements are translated into a list of key/core requirements on the Grid4All architecture and the Grid4All modules. This translation takes into account:

- Operational and functional requirements for Grid4All infrastructure
- How existing Grid and peer-to-peer functionality can be redesigned and implemented for Grid4All environment (features and requirements evolve over time and hence the infrastructure should permit addition of new features/services as needed)
- Selection of key applications pertinent to the targeted user population. These applications may need to be either adjusted or redesigned and implemented according to operational and functional requirements.

The subset of applications to be developed and the features to incorporate will be selected during the project as a result of the analysis of user requirements and their implications on the context of the scenarios.

3 Use scenarios

This section is a collection of user cases studied and described in detail to later derive requirements.

Each use case is slightly different from each other is terms of level of detail and style of writing since each case is contributed by a different organization.

This collection of use cases helps to identify the scope of important services required, identify a core set of services, and specify at a high level the functionalities required for these services and interrelations.

Each generic use case (collaborative learning, domestic users, small enterprises, communities) is structured to raise requirements both from the state of the art research and from a collection of concrete scenarios. Hence the structure of a use case is as follows. Each use case starts with an introductory section, a state of the art and a collection of scenarios.

Each scenario starts with an introductory section, and then it describes the roles involved and their needs. It follows with a description of the organizations involved, the use cases of how roles interact with the system. The description of the scenario concludes with the user requirements and the infrastructure requirements.

The list of scenarios, as agreed in Barcelona meeting (October 2006):

Collaborative learning: <UPC>

- Using Asynchronous Multimedia Collaborative Tools and Grid resources to support Collaborative Learning <UPRC>
- Collaborative simulation of computer networks <UPC>
- Distributed Collaborative Software Project Development <UPC>
- Live tutorized online sessions <Antares>

The grid for small and medium enterprises: <UPRC>

- Live sessions using collaborative tools <Antares>
- Collaborative software development <UPC>
- Earthquake prediction <UPC>

Domestic users <UPRC>

- Multimedia and entertainment spaces <UPC>
- Digital home health <UPRC>

The grid for for-profit and non-profit communities: <UPC>

- Communities <UPC>
- Emergency handling <UPC>
- Shared virtual libraries <UPC>

3.1 Collaborative learning

Education is considered to be one of the most important applications of grid computing in the near future [12][30][6]. In this way, some efforts have already been made in order to identify the potential uses of grid infrastructures within the context of education in general and collaborative learning in particular. This section first provides an overview of the grid-supported collaborative learning scenarios that can be found in the literature. Then, a number of new scenarios that could be supported by Grid4All infrastructure are introduced.

3.1.1 State of the art

Grid infrastructures have been traditionally employed in order to enable the sharing of computational resources. Given their typically large scale, computational grids allow applications to access to a large amount of computational resources that can be used in an aggregated way in order to tackle tasks requiring supercomputing capabilities. Interestingly, some of the collaborative scenarios conceived to be supported by grid infrastructures that can be found in the literature focus on the use of supercomputing capabilities to carry out data analysis or simulations whose results are then employed in collaborative learning activities.

An example of this kind of scenarios is presented in [3]. This paper introduces Quarknet, a grid-based system that allows to applying the project-based learning methodology in a collaborative way to investigate cosmic rays. With this aim, Quarknet includes a network of cosmic ray detectors in high schools across USA. The detectors capture the time and location of the arrival of cosmic rays. Supercomputing capabilities can then be employed to analyze the data in order to study issues such as particle decays.

Another interesting scenario is introduced in [23]. In this case, the authors propose to employ the supercomputing capabilities of grid infrastructures in order to make complex simulations of hemodynamics. The results of such simulations could then be collaboratively visualized and discussed within the context of a course on Fluid Mechanics.

Computational grids are also typically used to access specific hardware resources (i.e. electronic devices or computers with specific features) due to the fact that grid technology enables transparent access and seamless integration such type of resources. This possibility is also very useful within the context of collaborative learning in order to enable the realization of experiments with specific hardware resources that may not be available locally because of different reasons such as economic constraints (e.g. a given resource may be too expensive), safety issues (e.g. direct use of a resource may be dangerous for learners) or logistic limitations (e.g. there is not enough room for a resource).

The scenario described in [7] is a good example of the potential use of grid infrastructure to access specific hardware resources in a collaborative learning situation designed for a course on Computer Architecture. In this scenario, students collaborate to decide the best computing solution for a fictitious customer whose needs are provided by the educator. This situation involves a number of tasks such as benchmarking a set of real computers with different characteristics (CPU, memory, etc.) that can be accessed through a computational grid.

Another example of this type of scenarios can be found in [1]. In this case, the grid allows students to remotely access to expensive electronic devices in order to perform highly realistic experiments. Students can then compare and discuss the results of their experiments with other partners.

Computational grids can also be employed to enable and enhance different types of human-to-human interactions in a synchronous or asynchronous way via a shared space. The infrastructures deployed to be used with this purpose may include even high-end audio and visual technology needed to provide a rich user experience. This is clearly another application of grid computing that can be useful in the context of many collaborative learning scenarios in which interactions among participants take place in a virtual space.

Virtual Harlem [22] is a grid-based system that was designed to support a collaborative learning scenario of this kind. Students can use this system to make a virtual tour in a reconstruction of Harlem, New York, during the 1920s. These tours may include different activities such as visiting significant places or listening to the speech of political figures of that moment. Furthermore, students can create virtual annotations that can be retrieved by their partners.

VMS (Virtual Medical School) [28] is another collaborative learning system that uses grid technology to support the interaction among participants. In this sense, VMS provides a virtual class environment in which students can discuss clinical cases among them and also with their instructors. Interestingly, this environment uses a grid-based multipoint videoconferencing system to enable real time interactions between participants.

Another typical use of grid infrastructures is data storage. Again, the large scale of computational grids allows applications to access to the resources required in order to store large amounts of data. In the context of collaborative learning, this may be very useful in the case scenarios involving experiments or simulations that require or generate large amounts of data. Besides, it can be helpful to store different of types learning objects that can be employed to support collaborative learning.

The use of mass storage facilities provided by grid infrastructures in a collaborative learning scenario can be illustrated with the aforementioned Quarknet system. Indeed, Quarknet system uses such facilities in order to store the data gathered by cosmic ray detectors.

Another interesting example can be found in [16]. In this case, the proposed system enables the possibility of storing large amounts of learning contents in a data grid. Such contents can be retrieved and used in a virtual classroom in which students and educators can interact. This virtual environment is also supported by a grid infrastructure.

3.1.2 Scenario: Using Asynchronous Multimedia Collaborative Tools and Grid resources to support Collaborative Learning.

3.1.2.1 Introduction

The last few years have seen a rapid maturation of basic streaming media technologies. Several efforts have been made in order to avoid passive attendance. Interacting and collaborating around digitized multimedia can be facilitated by technologies that allow people to easily create and share annotations that are anchored to specific video segments.

According to cue summative learning method [26], the use of various channels of communication and stimuli helps in the retaining of knowledge and in the acquiring of new knowledge and skills. Digital video and audiovisual media in general can be a great resource for education thus the last decades we have witnessed a lot of research and various attempts to utilize it in classrooms and generally in education. Bowie [8] in an analytical review of studies that concern the educational use of films argued that audiovisual material:

- Is effective in discovery learning.
- Can be used for demonstrating the solution of a problem.
- Is appropriate for developing skills of attention and observation of details.
- Can influence positively learner’s self efficacy.
- Improves creativity, imagination and aesthetics.

Yet beneath the apparently unproblematic appeal of audiovisual media, there is a counterargument which states that these media (and video in particular) is a passive educational mean which creates passive learners. The conclusion of a relevant research was that audiovisual material wasn’t successful in classrooms [13], except of foreign languages.

According to [14] the ability to combine digital video seamlessly with other interactive elements such as communication and assessment tools, offers an opportunity to move our concept from video as a purely passive presentational tool to video as a focus for student activity and communication. Our intention is to use Asynchronous Multimedia Collaborative (AMC) and Media Sharing tools, in order to combine the educational use of digital video with web-based asynchronous collaborative learning activities so as to take advantage of their added value, restraining passiveness and inaction.

These tools can benefit a lot from the grid. To begin with, many companies could offer storage and other useful services concerning video streaming and annotations as also and computational resources for video processing. The middleware of the grid could also take care of several other issues such as user authentication and authorization, security and file access rights. Finally other companies or organizations may provide educational content such as digital videos and other multimedia content in order to be used by the AMC and Media Sharing tools.

For the better understanding of the usage of AMC and Media Sharing tools to support asynchronous collaborative learning consider the following scenario: A company wants to train AutoCAD to a team of new civil engineers and they will use AMC and Media Sharing tools in order to support this training. In this scenario the chief civil engineer plays the role of the educator and the civil engineers are the learners. They will use the AMC tools to communicate, discuss, and analyze and annotate a video tutorial probably provided by a company specialized in creating professional video tutorials and Media sharing tools in order to share other multimedia content .

3.1.2.2 Roles

In this scenario the main users are the educator and the learner (the civil engineer). They use AMC and Media Sharing tools and a set of services provided by the grid to analyze, annotate and comment on streaming video clips. They benefit from the grid since they can have access to shared video clips stored in a streaming server and use important services provided by various service providers, such as video processing, user authentication and authorization service.

- **Educator**. The educator is a civil engineer that plays the role of the administrator of the AMC tools. He is responsible to search for a streaming server to upload the video tutorial, create tasks, give instructions and generally organize the whole event.
- **Learner**. Civil engineers play the role of learner. They are responsible to use the AMC and Media Sharing tools and carry out the instructions of the educator in order to complete the experiment and reach into conclusions.
- **Service and content providers**. Various companies, organizations, research institutions, universities that provide content and services to the AMC and Media Sharing tools.

3.1.2.3 Organizations providing services

The following organizations may be interested in providing the services required to support this scenario:

- **Content Providers**. Companies, organizations, research institutions or Universities can provide educational videos and other multimedia content. Private companies may charge for the content they provide.
- **Storage Providers**. Universities or other research institutions may provide storage as a service, possibly with enhanced software features, searching and commenting. Private companies can also provide storage and charge for its usage.
- **Video Processing Providers**. Companies will offer video processing services such as encoding and decoding various video formats and charge per operation.

3.1.2.4 Use cases

In this scenario we demonstrate the use of an AMC and a Media Sharing tool, digital video and other multimedia content in order to train the basic features of AutoCAD to civil engineers. For the needs of the experiment they were organized in three groups of four. Firstly the educator used a GUI (Web Interface) in order to search for a Streaming server responsible to store and stream educational video clips and a File Server responsible to store other multimedia content such as text documents, images etc. This was possible through a search engine for grid services. He has also used an upload service to upload a pre-recorder video clip containing a tutorial for AutoCAD's basic features and a text document (containing other important information related to AutoCAD functionality) on the Streaming Server and the File Server respectively. He has also requested and used a video processing service in order to edit and encode the video clip to various video formats.

In the next step he used the AMC tool to login as “Educator” through an authentication service in order to collaborate with the learners and assign three tasks (one for each group).He has also used a Media Sharing tool in order to share files with important information for the three tasks. Each group used the AMC tool to log in as “Learner” enabling them to have certain File Access Rights. They have also requested the list of available video clips in order to select and watch the uploaded video clip. In order to get more information related to AutoCAD's features, the practitioners used the text documentation as also and other Multimedia content uploaded and stored by the educator in the file server. The practitioners used the AMC tool in order to communicate and collaborate and the Media Sharing tool to exchange various files so as to complete their tasks.

In the last step the small groups of learners submitted a text file containing a report of their conclusions and comments about the procedure they followed to complete their tasks. At the end the educator and the civil engineers discussed the report of each group and in addition they suggested other possible solutions. The educator has also posed questions concerning the AMC and Media Sharing tool and asked them for suggestions and comments.

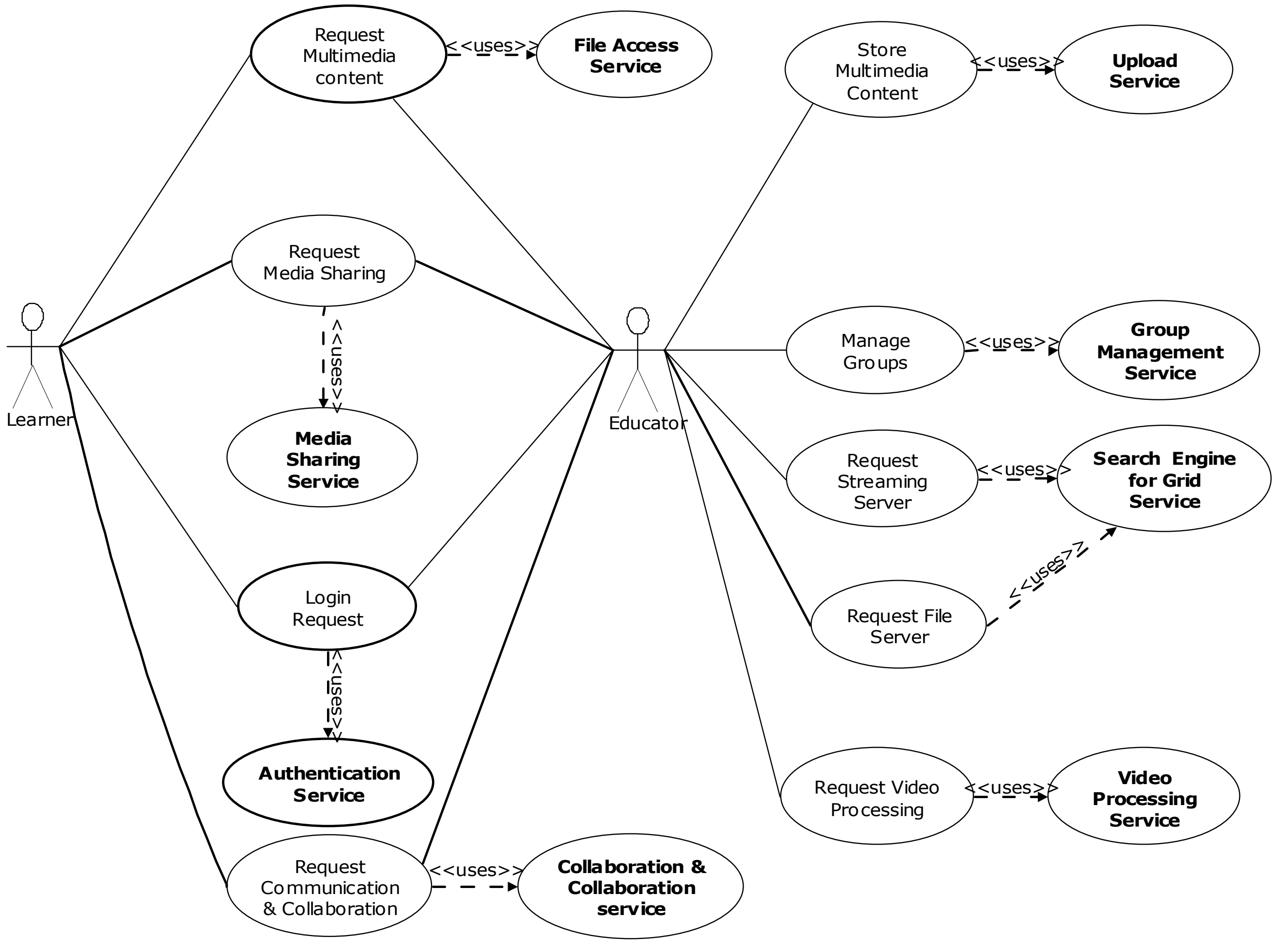


Figure 1: Use case diagram for both "Learner" and "Educator"

3.1.3 Scenario: Collaborative simulation of computer networks

3.1.3.1 Introduction

Simulators are nowadays usually employed with educational purposes in many Computer Network courses. Nevertheless, the actual use of simulators within the context of educational institutions is sometimes hindered by the following two problems. First, the computer resources available are sometimes not enough to carry out in a reasonable period of time the many simulations that can be launched at the same time by the students attending a laboratory session. This fact limits the complexity of the simulations that can be carried out and, as a consequence, the possibility of illustrating many network scenarios that could be of interest for students. Besides, the storage resources available are sometimes scarce. However, the amount of data generated by some simulations may be very large. Again, this implies that this type of simulations cannot be carried in order to foster the reflections of students on many network scenarios. Within this context, next subsections introduce a collaborative learning scenario in which students can launch simulations not limited in terms of the computer power required to carry out them or the storage resources needed to store the data generated thanks to the support provided by the computational grid.

