

HiBB 2011: 2nd Workshop on High Performance Bioinformatics and Biomedicine

Mario Cannataro

Bioinformatics Laboratory,
Department of Medical and Surgical Sciences,
University Magna Græcia of Catanzaro,
88100 Catanzaro, Italy
cannataro@unicz.it

Foreword

The availability of high-throughput technologies, such as microarray and mass spectrometry, and the diffusion of genomics and proteomics studies to large populations, are producing an increasing amount of experimental and clinical data. Biological databases and bioinformatics tools are key tools for organizing and exploring such biological and biomedical data with the aim to discover new knowledge in biology and medicine. However the storage, preprocessing and analysis of experimental data is becoming the main bottleneck of the analysis pipeline.

High-performance computing may play an important role in many phases of life sciences research, from raw data management and processing, to data integration and analysis, till data exploration and visualization, so well known high performance computing techniques such as Parallel and Grid Computing, as well as emerging computational models such as Graphics Processing and Cloud Computing, are more and more used in bioinformatics.

The huge dimension of experimental data is the first reason to implement large distributed data repositories, while high performance computing is necessary both to face the complexity of bioinformatics algorithms and to allow the efficient analysis of huge data. In such a scenario, novel parallel architectures (e.g. CELL processors, GPU, FPGA, hybrid CPU/FPGA) coupled with emerging programming models may overcome the limits posed by conventional computers to the mining and exploration of large amounts of data.

The second edition of the Workshop on *High Performance Bioinformatics and Biomedicine* (HiBB) aimed to bring together scientists in the fields of high performance computing, computational biology and medicine to discuss the parallel implementation of bioinformatics algorithms, the application of high performance computing in biomedical applications, and the organization of large scale databases in biology and medicine. As in the past, also this year the workshop has been organized in conjunction with Euro-Par, the main European (but international) conference on all aspects of parallel processing.

Presentations were organized in three sessions. The first session (Bioinformatics and Systems Biology) comprised two papers discussing the parallel

implementation of bioinformatics and systems biology algorithms on multicore architectures:

- On Parallelizing On-Line Statistics for Stochastic Biological Simulations
- Scalable Sequence Similarity Search and Join in Main Memory on Multi-Cores

The second session (Software Platforms for High Performance Bioinformatics) comprised two papers describing software environments for the development of bioinformatics workflows:

- Enabling Data and Compute Intensive Workflows in Bioinformatics
- Homogenizing Access to Highly Time-Consuming Biomedical Applications throughout a Web-Based Interface

Finally, the third session included a tutorial on:

- Distributed Management and Analysis of Omics Data.

This post-workshop proceedings includes the final revised versions of the HiBB papers and tutorial, taking the feedback from reviewers and workshop audience into account.

The program chair sincerely thanks the Euro-Par organization, for providing the opportunity to arrange the HiBB workshop in conjunction with the Euro-Par 2011 conference, the program committee and the additional reviewers, for the time and expertise they put into the reviewing work, and all the workshop attendees who contributed to a lively day.

October 2011

Mario Cannataro