Grid Characteristics and Uses: a Grid Definition

Miguel L. Bote-Lorenzo, Yannis A. Dimitriadis, and Eduardo Gómez-Sánchez [*]

School of Telecommunications Engineering, University of Valladolid
Camino Viejo del Cementerio s/n, 47011 Valladolid, Spain
{migbot, yannis, edugom}@tel.uva.es

Abstract. This paper discusses the concept of grid towards achieving a complete definition using main grid characteristics and uses found in literature. Ten definitions extracted from main literature sources have been studied allowing the extraction of grid characteristics while grid uses are defined in terms of the different types of application support provided by grids. A grid definition is proposed using these characteristics and uses. This definition may be very useful to determine the limits of the grid concept as well as to explore new application fields in grid computing. In this sense, the extracted characteristics are employed to determine the potential benefits a grid infrastructure may provide to Computer Supported Collaborative Learning applications.

1 Introduction

The origin of the grid concept may be established in the early 90's when great efforts were made to deploy several gigabit testbeds such as CASA [19] that linked super-computing sites across the USA. However, the term grid still did not exist a that time, and the terms metasystem and metacomputer where used instead referring to "the computing resources transparently available to the user via a networked environment" [21].

The successes of these testbeds inspired the 1995 I-WAY experiment [4] which provided a large-scale testbed (multiple supercomputers, advanced visualization de-vices and mass storage systems at 17 different places within North America connected through ATM networks) in order to allow mainly high-performance applications to be studied in a controlled environment. The I-WAY experiment in turn motivated some very important projects such as the National Technology Grid, which coined in 1997 the term grid, "a name derived from the notion of the electrical power grid", to call the "truly U.S. national-scale advanced computational infrastructure" they envisioned [22].

It was not until 1998 that grid computing entered the mainstream of research with both the establishment of the Global Grid Forum (www.gridforum.org) by early grid developers and practitioners, and the publication of the book [10] that laid the groundwork of the field and defined the term computational grid as

[*] This work has been partially funded by Spanish Ministry of Science and Technology projects TIC2002-04258-C03-02, TIC2000-1054 and Castilla y León Regional Government project VA 117/01.

Postproc. of the First European Across Grids Conference (ACG'03), Springer-Verlag LNCS 2970, pp. 291-298, Santiago de Compostela, Spain, Feb. 2004

"a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities" [6]. But this entrance turned into not only the intensification of the research on the field but also the conflation of the term grid that does not seem to have a clear and complete definition, i.e. a definition highlighting all main characteristics and uses of a grid. Instead, definitions found in literature usually highlight them only partially (and sometimes poorly).

Having a complete grid definition built using all main characteristics and uses may be considered important for several reasons. First, it is indispensable to exactly determine whether a given technology can be considered to be a grid or not [7]. In this sense, a grid definition can show how grid technologies can be found in application fields other than supercomputing. In addition, it allows studying the potential benefits that the use of a grid can report to application fields still unexplored by grid researchers. For example, the main project supporting this research is devoted to the use of grid infrastructure in Computer Supported Collaborative Learning (CSCL) applications.

The lack of a complete grid definition has already been detected in literature [7, 11]. Significantly, very recent works [14] challenge the already existing definitions of grid. However, such a complete definition has not been given yet.

Thus, the aim of this paper is to advance towards a complete grid definition made up of all main grid characteristics and uses found in grid literature. Hence, grid characteristics have been extracted from grid definitions provided in main literature sources while grid uses are defined in terms of the types of application support supplied by grids according to literature. Also, we perform an initial step towards the definition of the grid concept within the field of CSCL, in order to evaluate its suitability, limitations and potential benefits.

This paper is organized as follows. Section 2 examines ten literature definitions in order to find out the main characteristics that a grid is supposed to have. Section 3 categorizes the different types of support that grids can provide to applications. Both the list of grid characteristics and uses are employed in section 4 to build grid definition. Section 5 explores the field of CSCL within the context of grid, as it was de-fined in the previous sections and based on our research work in this domain. Conclusions may be found in section 6.