3.1.3.2 Roles

The following roles can be identified in the realization of this scenario:

- **Learners.** Learners participate in the collaborative learning situation with the aim of obtaining new knowledge and skills in relation with computer networks.
- **Educator.** The educator is responsible for providing the learning context used by learners in the situation (e.g. questionnaires, simulation files, etc.) as well as for answering the questions posed by learners during the realization of the situation.

3.1.3.3 Organizations providing services

The following organizations may be interested in providing the services required to support this scenario:

- The **local educational institution** in which the participants of the scenario are enrolled.
- **Other educational institutions** that may be willing to share their resources in exchange for the right to access to the resources owned by the local institution.
- **Companies** may also offer suitable resources for the support of the scenario in exchange for an economic compensation.

3.1.3.4 Use cases

This scenario includes the following sequence of activities:

1. The educator provides learners with a document containing a number of questions concerning some protocols that can be answered after carefully observing the behaviour of such protocols in different network scenarios.
2. Individually, students devise the network scenarios that could be simulated in order to be able to answer the questionnaire.
3. In pairs, students discuss the scenarios that they have devised individually and agree on a set of common scenarios that could be simulated in order to answer the questionnaire.
4. In groups of four, students discuss on the scenarios agreed in pairs and agree on a new set of common scenarios that will be simulated in order to answer the questionnaire.
5. In groups of four, students agree on a distribution of the network scenarios to be simulated so that each student is responsible for creating the corresponding simulation scripts.
6. Individually, students create the simulation script that they have been assigned. Once finished, the students upload the scripts to a shared repository.
7. In groups of four, students download the simulation scripts created by their partners. The correctness of all the scripts is then discussed. In case it is necessary, the scripts may be modified by any of the students. The modified scripts are also uploaded to the shared repository.
8. In groups of four, students agree on a script to be simulated. The simulation can be launched by any of the students. A trace file generated by the simulator is stored in the shared repository.
9. In groups of four, students use the trace file to collaborative visualize an animation of the simulation and discuss it in regards of the questions posed by the educator. The visualization can be started by any student. During visualization, any student can also stop, rewind, advance or resume the animation. Besides, students can make annotations in the animation that can be seen by their partners.
10. In groups of four, steps 8 and 9 are repeated until all simulations have been visualized and discussed by the students.
11. In groups of four, students agree on the answers to questions posed by the educator.

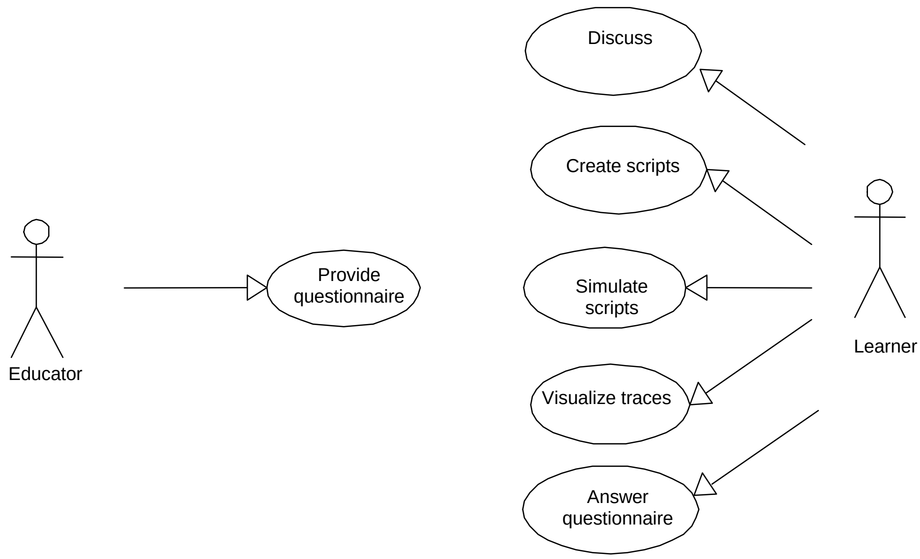


Figure 2: Use cases diagram for "Educator" and "Learner".

3.1.4 Scenario: Distributed Collaborative Software Project Development

3.1.4.1 Introduction

Software project development is the collection of techniques used to manage, develop and deliver various types of software products. This developing discipline traditionally includes technical issues such as the choice of software development methodology, how to estimate project size and schedule, how to ensure safety, what resources to reuse and which programming environment to use for the development.

The ability to translate solutions from the concept level to a proper document design, to code, test plans and other program artifacts, to describe interfaces, to use assertions, as well as to prepare and conduct team meetings in an effective way, are all communication skills related with collaborative software project development.

This collaborative grid based scenario consists in a group of students doing a software project in teams. In this situation, students could form a virtual organization with their own resources / services. In case they don't have enough resources they can obtain the resources externally (university, other group of students, a commercial resource provider, etc). To do the project, students must get together and collaboratively edit documents, develop software, make agreements, discuss, plan and schedule tasks, etc., either synchronously or asynchronously. In addition, the group could need to use some external software they don't have installed (e.g. due to license restrictions), deploy the developed software in a real o simulated environment in order to simulate its behavior, etc.

3.1.4.2 Roles

The following roles can be identified in the realization of this scenario:

- **Tutor**. The tutor would motivate students, clarify course objectives as well as explain the programming problem, project methodology and development process.
- **Technical Assistants.** They are familiar with all technical aspects of the project and share a common vision of the project course objectives. They would help student resolve technical problems related to the programming language and other artifacts used.
- **Students.** They carry out the project in team.
- **Subsystem responsible.** A student in charge of a project subsystem.
- **Project Coordinator.** A student responsible for coordinating tasks.

3.1.4.3 Organizations providing services

The main service providers in the basic scenario are:

- **Educational institutions.** One of the principal providers of the virtual environment formed of all resources and services for project managing, scheduling and collaborative tools is the University itself. It's very important to promote collaboration between all the groups developing projects so that in some moment the groups can provide services among themselves. In case that the services needed to develop a project cannot be obtained internally, it could be possible to contact external suppliers such as other educational institutions, commercial resource/services providers, software development industry and GRID community.
- A **storage provider** may offer storage as a specific service, due to the large amount of documentation generated by the software project, possibly with enhanced software features, such as version control, searching, commenting… The storage provider could charge for its service, or it could actually be one of the educational institutions that provided such a service.

3.1.4.4 Use cases

The activities of the scenario begin with the Tutor's (or Chief Instructor's) definition of objectives, environment and services/resources to be used in the elaboration of projects. This environment needs to support collaborative communication in synchronous and asynchronous forms.

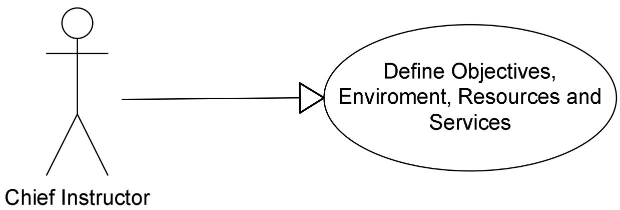


Figure 3: Use case diagram for "Chief Instructor".

Once the objectives and scope of the projects are defined, students interact with each other through the learning collaborative tool in order to form the teams and to define the project characteristics and the role they are going to play in the project development. In this first phase of the project the team formation and organization, the project planning, and the temporary monitoring of the deliverables and the documentation take place. We should pay special attention to issues related to member communication. The project plan must reflect the current needs in order to solve a specific problem and must allow us to develop a solution through a software life cycle.

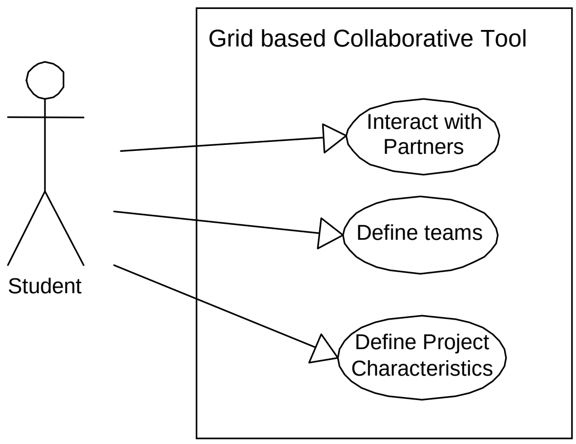


Figure 4: Use cases diagram for "Student".

The project plan must be based on a software development methodology, which will define the activities that must be carried out to complete the project and which should contain the following phases:

1. System/Information Engineering and Modeling. Work begins by establishing requirements for all system elements and then allocating some subset of these requirements to software. On the one hand, this system view is essential when software should function as an interface between other elements such as hardware, people and other resources. On the other hand, it is the basic and very critical requirement for the existence of software in any entity.
2. Software Requirement Analysis. The essential purpose of this phase is to find the need and to define the problem that needs to be solved.
3. System Analysis and Design. In this phase, the software development process, the software's overall structure and its nuances are defined. In terms of the client/server technology, the number of tiers needed for the package architecture, the database design, the data structure design etc. are all defined in this phase. A software development model is created. Analysis and Design are very crucial in the whole development cycle.
4. Code Generation. The design must be translated into a machine-readable form. The code generation step performs this task. If the design is performed in a detailed manner, code generation can be accomplished without much trouble. Programming tools like Compilers, Interpreters, Debuggers are used to generate the code.
5. Testing. Once the code is generated, the software program testing begins. Different testing methodologies are available to unravel the bugs that were committed during the previous phases. Different testing tools and methodologies are already available. Some companies build their own testing tools that are tailor made for their own development operations.

Once the project has initiated, students will undertake two different types of roles. Those being responsible for a specific part of the project (Subsystem responsible) and those that take care of the project control (planning and evolution), write reports related to the project and teammates' performance and moderate the collaborative tool function (Project Coordinator). The Tutor (Chief Instructor) and Technical Assistants take the role of supervisors checking the correct development of project activities and roles.

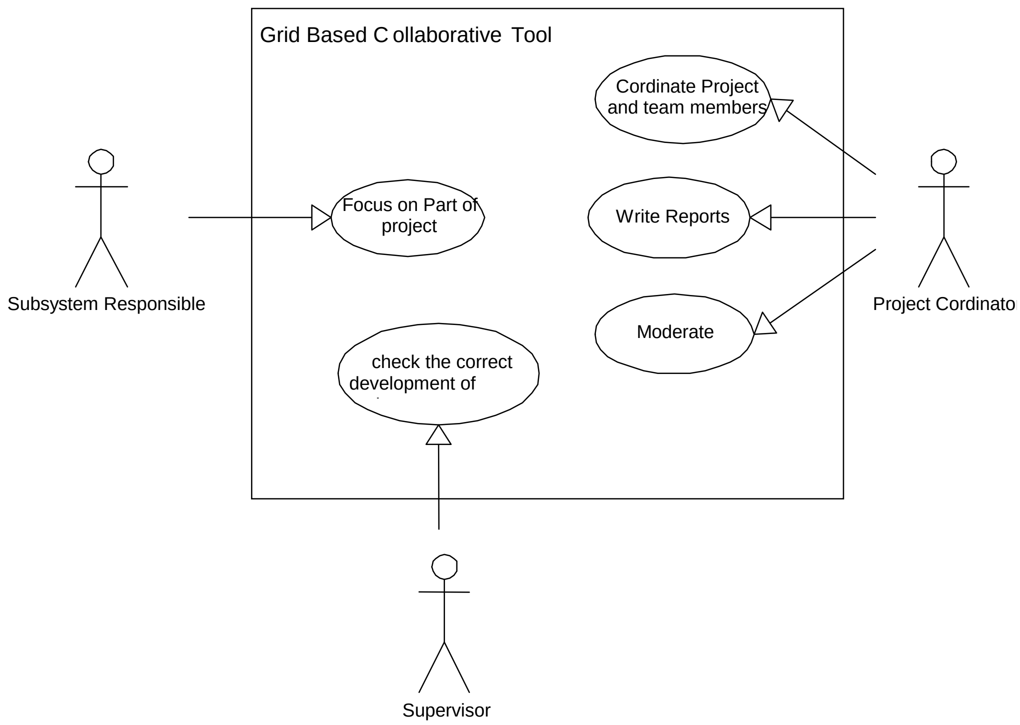


Figure 5: Use cases diagram for "Subsystem Responsible", "Project Coordinator" and "Supervisor".

