Editorial

Christophe Cérin

Université Paris 13, Sorbonne Paris Cité,
LIPN UMR CNRS 7030,
99 avenue Jean-Baptiste Clément,
F-93430 Villetaneuse, France
Email: christophe.cerin@lipn.univ-paris13.fr
Email: c.cerin@computer.org

Biographical notes: Christophe Cérin has been a Professor of Computer Science at the University of Paris 13, France since 2005. At Paris 13, he chairs the board for the cluster computing facility available to all campus scientists and also chairs the 'Expert Committee' in charge of recruiting and mentoring full time junior and senior professors in computer science. His current industrial experience includes serving as local chair for the Wendelin project related to Cloud and Big-Data. In 2015, he was also acting as Chairman for a joint research Lab with Chinese researchers in Beijing, Wuhan and Hangzhou. He is also acting as an editorial board member for the *IEEE Cloud Computing Magazine* and as an expert on cloud computing for ALECSO and ITU. His research focuses on high performance computing, including grid and cloud computing and he develops middleware, algorithms, tools and methods for distributed systems.

Welcome to this new issue of *IJBDI*. This time we gathered four papers, two of them are related to the acceleration of I/O operations, one is related to machine learning in a broad sense, one on the coupling of software engineering tools to have more confidence in the codes we may develop on GPU. A longer description of the authors work is now introduced to better appreciate them as relevant contributions to our journal.

The paper entitled 'A framework for collective I/O style optimisations at staging I/O nodes' by people from the University of Houston and from the HDF group investigate the potential of a technique that makes use of staging nodes to delegate I/O requests. The effort is accomplished not at the application level but at the system level, we mean that authors propose some kind of general middleware to accelerate I/O operations in the context of exascale systems. This middleware is called compactor and its name perfectly reflects some useful techniques among them collective buffering across requests from multiple processes, write stealing to service read requests at the staging node, and write morphing to optimise multiple write requests from the same process. Another strength of the paper is the evaluation part which is done with the popular PVFS2 file system using micro-benchmarks, the flash I/O benchmark and a parallel image processing application.

The paper entitled 'Towards cost-effective and high-performance caching middleware for distributed systems' by researchers from Argonne National Laboratory and from the Department of Computer Science, Illinois Institute of Technology, Chicago focuses on SSD drives, distributed over a system to form a big storage system. Authors design and implement a user-level caching system that offers SSD-like performance at a cost similar to a HDD. This means that with a fraction of SSD disks and a large number of HDD disks, they obtain very good performance (much better than in using only HDD disks), thanks to the caching and the scheduling policies that the author introduced in the paper. The key point is to put some intelligence into the SDD drives in order to cache data. In fact the paper covers a large set of techniques, past and present, but also implementation in state-of-the-art Linux-based systems.

The paper entitled 'Rainfall forecasting by relevant attributes using artificial neural networks – a comparative study' by researchers from National Institute of Technology, Silchar in INDIA is related to tuning effective parameters to improve the forecasting ability of a rainfall model. The quite classical neural network model is used throughout the paper with insights to artificial neural networks and variants such as multi-layer feed forward neural network, radial basis function neural network, focused time delay neural network (FTDNN) and nonlinear auto regressive exogenous input neural network. The performance model used in the evaluation is also the classical model of performance used in the literature. Experiments conducted by authors demonstrated the pro and cons of each approach and may serve as guidelines to tune specific tools.

The paper entitled 'Multi approach for real-time systems specification: case study of GPU parallel systems' by researchers from two different labs in Algeria is about the coupling of two systems to better schedule activities (jobs/tasks) on GPU architecture. The first system is MARTE, a UML-oriented tool for the specification of scheduling and timing constraints. The second one is the formal method event B, based on the B method from Abrial, to get a valid and proved specification. In some ways, the paper is a software engineering paper based on the

requirement analysis but with practical outputs in the field of high performance computing. This kind of paper establishes a bridge between different communities of researchers.

To conclude, I would like to thank all the reviewers and the editorial team for the work they have accomplished during the last few months.