Accelerating Bioinformatics Applications via Emerging Parallel Computing Systems

Juan A. Gomez-Pulido, Bertil Schmidt, and Wu-chun Feng

THIS special section on Advanced Parallel Computing Systems to Accelerate Bioinformatics provides a forum to publish recent advances in the improvement of handling bioinformatics problems on emerging parallel computing systems. These systems can be characterized by exploiting different types of parallelism, including fine-grained versus coarse-grained and thread-level parallelism versus data-level parallelism versus request-level parallelism. Hence, parallel computing systems based on multi- and many-core CPUs, many-core GPUs, vector processors, or FPGAs offer the promise to massively accelerate many bioinformatics algorithms and applications, ranging from compute-intensive to data-intensive. Such computing systems are increasingly ubiquitous, ranging from "big iron" datacenter supercomputers and datacenter cloud computing down to GPU-accelerated smartphones and laptops.

With the goal of shortest time to solution, or even real-time performance, across a gamut of compute-intensive and data-intensive applications in bioinformatics, faster execution time can be achieved via hardware, software, or a combination of both in these parallel computing systems. These computing systems may use different configurations, architectures, and technologies according to the specific problem to be tackled, combining the different levels of parallelism. Therefore, the design and use of advanced parallel computing systems are a challenge that opens up new possibilities to accelerate bioinformatics algorithms and applications. In pursuing these possibilities, there are many trends and open issues that deserve to be investigated.

This special section on Advanced Parallel Computing Systems to Accelerate Bioinformatics presents eight original, high-quality articles, clearly focused on theoretical and practical aspects of the design, implementation, and efficient application of hardware architectures to accelerate bioinformatics problems, including cutting-edge topics about hardware accelerators, architectures based on graphical processing units or FPGA reconfigurable devices, high-performance computing, grid and cloud computing, and real-world application cases.

- *J.A. Gomez-Pulido is with the Department of Technologies of Computers and Communications, University of Extremadura, Polytechnic School, 10003, Caceres, Spain. E-mail: jangomez@unex.es.*
- *B. Schmidt is with the Institute of Computer Science, Universität Mainz, 55128, Mainz, Germany. E-mail: bertil.schmidt@uni-mainz.de.*
- *W.-C. Feng is with the Department of Computer Science, Virginia Tech, VA 24061. E-mail: feng@cs.vt.edu.*

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TCBB.2015.2457736

The first article in this special section is entitled "FHAST: FPGA-Based Acceleration of BOWTIE in Hardware" by Edward B. Fernandez, Jason Villarreal, Stefano Lonardi, and Walid A. Najjar. This work shows how the intrinsic parallel capabilities of FPGA devices can be applied to accelerate important operations in the bioinformatics domain. Using this technology, the authors have tackled the problem of mapping short-sequence reads to a reference genome by designing a hardware accelerator for the popular BOWTIE mapping tool, which exploits concurrent access to the memory to assist in unblocking the analysis pipeline.

The second article is entitled "Parallelizing Epistasis Detection in GWAS on FPGA and GPU-Accelerated Computing Systems" by Jorge Gonzalez-Dominguez, Lars Wienbrandt, Jan Christian Kassens, David Ellinghaus, Manfred Schimmler, and Bertil Schmidt. High-throughput genotyping technologies allow the rapid collection of millions of single nucleotide polymorphisms (SNPs). Detecting epistasis in genome-wide association studies (GWAS) based on pairwise SNP-interactions is a highly time-consuming operation. Gonzalez-Dominguez et al. show how this important algorithm can be accelerated by several orders of magnitude on two types of massively parallel computing systems based on FPGAs and GPUs, respectively.

The third article is "Concurrent and Accurate Short Read Mapping on Multicore Processors" by Héctor Martínez, Joaquín Tárraga, Ignacio Medina, Sergio Barrachina, Maribel Castillo, Joaquín Dopazo, and Enrique S. Quintana-Ortí. Besides FPGA and GPU devices, a more traditional approach to accelerate bioinformatics applications is to better exploit the latent parallelism in multicore processors. This paper focuses on mapping RNA sequences using servers equipped with multicore processors and designing a parallel aligner that exploits a suffix array to rapidly map a large fraction of the RNA reads. This design detects splice junctions based on an adaptive division of RNA reads into small segments, reporting a performance in execution time that surpasses commonly used state-of-the-art aligners.

The fourth article is "Parallel Mutual Information Based Construction of Genome-Scale Networks on the Intel (R) Xeon Phi^{TM} Coprocessor" by Sanchit Misra, Kiran Pamnany, and Srinivas Aluru. In this article, the authors tackle the problem of handling network reconstruction at the whole-genome scale by using a parallel architecture based on the Intel Xeon Phi coprocessor, which takes advantage of multi-level parallelism, particularly with respect to the gene interactions in partially characterized pathways. The attractiveness of this single coprocessor approach lies in the wide availability to handle such biological systems,

1545-5963 © 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

as well as the applicability of some of its optimization techniques (such as random shuffle and mutual information) to other applications.