2 Main Grid Characteristics

Ten definitions extracted from main grid literature sources have been examined to find out the essential characteristics that a grid is supposed to have in order to be considered as such. As a result, a total number of ten characteristics has been identified. Both the definitions and the characteristics found in them, either explicitly or implicitly are shown in Table 1 together with their references. These characteristics may be described as follows:

- *Large scale*: a grid must be able to deal with a number of resources ranging from just a few to millions. This raises the very serious problem of avoiding potential performance degradation as the grid size increases.

- *Geographical distribution*: grid's resources may be located at distant places.
- *Heterogeneity*: a grid hosts both software and hardware resources that can be very varied ranging from data, files, software components or programs to sensors, scientific instruments, display devices, personal digital organizers, computers, super-computers and networks.
- *Resource sharing*: resources in a grid belong to many different organizations that allow other organizations (i.e. users) to access them. Nonlocal resources can thus be used by applications, promoting efficiency and reducing costs.
- *Multiple administrations*: each organization may establish different security and administrative policies under which their owned resources can be accessed and used. As a result, the already challenging network security problem is complicated even more with the need of taking into account all different policies.
- *Resource coordination*: resources in a grid must be coordinated in order to provide aggregated computing capabilities.
- *Transparent access*: a grid should be seen as a single virtual computer.
- *Dependable access*: a grid must assure the delivery of services under established Quality of Service (QoS) requirements. The need for dependable service is fundamental since users require assurances that they will receive predictable, sustained and often high levels of performance.
- *Consistent access*: a grid must be built with standard services, protocols and inter-faces thus hiding the heterogeneity of the resources while allowing its scalability. Without such standards, application development and pervasive use would not be possible.
- *Pervasive access*: the grid must grant access to available resources by adapting to a dynamic environment in which resource failure is commonplace. This does not imply that resources are everywhere or universally available but that the grid must tailor its behavior as to extract the maximum performance from the available re-sources.

3 Main Grid Uses

Opposite to what is often believed, the grid is not only a computing paradigm for providing computational resources for grand-challenge applications. Instead, it is an infrastructure that bonds and unifies globally remote and diverse resources in order *to provide computing support* for a wide range of applications. It is important to notice that grid uses are thus not defined in terms of applications (as usually found in the literature) but rather of the support the grid provides.

The different types of computing support offered by grids can be categorized according to the main challenges that they present from the grid architecture point of view. This categorization is the following:

- *Distributed supercomputing support* allows applications to use grids to couple computational resources in order to reduce the completion time of a job [18] or to tackle problems that cannot be solved on a single system [6]. The

Table 1. Grid definitions from the literature listed in chronological order, and features they mention among **(1)** large scale, **(2)** geographical distribution, **(3)** heterogeneity, **(4)** resource sharing, **(5)** multiple administration, **(6)** resource coordination, **(7)** transparent access, **(8)**, dependable access, **(9)** consistent access, and **(10)** pervasive access. Features marked with * are implicit rather than explicit. The definition that coined the term grid is highlighted in bold letters

Definition	Features
"The computing resources transparently available to the user via this networked environment have been called a metacomputer" [21]	7
"A metasystem is a system composed of heterogeneous hosts (both parallel processors and conventional architectures), possibly controlled by separate organizational entities, and connected by an irregular interconnection network" [16]	3,4*,5
"Metasystem is a wide-area environment in which users operate transparently, consisting in workstations, PCs, graphics-rendering engines, supercomputers and nontraditional computing devices such as televisions" [15]	2,3*,7
"Networked virtual supercomputers, or metacomputers, are execution environments in which high-speed networks are used to connect supercomputers, databases, scientific instruments, and advanced display devices, perhaps located at geographically distributed sites" [8]	2,3*
"The National Computational Science Alliance calls its prototype infrastructure the National Technology Grid, a name derived from the notion of the electrical power grid that transformed the U.S., and indeed the world, during the past century" [22]	
"Computational grids are large-scale high-performance distributed computing environments that provide dependable, consistent, and pervasive access to high-end computational resources" [9]	1,8,9,10
"A computational grid is a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities" [6]	8,9,10
"The real and specific problem that underlies the Grid concept is coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations" [11]	4,6,10
"A distributed network computing (NC) system is a virtual computer formed by a networked set of heterogeneous machines that agree to share their local resources with each other. A Grid is a very large scale, generalized distributed NC system that can scale to Internet-size environments with machines distributed across multiple organizations and administrative domains" [18]	1,2*,3,4,5,7*
"Grid technologies and infrastructure support the sharing and coordinated use of diverse resources in dynamic, distributed virtual organizations - that is, the creation, from geographically distributed components operated by distinct organizations with differing policies, of virtual computing systems that are sufficiently integrated to deliver the desired QoS" [12]	2,3,4,5,6,7*,8,10
"A Grid is a system that coordinates resources that are not subject to a centralized control using standard, open, general-purpose protocols and interfaces to deliver nontrivial qualities of service" [14]	5,6,8,9

