Chapter XXX

UNICORE: A Middleware for Life Sciences Grids

Piotr Bała
ICM University of Warsaw & N. Copernicus University, Poland

Kim Baldridge
University of Zurich, Switzerland

Emilio Benfenati
Istituto Mario Negri, Italy

Mosè Casalegno
Istituto Mario Negri, Italy

Uko Maran
University of Tartu, Estonia

Łukasz Mirosław
University of Zürich, Switzerland

Vitaliy Ostropytskyy
University of Ulster, UK

Katharina Rasch
Technische Universität Dresden, Germany

Sulev Sild
University of Tartu, Estonia

Robert Schöne
Technische Universität Dresden, Germany

Bernd Schuller
Research Centre Juelich, Germany

Nadya Williams
University of Zürich, Switzerland

ABSTRACT

This chapter provides an overview of Grid middleware and applications related to biomedical and life sciences disciplines. Various technologies, including web-based solutions, are presented. One of the solutions, the UNICORE framework, in its recent version implements key grid standards and specifications. The system architecture and capabilities, such as security, workflow and data management are described. Special attention is given to the idea of a ‘gridbean’, which expands the UNICORE use for different applications. Examples of gridbeans are provided and the capabilities of UNICORE are illustrated through specific examples built using this grid middleware. In particular, the Chemomentum workbench and its use for in-silico design and modeling in chemistry and life sciences are both described.
INTRODUCTION

The past decade has witnessed tremendous growth in our chemical and biological understanding at the molecular scale that impacts all facets of our daily lives. Inherently driven by information technologies (IT) and massive data acquisition, this growth in understanding requires a stronger coupling between computation and experiment. The expansion of high-throughput technologies—chemical design, microarray analysis, device technology, physical and chemical analysis processes, combinatorial chemical synthesis, and screening, as well as their computational counterparts, has generated an unprecedented volume and diversity of data that must be abstracted into manageable units. Historically, chemistry has evolved a sophisticated symbolic and iconographic way of projecting a myriad of chemical phenomena onto a model of molecular structure and associated properties. This rapid growth of both experimental and computational data has driven the proliferation of computational tools and the development of first-generation chemoinformatics for data storage, analysis, mining, management, and presentation. However, it is generally accepted that first-generation chemoinformatics tools have not been widely adopted, nor have they met the needs of the researchers. There is a general need for integrating data, information and knowledge that can be accomplished through innovative ways of computational access, resource integration, and creation of transparent toolsets.

The computational science challenge, therefore, is to provide new algorithms and data analysis tools that can exploit more efficiently the existing computational power for more detailed chemical analysis. The corresponding challenges in information technology and visualization include providing richer access to existing repositories of data, enabling transformation of that data into interpretable and predictable scientific information, and leveraging workflows to parallelize the efforts of many researchers for general access.

Over the past few years our team has been developing middleware and end-user interface technologies to enable novice and advanced users to access capabilities of sophisticated scientific applications running on large-scale grid computing resources. In this paper, we present some of these developments, from web-based tools to integrated toolkits such as NIMROD (Abramson, Lewis, Peachy 2001) and BOINC (Boinc 2008), to full-featured grid middleware such as Globus (Foster 2006) and UNICORE (Erwin, Snelling 2002). The main focus is on the latter, which is known from a number of successful deployments in the life sciences area.

BACKGROUND

Web Based Solutions

In the past, the primary access to computational resources was based on batch processing. The input files and command scripts were typically prepared using simple text editors, and the command line tools were used to transfer files from local workstations to remote servers located in the dedicated supercomputer centers, in the campus or departmental networks. The progress in web technology enabled replacement of some command line tools with a web browser. This opened the field for an easy access to distributed resources. A few successful examples of such interface tools are described here. All have differing capabilities, some offer very limited grid functionality yet clearly illustrate a change in the mode of access to the computational resources.

WebMO

WebMo (WebMo 2008) is a web based submission system for quantum chemistry software such as Gaussian (Gaussian03 2004), GAMESS (Schmidt et al. 1993), and MOPAC (Stewart 1990). Users can draw molecular structures in a 3D editor, run...
Related Content

A Novel Detection Approach for Cardio-Respiratory Disorders Using PPG Signals
[www.igi-global.com/article/a-novel-detection-approach-for-cardio-respiratory-disorders-using-ppg-signals/86048?camid=4v1a](www.igi-global.com/article/a-novel-detection-approach-for-cardio-respiratory-disorders-using-ppg-signals/86048?camid=4v1a)

Information Retrieval by Semantic Similarity
[www.igi-global.com/chapter/information-retrieval-semantic-similarity/26248?camid=4v1a](www.igi-global.com/chapter/information-retrieval-semantic-similarity/26248?camid=4v1a)

Data Provenance in Scientific Workflows
[www.igi-global.com/chapter/data-provenance-scientific-workflows/35687?camid=4v1a](www.igi-global.com/chapter/data-provenance-scientific-workflows/35687?camid=4v1a)

Reliability and Evaluation of Health Information Online
[www.igi-global.com/chapter/reliability-evaluation-health-information-online/26354?camid=4v1a](www.igi-global.com/chapter/reliability-evaluation-health-information-online/26354?camid=4v1a)