3.1.5 Scenario: Live tutorized online sessions

3.1.5.1 Introduction

Thanks to technology, which has made this possible, there is an ever growing demand for solutions to communication and intercommunication formats in which synchronized, but not necessarily spatial, presence is required.

This demand is a means of recreating methods of communication which are usually in person, such as conferences, seminars, tutorials, lessons, etc., in which intercommunication and live interaction between people is an essential part, and therefore, in the synchrony of time, but in which physical presence can perfectly be substituted by virtual presence.

In this scenario all those who take part in the communication situations are prominent, independently from the fact that there may be a direction in the group or the meeting. That is why anyone in the group should have access to the tools available although the direction role would have extended permissions.

3.1.5.2 Roles

The following roles will take part in this scenario:

Session director: is the person in charge of the control of the session. The direction may not imply prominence, is an organizational role, leading participation and directing interventions.

Participants: people taking part on the session. They've been called for the meeting by the session director and will have available communication and collaborative tools in order to guarantee a fluent participation.

3.1.5.3 Organizations providing services

Storage provider: store the results of the session.

Content provider: documents to be used in the session.

Educational institutions: to provide the pedagogic resources (teachers and tutors).

ISP: to supply enough bandwidth regarding the number of attendees to the session.

3.1.5.4 Use cases

Communication at e-learning environments is normally based on asynchronous tools in which interaction between the different roles involved on is not as fluent as it should be needed. Then, a virtual space where a live meeting can be performed would be appreciated.

Depending on the needs of the participants and also its role in the organization, different kind of meetings should be performed as, for example, tutorized sessions, work-group meetings, defense of a thesis without the need for the members of the jury to be in the same location, etc.

For this space to be useful, these objectives should be covered as a minimum:

1. Each component's voice must reach the others with an acceptable level of quality, both in volume, distortions and noises.
2. If those who are taking part have not met each other previously it is important for them to see each other's image by the rest with certain relevance, at least during each person's presentation time.
3. There should be a space on the screen that can be seen by all those who take part in which they should be able to identify each other.
4. As well as this identification, for each of the participants there should also be a visible state: listening, listening with the intention of speaking, and speaking. Logically, it should also show if one of those who is taking part leaves, because they could be questioned by another at any time.

5. Possibility of each person who takes part being able to show the rest of the group elements which have been previously created or stored in digital mediums.
6. To allow elements which have been created on the spot with the help of applications be it on the level of experiments, electronic board, optical pen, etc.
7. Possibility of reception on behalf of all candidates on acceptable levels of quality from broadband in the context of all the group's components.
8. Demand a sole standard of hardware in presentations which are frequent at the time.

3.2 Domestic users

Many Grid proponents take the position that grid computing represents a "next step" in the world of computing, and that grid computing promises to move the internet evolution to the next logical level. In this way, many efforts have already been made in order to take advantage of grid infrastructures to support various applications designed and developed for domestic users. In this section we will provide an overview of some grid – supported applications and scenarios concerning domestic users and then a number of new scenarios that could be supported by Grid4All.

3.2.1 State of the art

Grid infrastructures provide highly scalable, highly secure, and extremely high-performance mechanisms for discovering and negotiating access to remote computing resources in a seamless manner. This makes it possible for the sharing of computing resources, on an unprecedented scale, among an infinite number of geographically distributed groups.

A first scenario that can be often found in literature is the use of grid capabilities and infrastructure to support massive online multiplayer gaming. In this scenario a wide community of international gaming participants is occurring that require a large number of gaming computer servers instead of a dedicated game server .An example is the edutain@grid an EU specific targeted research project which aims to apply Grid technology in order to improve user experiences in both online gaming and distance e-learning systems. [11]

Another interesting scenario concerns meteorology. An example could be a virtual organization that requires resources such as weather prediction software applications to perform the mandatory environmental simulations associated with predicting weather. The organization will also require hardware resources so as to run the respective software, as well as data storage facilities to store the data generated from performing the simulations. A very interesting project related to this is the CrossGrid, which is an IST Programme R&D project that supports a Grid infrastructure for various meteorological applications, from data mining using large distributed databases, through atmosphere and ocean modelling for warning and diagnoses of the sea state to air pollution simulation. [10]

Media sharing is also an area that concerns domestic users and can benefit from grid infrastructure. An example is the GridCast project which is a collaborative industrial eScience project .GridCast aims to develop a prototype media grid that provides a media content sharing and media services environment for the BBC. GridCast main focus is to support various broadcasters such as BBC Northern Ireland, BBC Scotland etc in order to provide tailored broadcasting schedules to their communities. [15]

Another example of using grid infrastructure to support the area of Media Sharing is the Media Grid platform. Media Grid provides digital media delivery and processing services for a new generation of networked applications. It combines Quality of Service (QoS) and broadcast features with distributed parallel processing capabilities. Together these features create a unique software development platform designed specifically for networked applications that produce and consume massive quantities of digital media. [20]

Another interesting project is the European GEMSS Project which concerns the field of eHealth. This project aims to create medical Grid service prototypes such as maxillofacial surgery simulation, neurosurgery support, radio surgery planning, inhaled drug delivery simulation and their evaluation in a secure service oriented infrastructure for distributed on demand supercomputing the GEMSS testbed. [5]

Finally AstroGrid is another very interesting project. AstroGrid's main focus is to build a grid infrastructure that will allow a Virtual Observatory, unifying interfaces to astronomy databases and providing remote access as well as assimilation of data provides astronomers with a simple set of tools which enable access to a wide range of astrophysical data and applications which can be used to manipulate those data. The AstroGrid workbench is the main user interface for the astronomer to the Virtual Observatory - a world wide initiative linking all main providers of data in astronomy (e.g. ESO, ESA, NASA). [2]

The following two scenarios concern the eHealth and media sharing field respectively and could be supported by the Grid4All project.

3.2.2 Scenario: Multimedia and entertainment spaces

3.2.2.1 Introduction

Nowadays, home users need some virtual platform to express themselves a cross the Internet. Such platforms allow them to keep contact and exchange media with their friends and family. Amongst others, Windows Live Spaces, Yahoo! 360º, del.icio.us, Digg It and ma.gnolia have been developed with that aim [31]. Those applications have millions of home users from around the world that receive free resources when they join the community. Normally home users use the resources for media sharing purposes, blog edition, and remote bookmark storage. In addition, such platforms offer services to home users such as instant messaging, email, real time communication and multimedia file sharing services.

One drawback of those systems concerns the fact that users have to rely on that company owners would keep the application free of charge. Another aspect against such platforms is the limited amount (storage, process and communication) of resources a user can use.

The home users' community would benefit from an entertainment and media sharing application using the resources contributed by the members of the community. The application would provide them entertainment and communications spaces using the resources of the community so that they do not depend on third parties.

3.2.2.2 Roles

The primary customers are home users, particularly teenagers but not only. A secondary set of customers are specific communities of users with a common interest. A fundamental requirement is to integrate resources contributed by users and manage access to them. Application would provide services to users in order to manage personal spaces.

3.2.2.3 Organizations providing services

In this scenario services are provided by participants. Home users group in virtual organizations providing some of their own resources. Most users of these spaces are willing to collaborate to the community. They form a proactive community for entertainment and media sharing so that if the community does not have enough resources the community does not exist. As in all communities, there would be more involved members which would contribute with more stable resources.

3.2.2.4 Use cases

We model the scenario as follows:

Some friends decide to create a virtual community in order to keep in touch during the summer. They would exchange photos, videos, files, and talk with each others in either asynchronous or synchronous manner. They should also create a blog or some other web based content container.

Every one decides to contribute with his unused home pc. In addition all of them have a laptop which they would use during their vacations.

Each friend creates his personal spaces with varied multimedia content that wants to share with his friends. Once the space is created, the owner of the space can invite any of his friends. The owner can decide the permissions each of his friends has into his personal space. Once a user becomes invited to a space, the user has rights of access to some of the services of that space.

Although users can disconnect whenever, the content of the space remains available for his friends.

The application monitors the available resources. When there is a lack of resources community users are notified immediately. Given that all users are proactively collaborating they would provide resources to the community even they do not use as much as they contribute.

When a user logins into his personal space the user can access to the storage (space) management tools which allows him to change or add content. In addition the space offers messaging services, asynchronous messaging services, real time voice communication, video conference tools. Moreover a member can see the list of his friends, access to their spaces and use the services which he has access rights.

3.2.3 Scenario: Using grid resources and digital home health services to fight various diseases

3.2.3.1 Introduction

With the upcoming aging population, Europe faces the challenge of delivering high-quality healthcare to its citizens at affordable cost. While it is widely accepted that healthy and preventive lifestyle, together with early diagnosis, could systematically fight the origin various diseases, for cost-related reasons institutional points of care can currently offer only intermittent episodical treatment. In the following scenario we intent to use Asynchronous Multimedia Collaborative and Media Sharing Systems as also and streaming media to support people facing health problems, such as Stroke and Vascular Dementia in order to empower the users to take control of their own health status. Aging people (usually >55 years old), people facing health problems, health carers and other social and medical communities are possible users.

For the better understanding of this application consider the following scenario: A 50 year old man has a serious health problem that reduces the quality of his life.

The man wants to have a deep understanding of his problem and to take control his own health status by understanding various aspects of his problem such as proper nutrition.

In order to achieve that, he will be using various Asynchronous Collaborative Multimedia Systems and Media Sharing Systems as also and educational multimedia content concerning his problem. This is provided by companies, organizations and other special health centres.

Both patients and content providers can benefit a lot from the grid. To begin with many companies could offer storage and other useful services concerning video streaming and media sharing. In addition content providers could use grid's computational resources for video processing while grid's middleware could also take care of several other functions such as user authentication and authorization, security and file access rights.

3.2.3.2 Roles

In this scenario the primary user is a person who faces health problems. The primary user is likely to be either the patient himself or somebody who will help him, preferably a social worker, a nurse or a person (friend or a family) with basic computer knowledge.

The various content and service providers specialized in the area of Health are the second actors in the scene for this scenario. These are responsible to provide digital videos and other multimedia content concerning the health problem that the patient is facing. They are also responsible to provide help and answer questions by providing various communication services such as email, forums, chatting etc.

3.2.3.3 Organizations providing services

- **Content Providers**: Special health centers, private companies, health ministries, research institutions and universities can provide multimedia content. Private companies and special health centers may charge the content they provide.
- **Storage Providers**: Universities or other research institutions may provide storage as a service, possibly with enhanced software features, such as version control, searching, commenting. Private companies can also provide storage and charge for its usage.
- **Media Sharing Providers**: Companies, Universities or other research institutions may provide various media sharing services.
- **Media Processing Providers**: Companies will offer video, audio and image processing services such as encoding various video formats to other formats, editing and encoding audio etc.

3.2.3.4 Use Cases

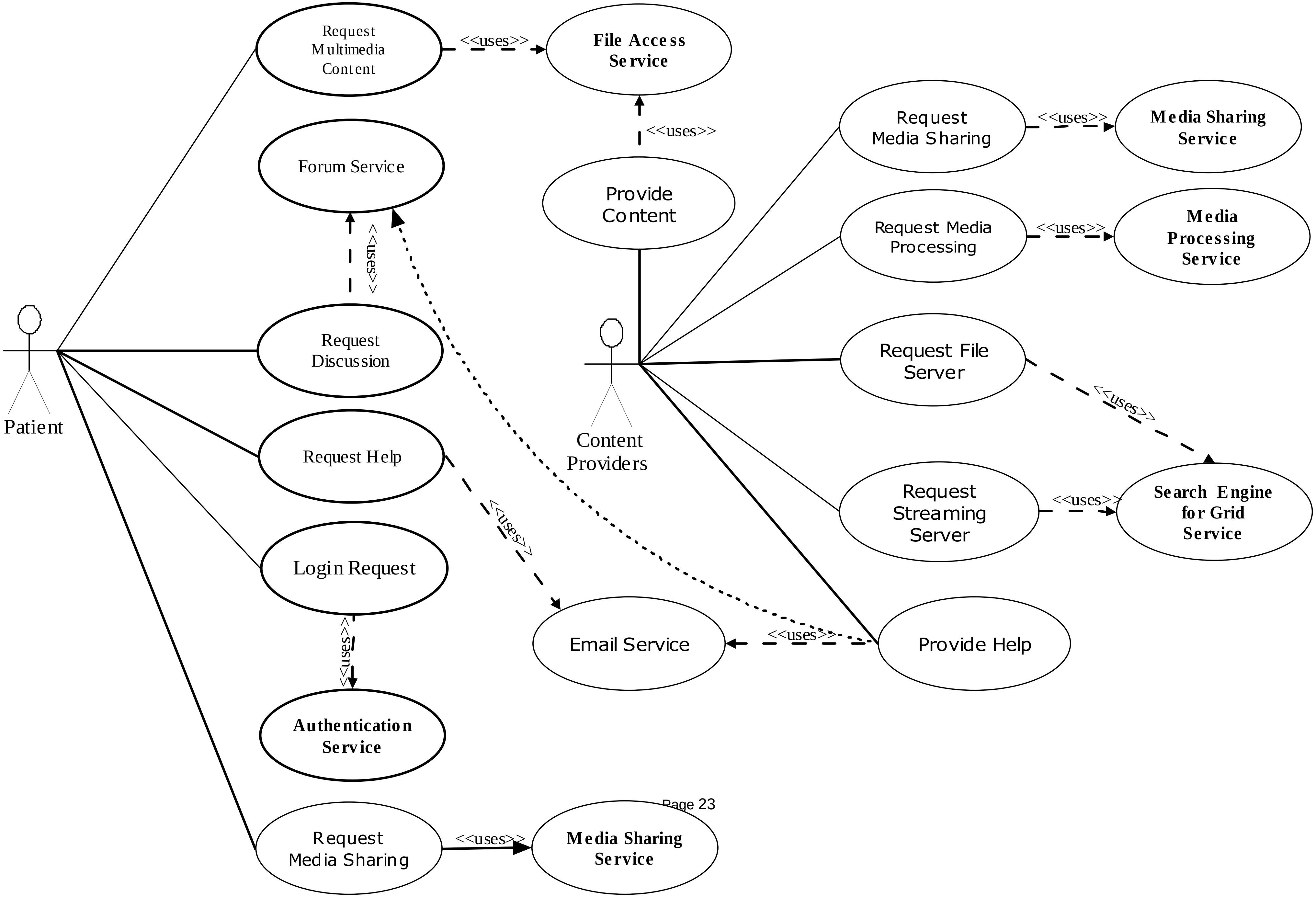
A 50 year old man facing a serious health problem and wants to have a deep understanding of his problem in order to improve his quality of life. In order to achieve that he is using various AMC and Media Sharing systems that are mostly based on P2P technologies as also and multimedia content concerning his problem provided by a private health centre. The man is trying to take control his own health status by understanding various aspects of his problem such as:

- *Proper Nutrition*
- *Reduce Inactivity and Isolation*
- *Improve Sleep Quality*
- *Reduce Stress and Depressive Symptomatology due to Negligence*
- *Reduce morbidity by early recognition of vital signs and diagnosis*

To begin with an employee of the private health centre stores multimedia content that concerns the patient's health problem in a file and a streaming server respectively. These servers are provided by third party private company. The employee has also asked for various services concerning media sharing and processing in order to be able to provide the content to the patient.

