RESEARCH PAPERS OF THE WROCŁAW UNIVERSITY OF ECONOMICS
No. 25 2008
Knowledge Acquisition and Management

Mieczysław L. Owoc
Wrocław University of Economics, Wrocław, Poland
mieczyslaw.owoc@ue.wroc.pl

INTELLIGENT PARADIGM IN GRID COMPUTING

Abstract: Grid computing offers many advantages essential for business, especially using large scale distributed information. The goal of this paper is presentation of intelligent challenges that come from the grid computing technology. The paper consists of three sections: a short characteristics of general assumptions and features of GC starts the contents, then chosen models of grid computing are demonstrated and finally usability of intelligent methods is investigated. Concluding remarks end the paper.

Key words: grid computing, intelligence, intelligent approach.

1. Introduction

No doubt information technologies seem to be one of the most dynamic streams of the general growth of computer science. In common opinion grid computing (GC) offers many advantages that meet business needs especially in large scale distributed information resources. There is a big amount of projects that are focused on grid computing, moreover leaders in information technologies very actively support and propagate GC concepts (for instance: IBM, Oracle and Hewlett-Packard). On the other hand there is a natural tendency to put intelligent methods into the more advanced solutions supporting information processing. From the very beginning artificial intelligence (AI) applications served different types of tasks in a smart way also involving integration and optimization aspects of business applications [Huang et al. 2006].

The aim of this paper is presentation of "intelligent challenges" that come from the grid computing technology. Investigation of intelligence and grid computing junction is very interesting because ultimate goal of GC acquires to apply most advanced techniques starting from planning up to implementation phases of modern information infrastructures. The main quest is to point out usability of intelligent paradigm in particular solutions of grid computing. There are a few approaches to this topic [*AI and Grid*…; *ICTS*…] but lack of research analyzing this problem in more global perspective. Therefore we should know what methods can be used in grid computing, what about usability of particular AI techniques in resolving optimisation tasks and in which areas of GC applications intelligent techniques can be applied. These questions are crucial for defining intelligent paradigm in GC.

The paper consists of three sections (apart from introduction and conclusion ones). A short characteristics of general assumptions and features of GC starts the paper. Then chosen models of grid computing are demonstrated. An essence of the paper is presented in the next section where usability of intelligent methods is investigated.

2. Assumptions of grid computing

Grid computing was invented to serve the needs of various scientific disciplines with large volumes of data and a need to process these rapidly in wide-area distributed systems. In the modern information world, where many applications have also been identified, it is also known as "utility"' computing because its distribution of information and computation services resembles the distribution of electricity or gas through a national grid. The inspiration for grid computing came from many sources, namely the web is to provide a broader range of services, not just information, big potential of commodity processors in PCs and other every day computing equipment which is idle much of the time, the need for better collaboration facilities among users of computers to achieve a common goal.

Discussing the nature of grid computing we should to point out the following properties [Owoc, Walasiński 2006; Goyal, Thom 2005]:

– a grid as a basic infrastructure, that means computers are spread but connected creating certain a network with typical components: hardware, software, databases and communication capabilities. Therefore grid computing treats all mentioned collections of IT resources holistically respecting flexibility of independent resource control;
– service-oriented of computing, where an system architecture can be agreed with a superior model for building applications. There are specific standards applied in this environment: XML-based Web Services, Internet protocols and distribute objects. This allows for virtualization and provisioning of application resources;
– innovative character of the combination of technology; grid establishes a common perspective and method for managing, referencing and accessing the particular IT resources available in an enterprise.

Therefore the ultimate goal of grid computing is to allow the sharing of computing and data resources for a number of workloads and to enable collaboration both within and across organizations. In order to fulfill these assumptions suitable models that are able to conform these features should be defined.

3. Solutions in the grid computing area

The general concepts of grid computing approach appeared in 1960s remind the grid as global connections of different and distributed resources that are able to co-operate and can be managed in an effective way. The global idea of such infrastructure is presented in Fig. 1.

