E-Learning Applications in the EELA Project

I. Dutra[1], F. Fernández-Lima[1], E. Silva[1], M. Diniz[1], A. Porto[1], D. Carvalho[1], P. Rausch[1], F. Fernández-Nodarse[2], D. López[3], M. Fernández[3], J. Cruz[4], L Enriquez[4], A. Acero[5], E. Montes[5], A.J. Rubio-Montero[5] and R. Mayo[5]
(On behalf of the EELA Project)

[1] Universidade Federal do Rio de Janeiro, Campus Ilha do Fundão, Rio de Janeiro, Brazil
[2] CITMATEL, Avenida 47 s/n entre 18A y 20, Miramar, Playa, La Habana. Cuba
[3] CUBAENERGIA, calle 20, 4111, Reparto La Sierra, Playa, Ciudad Habana, Cuba
[4] Laboratorio Experimental Multidisciplinar, Universidad Nacional Autónoma de México, Cuautitlán, Mexico
[5] CIEMAT, Avda. Complutense, 22, Madrid 28040, Spain
Email: ines@cos.ufrj.br, rafael.mayo@ciemat.es
phone: +34 913466180, *fax:* +34 913466537

Abstract

The EELA Project was created with the aim of supporting the increasing demand for Grid Infrastructures to foster the collaboration among groups of Latin America and Europe. As a starting point, several applications have been selected in four scientific fields. One of the most important areas among them is e-Learning due to its high impact and significance in the Latin America society. Newly applications are being developed by several groups from Latin America participating in the project. This article will shortly describe the EELA project itself and the current status and aims of the six e-Learning applications to be deployed.

1 Introduction

One of the objectives of the EELA project (http://www.eu-eela.org) is to identify and promote a sustainable framework for e-Science. This objective is reachable not only by deploying mature Grid applications, but also new ones coming from other European and Latin-American scientific communities.

E-learning has been a topic of interest in recent years. As a result, many e-learning platforms and systems have been developed basing mainly on teachers and students (or authors and learners). Teachers create content, which is stored under the control of a learning management system (LMS), typically in a database. Content can be updated and exchanged with other systems [1].

One of the great advantages of e-learning is the scheduling flexibility, where the student have the freedom to study according to his/her time constraints. A second advantage is related to distributed knowledge, where each site can have information about different contents. A third asset is that authors and learners can benefit from specific knowledge or resource not available locally.

In typical distributed e-learning systems, the user needs to know the locations of resources in order to access the resource or knowledge of interest.

Because of that, very often, the user may miss some knowledge or resource of interest, simply because he/she does not know that the resource of interest exists in some location. It could also be the case that the resource has access restricted to some determined users and/or machines.

For these reasons, e-learning in grid environments brings several benefits. A grid unifies many different resources such as computers with CPU and disk storages, remotely controlled instruments (microscopes, telescopes, etc.) and visualization environments using a single uniform interface. This characteristic makes it possible to the user to access any resource in the grid uniformly without having to know where the resource is located. Moreover, the grid works as a virtual social organization, where users have regulated rights and permissions to access any resource as a member of the social organization.

In this document the first Grid applications that will be deployed in the EELA infrastructure concerning e-Learning are described.

2 ACADIA

The application of long-distance education ACADIA is based on a platform that involves peer-to-peer technology and shared resources called DEDALUS. ACADIA provides a learning environment for students and teachers since the building of a lesson to the delivery of works and tests. It is intended to increase the interaction between teacher and students, and among students, opening a space for discussing questions, suggestions and criticism. The use of multimedia (images, videos, sounds and texts) stimulates creativity and collaboration among participants. The DEDALUS's architecture was developed to benefit from the application of concepts of object orientation and Java technology in order to guarantee easy expansion of the system without compromising its continuity. ACADIA in EELA intends to combine an application built using a peer-to-peer approach with a grid platform aiming to provide teachers and students with a powerful computational tool as well as to enable distance education in the grid technology. Architecturally, Acadia will act as a wrapper of the EELA grid middleware. More information can be found at [2].

3 CuGfL

The Learning Management System Cuba Grid for Learning (CuGfL) is a One-Stop-Centre for quality assured online learning content with the aim to promote and support the lifelong learning agenda in Cuba to accelerate the growth of K-Society. In order to incorporate the results obtained in the areas of digital content, digital repositories and remote laboratories the application is mainly focused on working on the Learning Management System (LMS), digital contents, multimedia repository and e-learning services in order to:

- Test and adapt the latest version of LMS, e-learning services, methodologies, infrastructures, and multimedia repositories, on grid.

- Customize and optimize the LMS and e-learning service redesigning the applications and service management according to the grid and IPv6 environment. Upgrade the e-learning tools, simulators, virtual laboratories, digital repository and multimedia servers mainly based on free software

According to the resources, the activities that will be studied are Remote laboratory, Multimedia Interactive Course, Semantic Grid and Collaborative Courses.

4 Video on Demanding (VoD) – Remote Lab

These e-learning applications [3] consist of a distributed interactive multimedia server (RIO) that is currently being employed in the distance learning consortium of public universities in the state of Rio de Janeiro (Brazil). The video stream is coupled with class slides and a brief index of the topics in the lecture. Students can interact with the slides, which are programmable entities and can launch other applications in the client's computer. The RIO server is made of storage units (storage servers) that may reside in different machines at different locations. For the EELA project the RIO server will be adapted to run over the GRID paradigm. The goal is to provide an underlying structure that allows dynamic reconfiguration of resources.