The domestic user (patient) is using an AMC system to log in system enabling him to access the multimedia content provided by the private health centre. He has also requested the list of available video clips and other health related multimedia content. In addition he is using and some other Media Sharing systems in order to share and exchange multimedia content with other patients facing similar health problems.

In addition he uses some features provided by the AMC system to reprocess and analyze the video clips, to collaborate and discuss his problem with other domestic users (patients) and the employees of the private health centre. Finally he uses various forum and email services provided by the private health centre in order to ask private questions, discuss and express his opinion for many aspects concerning his problem.


Request Multimedia Content
<<uses>>
File Access Service
<<uses>>
Forum Service
<<uses>>
Provide Content
Request Media Sharing
<<uses>>
Media Sharing Service
Request Media Processing
<<uses>>
Media Processing Service
Patient
Request Discussion
Content Providers
Request File Server
<<uses>>
Request Help
<<uses>>
Request Streaming Server
<<uses>>
Search Engine for Grid Service
Login Request
Email Service
<<uses>>
Provide Help
<<uses>>
Authentication Service
Request Media Sharing
<<uses>>
Media Sharing Service

3.3 Small and Medium Enterprises

- **Small enterprises and supply chains**: these organizations could borrow or share the resources of friends, collaborators or clients to provide services (e.g. a digital photo processing service using storage and processing resources of the local community; a local digital TV, video-club or radio station providing service, audio or video content, based on the resources contributed by the community and other resource providers)

SME constitutes a key component of the European economy with about 25 million SMEs, accounting for approximately two thirds of Europe's employment and GDP. Therefore, SME are a key component in the research and innovation system and also on creating and adopting new products and services. Due to their ability to adapt quickly to changes, SME are often better positioned towards innovative and more efficient ways to work.

SME are inserted in provision-supply chains as part of a social network or a services and resources ecosystem where they can borrow or share resources or services: collaborators, partners, customers, providers, and especially with some specific providers (utility computing centers).

In this section we have identified the potential uses of a grid infrastructure within the context of small and medium enterprises and the virtual organizations they belong to.

This section first presents an overview of the scenarios for (SME) that can be found in the literature. Then, we present a number of detailed scenarios that could be supported by the Grid4All infrastructure:

- Live sessions using collaborative tools
- Collaborative software development
- Earthquake prediction
- Multimedia and entertainment spaces

3.3.1 State of the art

Business and SME in particular can profit from the use of Grid considering their computing capacity is usually very limited, and this model helps them to adapt, grow and benefit from a shared computing infrastructure [19], [24]. Within the Open Grid Forum there is one are focused on industry applications [21]. However applications must be modified to work in this environment, but in many cases that effort is compensated by the savings in infrastructure.

There are some specific sectors or activities that can be more easily mapped or benefit from the use of Grid. The gaming industry is an example as presented in [9] and [27], where an SME may offer an online game but may require a computing infrastructure proportional to the demand which may be difficult to adjust. The same is applicable to cases where peak computing or storage capacity is required as in [29], or where specific applications require additional resources as in [4] or in domains where typical applications are computing intensive, as in [13], or in the medical domain as illustrated by [18] and [17].

3.3.2 Scenario: Live sessions using collaborative tools

3.3.2.1 Introduction

The phenomenon of the internet has created an indisputable environment which is perceived as the globalization of communication.

Technology has made possible to recreate communication methods, which are usually in person and where synchronous interaction between people is essential, where physical presence can perfectly be substituted by virtual presence.

This communication Scenario would be made up of situations, such as conferences, presentations etc., in which most of the weight rests upon one or several speakers although some of those who take part may also be involved, be it to clear doubts or ask questions.

The weight of communication in this Scenario is in one direction: transmitter-receivers (1→ N). That is why the objectives of this Scenario will logically be those which allow, on one hand, the transmitter to have all the facilities to be able to present his speech, and on the other hand, that this reaches the receiver via web with the same or similar relevance with which someone who is physically present where it is being held.

In any case, to make the most of the possibilities offered by technology, also being aware of the limitations, and to improve the cost/benefit relationship we will opt for some possibilities, although we will have to sacrifice others.

To be operative, this Scenario must cover, as a minimum, these specific objectives:

1. To emit the transmitter's voice, in such a way that it can reach the receiver with an acceptable level of quality, both in volume and distortion or noise.
2. To emit the transmitter's image towards the receiver with the adequate relevance, although the emission is fundamentally based on the voice.
3. To emit the image of physical objects to the level of detail which is necessary to be able to see what is being shown. This refers to what is written directly onto the blackboard to objects that are shown to be identified.
4. Show previously created or stored elements digitally to be shown during the event: presentations, videos, animations, etc.
5. Show elements which are made on the spot. For example, results from a data collection from an experiment which has just taken place, by way of animation programs, electronic board, optical pen, etc.
6. To make reception possible on acceptable levels of quality by way of broadband which is normally used by the receiver.
7. To make the reception of the Scenario possible by way of common general use applications, without the need of specific software and without having to install it previously to allow reception or to make it's download possible in case it is not already installed.
8. To make it possible for the receiver to download documents.
9. Demand a sole standard of hardware in presentations which are frequent at the time.
10. To make it possible for the receiver to ask the transmitter questions, either through an intervention protocol, or with the help or a third person as a moderator. Logically, both questions which are asked and their answers must reach all the receivers.
11. In events with various speakers, the possibility of each speaker taking part without having to be physically next to the rest of speakers.
12. Also, these objectives should compliment other derived objectives such as:
13. Record emissions to be able to gain access to them after the event.
14. Index emissions by date, speaker/s, theme, event, key words, etc.) to make the search easier later on.
15. Subtitle in other languages when it is appropriate for the contents, both for the inaccessibility of the languages or for certain users' difficulties. For example, hearing difficulties.

3.3.2.2 Roles

In order to guarantee an optimal session, these roles should be present:

Speaker: One or several speakers, transmitting from one or several places, taking part simultaneously or successively.

Attendants: The receptors of the speech.

Producer: Responsible of technical questions, has the control of the transmission. Marks the beginning and the end of the session, regulates video & audio qualities depending on the traffic on the net, sets turns of intervention when various speakers are involved and is able to publish messages for the audience regarding to transmission.

Moderator: Controls the attendants interventions flow selecting, filtering and directing questions to speakers. Moderator is also able to contact attendants for explanations (regarding its interventions).

3.3.2.3 Organizations providing services

Storage provider: where the documents to be presented are stored. To store the recorded files to be accessed after the emission

Content provider: Companies or institutions wanting to perform such a presentation.

ISP: to supply enough bandwidth regarding the number of attendees to the presentation.

3.3.2.4 Use cases

Now we will name some of the examples which represent this Scenario. They all have in common that they allow the reach of large or disperse groups from one or several speakers, situated in one or several places.

Case 1: Presentations

Case 2: Press conferences

Case 3: Implantation of protocol for risk prevention in the workplace

Case 4: New product launch events with simultaneous presentation in many sales outlets.

Case 5: Cultural events organized by councils.

Case 6: Conference cycles

Case 7: Fair and Congress presentations.

Case 8: Seminaries and round tables.

Case 9: Educational events for large or dispersed groups.

Case 10: Highly specialized educational events for not necessarily large groups but disperse groups.

3.3.3 Scenario: Collaborative software development

3.3.3.1 Introduction

Developing and debugging software in distributed systems is complex, since the state of the program must be traced in remote machines, and also because different parts of the software will be running in different environments, either operating systems or hardware, and often both [12]. To cope with these problems, some debugging platforms can deal with the communication tasks, hiding them to the programmer [6]. Other platforms emulate different types of hardware in order to allow developing programs for them, even not having physical instances of this hardware [3]. However, this solutions are often limited to certain types of languages, middleware or hardware platforms.

Sharing computational resources in a service oriented grid could allow developers to write, store, version, debug and run their code in different locations of the computational grid, which are selected according to their programming needs. The middleware of the grid would take care of security and accounting of the resources. Sharing policies could be defined within the virtual organizations to control the access to resources according priorities of owners, time reservation, etc. Some companies could offer storage and improved version control as interesting services in this virtual organization. Further, specific hardware vendors could become service providers, offering instances of their hardware as test banks where developers can test their programs. In other cases, industrial partners will collaborate without having to move to each other's premises.

To have a deeper understanding of this type of application, consider for following scenario: a hardware company is developing an autonomous robot for whatever specific task, and different software companies and academic institutions are developing artificial intelligence algorithms to control the robot behaviour, and coding them. Each of the latter care for a different algorithm (recognizing the environment, movement of wheels, grasping with arms…), and they can work in parallel at their own location. Whenever they want to test their code, they can do it in a robot available at the hardware company laboratory.

3.3.3.2 Roles

In this scenario the main users are software developers. They use the services provided in the grid to store their code, run it in the robot and debug it. They benefit from the grid since they do not have to move to the hardware company premises or establish *ad hoc* communications with the robot. Further, they can also benefit from collaboration with other developers in the virtual organization, by sharing code and using software tools available in the grid.

3.3.3.3 Organizations providing services

The main service providers in the basic scenario are:

- The hardware company: its service is actually a wrapper to the robot, and the relevant operations on it would be to install software, and to run it, autonomously or step-by-step, with the possibility to observe the program state. The hardware company benefits from the grid by having several developers working at the same time (eventually, even competitors could be trying to develop programs for the robot, so the hardware company could choose the best). Only one prototype of the robot is developed, and held at the company premises, achieving more industrial confidentiality during its test.
- A storage provider offers storage as a service, possibly with enhanced software features, such as version control, searching, commenting… The storage provider could charge for its service, or it could actually be one of the software development companies that provided such a service.

Other service providers could also appear in extended scenarios. For example, one could offer compilation services, with code optimization and compression. Though it is very likely that software developers have these facilities, it may not be the case in an open source project.

3.3.3.4 Use cases

One possible modelling of this scenario is as follows:

A hardware laboratory has developed a robot, and wants to sell it without much overhead because of the software controlling it. If calls for an open source project in the SourceForge website, and some programmers interested in artificial intelligence algorithms join it. Some of the members of the project know each other, but some join after the call, and never meet personally. The result of the project will be open source code, that this hardware laboratory will use, but available to any others.

To complete the project, SourceForge and the hardware laboratory form a virtual organization in a grid. The laboratory offers a service that is an execution environment for the robot, while SourceForge offers a Concurrent Version System (CVS) for storage and sharing (other tools could also be provided by SourceForge or others, but they are not relevant here).

When a programmer wants to test his code, the following interactions happen:

- The user request a reservation of a time slice of robot usage. This is necessary since the robot can only be used by a programmer at a time. If more robots were available, this would still be necessary but the scheduler could allocate tasks in different robots as convenient.
- At the allocated time slice, the code is copied from the CVS to both the robot execution environment and to the programmer graphical interface. A file transfer is required in this stage, and checking of appropriate security credentials is mandatory to avoid bad intentioned usage of the robot.
- Once the code is installed, the programmer manipulates the execution from a graphical debugger, running the code step-by-step, to a breakpoint, or completely, and observing the state of the robot. During this stage, the service that is wrapping the execution environment receives messages to carry out these operations, and informs of the changes in values of state variables.

- After the execution is complete, the traces of the experiment are stored in the CVS, associated to the specific versions of the code that has been run.

3.3.4 Scenario: "Earthquake prediction"

3.3.4.1 Introduction

Scientific community would benefit from a global data processing application. Such application enables them to accurately predict the exact movement of tectonic plates.

The movement of tectonic plates can be measured on a combination of real-time wide area seismic sensor instrumentation, GPS data, and large-scale simulation coupled with data modelling. This is an extremely difficult problem and it is far beyond the current capabilities of seismic simulations. At present, scientific communities do not have an application to study global tectonic plate's movements based on real data. In fact, most of research centres have their own data acquisition instruments and they make predictions using their own data. Each community has access to few data sources so that predictions are far from giving important information. Given the current situation, scientific community cannot predict when and where would be the next earthquake.

3.3.4.2 Roles

The primary customers are the seismologic communities and geologists. A secondary set of customers are the emergency management people, disaster recovery teams, governments, petrol industries and the mass media. They would benefit from a common data processing framework. A fundamental requirement is to integrate data from distributed and heterogeneous resources with analysis capabilities provided as services. Users would interact with services provided by the application.

3.3.4.3 Organizations providing services

The main interested organizations on earthquake prediction are Seismologic institutions, geophysics departments in Universities and research centers, civil defense institutions, petrol industries, and mass media enterprises would provide services for the earthquake prediction system. However, the importance of earthquake prediction for world population (economic and politic reasons) can bring different organizations to provide their resources. Even citizens of cities placed near 'edge faults' such as San Francisco or Tokyo would be interested on providing resources to the community.