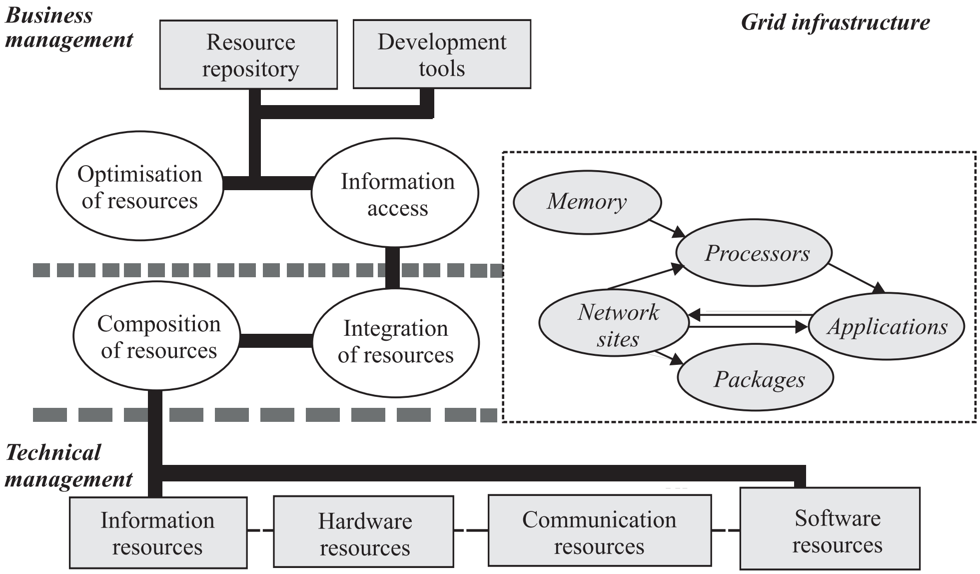

Fig. 1. General concept of grid computing

Source: own elaboration based on [Kelly 2004].

The lower part of the picture presents typical for this approach grid infrastructure. The common set of elements consist of the following resources: information structures, hardware and software and serve to information exchange and processing. One can assume that all defined categories refer to technical management.

In order to assure an efficient way of computer processing the particular resources should be selected and integrated (marked as oval shapes). The catalog of resources embraces: memories, processors, applications and packages for example.

Upper part of the picture presents components essential from the business point of view (marked as business management). Crucial processing refers to information access but very important is optimization of resources necessary for the performed tasks (marked as oval shapes). A repository plays the key role in integration and allows for the flexible usage of particular resources.

The general model of grid computing presented above can be implemented in many ways. Let us discuss some examples of real architectures proposed by key players in the market, stressing original approach to grid computing.

Grid computing presented by **IBM** is seen as mixture of technologies that allow for operating efficiency, utilization of assets and total operating costs. Going into details IBM efforts are focused on five grid areas [Kourpas 2006; *IBM…*]:

– Business Analytics Grid – to serve business planning and analysis that aimed at sharing of data and computing power,
– Engineering and Design Grid – to accelerate product design,
– Research and Development Grid – to enhance the R&D processes,
– Government Development Grid – to create new government services,
– Enterprise Optimization Grid – to improve utilization, efficiency and business continuity.

Grid solutions offered by IBM are oriented on different customers, starting from a new and representing small business one up to experienced and coming from huge sectors. For example: ***Grow*** and ***IBM Express*** enable application scheduling, efficient dynamic resource allocation and resource sharing, ***Clash Analysis in Automotive and Aerospace*** helps design engineers, ***Optimized Analytic Infrastructure*** serves financial market firms, ***Grid Medical Archive Solutions*** delivers storage systems supporting medical images archiving and many others. It is worthy to stress participation of IBM in creation of grid standards, namely Open Grid Services Architecture (OGSA) and an initiative of an ecosystem.

In case of **Microsof**t grid computing means usability of the DataGrid in .NET. In other words Microsoft platforms make data available as web services, for instance geospatial data, science data including also astronomy community and the like. There are original solutions that are able to serve some chosen governmental sectors, for example: US Geological Survey and Department of Agriculture. The corporation also supports Global Grid Forum as well as Open Grid Services Architecture middleware. Apart from the mentioned activities in the fields some addition embrace elaboration of grid standards and recently promotion of **Microsoft's collaboration solutions** like: Messenger, Share-Point, DISC, Groove, and Office (compare: [Gray 2002]).

