Chapter 6
High Performance Computing in Biomedicine

Dimosthenis Kyriazis
National Technical University of Athens, Greece

Andreas Menychtas
National Technical University of Athens, Greece

Konstantinos Tserpes
National Technical University of Athens, Greece

Theodoros Athanaileas
National Technical University of Athens, Greece

Theodora Varvarigou
National Technical University of Athens, Greece

ABSTRACT]

A constantly increasing number of applications from various scientific fields are finding their way towards adopting Grid technologies in order to take advantage of their capabilities: the advent of Grid environments made feasible the solution of computational intensive problems in a reliable and cost-effective way. This book chapter focuses on presenting and describing how high performance computing in general and specifically Grids can be applied in biomedicine. The latter poses a number of requirements, both computational and sharing / networking ones. In this context, we will describe in detail how Grid environments can fulfill the aforementioned requirements. Furthermore, this book chapter includes a set of cases and scenarios of biomedical applications in Grids, in order to highlight the added-value of the distributed computing in the specific domain.

INTRODUCTION

High performance distributed computing has emerged in the last years as a technology for large-scale, flexible and coordinated resource sharing. A successful example of high performance computing are Grids (Foster, 1999), which are increasingly considered as an infrastructure able to provide distributed and heterogeneous resources in order to deliver in a transparent way computational power to

DOI: 10.4018/978-1-60566-768-3.ch006

resource demanding applications (Foster, 2001), (Leinberger, 1999). The main objective of any distributed infrastructure is to serve as a means for providing resources for a set of purposes such as computational / processing, data storage / networking of file systems, communications and bandwidth, applications as services etc. Furthermore, a Grid-based environment enables the storage and distribution of data allowing access to various sources and analysis of them. Therefore, the information contained for example in medical records can be accessed and analyzed for various reasons (e.g. selection of the best treatment and prediction of its outcomes).

Exploitation of Grid technologies is imperative for medical applications due to a set of reasons such as the exponential increase of the required storage and computational resources, the heterogeneity of the required data (medical records, images, information obtained from sensors) with different preprocessing requirements and the large number of involved patients. Medical-related applications generally belong to those collaborative environments that are based on input from networked sensors and aggregation of acquired data under real-time conditions. With the simultaneous advent of technologies to support heterogeneous sources of information and computing resources (through Service Oriented Architectures - SOA (Sprott, 2004), Grids, etc), it is expected that in the years to come, there will be a great blooming in the development of infrastructures comprising multiplicity in resources both in number and nature.

To this end, a significant part of the value of Grid technology lies on the fact that Grids are in position to provide the fundamental management mechanisms for distributed data. This is one major reason that often many developed Grid-based systems were referenced as "data Grids", since the integration of data, infrastructures, digital libraries and persistent archives was a challenge forcing continued evolution of Grid technology. This challenge remains valid for medical applications, the requirements of which range from the transition from data handling, sharing and aggregation to the provision of knowledge as utility. Therefore, besides the main property of Grids - referring to their computational power, we will also describe in detail their added-value in terms of data sharing and aggregation.

Nowadays, high performance systems are realized following the SOA principles: as Grids and Clouds (Boss, 2007) are considered to be the most successful examples; and are finding their way into the medical sector both for computational reasons and for storage and aggregation of medical data. A set of medical scenarios will also be described in this book chapter to demonstrate the added value of distributed computing in the aforementioned cases. These refer to the use of Grids for computational reasons for a clinical trial simulation and for data aggregation and analysis for the simulation of possible therapeutic schemes and personalized medicine proposals.

The remainder of this chapter is structured as follows: the first section (namely *Background*) includes information related to high performance computing as realized through Service Oriented Infrastructures (SOIs) (Wikipedia, 2008) and Grids, while the following section focuses on Grid environments and the way these can be used either as computational infrastructures (Computational Grids) or as distributed, networking and sharing ones (Data Grids). Thereafter, in the subchapter called "*Biomedical Applications: Cases and Scenarios in Grids"* we discuss various cases and scenarios of biomedical applications in Grids. Finally, the last subchapter concludes with a discussion on future research and potentials for the current study.

BACKGROUND

As already mentioned, research efforts of the past years have led to the realization of high performance computing in environments that follow the