3.3.4.4 Use cases

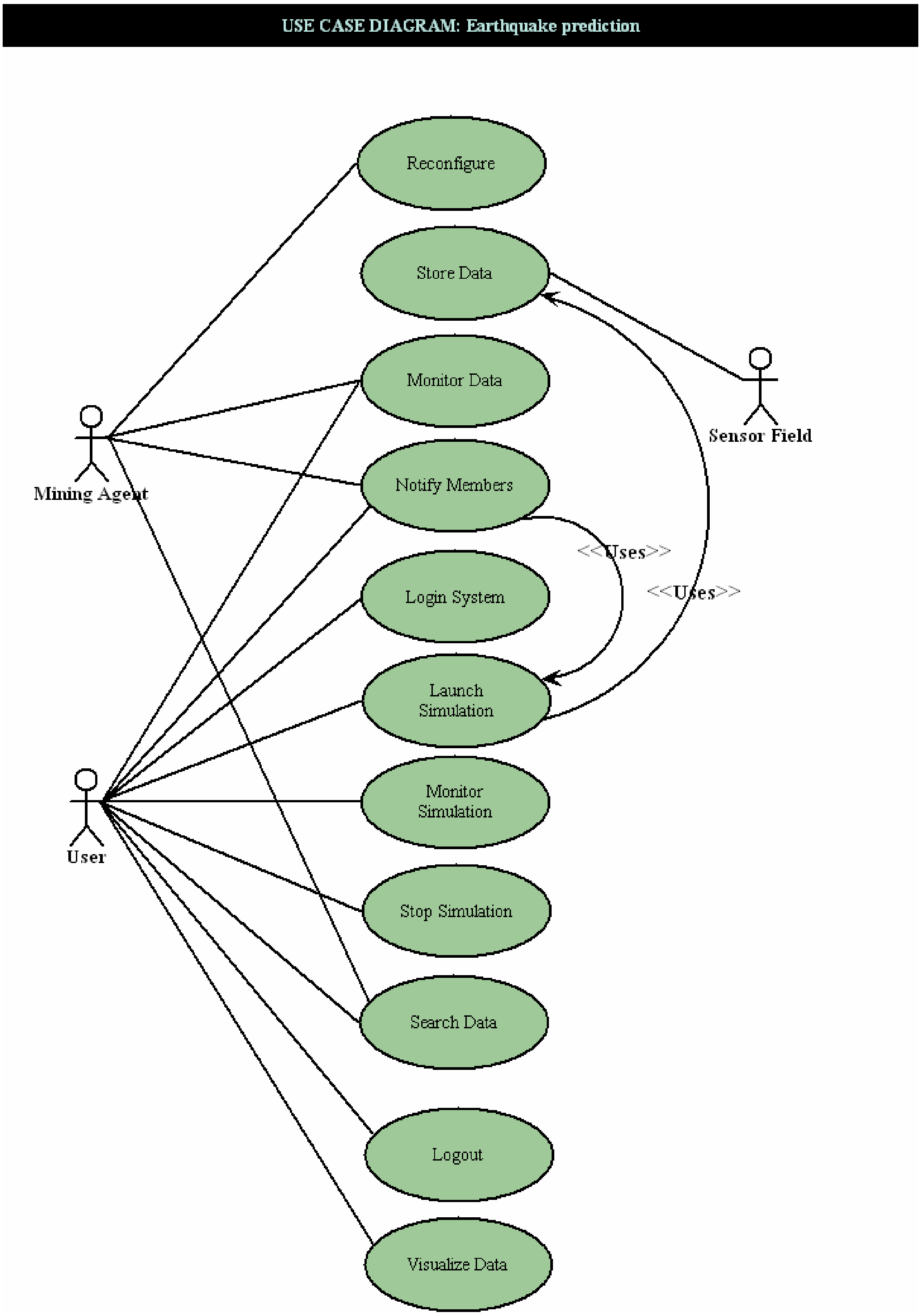
We model the scenario as follows:

Instrument data streams from Doppler radar, seismic stations, GPS devices, and sensors network components such as temperature and vibration detectors are collected from thousands of dispersed sensor fields around the world and sent to scientific community machines belonging to a Grid4All VO. The data is constantly monitored by data mining agents looking for dangerous patterns.

When one dangerous pattern is detected, VO members (scientists) are notified and a large number of simulations are launched automatically. Data mining tools are configured to scan the output of the simulations and compare the results against the evolving data stream from the instruments in different places. Data archives are searched for similar patterns. Some of the instruments are automatically reconfigured to refine the data streams.

When a significant event is detected, the application requires humans monitor the entire process and aide in the process by steering some of the simulations. (The simulations generate output files which can be visualized as animations.).Output files contain predictions about movements of tectonic plates and possible earthquakes. Whilst some scientist takes care about epicenter region, other scientists can detect possible replicas of an earthquake in other regions of the world.

The following use cases have been identified:



3.4 For-profit and non-profit communities

People belong or form communities (either for-profit or non-profit) to perform some complex tasks. These communities organize, constituting a VO around a common interest or mission, the set of participants and the required resources, contributed by some members or obtained from external sources. They may form specifically to handle some particular need (e.g. emergency handling, reaction to some event, or a specific campaign, etc.) or they may last longer (e.g. socio-political citizen movements, opinion groups, associations sharing common interest or goals, virtual libraries, etc.). They may be small (e.g. a few people involved in an emergency) or large (e.g. a political campaign).

In this section we have identified the potential uses of a grid infrastructure within the context of communities.

This section first presents an overview of the scenarios for communities that can be found in the literature. Then, we present a number of detailed scenarios that could be supported by the Grid4All infrastructure:

- Community computing:

  The opportunity and possibility of citizens to organize and collectively define a position, promote and collectively act is the basis of the information society, and a strategic line of action in the EU (*e*Democracy, *e*Participation, *e*Inclusion) and also on the United Nations on the Millenium Development Goals and the World Summit on the Information Society (WSIS) outcomes.

- Emergency handling:

  Emergency handling is one situation that requires an immediate organization of otherwise separate agents to handle a usually unplanned event (preparing for the unexpected) that may require a fast response involving shared information, communication and coordination in very short time. That immediate coordination is critical to success, and several projects in the area of civil protection and environmental emergencies (e.g. the 112 phone service in the EU, and several IST projects on the 5^{th} and 6^{th} framework research programme).

- Virtual libraries:

  Virtual digital libraries constitute an invaluable repository for many types of knowledge represented in digital format. This can be publicly accessible knowledge (such as the projects of the Wikimedia foundation) or with limited access (such as many books, movies, music, pictures, paintings, etc.), read only or open to change. Anyway, virtual digital libraries are a key resource for many activities where knowledge is involved and advanced services are required to find, access, annotate, or publish new knowledge.

3.4.1 Scenario: Communities

3.4.1.1 Introduction

Humans tend to organize in communities. Some are long-term communities, such as Sourceforge, Berlios or traditional NGOs, and others are ad-hoc and established in the short term, such as neighbors trying to prevent that a public park becomes a shopping center due to changes in the uses of land.

Since its beginnings, Internet has helped the creation and organization of communities, especially improving people's interaction and communication capabilities. E-mail, Usenet News, Web and Instant Messaging are four of the most well-known and successful examples of this. Internet has allowed for the creation of asynchronous virtual communities where members interact on a many-to-many basis despite organizational and technical barriers. Many-to-many interaction, uncommon in the physical world, has transformed the way people learn, work together, find other people with common interests and share information, etc.

The generalization of Internet has also led to the appearance of communities that allow users to provide their computational resources to projects that they think worth of it. The most well known example (and one of the pioneering one's, started in 1996) is Seti@Home, which allows PC owners volunteer idle cycles to help the search for extraterrestrial intelligence. This has been a very successful project: 1,200 CPU years per day, and around $1.5M year consumed power that users provided to project for free. FightAIDS, Folding@home or Entropia are other examples of such communities where users contribute resources either for money or altruistic purposes.

Grid4All can help communities by allowing the community members provide resources to the community, creating V.O. that will allow members to perform their activities. That will avoid communities to depend from third party entities that provide them the required resources. In case a community doesn't have enough resources for their daily activity or it needs extra resources to accomplish a specific activity, Grid4All will allow the community to obtain them from other V.O.s or from resource providers.

3.4.1.2 Roles

Different kinds of users are involved in the communities. Active members participate in the activities of the group. “Associated” members are interested on receiving information about the community and their activities, and maybe participate in sporadic activities. The main difference with the active members is in the degree of involvement. External users are not “officially” involved in the community, but they may be interested in knowing about the community or about any specific activity.

3.4.1.3 Organizations providing services

Resources are mainly provided by its members. Members of this kind of organizations will be willing to contribute resources to the community. It will also be members that share the goals of the community but that don’t want or cannot participate in the daily activities of the community. Like in the Seti@Home project they will contribute resources to the community.

It will exist a third kind of members that share the goals of the community and that not participate neither provide resources, but that may financially support the community. With these financial resources the community will be able to hire resources from resource providers or other communities.

3.4.1.4 Use cases

We present two different use cases that apply to this scenario.

<u>Use case 1</u>: long term community

Suppose a geographic area with natural, cultural and/or symbolic importance. A group of people decides to create a community to share their interest about the place and to help other people to know about it. They gather historical, cultural, natural, and any other information of interest about the place. They provide a platform to advertise any activity that is organized. They may also provide some repository for photos or videos about the place. They may have a web cam broadcasting 24 hours a day a very nice view of the place. They may organize photo or video rallies about the place. Every Saturday and Sunday morning a group of volunteers organize life radio broadcast on Internet about the history of the place and about its natural environment. Those radio programs are stored and any person can download any time in the future.

Other similar communities: to protect an animal species, an ecological organization, an association of schools, a civil rights organization, etc.

<u>Use case 2</u>: ad-hoc community

Imagine a very nice and quiet neighbourhood with only one public garden. After some not clear manoeuvres a building corporation gets a permit to build 1000 new houses and a new shopping centre. Neighbours decide to prevent it to happen. They start a community to organize actions and coordinate their activities.

Other similar communities: two schools doing an activity together, etc.

3.4.2 Scenario: Emergency handling

3.4.2.1 Introduction

Emergency handling is an important issue in modern societies, and in recent years it has become largely dependent on information technologies. Indeed, IT offers means to notify that an emergency has occurred, and to provide data on what happened, where and when. Further, IT allow different entities to collaborate sharing data or resources, or by coordinating their actions.

However, designing the infrastructure for adequate emergency handling involves arguable provisioning of resources. Indeed, dimensioning of the infrastructure is made for the worse case, committing communication and computational resources that will not be used almost any time, and replicating critical devices, even though they may be very expensive.

The computational grid can leverage this by joining public and private institutions in virtual organizations that share their resources in a safe, controlled and accounted environment. For example, consider a case of the public organization caring for 112 calls in an EU country. It can federate in a virtual organization with firemen, policemen, and other public servants to coordinate their actions, but also with private telephone companies providing massive SMS announcement in an area, with universities or companies providing disaster simulation and forecast, regional networks of surveillance cameras or sensors… In such a grid, the emergency coordinator will have a uniform and integrated access to all resources, from the moment the alarm is triggered. Further, the 112 organization only owns part of the resources, and uses third-party services on demand, with adequate agreements of availability and other quality of service indicators. Service integration, however, is made automatically by the middleware and access is transparent to the user, so their provisioning is very efficient and does not involve time consuming human negotiations. Finally, some institutions can benefit from lending their computational resources to different emergency organizations for a charge.

3.4.2.2 Roles

The main users of this application are emergency coordinators. There may be one or more, collocated or collaborating remotely. They will access a graphical interface in which data relating to the emergency will be shown pictorially, with charts, maps or whatever convenient. In the same interface, they will be able to request the invocation of other services, as described below.

Other human participants in the application, like firemen, are not actually *users* of the grid, but part of the service provision: they receive requests to perform some service and provide a result of the invocation.

3.4.2.3 Organizations providing services

The main organizations interested in providing services in this grid will be:

- The 112 organization may not only use other's services, but also provide some basic services, such as forecast systems or access to historic data. In fact, they may be shared with similar organizations in other countries or regions.
- Public servants that collaborate in emergency handling. Firemen are an important example. The service they offer is in fact just an interface to provide them with specific instructions or access their state. These services can be provided with fixed units in their central office, or mobile units in their cars.
- Telephone companies may offer services of messaging to population, restricted to specific regions, by broadcasting SMSs from the stations in those regions. Normally, different providers of this type will compete, and we may automatically choose that with higher quality of service in the region of the emergency.
- Surveillance companies or public surveillance networks, with cameras at public buildings, industries, highways, airports… Their service will mostly allow accessing to streaming data, but also to manipulate the cameras.
- Computation providers will offer disaster simulation programs, and run them in their premises, or even use third party services to acquire computation cycles. They must guarantee their availability on demand, but they can also offer services to other customers, simply by managing priorities with an appropriate scheduling system.
- Storage providers may be used to store logs of the emergency to later study the way they were handled, or to derive new models for simulation and forecast. These providers may charge for this service, but often the 122 organizations will also provide the storage.

3.4.2.4 Use cases

The following is a sample scenario for which this grid application is useful. It assumes that the virtual organization already exists and users have to log into the system, but once they do so, they have access to the right resources according to what was agreed during the formation of the virtual organization.

When an emergency alert triggers the following interactions happen:

- The emergency coordinators logs into the system, providing his password. The user will not have to identify himself anymore during this session.
- The user types in a text box some keywords describing a tool he wants to launch, like "geographic representation". Tools are searched for in indices in the grid, and a list is provided.
- The user selects one that he knows will show the emergency in a map. This service is deployed (it runs somewhere in the grid), and processes the alarm signal. If this signal is, for example, a fire, this service contacts temperature sensors in the region, if available, to show the fire in the map. The user can zoom in and out in the map.

- Because of the type of emergency, links to related tools are offered to the service coordinator. He decides to open the service of the firemen, and a graphical interface opens showing the activity of firemen in the region. The emergency coordinator can send orders to them, or recruit firemen from other regions.
- Besides, he determines if there are surveillance cameras from spots near the emergency. If the fire is close to a populated area, this may be the case, and no matter who is the provider of this service (in each region one or more surveillance networks may exist), he can transparently access to their streaming to have a richer insight of the emergency.
- Considering the fire is close to human populations, the coordinator requests a fire forecast service. The service is searched for in the grid. Once one is located, it retrieves local information about the area of the emergency (nearby forests, rivers…). Possibly other data are retrieved from temperature sensors. All this service coordination is done automatically, without human intervention. After the simulation service is launched, the coordinator sets the parameters of the simulation and runs it. The computation happens in one or more locations in the grid, but the coordinator is unaware of this.
- According to the results of the simulation, the coordinator takes decisions, such as sending additional firemen cars, or notifying population. For example, to do the latter, the emergency coordinator request service providers able to broadcast SMS messages to the population in the area, and selects one of them. Then, messaging is requested.

Finally, it is worth noting that during the emergency or after it, all data should be transferred to a storage service so that it can be later analyzed.

3.4.3 Scenario: “Shared virtual libraries”

3.4.3.1 Introduction

Digital libraries are becoming an important tool for academic and research institutions. They go further than the utility of traditional libraries, providing more powerful search mechanisms, permitting ubiquitous access, replication of documents to serve many users simultaneously, and storing huge amounts of information without the need of a large physical room.