The goal of **Hewlett-Packard** (HP) grid computing is supporting grid services and management of infrastructure resources for the so-called an adaptive enterprise. Actually, HP delivers grid-enabled services, solutions and products to help enterprises better manage and capitalize on change [*Hewlett*...]. Taking into account all possible profits for heterogeneous environments the company offers the following solutions:

– Enterprise Grid Consulting – to show customers when applying the concepts of grid computing to commercial environments,
– Grid Software Infrastructure – to build on the HP OpenView platform the capabilities of the software up through the management of Web services to deliver comprehensive real-time business process intelligence and enable immediate IT resource response in the context of Web services or grid services,

– HP Utility Data Center (UDC) – to deliver many grid capabilities to commercial customers compatible with OGSA standards,
– Grid Resource Topology Designer – to design and “draw” resource needs including GC infrastructure. The Grid Resource Topology Designer, working with the HP UDC, automatically decides on the appropriate resources to deploy to fulfill the service-level request,
– Web Services Management Framework (WSMF) – to formalize a logical architecture for the management of resources, including grid and Web services. WSMF can be also input for creating a standard management interface for all IT resources and services.
– Smart Framework for Object Groups (SmartFrog) – to enable administrators easily configure resources on the distributed computers that make up the grid.

The last version of the database server released by Oracle Corporation are marked with “g” symbol which denotes exactly grid computing. Holistic approach representing in Oracle allows for management of IT resources in three-dimensional aspects including: infrastructure, applications and information [Nach 2005; Shimp 2005]. In this solution two core unique tenets should be stressed: virtualization (abstraction of all available components) and provisioning (denoting determining by a system how to meet specific customer requirements taking into account optimizing of the whole infrastructure.

Three-layers model has been elaborated for this purpose. The general concept of Oracle infrastructure is presented in Fig. 2.

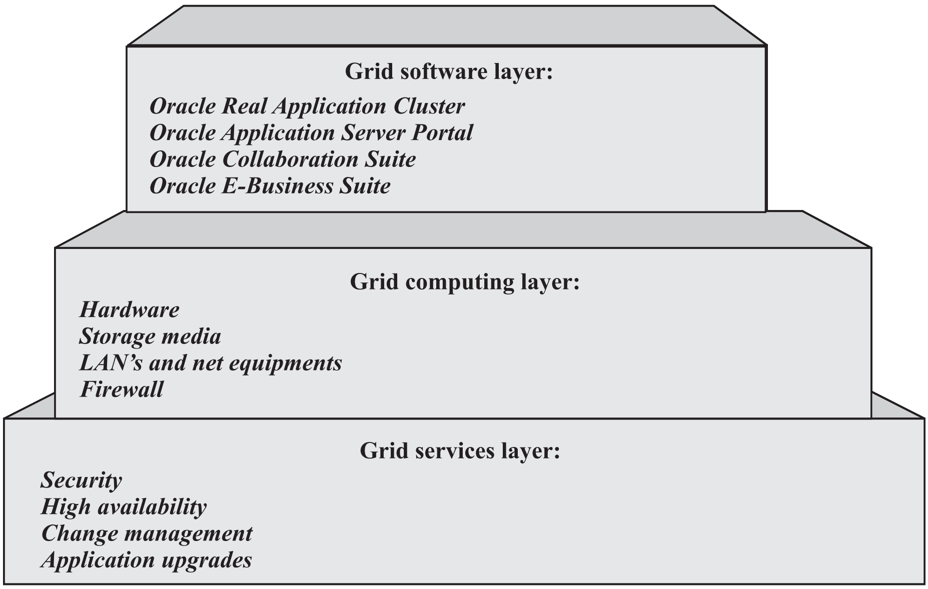

Fig. 2. Oracle grid computing infrastructure

Source: own elaboration based on [Baum 2004].

Software layer is responsible for access to grid technology via specialized packages. Some of them serve collaboration tasks *Oracle Collaboration Suit*, some manage clusters *Oracle Real Application Cluster* while the others support e-business (*Oracle Application Server Portal* and *Oracle E-Business Suit*). General purpose of this layer is administration and balancing of the available resources.

Grid computing layer (located in the middle) is created for performing user's and system's jobs according to open standards. Particular processes are placed in some nets of grid infrastructure and necessary hardware and network components support these processes.