The fifth article is entitled "Large-Scale Tissue Morphogenesis Simulation on Heterogenous Systems Based on a Flexible Biomechanical Cell Model" by Anne Jeannin-Girardon, Pascal Ballet, and Vincent Rodin. The research addressed in this article stems from the need to address the complexity of biological tissue morphogenesis, for which the authors propose a virtual cell model coupled to an agent-based formalism and a simulator built from the OpenCL framework. This way, this paper shows again how heterogeneous devices such as CPU or GPU can be used to drive efficient parallel simulations. In order to validate the implementation of the cell model, the authors presented two case studies: (1) cellular proliferation controlled by cell signaling and (2) limb growth in a virtual organism.

The sixth article is entitled "An Application Specific Instruction Set Processor (ASIP) for Adaptive Filters in Neural Prosthetics" by Yao Xin, Will X.Y. Li, Zhaorui Zhang, Ray C.C. Cheung, Dong Song, and Theodore W. Berger. This article elaborates on the advantages of using application-specific instruction set processors for adaptive filters in neural decoding activity for neuroprosthetic design, instead of other hardware approaches based on low-power, general-purpose processors or application-specific integrated circuits. The proposed processors accelerate the most time-consuming matrix/vector operations among three representative adaptive filters, outperforming pure software implementations on current CPUs in execution time and clock cycles.

The seventh article is entitled "Boosting the FM-Index on the GPU: Effective Techniques to Mitigate Random Memory Access" by Alejandro Chacon, Santiago Marco-Sola, Antonio Espinosa, Paolo Ribeca, and Juan Carlos Moure. Many accelerator implementations based on CPUs and GPUs share a common feature: memory access limits efficiency. This limitation is especially clear in high-throughput sequencing machines that generate copious amounts of short reads, thus creating memory-intensive access patterns to process these short reads. The authors of this article tackle this limitation by reducing the number of memory accesses on the GPU by combining several strategies, mainly using a modified k-step FM-index algorithm, speeding up the current multi-core CPU version and the implementation on the GPU provided by the new Nvidia NVBIO bioinformatics library.

The eighth and final article in this special section is entitled "Efficient and Accurate OTU Clustering with GPU-Based Sequence Alignment and Dynamic Dendrogram Cutting" by Thuy-Diem Nguyen, Bertil Schmidt, Zejun Zheng, and Chee-Keong Kwoh. De novo OTU (operational taxonomic units) clustering is a popular technique to perform taxonomic profiling of a microbial community by grouping 16S rRNA amplicon reads. Nguyen et al. introduce a novel dendrogram-based clustering pipeline which is able to produce more accurate OTU groupings compared to most existing approaches.

Finally, we thank all of the authors who submitted their manuscripts to this special section for considering the *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, and the reviewers for their hard work with the reviewing process.

Juan A. Gomez-Pulido
Bertil Schmidt
Wu-chun Feng
Guest Editors

Juan A. Gomez-Pulido received the PhD degree in physics, electronics specialty, from the Complutense University, Madrid, Spain, in 1993. He is currently a professor of design of processors in the Department of Computers and Communications Technologies, University of Extremadura. He has authored or co-authored 60 ISI journals, and almost 100 book chapters, and 300 peer-reviewed conference proceedings. He has participated in 19 funded research projects, leading some of them. His research lines fall within hot topics on reconfigurable and embedded computing based on FPGAs, multiobjective optimization, and wireless sensor networks. Currently, he is designing parallel reconfigurable architectures applied to several bioinformatics problems, such us bicluster processing and gene selection for cancer classification.

Bertil Schmidt (M'04-SM'07) is a tenured full professor and a chair for Parallel and Distributed Architectures at the University of Mainz, Germany. Prior to that, he was a faculty member at Nanyang Technological University (Singapore) and at the University of New South Wales (UNSW). His research group has designed a variety of algorithms and tools for Bioinformatics mainly focusing on the analysis of large-scale sequence and short read datasets. For his research work, he has received a CUDA Research Center award, a CUDA Academic Partnership award, a CUDA Professor Partnership award and the Best Paper Award at IEEE ASAP 2009. Furthermore, he serves as the champion for Bioinformatics and Computational Biology on gpucomputing.net. He is a senior member of the IEEE.

Wu-chun Feng (M'98-SM'04) received the BS and MS degrees in computer engineering from Penn State University in 1988 and 1990, respectively, and the PhD degree in computer science from the University of Illinois at Urbana-Champaign in 1996. He is a professor and Elizabeth & James Turner fellow in the Department of Computer Science at Virginia Tech and holds additional appointments in Electrical & Computer Engineering, Health Sciences, and the Virginia Bioinformatics INstitute. He directs both the Synergy Laboratory and the Synergistic Environments for Experimental Computing (SEEC) Center. Previous professional stints include IBM T.J. Watson Research Center, NASA Ames Research Center, Vosaic, University of Illinois, Purdue University, The Ohio State University, Orion Multisystems, and Los Alamos National Laboratory. His research interests lie in high-performance computing with specific interests in accelerator-based parallel computing, green supercomputing, dynamic multicore and many-core scheduling, and bioinformatics.. He is a senior member of the IEEE.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.