Chapter VII

Compliance and Creativity in Grid Computing[1]

Anthony E. Solomonides, University of the West of England, Bristol, UK

Abstract

Grid computing is a new technology enhancing services already offered by the Internet offering rapid computation, large-scale data storage, and flexible collaboration by harnessing together the power of a large number of commodity computers or clusters of basic machines. The grid has been used in a number of ambitious medical and healthcare applications. While these have been restricted to the research domain, there is a great deal of interest in real applications. There is some tension between the spirit of the grid paradigm and the requirements of healthcare applications. The grid maximises its flexibility and minimises its overheads by requesting computations to be carried out at the most appropriate node in the network; it stores data at the most convenient node according to performance criteria. A healthcare organization is required to maintain control of its patient data and be accountable for its use at all times. Despite this apparent conflict, certain characteristics of grids help to resolve the problem: "grid services" may provide a solution by negotiating ethical, legal, and regulatory compliance according to agreed policy.

Introduction: The Computing Context

This chapter will develop an argument concerning "healthgrid," the application of grid computing to biomedical research and healthcare. The issues that will concern us arise in the first place from a number of exemplar projects and research prototypes in the field.

Let us first consider the concept of grid computing. "Distributed computer systems" predate even the Internet and the World Wide Web ("the Web"). By means of a network of interconnections, computers are able to share a workload that would ordinarily be beyond the capacity of any one of them; they may also distribute data to different locations according to need or frequency of use. On the other hand, since the explosion of the Web in every conceivable statistic—users, nodes, volume of information—we are familiar with its ability to serve information and misinformation in equal measure. The grid combines the technical features of distributed systems and the Web, but efforts are also being made to ensure that it is not beset by the same problems of abuse, misuse and contamination as the Web has been.

The ideal grid, envisaged as a servant of a new paradigm of scientific research called "e-science," would provide transparent processing power, storage capacity and communication channels for scientists who may from time to time join the grid, do some work and then leave, so that the alliances they form in their scientific endeavours might be described as "virtual organizations" or VOs for short. Different sciences have different needs, and the grid concept has become differentiated: particle physics generates enormous amounts of data which must be kept, but not necessarily instantly processed; on the other hand, data in bioinformatics is not large by comparison—it is, of course, in plain terms, large—but requires intensive processing.

In extending the application of grid computing to biomedical and healthcare applications, another feature becomes pre-eminently necessary: that of collaboration. Indeed, from the outset, grid computing has been associated with smoothing the process, or "workflow," in data- and computation-intensive science; this is the notion, which led to the term e-science. Were it not that the term "e-health" was already in use and generally accepted to mean something at once broader and narrower—"health services and information delivered or enhanced through the Internet" (Oh, Rizo, Enkin, & Jadad, 2005)—the term "healthgrid" would not have had to be invented.

An important consequence of the fluidity of collaboration in grid computing has been in the choice of "architecture" for grid systems. "Architecture" is used loosely in computer systems to describe the manner in which hardware and software have been assembled together to achieve a desired goal. Favoured also in the commercial application of the Web, the so-called "service-oriented architecture" (SOA) has been widely adopted in grid applications. In effect, it means that needed services—software applications—once constructed, are provided with a description in an agreed language and made available to be "discovered" by other services that need them. A

Related Content