The last layer – *grid services* – is responsible for assuring automatic management of access, efficiency and processing security. For example some computing components can be activated during idle time, mobile processing or moving of data can be starting.

There are plenty of very specialized tools and technologies that support many tasks of grid computing in Oracle environments, for example: Oracle Automatic Storage Management, Oracle Enterprise Manager.

The solutions presented in this section are different but all of them are able to manage IT infrastructure in an automatic manner (see: [Neurona]). Therefore applying of intelligent methods seems to be a must.

4. Placement of intelligent methods in GC

To develop efficient mechanisms that address the complex grid applications some specific assumptions should be taken into account [Gil et al. 2004]. First, we need expressive and extensible methods of describing Grid at all levels. Second, we need flexible mechanisms to explore trade-offs in the Grid complex decision that incorporate heuristics and constrains into the process. Concluding, we need more flexible and knowledge-rich Grid infrastructure. In other words, some artificial intelligence paradigm useful in this context seems to be very obvious. Let us regard the essential quests.

- Knowledge capture – the main problem is how to describe Grid entities, their relationships and capabilities.
- Usability – apart of Grid description (as above) it is very difficult to provide automated workflow generation techniques that would incorporate the necessary knowledge to access grids in very unstable conditions.
- Robustness – as a consequence of dynamic grid environment – the resources are highly heterogeneous and shared by many users. As before, we need knowledge about workflow history, the current status of their subtasks and taking decisions in particular conditions.
- Scale – we may expect still growing amounts of data, performing transactions and increasing number of users. In order to perform analysis for different pur-

poses a need of managing complex workflow pools that balance access to resources, provide or reserve new capabilities requires very smart methods.

General idea of functioning Grid infrastructure is presented in Fig. 3.

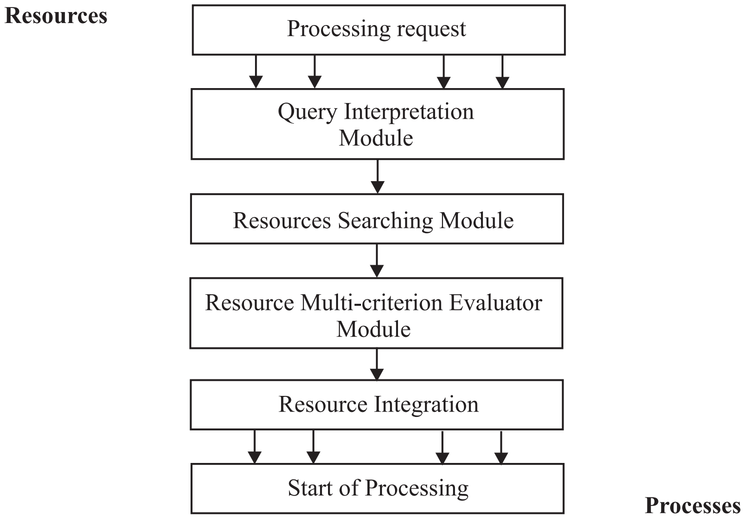

Fig. 3. Steps in grid computing

Source: own elaboration based on [Shimp 2005].

In order to simplify processing cycle, let us assume service of user querying. We have many resources (top left) that could be used to answer the query. Apart of necessary steps from formal point of view (query interpretation, data searching) there are two modules essential for successful performing: "Resource Multi-criterion Evaluator" and "Resource Integrator". This is a place for including artificial intelligence techniques supporting grid computing.

Let us stress that achieving an efficient, secure and flexible connection between resources and users can be complicated by a number of factors:

– frequent mergers and acquisitions, bringing extreme complexity,
– use of separate data systems for trading separate assets,
– inherent complexity driving financial infrastructures: servers, networks, applications.

Therefore more advanced technologies (including artificial intelligence methods) should be applied. There are several solutions that confirm usability of intelligent techniques in grid computing environments.

Pegasus [Gil et al. 2004] **–** a system that integrates an AI planning system into a grid environment**.** After submitting by an user application-level description of the desired data product, the system generates a workflow by selecting appropriate components, assigning the necessary computing resources and overseeing the successful execution with an optimization phase. Pegasus was used in different domains; genomics, neural tomography and particle physics. Knowledge sources and intelligent reasoners are available as grid services.

