INTRODUCTION TO THE SPECIAL ISSUE ON CLOUD FOR HEALTH

Dear SCPE readers,

It is a pleasure to present this special issue covering subjects related to applying cloud computing to bioinformatics, biomedicine, and health, including solutions to area problems and architectural adaptation of cloud systems to fit those problems. The special issue includes papers selected from the workshops C4Bio 2014 and CCGrid-Health 2014 that, after further extension and additional review, were selected for publication. C4Bio 2014 and CCGrid-Health 2014 were held in Chicago, during May 26-29th 2014 within the framework of the 14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid 2014).

The special issue begins with the review paper "Cloud Computing in Healthcare and Biomedicine". Calabrese and Cannataro explore the current academic and commercial cloud-based solutions developed in the healthcare and biomedicine domains, designed in order to satisfy the specific needs (e.g. massive scalable computing and storage, data sharing, on-demand access anytime and anywhere) in these fields. Subsequently, the paper summarizes main issues regarding the use of Cloud Computing in healthcare and biomedicine.

The article "Selected Approaches and Frameworks to Carry out Genomic Data Analysis on the Cloud", by Church and Goscinski, examines commonly used cloud-based genomic analysis services, introduces the approach of exposing data as services and proposes two new solutions, namely HPCaaS and Uncinus, which aim to automate service development, deployment process and data provision. By comparing and contrasting these solutions, the authors identify key mechanisms of service creation, execution and data access required to support non-computing specialists employing clouds.

Ahmad, Hasan, and Siddique propose to formally reason about molecular pathways of Zsyntax based deduction using the HOL4 theorem prover and developing a biologist friendly graphical user interface. "On the Formalization of Zsyntax with Applications in Molecular Biology" shows additionally a use case with the formal analysis of cancer-related molecular pathway, i.e., TP53 degradation and metabolic pathway, known as Glycolysis.

The article "Accelerating Comparative Genomics Workflows in a Distributed Environment with Optimized Data Partiotioning and Workflow Fusion", by Choudhury et al., discusses different strategies for alleviating the computational constraints experienced by several data-intensive bioinformatics applications, presenting a workflow-based framework to parallelize the compute-intensive tasks of genome alignment and variant calling, the two most important stages in most comparative genomics applications. This framework provides workflow fusion to merge multiple sequential workflows into a single optimized workflow that benefits from caching, eliminating choke points, and stacked partitioning.

In "Performance Comparison and Tuning of Virtual Machines For Sequence Alignment Software", Estrada et al. explore the performance cost of virtualization for the fast growing application domain of genomics to inform the feasibility of running an NGS pipeline in a cloud, and in doing so consider two prevalent short-read sequence alignment programs, BWA and Novoalign. They execute these applications in three separate open-source system virtualization solutions (KVM hypervisor, Xen paravirtualized hypervisor, and Linux Containers) and compare the runtime in each environment against the runtime of the same system without virtualization and measure the relative performance of each hypervisor, presenting tuning suggestions for cloud implementors and users.

The article "Extending XNAT towards a Cloud-based Quality Assessment Platform for Retinal Optical Coherence Tomographies", by Jansen et al., presents a platform designed for automatic quality assessment of retinal OCTs and provided as a cloud-based service employing OpenStack. It extends the image management platform XNAT by services to calculate and store quality measures. It is also extensible regarding new quality measure algorithms, allowing the developer to upload, compile and test code for the system's architecture.

Finally, the paper "A Tool for Managing the X1.V1 Platform on the Cloud", by Marzini et al., proposes a framework for managing the execution of the X1.V1 platform on the Cloud. This framework enables an easy, quick, and secure management of the Cloud resources allocation and reallocation to X1.V1 Virtual Machines, in order to enhance the platform performance, optimize resource utilization and, consequently, reduce the whole services cost.

We would like to thank all the reviewers for their effort, cooperation, and valuable feedback and all the authors who submitted papers to C4Bio 2014 and CCGrid-Health 2014 and to this Special Issue.

Jesus Carretero and Javier García-Blas, Universidad Carlos III de Madrid, Spain
Sandra Gesing, University of Notre Dame, USA