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 del Cementerio s/n, 47011 Valladolid, Spain
{migbot, yannis, edugom}@tel.uva.es
http://pireo.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.

1 Introduction

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 [2]. 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 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 [2],[6]. Significantly, very recent works [7] challenges 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. In order to achieve such a goal, 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.

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. Conclusions and future work may be found in section 5.

* This work was supported by Spanish CICYT project TIC2002-04258-C03-02

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 amount of ten characteristics have 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: (1) *Large scale*: a grid must be able to deal with a number of resources ranging from just a few to millions. (2) *Geographical distribution*: grid's resources may be located at very distant places. (3) *Heterogeneity*: a grid hosts both software and hardware resources that can be very varied. (4) *Resource sharing*: resources in a grid belong to many different organizations that allow other organizations (i.e. users) to make use of them. (5) *Multiple administrations*: each organization may establish different security and administrative policies under which their owned resources can be accessed and used. (6) *Resource coordination*: resources in a grid must be coordinated in order to provide aggregated computing capabilities. (7) *Transparent access*: a grid should be seen by the user as a single virtual computer. (8) *Dependable access*: a grid must assure the delivery of services under established Quality of Service (QoS) requirements. (9) *Consistent access*: a grid must be built with standard services, protocols and interfaces thus hiding the heterogeneity of the resources while allowing its scalability. (10) *Pervasive access*: the grid must grant access to available resources by adapting to a dynamic environment in which resource failure is commonplace.

3 Main grid uses

Opposite to what it is often believed, the grid is not only a computing paradigm for just 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 types of applications (as it is usually found in literature) but in terms of types of support to applications.

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: (1) *Distributed supercomputing support* allows applications to use grids to couple computational resources in order to reduce the completion time of a job [10] or to tackle problems that cannot be solved on a single system [1]. (2) *High-throughput computing support* allows applications to use grids to put unused processor cycles to work in generally loosely coupled or independent tasks [1]. (3) *On-demand computing support* allows applications to use grids to retrieve resources that cannot be cost-effectively or conveniently located locally [1]. (4) *Data-intensive computing support* allows applications to use grids to synthesize new information from distributed data repositories, digital libraries and databases [1]. (5) *Collaborative computing support* allows applications to use the grid to enable

and enhance human-to-human interactions [1] in a synchronous or asynchronous way [6] via a virtual space. (6) *Multimedia computing support* allows applications to use grids to deliver contents assuring end-to-end QoS [10].

Table 1. A review of the main concepts found explicitly (E) or implicitly (I) in grid definitions extracted from the literature. Concepts are numbered following the description in text.

	1	2	3	4	5	6	7	8	9	10
"The computing resources transparently available to the user via this networked environment have been called a metacomputer" ([11])							E			
"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" ([8])			E	I	E					
"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" ([9])		E	I				E			
"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" ([3])		E	I							
"Computational grids are large-scale high-performance distributed computing environments that provide dependable, consistent, and pervasive access to high-end computational resources" ([4])	E							E	E	E
"A computational grid is a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities" ([1])								E	E	E
"The real and specific problem that underlies the Grid concept is coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations" ([6])				E		E				E
"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" ([10])	E	I	E	E	E		I			
"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" ([5])		E	E	E	E	E	I	E		E
"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" ([7])					E	E		E	E	

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 new definition of the grid concept can be built as follows: "A grid is a large-scale

geographically distributed hardware and software infrastructure 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".

4 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 project supporting this work in order to determine the possible advantages of using a grid infrastructure for CSCL applications.

References

1. Foster, I.: Computational Grids. In: Foster, I., Kesselman, C. (eds.): The Grid: Blueprint for a Future Computing Infrastructure. Morgan Kaufmann, San Francisco, USA (1998) 15-52
2. Foster, I.: What Is the Grid? A Three Point Checklist. Grid Today. 1 (6) (2002)
3. Foster, I., Kesselman, C.: Globus: a Metacomputing Infrastructure Toolkit. International Journal of Supercomputer Applications. 11 (2) (1997) 115-128
4. Foster, I., Kesselman, C.: The Globus Project: a Status Report. Proceedings of IPPS/SPDP'98 Heterogeneous Computing Workshop (1998) 4-18
5. Foster, I., Kesselman, C., Nick, J. M., Tuecke, S.: Grid Services for Distributed System Integration. Computer. 35 (6) (2002) 37-46
6. Foster, I., Kesselman, C., Tuecke, S.: The Anatomy of the Grid: Enabling Scalable Virtual Organizations. International Journal of Supercomputer Applications. 15 (3) (2001) 200-222
7. Grimshaw, A.: What Is a Grid? Grid Today. 1 (26) (2002)
8. Grimshaw, A. S., Weissman, J. B., West, E. A., Loyot Jr., E.: Metasystems: an Approach Combining Parallel Processing and Distributed Heterogeneous Computing System. Parallel and Distributed Computing. 21 (3) (1994) 257-270
9. Grimshaw, A. S., Wulf, Wm. A.: The Legion Vision of a Worldwide Virtual Computer. Communications of the ACM. 40 (1) (1997) 39-47
10. Krauter, K., Buyya, R., Maheswaran, M.: A Taxonomy and Survey of Grid Resource Management Systems for Distributed Computing. International Journal of Software: Practice and Experience. 32 (2) (2002) 135-164
11. Smarr, L., Catlett, C.: Metacomputing. Communications of the ACM. 35 (6) (1992) 44-52