INVOLVING GRID COMPUTING IN AGRICULTURAL RESEARCH

IMPLICAREA TEHNOLOGIEI GRID COMPUTING ÎN CERCETAREA AGRICOLĂ

M. CĂLIN*, M. CRAUS, F. FILIPOV*, C. CHIRUȚĂ***

** University of Agricultural Sciences and Veterinary Medicine, Iasi, Romania*
*** Technical University, Iasi, Romania*
Corresponding author: Marius Calin, e-mail: mcalin@univagro-iasi.ro

***Abstract:** The defining features of the "Grid computing" concept are enumerated. As well, the impact of Grid computing on different branches of agriculture is highlighted. Four universities and one research institute of Iasi, Romania, are developing an academic grid together. In the paper, the structure of this academic network is briefly described. In this context, the positive effect that the academic grid is expected to have on the multi-disciplinary agricultural research is emphasized. An application in Soil Science is given as an example.*

***Rezumat:** În lucrare sunt enumerate câteva elemente definitorii ale conceptului "Grid computing" şi este subliniat impactul pe care utilizarea reţelelor de acest tip îl poate avea în diferite ramuri ale ştiinţelor agricole. Este prezentat pe scurt un proiect de Grid dedicat cercetării ştiinţifice, aflat în curs de dezvoltare, la realizarea căruia participă principalele patru universităţi şi un institut de cercetare, din centrul universitar Iaşi. Este detaliată contribuţia USAMV Iaşi în acest proiect, cu accent asupra serviciilor Grid pentru cercetare în ştiinţa solului.*

***Keywords:** Grid computing, Soft computing, Agriculture, Soil Science*
***Cuvinte cheie**: Grid computing, Soft computing, agricultură, ştiinţa solului*

INTRODUCTION

Complex applications that need an increasing number of computational resources require the exploitation of various resources which are geographically distributed. Such applications use networks that interconnect supercomputers, large databases, storing devices, advanced visualization devices and scientific instruments [1].

The term *Grid Computing* has been used for the first time during the middle of the last decade to denominate an advanced infrastructural proposal for parallel/distributed computing that implies using component organized software that runs on a large number of computers. Consecrate names in the world of grid systems are: NASA's Information Power grid, Distributed ASCII Supercomputer (DAS-2 – integrates clusters from five German universities), Teragrid (interconnects almost all universities in the USA).

As one can see, grid systems embed resources from several institutions that share computing resources and supply with reliable, secure and high performance services. Such collaborative work becomes a necessity in various fields of science, engineering and business and it can be employed in *virtual organizations*, using Grid technologies.

Grid technologies have the potential of dramatically changing the use of computers in solving problems. If the Internet revolutionized human communication, Grid systems will bring major changes in activities related to research, design and production. These changes can be achieved by combining scientific, technical and human resources into huge, yet flexible, high performance entities.

GRID COMPUTING SYSTEMS IN AGRICULTURE

By its nature, agriculture is a field in which research and production require large scale coordination of efforts both at a geographical level as well as from a multi-disciplinary point of view. In this respect, the benefits that Grid computing can bring were accounted and powerful Grid projects were developed worldwide to integrate different agricultural branches with other fields of activity. Some relevant achievements in this direction are subsequently pointed out.

In the USA, the Colorado State University is developing a national animal identification system that will rely on Grid computing technology to process massive amounts of animal tracking data [2]. The Colorado Department of Agriculture and private companies involved in agriculture and computer hardware are also participating in this project with funding and sponsorship. As the project's designers state, *"the Grid computing identification system will enable researchers to track the travels of a specific animal in seconds or minutes, instead of the weeks it currently takes. In case of an outbreak of a disease such as bovine spongiform encephalitis, or BSE (also called mad cow disease), having a national database of cattle would allow officials to find out where a cow has been at all times from birth to death. Officials could identify all of the cow's co-residents for immediate testing, limiting the spread of the disease."* The project uses Colorado State's new Colorado Grid Computing Initiative, or COGrid, which is planed to be expanded to become an interconnected "Grid of Grids" that would become a huge computing network including academic, industrial, governmental, education and public users.

