Review of the Platform for Advanced Scientific Computing Programme (PASC)

Lugano, 3rd and 4th December 2015

1 Introduction

This report presents the outcome of the review of the PASC program undertaken the 3rd and 4th of December 2015 in Lugano by a panel of four international experts:

- Bruce Harmon, Ames Laboratory and Iowa State University.
- Thomas Lippert, Jülich Supercomputing Centre.
- Jeffrey Nichols, Oak Ridge National Laboratory (chair).
- William Tang, Princeton Plasma Physics Laboratory and Princeton University.

They were supported by the program director Prof. Thomas Schulthess and by the program manager Dr. Benedetto Lepori.

The experts received detailed self-assessment reports from the PASC funded initiatives. During the review, general presentations on the overall program strategy, two of the codesign networks (climate and materials) and of the PASC conference series were held. This was followed by parallel sessions were the initiatives' PIs' had the opportunity to present their project and to discuss critically with the panel. The review was concluded by a plenary debriefing session. The review program and the list of PASC funded initiatives are annexed to this report.

The report is organized as follows. Section 2 introduces the program goals and the review objectives, section 3 presents an overall assessment, while section 4 formulates some recommendations for the next program phase. Finally, section 5 includes more in-depth assessments of the individual initiatives.

2 Goals

Advanced computer technology is rapidly driving advances in discovery, design, and development. From cosmology to medicine to the microscopic interactions on modern computer chips, the ability to model, analyze, and predict fundamental interactions leads to understanding, control, and development of future technologies. Therefore, advanced computing has become a strategic imperative.

The PASC (http://www.pasc-ch.org/) overarching goal is to enhance and empower Swiss computational sciences in the emerging high performance computing, by networking Swiss universities with the CSCS. It is complementary to the supercomputing-hardware-focused elements of the Swiss High-Performance and Networking (HPCN) initiative. The PASC consolidates and builds on the achievements of the current High-Performance and High-Productivity Computing (HP2C) project, which supported 13 large-scale projects in the period 2009-2013.

PASC aims to promote joint effort to address key scientific issues in different domain sciences through interdisciplinary collaborations among domain scientists, computational scientists, software developers, computing centres and hardware developers. Thus, PASC builds on the principle of co-design, namely that software codes exploiting the potential of the next generation of computing architectures need to be jointly and interactively developed by these participants throughout the whole enterprise.

An equally important aim of the PASC program is to set up sustaining structured programs that will thrive beyond 2016, when the current structuring program ends. This is accomplished by organizing the PASC program into three pillars:

1. An application support network pillar consisting of experts at participating universities that assist researchers in adopting HPC platforms.
2. A co-design pillar consisting of projects in which interdisciplinary teams are created with members from various application domains that have similar algorithmic and architectural needs – with respect to computer science, applied mathematics and computer engineering – in order to develop efficient (in terms of time and energy/cost to solution) simulations strategies.

a) An institutional computing systems pillar to organize, and where necessary help fund the provisioning of institutional computing resources that complement the supercomputing systems of the CSCS user lab.

The 2015 PASC review panel was asked to:

- assess the current state of development of the PASC program comparing it with similar initiatives at the international level, both concerning individual application domains and computing infrastructure, and
- provide directions for the new phase of development and support of the computing infrastructure in Switzerland and identify the needs for development and networking activities in this context.

The review panel will give an overall impression of the program elements, a short assessment of the individual projects, and finally some advice/recommendations.

3 Overall Impression of the PASC Program

The panel's impression is that the PASC program has contributed in a significant way to the highly productive Swiss computing and computational science programs. For example, the number of publications per year from the whole CSCS User Lab has now well exceeded over 250 with over 10% published in high impact journals.

The hardware procured and deployed at CSCS is first class – it is the most powerful supercomputer in Europe and the 7th ranked system on the top 500 list of supercomputers. It is extremely challenging to recursively deploy powerful supercomputers that are utilized efficiently by a broad range of computational scientists during a time when technology is rapidly changing *and* while the conduct of science is also rapidly changing – being informed by multiple streams; experimental data, data analytics and modeling and simulation.