main problems raised by applications requiring this support are the need to co-schedule the use of scarce and highly expensive resources, the scalability of protocols and algorithms to a large number of nodes, latency-tolerant algorithms as well as achieving high levels of performance [6]. Typical applications that require distributed supercomputing are weather forecasting and military scenario simulations.

- *High-throughput computing support* allows applications to use grids to put unused processor cycles to work in generally loosely coupled or independent tasks [6]. 'Parameter sweep' applications such as Monte Carlo simulations are well suited for high-throughput computing.
- *On-demand computing support* allows applications to use grids to retrieve re-sources that cannot be cost-effectively or conveniently located locally [6]. Challenging issues in order to provide on-demand computing support are resource lo-cation, scheduling, code management, configuration, fault tolerance, security, and payment mechanisms [6]. A financial application allowing users to perform accurate stock market analysis and price prediction employing their home desktop computer is a representative example of application requiring on-demand computing.
- *Data-intensive computing support* allows applications to use grids to synthesize new information from distributed data repositories, digital libraries and databases [6]. Challenges for the data-intensive computing support include the scheduling and configuration of complex, high-volume data flows [6]. The creation of a new database using data mined from a number of online databases would be an example of data-intensive computing application.
- *Collaborative computing support* allows applications to use the grid to enable and enhance human-to-human interactions [6] in a synchronous or asynchronous way [11] via a virtual space. The real-time requirements imposed by human perceptual capabilities as well as the wide range of many different interactions that can take place are one of the most challenging issues of collaborative computing support [6]. Typical examples of applications that may use a collaborative computing infrastructure provided by grids are groupware applications and multiconferencing applications.
- *Multimedia computing support* allows applications to use grids to deliver contents assuring end-to-end QoS [18]. Main challenges for the multimedia computing sup-port derive from the need to provide QoS across multiple different machines. Videoconference applications is a typical example of application requiring multimedia computing support.

4 A Grid Definition

According to the list of grid characteristics extracted from literature and to the different categories of support provided by grids that have been identified in this paper, a grid can be defined as *a large-scale geographically distributed hardware and software infra-structure composed of heterogeneous networked resources owned and shared by multiple administrative organizations which are* *coordinated to provide transparent, dependable, pervasive and consistent computing support to a wide range of applications. These applications can perform either distributed computing, high throughput computing, on-demand computing, data-intensive computing, collaborative computing or multimedia computing.*

5 CSCL as an Application Field within the Grid Context

The definition just presented can help to determine the suitability, limitations and potential benefits of the grid within a certain domain. One relevant field is Computer Supported Collaborative Learning (CSCL). CSCL is a discipline devoted to research in educational technologies that focus on the use of information and communications technology (ICT) as a mediational tool within collaborative methods (e.g. peer learning and tutoring, reciprocal teaching, project or problem-based learning, simulations, games) of learning [5, 23]. Therefore, it seems appropriate to explore and apply grid technology, as defined above, to this multidisciplinary and complex field. Though CSCL is closely related to groupware or CSCW (Computer Supported Collaborative Work), there are substantial differences in the context and objectives of each field. Thus, works published in CSCW literature, such as [1], focus much more on the technological support instead of a deep analysis of the domain.