Another US example of Grid system with agricultural purpose is the partnership between the US Cornell Theory Center (CTC) and Cornell's College of Agriculture and Life Sciences (CALS). CALS focuses on using CTC's computational infrastructure and expertise to develop science-based technologies support of farmers [3]. The project also has different sources of funding and support, the most important of them being a division of the United States Department of Agriculture. The fields that benefit from the use of the Grid cover domains involved in precision agriculture: crop and soil sciences, economics, climate modelling, and data-mining.

In Japan, the National Agriculture Research Center, Tsukuba, is developing a project named *GRID for Agricultural Decision Support* [4]. The basic premise for starting this project was that in agriculture one must combine data from various different databases such as weather data, soil data, crop data and market data. This data is available on geographically widespread computers. Moreover, the respective database applications were built and are running under different software engines and on different platforms. The only realistic way to integrate this diverse information within a unique decision-support tool is to virtualise and exploit it under a Grid computing system. As a typical example, the project authors discuss the possibility of combining data from a weather database, which is available on a server, with several different application programs located on other remote severs, such as a rice growth prediction model and a pest prediction model, that both need the weather data to make the predictions for a specific area.

As genetic research has an important impact on agricultural sciences, a European Grid computing initiative in this field must be mentioned. The project is developed by the European Bioinformatics Institute (EBI). EBI is a non-profit academic organization that forms part of the European Molecular Biology Laboratory (EMBL). Following its mission, EBI ensures that the growing body of information from molecular biology and genome research is placed in the public domain and is accessible freely. The EBI serves researchers in molecular biology, genetics, medicine and agriculture from academia, and the agricultural, biotechnology, chemical and pharmaceutical industries. The data in EBI's databases containing information relevant to molecular biology is doubling every year, Moreover, the researchers involved in EBI programs are distributed across multiple geographic locations and work on diverse computing platforms

In 2002, the EBI decided to build an enterprise Grid solution [5]. The Institute, which manages databases of biological data, including nucleic acid, protein sequences and macromolecular structures, will use it to facilitate resource sharing and increase collaboration among international researchers who use its numerous databases and information services. The Grid is designed to run under software from Platform Computing, one of the world leaders in distributed computing software.

In some Asian emerging countries the decision to develop Grid computing systems was also made. A relevant example is the National Grid Project of Thailand [6] that started in 2000 with substantial financial support from the government. As agriculture plays an important role in its economy, one major feature of the Grid is to support production, research and development in this area.

AN ACADEMIC GRID COMPUTING PROJECT

In 2006, four faculties and a research institute in Iasi, Romania started the research project named *Academic Grid for Complex Applications*. The acronym of the project is GRAI and it runs under a CEEX grant (the excellence research framework created by the Romanian Ministry of Education and Research following the EU FD7 model). The five participants in the project are:

- The Technical University of Iasi, Faculty of Automatic Control and Computer Engineering (which also holds the leadership of the GRAI project);
- Institute for Computer Science, Romanian Academy, in the location of the Faculty of Electronics and Telecommunication;
- The “Al. I. Cuza” University of Iasi, Faculty of Computer Science,
- The University of Medicine and Pharmacy Iasi, Faculty of Biomedical Engineering
- The University of Agricultural Sciences and Veterinary Medicine of Iasi, Faculty of Horticulture.

As its title shows, the GRAI project aims to develop a grid computing structure for research and for other academic purposes. To achieve them, two main directions must be followed:

1. Development of a grid computing system that would interconnect the scientific and computational resources of the five partners.
2. Development of grid services and specific applications based on them.

In this effect, the grid will be structured on four layers: the Computational Resources layer, the Service layer, the Application layer and the Information layer.