Content and videos are stored in different academic institutions. This distributed software is not yet running on the context of EELA, but an investigation on the efforts to port it to the EELA infrastructure is being made. The software deployed on EELA will have the following characteristics: each EELA site will have an UI with the front-end multimedia server, used by clients to retrieve content (video, text, etc). This UI will be connected to other machines in a local network that will work as storage elements for the contents. This way, clients can use any UI available and registered on the EELA virtual organization to retrieve any videos or execute any experiments that may be located elsewhere.

As part of this e-Learning initiative, a prototype virtual lab is being built for the study of the mechanical oscillations. The prototype will be tied to the multimedia server clients' environment. This way, students watching a video course would be able to interact with the class slides and, from those, access the prototype virtual lab while guided by the video-lecturer

5 LEMDist

LEMDist [4] is not only an application, but a project itself. Thus, its goal is to get web access to laboratory equipment and another web service to help e-science and e-learning users. The project is aiming to build technological support for a different pedagogical approach in natural science teaching, and heterogeneous problem solving environment for scientific work. The system is based on applications with access to distributed computer-enhanced instrumentation, remote access to simulation capabilities with high performance computing support, interactive visualization, distributed data analysis and access to heterogeneous data sources systems.

To do this, a five layers architecture has been identified based on the gLite environment: Access layer, Service layer, Grid layer, Administration and security layer and Resource layer.

The research in the remote access will be focused on several devices. The most important will be a 200 Mhz Nuclear Magnetic Resonance (NMR) Varian spectrometer, a Capillarity Electrophoresis (CE) equipment and the associated computer controlled laboratory apparatus through a serial interface. Other fields of work to be tackled in the future will be chemical and food engineering and Industrial chemistry

6 Parallel Inductive Logic Programming (PILP)

This application is intended to discover hidden data from relational databases. It uses a technique called Inductive Logic Programming, where, given background knowledge -a set of positive examples, a set of negative examples and a language bias- the objective is to generate first order rules that (almost) perfectly describe all positive examples and none of the negative examples. Several domains when applying ILP have been studied: drug discovery, analysis of mammograms, link discovery, among others.

Traditionally, machine learning algorithms execute by splitting the dataset into a training set and a test set, but if the number of examples is small, this means that we will not be making the best advantage of the examples. Thus, another technique is proposed. In Cross-validation, a split of a full training data set of positive and negative examples in several different subsets is made. In k-fold cross-validation, the data is divided into k subsets of (approximately) equal size. The data is then trained k times, each time training on k-1 subsets, and then using the omitted k-th subset to evaluate model accuracy (through whatever error criterion is interesting to the user). If k equals the sample size, this is called “leave-one-out” cross-validation. “Leave-v-out” is a more elaborate and expensive version of cross-validation that involves leaving out all possible subsets of v cases. It also produces more experiments.

Often, machine learning systems are used to produce classifiers, i.e., programs that classify an input according to a model. Ensembles are classifiers that combine the predictions of multiple classifiers to produce a single prediction. The main advantage is that the ensemble is often more accurate than its individual components. Moreover, we can use parallelism both in generating and in actually evaluating the ensemble (bagging [5] and boosting [6]). It is possible to obtain an ensemble simply by doing learning several times, but always starting from different starting points: we call this methodology different seeds. For the combination of ensembles and cross-validation it is necessary to run the experiments in parallel, so the grid is a very convenient environment for these kinds of experiments that demand many resources. Some experiments have already been done for predicting carcinogenesis in rodents

7 SATyrus

SATyrus is a novel approach to the specification and solving of optimization problems. In the scope of the SATyrus architecture, a given target problem is specified, using a logical style declarative language, as a set of pseudo-Boolean constraints. Then, this set of constraints is compiled, following a satisfiability (SAT-based) mapping, into an energy function representing the space state of solutions of the target problem. Among other possible computational intelligence models that could have been adopted (e.g., genetic algorithms, artificial immune systems, etc.), higher-order Hopfield networks of stochastic neurons were chosen to map and minimize the resulting energy function. This choice is justified by the linear cost of representing any energy function produced by SATyrus as symmetric neural networks.

SATyrus' architecture is composed of a concatenation of three basic modules: compiler; mapper and (neural) engine. A problem specification, written in a customized language, is fed to the compiler as a set of pseudo-Boolean constraints. The object code produced by the compiler consists of a “pre-energy function”, which can be fine-tuned by the user. Such a tuning, achieved through the adjustment of a “penalty scale” (modulating the whole set of constraints), has the purpose of enabling experiments over the convergence behaviour of SATyrus' engine [7].

8 Conclusions

The EELA e-infrastructure permits various collaborative groups in Latin America more powerful computational resources than those available on their centers. This achieves that the lines of investigation stated in the document can be feasible as the computational requirements for them can be met. As well of creating a network where Latin American researches can participate in European initiatives and vice versa reducing the digital gap.

The EELA project currently has six e-Learning pilot applications to be run in the first quarter of 2007. This project even though its in early phases will intend to bring more research lines into its e-infrastructure as well as creating a bigger network of collaborative partners.

References

1. V. Pankratius, G. Vossen. Towards E-learning Grids: Using Grid Computing in Electronic Learning. Proceedings of the IEEE knowledge Grid and Grid Intelligence (2003). 4-15
2. http://labase.nce.ufrj.br/projetos/pesquisa.html
3. http://trindade.land.ufrj.br/~vod/rio
4. http://hydra.dgsca.unam.mx/lemdist
5. L. Breiman, Bagging Predictors. Machine Learning 24 2 (1996).
6. Y. Freund, R. Shapire. Experiments with a new Boosting algorithm, AAAI 96 (1996).
7. P.V.M. Lima et al. SATtyrus: A SAT-based Neuro-Symbolic Architecture for Constraint Processing. Proc. 5th Hybrid Intelligent SystemsConference (2005).