The PASC program has created 5 domain science "networks"; Materials, Climate, Life Sciences, Physics and Solid Earth Dynamics. Several of these networks are quite powerful. In particular, the Materials, Climate and Life Sciences are functioning at a very high level. A few examples that demonstrate the success of these networks are:

- Successful development and deployment of a Swiss National Center on Computational Design and Discovery of Novel Materials; **12-year timespan (2014-2026); 11 Swiss Institutions** (two federal and five cantonal universities, the Swiss Supercomputing Center, IBM, two experimental Labs). **First phase (2014-2018): 33 PIs, 34.4M CHF** (18M SNSF, 6.6M EPFL, 9.8M others)
- Frontier climate research has been enabled with impact well beyond individual projects with over 4 research groups involved at ETH, an HPC group at MeteoSwiss, strong presence in the C2SM and COSMO consortiums, and strong collaborations with CS and CSE vendors
- Joint project of MeteoSwiss and CSCS for replacement of an operational weather forecasting system resulted in the first weather service to move to GPUs and allowed for an increase in the problem size with 40 times at the same operational costs
- Development and deployment of new state of the art tools, solvers, and applications using GPUs at unprecedented scales (Gordon Bell finalists) resulting in high-throughput simulations of microfluidics at cell resolution.

Applications that are scalable with respect to system architecture are facilitated by employing scalable algorithms and solvers along with scalable performance tools and system software environments (e.g., compilers). The PASC program includes 20 co-design projects that address these needs. Nearly all the projects have reached a mature level where serious results were presented. Even a few unforeseen issues in some projects had been analyzed and ameliorated. Nevertheless, there remains significant room for improvement with respect to associated risk mitigation challenges in planning [see Recommendation 6]. The wide range of projects was impressive and indicative of high performance computing pushing science and engineering into the future. However, there remain significant challenges in improving scalability, portability, and connections to future architectures at extreme scales [see Recommendations 1,3, and 4]. In addition, specific metrics are needed to better describe the level of success of the PASC projects [see Recommendation 7 and 8].

Co-design is facilitated within the various domain science "networks" but also reinforced by a strong PASC conference series. This 3-day annual meeting serves as a community conference platform for Computational Science and HPC in Switzerland (and Europe). It is the only conference in Europe which is supported by ACM SIGHPC. It has grown to over 350 participants from academia, industry and CSCS with broad representation from around the world. The last conference was comprised of 5 plenary sessions; 1 public lecture; 126 mini-symposium presentations, 21 contributed talks, 77 posters, an information event dedicated to users of the Swiss National Supercomputing Centre (CSCS), an inter-PASC Networks discussion, a poster session, and exhibition booths. It is extremely important for the global scientific high performance computing community to exchange ideas and share results. To some extent, it helps the global community to be more efficient; but it also spreads ideas and generates insights among talented people who can advance science, engineering, and algorithms … to solve urgent global problems. Of course, solutions to those problems help generate demand for more advanced computing.

However, the concept of co-design with respect to the PASC program should be better defined and associated areas of improvement recommended (see Recommendation 2).

4 Recommendations

While the panel was (impressed with all facets of the PASC program, there are a few places where recommendations can lead to more effective outcomes both in the current program and in future similar initiatives.

a) General recommendations on the overall program strategy

1) Whereas much is being done in developing new tools and libraries, it may be prudent to put additional emphasis on capability (more scalability) or high-throughput modeling and simulation systems.
2) Efforts should be made to close the co-design loop. Lots of co-design efforts are unidirectional, for example taking information from the vendors to incorporate into PASC development activities. More effort should be made in creating mini-apps or proxy-apps representative of the CSCS workload and getting these in front of the vendor community in order to facilitate co-design developments on their end. In addition, an accompanying explanation is needed as to how such a "proxy/mini app co-design code" can actually enable delivery of discovery science -- either by itself or in contributing to significant improvements on its associated more comprehensive "flagship" production code. This should also include a description of how success in terms of significant new discovery science capability would be measured (e.g., metrics including associated improvements in time to solution and energy to solution). Specifically, co-design proxy/mini-app approaches need to accountably produce impactful integration of new software into associated “flagship” production codes.
3) The number of potential accelerated system architectures in the international HPC complex should drive more efforts in software “portability”. These efforts should include identification of modern programming approaches (OpenMP4, OpenACC, etc.) that would best help enable code portability.
4) More consideration and thought could be expended toward looking longer into the future and anticipating the kind of hardware architectures and related software that might be anticipated. This will require maintaining high level discussions with vendors and the with leading software experts.