The Computational Resource layer. The network will be geographically situated (figure 1) in the five locations of the project partners:

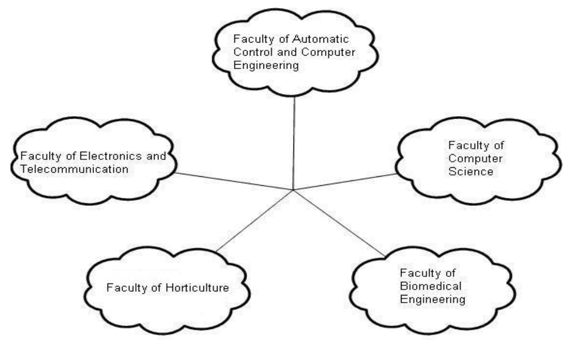

Figure 1. The GRAI network

Each of the five locations will have a grid node that includes a high performance server and a group of workstations (figure 2). These workstations will be used both as computing support within the grid and in developing grid services and applications.

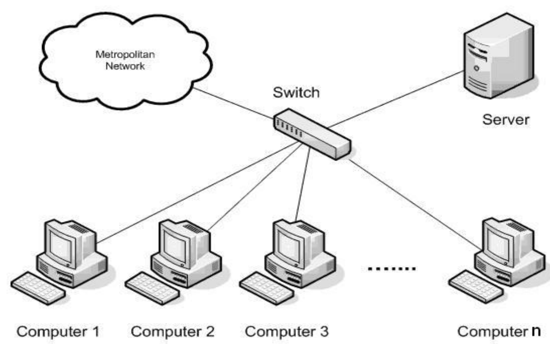

Figure 2. Node structure

The Service Layer. This layer will be divided into two sub-layers. One of them is the *Basic Services* sub-layer, designed to develop and provide software tools that implement algorithmic methods in modern branches of Computer Science such as *Knowledge Discovery*, *Image Processing*, *Graph and Combinatorial Optimization*, *Evolutionary Computing*. The other is the *Problem Oriented Services* sub-layer which would benefit of the former layer to solve problems in more specific areas such as *Abstract Population Modeling, Optimization and Forecasting in Logistics and Economics, VLSI Design.*

The Application layer. This layer is designed for users that develop applications based on grid services. This layer will intensively use the GRAI Grid resources. Numerous such applications were identified in different branches of agricultural sciences. This is why the partnership of the

University of Agricultural Sciences and Veterinary Medicine of Iasi would be most active within this layer.

The Information layer. This layer is designed for potential users. At this layer, users may query the informational offer of the GRAI Grid.

THE ROLE OF GRAI IN DEVELOPPING MULTI-DISCIPLINARY RESEARCH IN SOIL SCIENCE

Numerical processing in agricultural research often requires (and sometimes is limited to) statistical methods. Great amounts of data result from direct measurements and laboratory analysis. In numerous cases this data is stored without concern for a unitary representation. Therefore, an attempt to diversify the services offered in this area would be welcome.

Like in many other fields, in agriculture and soil science the researcher/expert has to make decisions based on the available numerical data. A software tool would be useful to handle this data in order to make the appropriate decision [7]. Such tools are called Decision Support Systems (DSS).

There are various types of assistance that a DSS can offer. It can be used mainly for filtering useful information from a large amount of data. It can be a help in evaluating alternatives or it can be a tool to compare goals with achievements. This can be done in an algorithmic way or by using a knowledge base. Historically, the first DSS approaches were of two types: one type that used Operational Research mechanisms based on optimization algorithms, and another type that were focused exclusively on information administration (databases, file and data flow administration). The knowledge base approach came subsequently, together with the development of Expert Systems. Another perspective in DSS approach emerged after the inception of Fuzzy Sets Theory and Fuzzy Logic. The term Fuzzy DSS was introduced to identify those decision support systems that dealt with imprecise, approximate or linguistic qualifiers [8, 9].

One can notice that DSS structures can be very different from case to case, as they must be adapted to the needs of each considered situation. In many areas of agricultural sciences the creation of such systems is welcome. In particular, Fuzzy DSS concepts can be used because in agriculture, as well as in biology, one is likely to be put in a situation where one must formalize and handle linguistic terms within the decision process.

The GRAI grid can offer a long term development environment for DSSs applied to agricultural decision problems because, by its nature, it offers a favourable framework for multi-disciplinary research. A service oriented towards supporting decisions in soil science applications is proposed as a pilot project. At national level there is already a great deal of calculus methods and reported scientific work in the area [10]. IT solutions exist as well [11]. At the University of Agricultural Sciences of Iasi previous multi-disciplinary research programs approached the domain [12, 13, and 14]. New algorithms for different situations are currently under study. These algorithms are meant to be useful in practical applications. The project partner USAMV Horticulture will design and implement a DSS whose goal is to support decisions regarding the durable exploitation of pedoclimatic resources in horticulture. This service will be implemented on the grid.