CSCL applications are inherently distributed, since students have to collaborate through networked computers, though we may distinguish cases in which students are collocated (and collaboration also happens face to face) and those in which collaboration is strictly carried through the computer [3]. Besides, CSCL applications may involve synchronous or asynchronous interactions, or a mixture of them [5]. In addition, Component Based Software Engineering (CBSE) offers a flexible paradigm for distributed applications development that is being exploited for CSCL [2]. For applications with these characteristics, the computational grid can offer a suitable infra-structure that facilitates their deployment and enhances their performance. Particularly, the analysis of grid pints out several benefits for the use of grid technologies to deploy CSCL applications.

The *large scale* of the grid may allow the of a high number of single/group participants collaborating within the CSCL application. Though in purely synchronous collaboration the number of participants is usually low, in applications involving only asynchronous interactions, or a mixture, a large number of participants may be involved. For example, applications allowing children from different schools to collaborate asynchronously in order to publish and consult an electronic magazine using CSCL tools may involve a large number of users at different locations in a wide area. Moreover, the wide *geographical distribution* of grid resources should enhance the participation of users from very distant places. Again, many CSCL applications have their users co-located, but others such as distance universities tools may have users spread along one country or even a larger area. Besides, the *heterogeneous nature* of grid *shared resources* may allow users to participate employing heterogeneous devices such as computers, PDAs or electronic black-boards, as well as data or software resources such as simulators or other domain packages. This allows the use of locally unavailable computational resources for processing data (such as in Punch [17]), for visualization (like in CoViS [20]) or for content delivery (a secondary but necessary service in CSCL). Therefore, an efficient use of grid resources may benefit such needs of the CSCL domain. Concerning performance, *dependable access* provided by the grid infrastructure that offers either guaranteed or best effort service can improve application performance over *ad hoc* deployment decisions, which becomes specially important when the application is run over a wide area network, such as an synchronous brainstorming. Finally, one grid feature that will appeal CSCL application developers is *transparent access*, since it eases some serious problems such as automatic application deployment and configuration, that can be passed onto the grid scheduler and per-formed with some optimality criteria. This is especially valid either for component or web services based implementations of CSCL applications, towards which the grid seems to be shifting, as promoted by the Open Grid Services Architecture (OGSA) [13].

In summary, the complex, dynamic and contextualized nature of the CSCL field is especially motivating with respect to the use of grid. CSCL applications call for some infrastructure that allows the use of heterogeneous distributed resources in a scalable manner, and that facilitates the deployment in order to improve performance. The computational grid, as defined above, offers this support and can then be used satisfactorily for the CSCL domain. However, further study is necessary in order to precisely determine how CSCL applications can benefit from the grid, especially by providing a clear correspondence between components (or web services) present in a formal architectural framework and the corresponding grid services, eventually within the OGSA.

6 Conclusions and Future Work

Grids are mainstream technology that do not have a clear and complete definition in the literature. Establishing a complete grid definition is considered an important goal in order to allow determining the limits of the grid research field as well as exploring new fields of application in grid computing.

To tackle this problem, ten definitions extracted from main literature sources have been studied allowing the extraction of the ten main characteristics of grids. In addition, grid uses have been defined in terms of the different types of application support provided by grids according to the literature. Both grid characteristics and uses have been used to build a grid definition.

The establishment of a grid definition in this paper is not an isolated effort. On the contrary, this definition is currently being used within the framework of the main project supporting this work in order to explore the possibility of using CSCL applications within grid computational environments. In this sense, the analysis of the characteristics extracted in section 3 shows that the support provided by grid technologies may offer a number of benefits to CSCL applications, although a formal study is still necessary together with an analysis of experimental data within real systems.