However, digital libraries could go one step further and provide means to process the information, by having document formatters (to generate, for example, PDF files), document translators, *voicers* that read documents aloud, comparers that seek for related documents and organize them in taxonomies, etc. Digital libraries could even take readers opinions and comments to facilitate the searches of other users with similar profiles, as already done in some internet shops.

In order to share both information and computational resources, and be able to support a large amount of users collaborating in the library, all public universities and research institutions of a country can join in a grid-based virtual library. This way, access can be granted to any of the resources of the library of a given institution to a given user, according to the policies that govern the virtual organization. Moreover, storage can be saved by having certain documents only in one institution’s premises, letting others access them. Alternatively, availability can be improved by having replicated copies of documentation. Scalability in the number of users can be achieved this way, because of the number of different sources of information. In either case, institutions with low budget can access more resources than they could afford on their own, thus homogenizing access to information among public universities. Furthermore, specific computational resources, like voicing programs or document formatters can be accessed transparently, even if they are provided by third parties, for free or charging.

3.4.3.2 Roles

The main roles in this scenario are library users (readers) and library administrators. Readers are students, teachers or researchers at any of the public institutions. They are given access rights to their institution virtual library, and will have transparent access to other libraries as agreed when forming the virtual organization between institutions. Library administrators upload, remove or index documents using services of the virtual library. If agreed that way, new documents can be stored anywhere in the grid, transparently to the administrators. Library administrators could also be in charge of moderating document annotation or discussion forums.

3.4.3.3 Organizations providing services

The main organizations interested in providing services in this grid will be:

- Academic and research institutions: they provide a library service, which allows searching, retrieving, and (to administrators) introducing documents. Library services of each of the institutions are orchestrated to provide the view of a single virtual library.
- Publishing companies: they provide a similar library service, but charge for the access to their products, either on a per-document basis, or monthly or yearly subscriptions.
- Software companies: they can be interesting in providing services such as format transformation tools or voicing tools. They can charge per operation (thus accounting is needed) or with a licence for a period.
- Storage providers: these will most likely be the actual academic and research institutions sharing their libraries, but could also be private companies charging for the usage of storage, but committing to certain level of quality of service.

3.4.3.4 Use cases

Consider the following scenario as a typical workflow with this grid:

- The reader logs into the system and types some keywords describing a document he wants to search.
- The search service, before proceeding with the query, can contact a personalization service to ask further preferences from the user profile. For example, if a user typed “grid” but often look for documents on electric power production and distribution, papers on the computational grid are not relevant to him.
- After this, the search service contacts one or several document indices, and searches for documents. There is a broad number of possible approaches in this step, either consulting all document indices in parallel, making successive searches until a reasonable number of hits is achieved, making searches first by coarse grain categories… These choices will affect the efficiency and the precision of searches, but they are not relevant at this point.
- The search service provides the documents found. Normally, it will show the title, and some other data like authors, dates or abstracts, if available. These can be retrieved from the document indices, or they may be metadata retrieved by the storage providers.
- The reader selects a document and the preferred format. The search service selects the service provider that offers that particular format, if available. Otherwise, looks for a service changing from any of the available formats to the desired final format. When this service is located, the appropriate source file is retrieved from the storage provider, then transferred to the format transformation service, and the result handled to the user. The result may also be stored with a storage provider, so that format transformation will not be required again for this document.
- Sometimes, the reader wants to hear the document. To do so, the metadata of the file is checked to find out its language, and a voicing service for that language is sought. Ideally, several will be available, and the scheduler would choose the one closer to the user, in order to provide him a better quality of service, and to consume less communication resources. Once the voicing service is selected, bandwidth is allocated, and streaming can start.

Other relevant scenario concerns the library administrator, though the interactions can be simpler:

- The administrator requires adding a new document in a graphical environment. She provides the file and the document metadata, and they are stored in some storage provider (the selection of provider should be transparent to the user), and the document is registered in a document index.
- The administrator requires removing a document in her graphical interface. The document is sought with similar interactions as above, but before removing that document the system must check that the user is actually allowed to do it (for example, the library administrator of one particular university may be treated as a plain reader when accessing the contents of other libraries in the grid).

4 User requirements

This section compiles the user requirements (functional and non functional) from the collection of use scenarios in the previous section. At the end of the section there a classification of them to distinguish between those who are common and important for several scenarios, from those that are specific to a scenario and those who are also specific to a particular application.

This list of user requirements is summarized in a table at the end of this section.

4.1 Collaborative learning

4.1.1 Scenario: Using Asynchronous Multimedia Collaborative Tools and Grid resources to support Collaborative Learning

The user may:

- Request to login as educator or learner.
- Need to manage groups (Educator).
- Search for a streaming server to upload - store and stream educational videos (Educator)
- Search for a file server to upload and store other Multimedia content (Educator).
- Request for video processing. (Educator)
- Need to share files with other users. (Educator, Learner)
- Need to communicate and collaborate with other users. (Educator, Learner)
- Need to have access to multimedia content. (Educator, Learner)

4.1.2 Scenario: Collaborative simulation of computer networks

The following user requirements can be identified in this scenario:

- **Role management.** Users must be allowed to perform different types of tasks according to the specific role they play in the scenario.
- **Group management.** The educator must be able to create, populate, modify or eliminate the groups of students that will participate in the experience.
- **Awareness.** Participants must be provided with information that helps them to track what their group partners are doing with the aim of improving the efficiency of collaboration.
- **Responsiveness.** Simulations must be performed in a short period of time so that the proposed collaborative learning scenario can be carried out in a single laboratory session. This requirement must be accomplished even in the case of highly complex simulations. In this way, students will be able to study networking scenarios that can only be illustrated with complex simulations.
- **Single sign-on authentication.** Users should not be bothered with multiple authentication procedures in order to participate in the proposed scenario.

4.1.3 Scenario: Distributed Collaborative Software Project Development

Basic user requirements are:

- **Version management.** Users require to store and share pieces of code, but they must be properly versioned.
- **Role management.** Users must be allowed to perform different types of tasks according to the specific role they play in the scenario.
- **Group management.** The tutor (educator) must be able to create, populate, modify or eliminate the groups of students that will participate in the experience.
- **Awareness.** Participants must be provided with information that helps them to track what their group partners are doing with the aim of improving the efficiency of collaboration.

- **Step-by-step execution.** Communication between the debugger and the execution environment will be driven by events, and the debugger will be able to access the state of the execution environment.
- **Single sign-on authentication.** Users must be identified as belonging to the virtual organization, and given the adequate access rights. They should not be bothered with multiple authentication procedures in order to participate in the proposed scenario.

Other service requirements are:

- **Discovery services.** These services have to provide a standard interface. They must allow metadata-based queries against the registries.
- **Service Group services.** They are used for the realization of registries and repository.
- **Monitoring services.** Among them, we need to have:
    - *Resource Monitoring services*, which monitor the status of the resources.
    - *Application/Service Monitoring services*, which monitor the domain specific parameters of the SLA.
    - *Learning Activities Monitoring services*, which are used to monitor the job list of a learner.
- **Resource Management services.** They provide load balancing, reservation and scheduling of jobs. Organizations have their own resource managers that dynamically allocate the resources. The VO has a VO Resource Manager, which manages organizations and asks for the reservation of organization resources.
- **Fault Management services.** These services have to manage the faults. On the basis of recovery policies, they take appropriate decisions invoking the Resource Manager. Their interfaces act as a notification sinks that are able to obtain notification events from the systems that generate faults.
- **Orchestration service**. Organizations provide services that are composition of other services. Composed services are developed as workflows, which are parsed and executed by an engine. The orchestration service acts as the workflow engine. It exposes an interface that takes in input BPEL4WS scripts enhanced with metadata used to manage the grid features of the services to compose.
- **Security services.** They correspond to AAA, Trust management, Certification and Digital and Intellectual rights functionalities. They must be able to manage, and interact with proprietary security systems.
- **Collaboration/Communication services**. They are used to support collaboration and communications between services. They implement the Notification/Messaging and the P2P functionalities,
- **VO services**. They provide the functionalities for creation, management, join and resign of a VO and of groups inside a VO
- **Data, Information and Knowledge Management services.** These services have to provide the functionalities for storing, retrieving and managing of data, information and knowledge. Their interfaces should provide transparency to the user, e.g. he should be able to perform a query against a data, information or a knowledge structure in the same way. Furthermore, they need to manage replicas, ownership and consistency among replicas.

4.1.4 Scenario: Live tutorized online sessions

- Reach the assistants with an acceptable level of quality, volume, distortions and noises.
- See each participant image with certain relevance, at least during each person's presentation time.
- To know during the session each participant state: listening, listening with intention of speaking and speaking. It is also necessary to know if someone leaves the session.
- Be able to show the rest of the group elements which have been previously created and / or stored in digital mediums.

4.2 Domestic users

4.2.1 Scenario: Multimedia and entertainment spaces

- **Discovery:** There is the requirement of a discovery service in order to find the appropriate resources for each user.

- **Data sharing:** There is the requirement of a data sharing and distributed storage service**.** The service must ensure the availability of stored data even in presence of dynamism. The mechanism can exploit redundancy to achieve its goal.
- **Roles**: Although home users do not need privileges over other home users, some roles must be defined in order to protect access to some content in the spaces.
- **Events:** Information stored in the spaces changes dynamically so that other interested members should be notified using any events mechanism.
- **Policy:** Policies control which members of the VO have access to spaces.
- **Authentication, Authorization, and Accounting:** These are all essential for management of the individuals in the VO and establishing their privileges.
- **Instantiate new services**: Administrator users should be interested on providing new services to the community.
- **Scheduling:** Although scheduling is not a main requirement, some users could be interested on scheduling task.
- **Notification/Messaging:** Notification and messaging are critical in this very dynamic scenario. It is completely event driven.
- **Logging**: Logging is required to keep track of what happed in the past. This enables performance optimization.
- **Fault tolerance**: Fault tolerance is not critical; however users could not lose their data.

4.2.2 Scenario: Using grid resources and digital home health services to fight various diseases

The primary users (patients) may:

- Request to login in digital libraries
- Need to have access to a multimedia content list in order to have access to health related content
- Need to be able to share multimedia content with other patients.
- Need to be able to discuss with other patients and health careers through forums
- Email the content providers for help or feedback concerning his problem.

The secondary users (Content Providers) may:

- Need to provide patients with health related multimedia content and especially educational videos
- Search for a streaming server to upload - store the stream the educational videos
- Search for a file server to upload - store the multimedia content.
- Need to process and analyze the multimedia content.
- Need to provide help and answer questions raised by the patients

4.3 Small and Medium Enterprises

4.3.1 Scenario: Live sessions using collaborative tools

- Receive the speaker's voice with an acceptable level of quality, both in volume and distortion or noise.
- Receive the speaker's image towards the receiver with the adequate relevance, although the emission is fundamentally based on the voice.
- Receive the image of physical objects to the level of detail which is necessary to be able to see what is being shown. This refers to what is written directly onto the blackboard to objects that are shown to be identified.
- See previously created or stored elements digitally to be shown during the event: presentations, videos, animations, etc.
- See elements which are made on the spot. For example, results from a data collection from an experiment which has just taken place, by way of animation programs, electronic board, optical pen, etc.
- Receive the transmission with acceptable levels of quality by way of broadband which is normally used by the receiver.

- Receive the transmission using common applications, without the need of specific software and without having to install it previously to allow reception (or to make its download possible in case it is not already installed).
- Be able to download documents.
- Be able to ask the transmitter questions, either through an intervention protocol, or with the help or a third person as a moderator. Logically, both questions which are asked and their answers must reach all the receivers.
- Be able to access to a recorded version after the emission.

4.3.2 Scenario: Collaborative software development

- **Version management:** users require to store and share pieces of code, but they must be properly versioned.
- **Step-by-step execution:** communications between the debugger and the execution environment will be driven by events, and the debugger will be able to access the state of the execution environment.
- **Authentication, Authorization and Delegation:** users must be identified as belonging to the virtual organization, and given the access rights of the correct, existing user in systems such as the SourceForge CVS. Services can be delegated the right to act on behalf of the user, i.e. the CVS could initiate a transfer to the robot execution environment with the rights of a given programmer.
- **Accounting:** In this scenario will only serve to achieve fair sharing of the execution environment, but in similar scenarios with other type of providers could also serve for payment.
- **Logging**: Logging is required to keep track of the usage of services. For example, if an execution in the robot break it down, it is necessary to find out what was the cause.

4.3.3 Scenario: Earthquake prediction

- **Discovery:** Very large number of simulations and data mining tasks are dynamically launched. There is the requirement of a discovery service in order to find the appropriate resources.
- **Data sharing:** Very large databases of data history must be accessed constantly. This information is distributed over hundreds of different databases. The evolving real-time seismic data is tracked against the historical information by data mining services and used to control the boundary conditions on the simulations.
- **Roles**: There is the need to define different roles. It is very important to control the access rights to different parts of the instrumentation, data or computing resources.
- **Monitoring:** The large simulations must be monitored constantly to make sure they have the compute resources to continue. The entire grid of instruments and compute/data grid must be constantly monitored.
- **Policy:** Policies control which members of the VO have access to the databases, instruments and the simulations. Policy also defines who must be notified when an earthquake is predicted. The notification process is automatically executed.
- **Multiple security infrastructures**: Security controls who can control the on-line instruments.
- **Authentication, Authorization, and Accounting:** These are all essential for management of the individuals in the VO and establishing their privileges.
- **Instantiate new services**: Many of the services are simulation and data mining transient services. These must be instantiated on-the-fly by agents that are monitoring the data.
- **Advanced Reservation:** This is required for many of the scheduled data analysis tasks. However, the most important tasks have to be scheduled dynamically.
- **Scheduling:** Dynamic scheduling is an essential component of this scenario. Compute resources must be provisioned on-demand to satisfy the need to complete a prediction.
- **Notification/Messaging:** Notification and messaging are critical in this very dynamic scenario. It is completely event driven.
- **Logging**: Logging is required to keep track of what happed in the past. This enables performance optimization.
- **Fault tolerance**: Better than real-time prediction requires extreme fault tolerance.

- **Disaster Recovery**: Must be very fast. This may require that all computations be mirrored and very distributed.
- **The self-healing capabilities**: The entire analysis/simulation/prediction scenario must be able to correct for its own errors.

