SPPEXA Workshop (4th Edition) on

Parallel Programming Models - Productivity and Applications for Exascale and Beyond

University of Versailles - March 21, 2019

The three fields, parallel programming models, HPC tools to foster productivity, and applications in numerical computing form the common interests of the several SPPEXA projects and comprised the themes of the three previous editions of the workshop hold in Japan (Tokyo 2017), France (Versailles 2017) and Germany (Aachen 2018). Through the different editions the workshop has become a great opportunity to discuss future research on software for Exascale supercomputers and beyond. The goal of this 4th edition of workshop is to foster collaborations between attendance from Europe and Japan beyond the SPPEXA projects. In addition to the invited presentations (*Barbara Chapman, Michel Daydé and Michael A. Heroux*), the agenda contains contributions from the SPPEXA PIs on the recent progress. The agenda and the social interactions are expected to make new contacts among the participants to build upon in the future.
Chair: Nahid Emad, Co-chair: Thomas Dufaud, Local Organizer: Isabelle Moudenner

Program of the workshop

09:40-09:50	Nahid Emad, UVSQ	***Welcome introduction***
Chair: Miwako TSUJI (Riken/Kobe, Japan)		
09:50-10:00	Severin REIZ, TUM, Germany	***Some words from SPPEXA***
10:00-10:20	Taisuke BOKU, Center for Computational Sciences, University of Tsukuba, Japan	***MUST system applied to high level language approach in MYX project***
10:20-10:40	Joseph SCHUCHART, HLRS Stuttgart, Germany	***A distributed task scheduler in the DASH project***
10:40-11:00	Michel DAYDÉ, Informatics Research Institute of Toulouse – Toulouse University – CNRS, France	***Multidisciplinary High Performance Data Analysis Forum: summary and recommendations***
11:00-11:30	Cafe break	
11:30-12:15	Barbara CHAPMAN Stony Brook University, USA	***Programming for the Future: Are We There Yet?***
12:15-13:30	Lunch	
Chair: Thomas DUFAUD (UVSQ, France)		
13:30-14:15	Michael A. HEROUX, Sandia National Laboratories, USA	***Making Reproducibility Indispensable: Changing the Incentives that Drive Computational Science***
14:15-14:35	Matthias MÜLLER, RWTH Aachen University, Germany	***Overview of MYX (results and perspectives)***
14:35-14:55	Gerhard WELLEIN, University of Erlangen-Nuremberg, Germany	***Overview of ESSEX II (results and perspectives)***
14:55-15:15	Jose GRACIA, HLRS, Germany	***Overview of Smart-DASH (results and perspectives)***
15:10-15:40	Cafe break	
Chair: Joachim PROTZ (RWTH Aachen university, Germany)		
15:40-16:00	Kengo NAKAJIMA, University of Tokyo, Japan	***Numerical Library with High-Performance/Adaptive-Precision/High-Reliability: Extension of ppOpen-HPC towards the Post Moore Era***
16:00-16:20	Christian TERBOVEN, RWTH Aachen University, Germany	***POP: A Centre of Excellence in HPC (Results and service offerings)***
16:20-16:40	Benjamin UEKERMANN, TUM, Germany	***Scalable and Easy-to-use Coupling of Multi-Physics Simulations***
16:40-17:30	B. Chapman, M. Heroux, M. Müller, S. Petiton, K. Nakajima (moderator: Nahid Emad)	Panel on ***the convergence of big data analytics & high scale computing, Future collaboration***
17:30-17:40	Closing remarks	

ABSTRACTS

Title: ***Programming for the Future: Are We There Yet?***

Presenter: Barbara Chapman, Stony Brook University, USA

Abstract: A significant effort is being expended internationally to develop programming interfaces that meet the needs of extreme-scale applications, and there has been much progress in the last few years. Yet our computer hardware is expected to grow increasingly complex in the future in order to provide highest levels of performance with acceptable power consumption. What open challenges does this imply, and what research is needed if we are to ensure that future platforms can be utilized with reasonable effort?