First Person Shooter [*IncrediBuild…*] is elaborated to show intelligent behavior in a game that covers spatial and temporal environment. There are plenty of complex and time consuming processes where geometric operations are performed on millions of triangles. To speed up the process is divided into many independent chunks which allows maximum parallelism to be applied using special Xoreax Grid Engine. Comparing to different approaches to the problem solution one can say joining artificial intelligence techniques and grid computing applied in the described game allows to reduce significantly time of processing.

Knowledge grid architecture [Zhuge 2004] consists of theories, models, and methods for supporting knowledge capture and representation. Thanks to the Knowledge Grid, it will be possible to effectively capture and conveniently publish knowledge in a machine- processable form that could directly, or after simple transformation, be understood by humans. Interesting in such a context are additional properties of this approach. First is knowledge integration. Integrating knowledge resources could support analogies, problem solving, and scientific discovery. Second is scalable network platform. The Knowledge Grid should enable a user, a machine or a local network to freely join in and leave the platform.

Examples of successful joint AI and GC solutions refer to very far application areas. This confirms usability of the approach and spreads of intelligent paradigm in grid computing context.

5. Conclusions

The paper has reviewed state-of-the-art literature in Grid computing and its connections to Artificial Intelligence. The basic findings can be formulated in the following way:

– different networks components are served more flexible and efficient via grid computing. Knowledge necessary to support this infrastructure should be acquired in a very smart way;
– corporations supporting the grid computing approach offer different solutions but almost everywhere the need of intelligent techniques is obvious (e.g. where tasks are performed in an automatic way);
– intelligent support can be applied at different areas of grid computing architectures. The most frequent fusion of AI and GC refers to optimization procedures, especially in case of planning.

References

AI and Grid Computing Symposium. http://comma.doc.ic.ac.uk/aisb2002/.

Baum D. (2004). Technology on-demand. A new frontier. *Oracle Magazine*, May/June 2004, pp. 26-35.

Gil Y., Deelman W., Blythe J., Kesselman C., Tangmunarunkit H. (2004). Artificial intelligence and grids: Workflow planning. *IEEE Intelligent Systems*, January/February 2004.

Goyal B., Thom B. (2005). Oracle database 10g: The database for the grid. *An Oracle White Paper*, January 2005. http:www.oraclewhitepapers.com.

Gray J.(2002). *Microsoft and Grid Computing*. http://research.microsoft.com/~Gray/papers/Microsoft_and_Grid_Computing.doc.

Hewlett Packard and Grid Computing. http://www.hp.com/techservers/grid/index.html.

Huang W., French T., Maple C., Bessis N. (2006). *Can intelligent optimisation techniques improve computing job scheduling in a grid environment? Review, problem and proposal*. http://www.all-hands.org.uk/2006/proceedings/papers/596.pdf.

IBM and Grid Computing. http://www-03.ibm.com/grid/.

ICST QShine, Hong Kong'2008 – the Fifth International ICTS Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness. http://www.qshine.org/.

IncrediBuild by Xoreax Software – Grid Computing for Artificial Intelligence. http://www.xoreax.com/technology_case_studies_id.htm.

Kelly D.A. (2004). Core to your business. *Oracle Magazine* May/June 2004, pp. 26-35.

Kourpas E. (2006). Grid computing: Past, present and future. An innovation perspective. *IBM White Paper*. http://www-03.ibm.com/grid/pdf/innovperspective.pdf.

Nach M. (2005). Grid computing with Oracle. *An Oracle Technical White Paper*. Oracle Corporation, March 2005.

Neurona. https://neurona.dev.java.net/.

Oracle and Grid Computing. http://www.oracle.com/technologies/grid/index.html.

Owoc M.L., Walasiński T. (2006). *Przetwarzanie siatkowe jako przełom w technologiach informatycznych*. Studia Informatica '2006, Uniwersytet Szczeciński, Szczecin.

Shimp R.G. (2005). Oracle grid computing. *An Oracle Business White Paper*. Oracle Corporation, February 2005.

Zhuge H. (2004). *The Knowledge Grid*. World Scientific Publishing Co., Singapore.