CONCLUSIONS

Grid computing networks are long term projects which finally bring undoubted benefits that can be measured through financial effect, scientific and social impact. However, they have an initial costly investment phase and the discussed worldwide examples reveal that both government agencies and private companies participated with money and equipments software in building an appropriate infrastructure.

In several American, European and Asian countries, the usefulness of Grid computing for agricultural research and production was already proved.

The perspective of integration with the European Community demands the development of such entities that are already active in many European Countries. The development of the GRAI project follows this line of action.

LITERATURE

1. ATTIYA, H., WELCH. J., *Distributed Computing: Fundamentals, Simulations and Advanced Topics*, The McGraw-Hill Companies, London, 1998.
2. xxx., *Grid Powers Nat'l Animal Tracking System at Colorado State*. GRIDToday, http://www.gridtoday.com/grid/392144.html, 2005.
3. EMMEN, A., *Cornell Theory Center and College of Agriculture and Life Sciences partner*. Primeur Monthly, http://www.hoise.com/primeur/04/articles/monthly/AE-PR-08-04-23.html, 2004.
4. NINOMIYA, S. *Network Computing for Agricultural Information Systems -GRID for Agricultural Decision Support*. Proc. 3rd AFITA;Asian Agricultural Information Technology and Management. Ed. F. Mei:26-30, 2002
5. xxx., *European Bioinformatics Institute Enters GRID with Platform Computing* http://www.platform.com/Newsroom/Press.Releases/2002/PR.01-22-2002.htm, 2002.
6. VANNARAT, S., *Computing and Information Grid Development in Thailand.* International Symposium on GRID Computing, 25-27 April 2005, Taipei, Taiwan, 2005.
7. BENCHIMOL, G., P. LEVINE POMEROL, J-C., *Sisteme expert in intreprindere.* Editura Tehnica, Bucuresti, 1993.
8. REUSCH, B., *Mathematics of Fuzzy Logic*, In: *Real World Applications of Intelligent Technologies* (Zimmermann H-J, M. Negoita and D. Dascalu, (Eds)). The Publishing House of the Romanian Academy, 1996.
9. ZIMMERMANN, H-J. *Fuzzy Decision Support Systems*, In: *Real World Applications of Intelligent Technologies* (Zimmermann H-J, M. Negoita and D. Dascalu, (Eds)). The Publishing House of the Romanian Academy 1996.
10. CANARACHE, A., DUMITRU S., *Estimating workability and trafficability in Romania.* Proc. of the 15th International Conference of the International Soil Tillage reseach Organization, "ISTRO – 2000", Temple, Texas, USA, 2000.
11. CANARACHE, A., SIMOTA C., *Databases and Simulation Modelling in Compaction and Erosion Studies*, Advances in Geoecology Volume no. 35, pp 495-506, Catena Verlag, 2002.
12. CALIN, M., FILIPOV, F., CHIRUTA, C., TOMITA, O., *Fuzzy Modelling of Linguistic Terms Used to Estimate the Calcium Carbonate Contents in Soil Horizons.* Lucrări Ştiinţifice, seria Agronomie Volume no. 49**,** 2004.
13. CALIN, M., FILIPOV, F., CHIRUTA, C., TOMITA, O., *Using Fuzzy Sets Support in Computing the Optimum Value of Carbon/Nitrogen Ratio in Compost*, Proc. of ESNA XXXVI Annual Meeting, Sept. 10-14 Iasi, Romania, 2006.
14. FILIPOV F., CALIN M., TOMITA O., LUCIA DRAGHIA CHIRUTA C., *A method to calculate the weight of components in a nutrient substratus for horticultural plants that has optimal base saturation.* Lucrari Stiintifice, seria Horticultura, Volume XLVII, (47), pp. 1305 – 1310, 2004.