Title: ***Multidisciplinary High Performance Data Analysis Forum: summary and recommendations***

Presenter: P.-H. Cros (Informatics Research Institute of Toulouse – Toulouse University), **M. Daydé** (Head of Informatics Research Institute of Toulouse – Toulouse University – CNRS), O. Marques (Lawrence Berkeley National Laboratory)

Abstract: The Multidisciplinary High Performance Data Analysis for Societal Challenges Forum held in Saint Girons (France) on June 17-19, 2018 was organized by the Informatics Research Institute of Toulouse (IRIT) and the Lawrence Berkeley National Laboratory (LBNL), with the support of the LabEx CIMI (Centre International de Mathématiques et Informatique de Toulouse).
The participants included decision-makers in research laboratories and academy, all faced with problems related to societal challenges. The forum had an international dimension with the participation of representatives from the USA, Japan and Europe.
The planning of the forum was motivated by the realization that the digital transformation is of great concern for public and private research organizations. This forum will be convened yearly, such that participants will be able to take stock, share experiences, and propose dynamic approaches for organizing R&D communities.
The first Forum sought to foment discussions about ongoing or yet to-be-implemented mechanisms for data collection and processing related to the societal challenges. It is important to assess their potential impact, particularly on the management of research teams working on societal challenges.
We will summarize the outcome for the Forum in recommendations that will be presented during the talk.

Title: Making Reproducibility Indispensable: Changing the Incentives that Drive Computational Science

Presenter: Michael A. Heroux (SNL, USA)

Abstract: Reproducibility is essential for scientific progress and engineering advances. Even so, many published computational results lack sufficient capture and description of companion information that would enable subsequent confirmation and extension of the results. Certainly, most scientists intend to publish correct results, but without sufficient rigor in computational processes and practices, risk is unnecessarily high that results will occasionally be wrong and will always be costly to confirm and extend.

The reasons for inadequate reproducibility are fundamentally matters of incentives and costs. In recent years, because of the availability of improved software platforms from GitHub, GitLab, Atlassian and others, and container environments such as Docker, the cost of capturing and describing the computing environment used to produce scientific results has dramatically decreased. Furthermore, new workflows and skill-building opportunities are available for those who are interested in improving their practice. What needs further attention is our incentive system.

In this presentation, we discuss efforts to improve computational reproducibility by fostering and promoting changes to our incentive systems. We talk about efforts to increase reproducibility expectations by publishers, funding agencies, employers and the broader computational scientific community. By improving incentives to produce reproducible results, providing recognition for those who lead the community and providing conduits for effective exchange of best practices, we can expect and make reproducibility indispensable.

Mike Heroux's Bio: Mike Heroux is a Senior Scientist at Sandia National Laboratories, Director of SW Technologies for the US DOE Exascale Computing Project (ECP) and Scientist in Residence at St. John's University, MN. His research interests include all aspects of scalable scientific and engineering software for new and emerging parallel computing architectures.

He leads several projects in this field: ECP SW Technologies is an integrated effort to provide the software stack for ECP. The Trilinos Project (2004 R&D 100 winner) is an effort to provide reusable, scalable scientific software components. The Mantevo Project (2013 R&D 100 winner) is focused on the development of open source, portable mini-applications and mini-drivers for the co-design of future supercomputers and applications. HPCG is an official TOP 500 benchmark for ranking computer systems, complementing LINPACK.

Mike's most recent interests are focused on improving scientific software developer productivity and software sustainability. He leads the IDEAS project, dedicated to engaging scientific software teams to identify and adopt practices that improve productivity and sustainability.

Title: ***MUST system applied to high level language approach in MYX project***

Presenter: Taisuke Boku (Center for Computational Sciences, University of Tsukuba, Japan)

Abstract: In MYX Project as a Germany-France-Japan tri-country international project, we applied MUST MPI library and run-time for our original high level language named XcalableMP and workflow framework named YML to use the system for fault aware programming on distributed memory highly parallel computing system. In this talk, I will show how we applied it to these language framework and how it works and the result of MYX project in high level language environment.