4.4 For-profit and non-profit communities

- **Roles:** some roles must be defined in order to protect access to some content in the community. In addition, the possibility to define different levels of membership should be necessary.
- **Events:** Information stored in the community changes dynamically so that other interested members should be notified using any events mechanism.
- **Policy:** Policies control which members of the VO have access to the community.
- **Authentication, Authorization, and Accounting:** These are all essential for management of the individuals in the VO and establishing their privileges.
- **Scheduling:** Although scheduling is not a main requirement, some users could be interested on scheduling task.
- **Notification/Messaging:** Notification and messaging are critical in this very dynamic scenario. It is completely event driven. Asynchronous communication should also be needed.
- **Logging/awareness:** Logging is required to keep track of what happed in the past.
- **Fault tolerance:** users should not loose their data.
- **Public/private information:** It will exist information only available to community members and information available to any Internet user.
- **Dissemination of information:** some dissemination service is required to allow community members subscribe to receive news, changes in the community, etc.
- **Management of data (text, pictures or movies):** community may handle different types of data. Some format converters may be required.

Audio and video broadcast:

- **Presence:** to know if a certain member (or group of members) of the community is connected.
- **Data sharing:** There is the requirement of a data sharing and distributed storage service. The service must ensure the availability of stored data even in presence of dynamism. The mechanism can exploit redundancy to achieve its goal.
- **Chat:** synchronous discussions among community members.
- **Discussions:** asynchronous discussions among community members.
- **Instantiate new services:** users should be interested on providing new services to the community.

Agenda:

- **Resource handling:** a member must be able to provide and withdraw resources to and from the group. The system must notify to the community members if there are not enough resources.
- **Financial support:** users need mechanisms to provide financial support to the community.
- **Exception handling:** community members need some mechanisms to know about exceptions that the system can not solve automatically. This information will help them take decisions about the system. Examples of exceptions: not enough resources, resources not temporally available, etc.
- **Conflicts:** in case the system detects a conflict that can not be automatically solved, community members need some mechanisms to know about it and solve it. Examples of conflicts: concurrent action in a stored data; concurrent contradictory orders to a deployed service, etc.

4.4.1 Scenario: Emergency handling

- **Discovery:** The coordinator should be able to discover adequate services from a simple keyword interface and then select one.
- **Monitoring:** Sensors, cameras and other devices will be monitored by the coordinator, directly or through other composed services.
- **Responsiveness:** It is critical that all requests are served in a short period of time, so that the response to the emergency is timely.
- **Single sign-on authentication:** Users should not be bothered with multiple authentication procedures.

4.4.2 Scenario: Shared virtual libraries

- **Document search:** Any user should be able to search for documents, independently of their format or location.
- **Personalization:** The searches of each user may be personalized by using explicit user preferences or his usage history.
- **Document addition and removal:** The library administrator must be able to add and remove documents to her institution library.
- **Format conversion:** there must be format converters (including voicing services) that permit that the user retrieves the file in his preferred format.
- **Single sign-on authentication:** Users should not be bothered with multiple authentication procedures.
- **Availability:** Users want to have permanent access to all documents. Particularly, once they know a document is in the virtual library, they will assume it is always there, and it will always be available.
- **Responsiveness:** It is not as critical as in other scenarios, yet users expect they can get answers in a short time.

4.5 Summary and discussion

The user requirements raised in each scenario are diverse. To summarize, the collection of requirements have been grouped in several categories: infrastructure, people and groups, communications, data, application-specific, and operational. These requirements have been also classified in terms of generic, storage, sharing and application-specific requirements. The following table presents the result of this classification.

User requirements	Generic	Storage	Sharing	App.spec.
INFRASTRUCTURE RELATED				
Single sign-on authentication, Authorization, Accounting, Delegation	x			
Scheduling	x			
Policy	x			
Logging	x			
Accounting	x			
Instantiate new services	x			
Discovery	x			
Monitoring	x			
Resource handling	x			
Advanced Reservation	x			
Financial support	x			
PEOPLE AND GROUPS RELATED				
Roles	x			x
Group management	x			x
Request to login as educator or	x			x
COMMUNICATION RELATED				
Notification/Messaging	x			x
Events	x			
Awareness and presence	x			x
Dissemination of information		x		x
DATA RELATED				
Data sharing, public/private		x	x	
Conflicts			x	
Version management		x	x	
Access to multimedia content		x		
Management of multimedia data		x		
Document search		x		
Document addition and removal		x		
SPECIFIC APPLICATION RELATED				
Chat				x
Audio and video broadcast				x
Audio and video communication				x
Agenda				x
Discussions				x
High performance processing				x
Request for video processing				x
Search for a streaming server to upload - store and stream videos				x
File server to upload and store Multimedia content				x
Step-by-step execution of simulations				x
Format conversion				x
OPERATIONAL				
Responsiveness	x			
Personalization	x			
Multiple security infrastructures	x			
Communicate and collaborate among people with high quality	x			
Resilience related:	x			
Resilience	x			
Self-healing capabilities	x			
Disaster Recovery	x			
Availability	x			
Exception handling	x			

5 Infrastructure requirements

This section presents and summarizes the translation of user requirements into a list of key/core requirements on the Grid4All architecture and Grid4All modules. For each scenario, the operational and functional requirements for the infrastructure have been defined, and finally they are summarized in the last subsection.

The selection of key applications to develop or adapt as demonstrators for this project is left for D4.2 (associated to T2.2).

5.1 Collaborative learning

5.1.1 Scenario: Using Asynchronous Multimedia Collaborative Tools and Grid resources to support Collaborative Learning

This scenario requires the following services to be provided by the grid infrastructure:

- **File Access Service.** The file access service is required to check if the user has the proper rights to access a requested video clip, annotation or text file.
- **Search Engine for Grid Service.** This service is responsible to search for available Grid Services such as:
    - *Search for Streaming Server.* A Streaming server responsible to store and stream video-clips is required.
    - *Search for File Server.* A file server responsible to store Multimedia content and other important files (e.g. additional related text files, images etc) is also required.
- **Authentication and authorization Service.** An authentication and authorization service is required in order to authenticate users.
- **Video Processing Service.** A service responsible to process the video-clips is required. Processing includes converting to various video formats and encoding to various bitrates to support different bandwidth etc.
- **Upload Service.** An Upload Service for both annotations and video-clips is also required. This service is responsible for the successful uploading of annotations, video clips and other files to the File and Streaming server respectively.
- **Security.** Authentication is required to access both the AMC tool and the videos stored in the streaming server provided by the grid.
- **Storage.** Data including the educational videos and other files needs to be physically stored.
- **Accounting.** An accounting service is required so that service and content providers can charge the academic institutions for their requests.

5.1.2 Scenario: Collaborative simulation of computer networks

This scenario imposes the following requirements on the grid infrastructure:

- **Virtual organization management.** This functionality is required in order to enable the creation and management of the virtual organizations in whose context this scenario will take place. A course, class, group and all associated assets.
- **Authentication and authorization.** These functionalities are needed in order to establish the privileges of each user according to their roles.
- **Delegation.** Delegation mechanisms must be implemented in order to enable single sign-on authentication.
- **Service discovery.** The infrastructure must provide this functionality in order to enable the discovery of suitable resources to carry out the simulations as well as to store the data files employed in this scenario.
- **Scheduling.** Scheduling facilities are required in order to decide which resources can be employed to assure that simulations are carried out in a period of time as short as possible.

- **Monitoring.** Simulations must be monitored in order to make sure that they progress adequately so that the results will be available in the expected period of time. If that is not the case, the scheduling of the simulation can be changed (e.g. by adding extra computational resources).
- **Fault tolerance.** Fault tolerance mechanisms should be implemented by the infrastructure in order to easily recover from failures concerning the resources employed for simulations and data storage.
- **Accounting.** This functionality is required in the case that companies act as service providers in the scenario in order to enable them to charge other organizations for the use of their services.

5.1.3 Scenario: Distributed Collaborative Software Project Development

This scenario imposes the following basic requirements on the grid infrastructure:

- **Virtual organization management.** This functionality is required in order to enable the creation and management of the virtual organizations in whose context this scenario will take place.
- **Authentication and authorization.** These functionalities are needed in order to establish the privileges of each user according to their roles.
- **Delegation**. Delegation mechanisms must be implemented in order to enable single sign-on authentication.
- **Service Coordination.** Collections of services need to be carefully coordinated. Resource brokers must ensure compute and data storage resources. Network bandwidth must be available for on-time simulation and analysis. So these different types of brokers must be carefully coordinated.
- **Service discovery.** The infrastructure must provide this functionality in order to enable the discovery of suitable resources to carry out the software project development as well as to store the data files employed in this scenario.
- **Security**. Authentication is required by all members of the VO. However careful authorization polices with govern who has access to specific resources such as data or instruments. For example, not every VO member can be allowed to control an instrument.
- **Scheduling**. Scheduling facilities are required in order to decide which resources can be employed to assure that each project subsystem is carried out in a period of time as short as possible.
- **Monitoring**. Monitoring is needed by the resource brokers in order to provision the needed resources on time. Project subsystems must be also monitored in order to make sure that they progress adequately so that the results will be available in the expected period of time. If this is not the case, the scheduling of the subsystems can be modified (e.g. by adding extra computational resources).
- **Messaging**, Queuing and Logging. Messaging and event systems are needed because of the type of communication between elements of the subsystems as well as of the very dynamic "demand driven" nature of the application workflow. Logging services are needed to understand what went wrong.
- **Reservation Services**. Dynamic and on demand reservation (of resources). Any particular scheduling policy could be applied (FIFO, priority based, equal usage among programmers…).
- **Brokering and Scheduling Services.** Compute and data resource brokering services are needed. Scheduling and co-scheduling services will be needed.
- **Fault tolerance**. Fault tolerance mechanisms should be implemented by the infrastructure in order to easily recover from failures concerning the resources employed for project subsystems and data storage.

5.1.4 Scenario: Live tutorized online sessions

This scenario requires the following services to be provided by the grid infrastructure:

- **Authentication and authorization Service.** An authentication and authorization service is required in order to authenticate users and establish privileges of each user according its role on the session.
- **Storage.** All documents showed and used during the meeting should be stored on a server in order to access them during and after the session.
- **Accounting.** An accounting service may be required to enable charges associated to the possible use of external services.

5.2 Domestic users

5.2.1 Scenario: Multimedia and entertainment spaces

- **Name Resolution and Discovery**: Severe store modelling must be able to discover data resources and data catalogs from metadata descriptions. This is part of discovery.
- **Service Coordination**: Collections of services need to be carefully coordinated. Resource brokers must assure compute and data storage resources. Network bandwidth must be available for real time multimedia. So these different types of brokers must be carefully coordinated.
- **Security**: Authentication is required by all members of the VO. Spaces must be secure against non-authorized users.
- **Policy enforcement**: Policy issues involve services access and spaces security. Policies will determine when a particular user exceeds her rights.
- **Data Management**: metadata catalogs, directory and index services, wide access to data archives, management or virtual repositories of data.
- **Messaging, Queuing, Monitoring and Logging:** Messaging and event systems are needed because of the very dynamic "demand driven" nature of the application workflow. Logging services are needed to understand what went wrong. Monitoring is needed for resource allocation purposes.
- **Events:** Actions occurred within a particular space are notified to authorized users by an events mechanism. Information stored in the spaces changes dynamically so that other interested members have to be notified using any events mechanism.
- **Provisioning and Resource Management**: Resource requirements change on a very dynamic basis. In the case of real time multimedia must be possible to provision very large amounts of compute, bandwidth and data resources.
- **Reservation Services**: Reservation is not a main requirement, however real-time multimedia processing can be provided if task are scheduled in respect to resources.
- **Brokering and Scheduling Services**: Compute and data resource brokering services are needed. Resources must be managed dynamically. Services must be allocated to dynamic resources.
- **Fault Handling Services**: System redundancy is the best option.

5.2.2 Scenario: Using grid resources and digital home health services to fight various diseases

This scenario imposes the following requirements on the grid infrastructure:

- **File Access Service:** The file access service is required to check if the user has the proper rights to access a requested video clip or other multimedia content. Additionally this service is also required by the content providers in order to set access rights to the content they provide.
- **Search Engine for Grid Service:** This service is responsible to search for available grid services.
  - **Search for Streaming Server**: A Streaming server responsible to store digital video and to serve the users is required.
  - **Search for File Server**: A file server responsible to store multimedia content is also required.
- **Authentication Service**: An authentication service is required in order the main users to connect to various digital video libraries provided by the content providers
- **Media sharing:** There is the requirement of a data sharing and distributed storage service. The service must ensure the availability of stored data even in presence of dynamism. The mechanism can exploit redundancy to achieve its goal.
- **Media Processing Service**: A service responsible to process the Multimedia content is required. This service may be required by the content providers. Processing includes editing and encoding of digital video, audio and image.
- **Forum Service:** A forum service that will enable patients, content providers and health careers to communicate is required.
- **Security**: Authentication is required to access the multimedia content stored in digital libraries provided by the content providers.
- **Storage**: Multimedia content needs to be physically stored.

- **Accounting**: An accounting service is required so that service and content providers can charge the academic institutions for their requests.

5.3 Small and Medium Enterprises

5.3.1 Scenario: Live sessions using collaborative tools

This scenario requires the following services to be provided by the grid infrastructure:

- **Authentication and authorization Service.** An authentication and authorization service is required in order to authenticate users and establish privileges of each user according its role on the session.
- **Storage.** All documents showed and used during the meeting should be stored on a server in order to access them during and after the session. Also recorded versions should be stored.
- **Accounting.** An accounting service may be required to enable charges associated to the possible need of external services.

5.3.2 Scenario: Collaborative software development

- **Service Coordination**: This is a simple case in which only the CVS and the execution environment need to be coordinated, for the installation of code or for the storage of execution traces.
- **Concurrent access:** the CVS must allow many users access different pieces of code at the same time, or even the same piece of code, with the appropriate access rights.
- **Security**: Authentication is required to access the CVS or the robot execution environment.
- **Messaging and logging:** Messaging and event systems are needed because of the type of communications between elements of the systems. Logging services are needed to understand what went wrong.
- **Monitoring and reservation**: Monitoring is needed in order to attend the reservation demands. Any particular scheduling policy could be applied (FIFO, priority based, equal usage among programmers…).
- **Storage:** Data needs to be physically stored. The CVS system could actually be a composed service using third party storage.