References

1. Amin, K., Nijsure, S., and von Laszewski, G. Open Collaborative Grid Services Architecture (OCGSA). In *Proc. Euroweb'02*, Oxford, UK, pp. 101–107, 2002.
2. Asensio, J.I., Dimitriadis, Y.A., Heredia, M., Martínez, A., Álvarez, F.J., Blasco, M.T. and Osuna, C. From collaborative learning patterns to component-based CSCL application. In *Proc. ECSCW'03 workshop "From Good Practices to Patterns"*, Helsinki, Finland, 2003.
3. Crook, C. *Computers and the Collaborative Experience of Learning.* Routeledge, London, UK, 1994.
4. DeFanti, T., Foster, I., Papka, M., Stevens, R., Kuhfuss, T. Overview of the I-WAY: Wide Area Visual Supercomputing. *Int. J. Supercomp. App.*, **10**(2):123–130, 1996.
5. Dillenbourg, P. *Collaborative Learning: Cognitive and Computational Approaches.* Elsevier Science, Oxford, UK, 1999.
6. Foster, I. *Computational Grids*, pp. 15–52. In [10], 1998.
7. Foster, I. What Is the Grid? A Three Point Checklist. *Grid Today*, **1**(6), 2002.
8. Foster, I. and Kesselman, C. Globus: a Metacomputing Infrastructure Toolkit. *Int. J. Supercomp. App.*, **11**(2):115–128, 1997.
9. Foster, I. and Kesselman, C. The Globus Project: a Status Report. In *Proc. IPPS/SPDP'98 Workshop on Heterogeneous Computing*, pp. 4–18, 1998.
10. Foster, I. and Kesselman, C. *The Grid: Blueprint for a Future Computing Infrastructure.* Morgan Kaufmann, San Francisco, CA, 1998.
11. Foster, I., Kesselman, C., and Tuecke, S. The Anatomy of the Grid: Enabling Scalable Virtual Organizations. *Int. J. Supercomp. App.*, **15**(3):200–222, 2001.
12. Foster, I., Kesselman, C., Nick, J., and Tuecke, S. Grid Services for Distributed System Integration. *Computer*, **35**(6):37–46, 2002.
13. Foster, I., Kesselman, C., Nick, J., and Tuecke, S.. The Physiology of the Grid: an Open Grid Services Architecture for Distributed Systems Integration. Global Grid Forum technical report, 2002.
14. Grimshaw, A. What is a Grid? *Grid Today*, **1**(26), 2002.
15. Grimshaw, A. and Wulf, W. The Legion Vision of a Worldwide Virtual Computer. *Comm. of the ACM*, **40**(1):39–47, 1997.
16. Grimshaw, A., Weissman, J., West, E., and Loyot Jr., E. Metasystems: an Approach Combining Parallel Processing and Distributed Heterogeneous Computing System. *Parallel and Distributed Computing*, **21**(3):257–270, 1994.
17. Kapadia, N., Figueiredo, R., and Fortes, J. PUNCH: Web portal for running tools. *IEEE Micro*, **20**(3):38–47, 2000.
18. Krauter, K., Buyya, R., and Maheswaran, M. A taxonomy and survey of grid resource management systems for distributed computing. *Int. J. of Software Practice and Experience*, **32**(2):135–164, 2002.
19. Lyster, P., Bergman, L., Li, P., Stanfill, D., Crippe, B., Blom, R., and Okaya, D. CASA Gigabit Supercomputing Network: CALCRUST Three-Dimensional Real-Time Multi-Dataset Rendering. In *Proc. Supercomputing'92*, Minneapolis, 1992.
20. Ramamurthy, M., Wilhelmson, R., Pea, R., Louis, M., and Edelson, D. CoVis: A National Science Education Collaboratory. In *Proc. American Meteorological Society 4th Conference on Education*, Dallas, TX, 1995.
21. Smarr, L. and Catlett, C. Metacomputing. *Comm. of the ACM*, **35**(6):44–52, 1992.
22. Stevens, R., Woodward, P., DeFanti, T., and Catlett, C. From the I-WAY to the National Technology Grid. *Comm. of the ACM*, **40**(11):50–60, 1997.
23. Wasson, B. Computer Supported Collaborative Learning: an Overview. Lecture notes IVP 482, University of Bergen, Norway, 1998.

Postproc. of the First European Across Grids Conference (ACG'03), Springer-Verlag LNCS 2970, pp. 291-298, Santiago de Compostela, Spain, Feb. 2004