Title: ***A distributed task scheduler in the DASH project***

Presenter: Joseph Schuchart (HLRS Stuttgart, Germany)

Abstract: This talk will provide an update on the efforts to integrate task-based parallelization into the DASH PGAS abstraction. In order to achieve maximum concurrency while retaining a one-sided programming model, DASH employs the notion of data dependencies similar to OpenMP but in a global context, with dependencies spanning across process boundaries while requiring only the discovery of the process-local task-graph. The talk will provide an overview of this new programming model and cover recent developments within the distributed scheduler as well as updated performance results.

Title: ***Overview of Smart-DASH (results and perspectives)***

Presenter: Jose Gracia (HLRS, Germany)

Abstract: This talk provides an overview on DASH (www.dash-project.org), a C++ template library that offers distributed data structures and parallel algorithms and provides a PGAS (partitioned global address space) abstraction. DASH data structures are modeled after the familiar C++ STL container classes but can utilize the memory of multiple compute nodes. DASH allows convenient element-wise global data access using one-sided remote memory operations and also supports a strong notion of data locality that makes it easy to realize the owner-computes idiom. DASH also offers parallel variants of many STL algorithms that harness the computing power of multiple compute nodes. This talk will also provide an update of some of the efforts in the SmartDASH project, including support for halo exchange in N-dimensional arrays and an update on the DASH algorithms.

Title: Numerical Library with High-Performance/Adaptive-Precision/High-Reliability: Extension of ppOpen-HPC towards the Post Moore Era

Presenter: Kengo Nakajima (University of Tokyo & RIKEN R-CCS)

Abstract: Power consumption is one of the most critical issues in supercomputing towards the Post Moore Era. Although computations with lower precision, such as approximate computing, is one of the reasonable choices for reducing power consumption, such computations may provide results with less accuracy. Therefore, accuracy verification is very important for computations with lower precision. In the present work, we are focusing on preconditioned iterative solvers with sparse coefficient matrices derived from real applications, and are developing a practical method for accuracy verification in adaptive/trans-precision computations based on iterative refinement. Method for automatic selection of the optimum precision, which minimizes computation time and power consumption under certain target accuracy, will be developed, and it is implemented to "ppOpen-HPC (https://github.com/Post-Peta-Crest/ppOpenHPC)", which is an open source infrastructure for development and execution of optimized and reliable simulation code on post-peta-scale (pp) parallel computers based on many-core architectures, and it consists of various types of libraries, which cover general procedures for scientific computation. "ppOpen-HPC" has been supported by Japanese Government (FY.2011-2018). In this talk, we will overview the strategy for the new numerical library with High-Performance/Adaptive-Precision/High-Reliability, and show preliminary results.

Title: POP: A Centre of Excellence in HPC (Results from phase I and service offerings in phase II)

Presenter: Christian Terboven, (RWTH Aachen University, Germany)

Abstract: Developers and users of HPC applications can count on free advice from experts to analyse the performance of their scientific codes within the EU Performance Optimization and Productivity (POP) Centre of Excellence. The objective of POP is to provide performance measurement and analysis services to the industrial and academic HPC community, help them to better understand the performance behaviour of their codes and suggest improvements to increase their efficiency. This short talk will inform about opportunities to get help in performance optimization by applying for services offered in the POP project and will present key findings and success stories from performance assessments already performed in the first 18 months of operation.

Title: Scalable and Easy-to-use Coupling of Multi-Physics Simulations

Presenter: Benjamin Uekermann (TUM, Germany)

Abstract: The SPPEXA project ExaFSA studies fluid-structure-acoustic interaction as a representative, yet challenging example of a surface-coupled multi-physics problem. The deeper research question is how to make a multi-physics coupling approach at the same time scalable and easy-to-use. Scalable is this context means to not deteriorate the scalability of the coupled single-physics solvers by the coupling itself. Easy-to-use means to allow for minimally-invasive integration into existing legacy codes. This talk discusses how both concepts are combined in the coupling library preCICE.