5.3.3 Scenario: Earthquake prediction

- **Name Resolution and Discovery**: Severe store modelling must be able to discover data resources and data catalogs from metadata descriptions. This is part of discovery.
- **Service Coordination**: Collections of services need to be carefully coordinated. Resource brokers must assure compute and data storage resources. Network bandwidth must be available for on-time simulation and analysis. So these different types of brokers must be carefully coordinated.
- **Security**: Authentication is required by all members of the VO. However careful authorization polices with govern who has access to specific resources such as data or instruments. For example, not every VO member can be allowed to control an instrument.
- **Policy enforcement**: Policy issues primarily involve access to instruments. Also, policies will determine when a particular running system of services will be allowed to preempt resources for what "it perceives" as a critical need for public safety. This is particularly relevant in this scenario wehre instruments are involved.
- **Data Management**: metadata catalogs, directory and index services, wide access to data archives, virtual data management.
- **Messaging, Queuing, Monitoring and Logging:** Monitoring is needed by the resource brokers in order to provision the needed resources on time. Messaging and event systems are needed because of the very dynamic "demand driven" nature of the application workflow. Logging services are needed to understand what went wrong.
- **Events:** When a dangerous pattern is detected, an event is triggered. The system need to start the simulations and other data mining applications.

- **Reservation service brokering and scheduling**: Workflow engines have to handle simulations. Allocate resources when needed.
- **Provisioning and Resource Management**: Resource requirements change on a very dynamic basis. In the case of emergencies it must be possible to provision very large amounts of compute, bandwidth and data resources.
- **Reservation Services**: Dynamic and on demand reservation. Simulation can be scheduled in advance or triggered when something occurs.
- **Brokering and Scheduling Services**: Compute and data resource brokering services are needed. Scheduling and co-scheduling services will be needed.
- **Fault Handling Services**: System redundancy is the best option.

5.4 For-profit and non-profit communities

5.4.1 Scenario: Communities

- **Name Resolution and Discovery:** Resources are dynamically provided by members of the community, such dynamism requires naming services and discovery mechanisms in order to find appropriate resources that fit users' needs.
- **Service Coordination:** Collections of services need to be carefully coordinated. Resource brokers must assure compute and data storage resources. Network bandwidth must be available for real time multimedia. So these different types of brokers must be carefully coordinated.
- **Decentralized security:** Authentication is required by all members of the VO. The community must be secure against non-authorized users.
- **Policy enforcement:** Policy issues involve services access and community security. Policies will determine when a particular user exceeds her rights.
- **Data Management:** metadata catalogs, directory and index services, wide access to data archives, virtual data management.
- **Messaging, Queuing, Monitoring and Logging:** Messaging and event systems are needed because of the very dynamic "demand driven" nature of the application workflow. Logging services are needed to understand what went wrong and to know what is going on in the group. Monitoring is needed for resource allocation purposes.
- **Events:** Actions occurred within a particular community are notified to authorized users by an events mechanism. Information stored in the community changes dynamically so that other interested members have to be notified using any events mechanism.
- **Provisioning and Resource Management:** Resource requirements change on a very dynamic basis. In the case of real time multimedia must be possible to provision very large amounts of compute, bandwidth and data resources.
- **Reservation Services:** Reservation is not a main requirement, however real-time multimedia processing can be scheduled to enhance performance of the system.
- **Brokering and Scheduling Services:** Compute and data resource brokering services are needed. Resources must be managed dynamically. Services must be allocated to dynamic resources.
- **Fault Tolerance Services:** System redundancy is the best option.
- **Accounting and Pricing mechanism:** Communities pay money to rent resources. Price of resource must be determined. Accounting and pricing mechanisms are required.
- **Presence:** to know if a certain member is connected (or group of members).
- **Exception handling:** notify about the occurrence of an exception. Also provide mechanisms to solve it.
- **Conflicts:** notify about the occurrence of a conflict when participants interact in a many-to-many basis. Also provide mechanisms to solve it.

5.4.2 Scenario: Emergency handling

This scenario imposes the following requirements on the grid infrastructure:

- **Service Coordination**: in this example, service coordination is critical, since there are many services (disaster graphical map, disaster simulator…) consuming data from other services (local map storage, local sensors…), or using computational capabilities of other providers.
- **Authentication and authorization:** These functionalities are needed in order to establish the privileges of each user according to their roles.
- **Delegation:** Delegation mechanisms must be implemented in order to enable single sign-on authentication.
- **Service discovery:** The infrastructure must provide this functionality in order to enable the discovery of suitable resources according to the keywords typed by the coordinator, but also automatically from the type of emergency signal.
- **Scheduling:** Scheduling facilities are required in order to decide which resources can be employed, guaranteeing appropriate treatment of priorities, so that emergency needs are served first.
- **Fault tolerance:** Fault tolerance mechanisms should be implemented by the infrastructure in order to easily recover from failures concerning the resources employed for simulations and data storage.
- **Accounting:** This is required so that service providers can properly charge the 122 organization for the use of their services.

5.4.3 Scenario: Shared virtual libraries

This scenario imposes the following requirements on the grid infrastructure:

- **Service Coordination**: this is particularly important since document indices, storage and format conversion will be separate services, but they will have to interact.
- **Storage and replication**: documents will be given to storage providers, but sometimes they may be replicated to improve availability or reduce the retrieval time.
- **Authentication and authorization:** These functionalities are needed in order to establish the privileges of each user according to their roles.
- **Delegation:** Delegation mechanisms must be implemented in order to enable single sign-on authentication.
- **Service discovery:** Some services have to look for others to complete their functions, as for example for a given formatter according to the desired output format, or for a voiced of a certain language and with specific levels of quality of service.
- **Scheduling:** Scheduling is required to reserve resources or to plan complex searches.
- **Accounting:** This is required so that some service providers can charge the academic institutions for their requests.

5.5 Summary and discussion

The following table provides a summary of the infrastructure requirements for each scenario:

Scenarios	Collab. learning				Domestic		SME			Communities		
Infrastructure requirements	Async mm	Collab sim	SwDevPrj	LiveTutor	MM&Entr	Health	LiveSess	SoftDev	EQPred	Commts	EmgHdlg	SVLibs
VO related												
Virtual organization management		x	x			x						
Accounting	x	x		x		x	x				x	x
Authentication, authorization, delegation	x	x	x	x	x	x	x	x	x	x	x	x
Presence						x	x			x		
Storage related												
Concurrent access (& conflict handling)				x			x	x		x		
Data Management (& file access and storage)	x			x	x	x	x	x	x	x		x
Logging					x		x	x	x	x		
Messaging related												
Events and messaging			x		x	x			x	x		
Component management related												
Fault and exception handling		x	x		x	x			x	x	x	
Monitoring		x	x		x	x		x	x	x		
Name Resolution and Discovery					x				x	x		
Allocation related												
Provisioning and reservation			x		x	x	x	x	x	x		
Scheduling		x	x		x	x			x	x	x	x
Service related												
Service Coordination			x		x			x	x	x	x	x
Service discovery	x	x	x			x					x	x

The infrastructure requirements are grouped and summarized across differences in granularity and terminology on each description of scenario, but it gives an overview of the diversity of requirements raised by the collection of scenarios. The mark “x” or the lack of it cannot be seen as a strict judgement. It reflects the judgement of the author that a certain requirement is strongly relevant or little relevant to a given scenario.

6 Conclusions

The generic use cases proposed in the description of work for Grid4All have been explored in detail for some selected scenarios:

- Collaborative learning: "Using Asynchronous Multimedia Collaborative Tools and Grid resources to support Collaborative Learning", "Collaborative simulation of computer networks", "Distributed Collaborative Software Project Development" and "Live tutorized online sessions".
- Domestic users: "Multimedia and entertainment spaces", "Using grid resources and digital home health services to fight various diseases".
- Small and medium enterprises: "Live sessions using collaborative tools", "Collaborative software development", "Earthquake prediction".
- For-profit and non-profit communities: "Communities", "Emergency handling", "Shared virtual libraries".

From each of these scenarios, user requirements and infrastructure requirements have been derived. These requirements are summarized in tabular form at the end of each section.

The summary of user requirements reflects how the selection of scenarios covers the different aspects of application and infrastructure, and shows a non exhaustive but representative collection of requirements, some of them specific to a single application, some other specific to a group of applications, and many of them common to several scenarios across different main use cases.

The summary of infrastructure requirements shows that the focus on providing certain infrastructural functions within the project is really needed and the needs are shared across several application scenarios considered in the project and exemplified by the 12 scenarios studied in task T4.1.

7 References

[1] Allison, C., Cerri, S.A., Ritrovato, P., Gaeta, A. et al., "Services, semantics and standards: elements of a learning grid infrastructure," Applied Artificial Intelligence Journal, vol. 19, no. 8, 2005.

[2] AstroGrid, URL=http://www2.astrogrid.org/science

[3] Bardeen, M., E. Gilbert, T. Jodran, P. Nepywoda et al., "The QuarkNet/Grid Collaborative Learning e-Lab," Future Generation Computer Systems, vol. 22, no. 6, 2006, pp. 700-708.

[4] Baset, S.A., Schulzrinne, H. An Analysis of the Skype Peer-to-Peer Internel Telephony Protocol, Arxiv preprint cs.NI/0412017, 2004.

[5] Bengner, S. et al. (2002). Numerical simulation for eHealth: Grid enabled Medical Simulation Services

[6] Bote-Lorenzo, M. L. and Ardaiz-Villanueva, O. "Special section: Collaborative and learning applications of grid technology (Guest Editorial)," Future Generation Computer Systems, vol. 22, no. 6, 2006, pp. 699-699.

[7] Bote-Lorenzo, M. L., D. Hernández-Leo, Y. A. Dimitriadis, J. I. Asensio-Pérez et al., "Towards reusability and tailorability in collaborative learning systems using IMS-LD and grid services," International Journal on Advanced Technology for Learning, vol. 1, no. 3, 2004, pp. 129-138.

[8] Bowie, M. M. “Instructional film research and the learner,” Proceedings of the Annual Convention of the Association for Educational Communications and Technology, 1986.

[9] Butterfly (2003). Butterfly.net, Butterfly.net: Powering Next-Generation Gaming with On-Demand Computing, IDC case study. URL: http://www-1.ibm.com/GRID/pdf/butterfly.pdf

[10] CROSSGRID, URL=http://www.crossgrid.org/products/applications/meteo_pollution.html

[11] EdutainGrid, URL=http://www.edutaingrid.eu/

[12] Fox, G. "Education and the enterprise with the grid," Grid computing: making the global infrastructure a reality, F. Berman et al., eds., John Wiley & Sons, 2003, pp. 963-976.

[13] Gaynor, M., Moulton, S. L., Welsh, M., LaCombe, E., Rowan, A., and Wynne, J. 2004. Integrating Wireless Sensor Networks with the Grid. *IEEE Internet Computing* 8, 4 (Jul. 2004), 32-39. DOI= http://dx.doi.org/10.1109/MIC.2004.18

[14] Gullapalli, S., Dyke, S., Hubbard, P., Marcusiu, D., Pearlman, L., and Severance, C. 2004. Showcasing the Features and Capabilities of NEESgrid: A Grid Based System for the Earthquake Engineering Domain. In *Proceedings of the 13th IEEE international Symposium on High Performance Distributed Computing (Hpdc'04) - Volume 00* (June 04 - 06, 2004). HPDC. IEEE Computer Society, Washington, DC, 268-269. DOI= http://dx.doi.org/10.1109/HPDC.2004.31

[15] Harmer, T. et al., GridCast: A Service Architecture for the Broadcasting Media

[16] Ho, H., Yang, C. and Chang, C., "Building an e-learning platform by Access Grid and Data Grid technologies," Proceedings of the International Conference on e-Technology, e-Commerce and e-Service (EEE'04), pp. 452-455.

[17] IBM, IBM Grid Medical Archive Solution designed to provide storage optimization with multi-site availability to patient diagnostic images, URL: http://www-1.ibm.com/grid/pdf/gmascustomerbrief.pdf

[18] IBM, Whitepaper: Optimizing Medical Imaging With Grid, URL: http://www-1.ibm.com/grid/pdf/wp_opt_med_imging.pdf

[19] Joch, A. (2004). Grid gets down to business. Network World, December 2004 www.networkworld.com/power/2004/122704techgrid.html

[20] Media Grid, URL=http://mediagrid.org

[21] Open Grid Forum, Industry applications. URL: http://www.ogf.org/gf/group_info/areasgroups.php?area_id=10

[22] Park, D. S., J. Leigh, A. E. Johnson, B. Carter et al., "Distance learning classroom using Virtual Harlem," Proceedings of the 7th International Conference on Virtual Systems and Multimedia, pp. 489-498.

[23] Pereira, L.A., R. N. Melo, F. Porto, and B. Schulze, "A workflow-based architecture for e-learning in the grid," Proceedings of the 1st International Workshop on Collaborative Learning Applications of Grid Technology (CLAG 2004), held in conjunction with the International Symposium on Cluster Computing and the Grid 2004 (CCGRID 2004), 2004, pp. 208-209.

[24] Rosenberg, D. (2005). Building the business case for grid. IBM Developer Website, June 2005 01 Jun 2005 http://www-128.ibm.com/developerworks/grid/library/gr-bizcase/

[25] Salomon, G. Interaction of Media, Cognition, and Learning, Lawrence Erlbaum Associates, 1994, p. 16.

[26] Severin, W. J. “Another look at cue summation,”. Audio Visual Communications Review, vol. 15, 1967, pp. 233-245.

[27] Shaikh, A., Sahu, S., Rosu, M., Shea, M. Saha, D. Implementation of a service platform for online games, Proceedings of 3rd ACM SIGCOMM workshop on Network and system support for games table of contents, Pages: 106 – 110, 2004.

[28] Shyu, F. M., Y. F. Liang, W. T. Agnes, J. J. Luh et al., "A problem-based e-learning prototype system for clinical medical education," Medinfo, vol. 11, no. 2, 2004, pp. 983-987.

[29] Simson Garfinkel, Commodity Grid Computing with Amazon's S3 and EC2, ;login:, The USENIX Magazine, February 2007, Volume 32, Number 1.

[30] Smarr, L. "Grids in context," The Grid 2: blueprint for a future computing infrastructure, I. Foster et al., eds., Morgan Kaufmann Publishers, 2004,

[31] Virtual spaces on Internet for social interaction. URL= http://spaces.live.com/, http://360.yahoo.com/, http://ma.gnolia.com/, http://del.icio.us/, http://digg.com/

[32] Young, C. and M. Asensio, M, "Looking through Three 'I's: the Pedagogic Use of Streaming Video," Proceedings of the Third International Conference on Networked Learning, 2002, pp. 628-635.