b) Specific recommendations concerning program activities

5) While lots of intra-network collaboration is evident there doesn’t seem to be very much inter-network activity. Some of the algorithm and tool developments could be leveraged across the various scientific domains.
6) There needs to be more risk mitigation for key personnel. The review panel saw a few efforts where loss of a key individual brought a particular project to a halt or had to be re-directed.
7) In a few instances it was unclear what success looked like. Each project ought to have very clear and tangible metrics for success. These should be imbedded within a clearly-articulated project roadmap.
8) The way science is being conducted is rapidly changing – being informed by multiple streams; experimental data, data analytics and modeling and simulation. More emphasis on workflows is warranted (here are some signs of this – in particular work like AiiDA found in the Materials network).

c) Specific recommendations for the next program phase

9) A stimulus is needed to spark some of the networks – the materials and climate networks are working great. Life Sciences is productive but it is not clear that it's productivity is greater than the collection of individual projects; not a lot of synergy/leverage. There is nothing evident in the Physics or Solid Earth Dynamics networks.
10) Additional focus on workflows to address science complexity issues (multiple streams of experimental data, data analytics and modeling and simulation)
11) It would be advisable to formalize the relationship of the "networks" with CSCS. The CSCS technology roadmap is informing the applications, tools, libraries, system software, etc being developed in the networks. But, the network's user requirements need to drive the future technology investments of the center. This could be formally accomplished by turning the networks into joint institutes or using an end-station approach (giving block CPU allocations to the networks to administer).

4.1 Closing Remarks

The review arrangements were well handled and greatly appreciated. The panel was highly impressed by program director Thomas Schulthess for his insights, leadership skills and determination. The panel would like to especially commend the excellent "vision" that he has provided for the future of the PASC that includes a realistic picture of exciting opportunities with associated major challenges. Benedetto Lepori was superb at logistics and graciousness, and helped keep the review moving in the right direction with a positive attitude. The panel also appreciated the interaction with two university presidents (USI and University of Zurich), which are members of the program Steering Committee; their presence at the review demonstrated the high level of institutional commitment to the field in Switzerland.

Finally, the panel would like to commend Professor and President Piero Martinoli for his appreciation, serious involvement, and promotion of the PASC enterprise. Leadership commitment and enthusiasm are essential for the success of such an ambitious adventure into a promising but risky frontier. We would like to express our sincere admiration for his nurturing of the supercomputing enterprise in Lugano and in the whole of Switzerland.

PASC Review 2015. Program

Date: 3rd and 4th December 2015

Place: Lugano, Università della Svizzera italiana.

Goal of the review

- To assess the current state of development of the PASC program comparing it with similar initiatives at the international level, both concerning individual application domains and computing infrastructure
- Provide directions for the new phase of development and support of the computing infrastructure in Switzerland and identify the needs for development and networking activities in this context.

Members of the review board

- Bruce Harmon, Ames Laboratory and Iowa State University
- Thomas Lippert, Jülich Supercomputing Centre
- Jeffrey Nichols, Oak Ridge National Laboratory (chair)
- William Tang, Princeton Plasma Physics Laboratory

Program

2nd of December

19.30 Dinner between program direction and review board

3rd of December

10.00 Welcome coffee and Registration Desk **(Aula magna)**

10.30 AM – 12.30 AM Opening session (Plenary) **(Aula magna)**

Introduction and objectives of the review (Piero Martinoli, president of the Università della Svizzera italiana and of the PASC steering committee)

Presentations:

- General presentation of the program structure (Thomas Schulthess, program director)
- PASC conference series (Olaf Schenk, responsible of the PASC conferences)

Network in computational sciences: some examples of the broader embedding of PASC:

- Material sciences (Nicola Marzari, EPFL)
- Climate (Oliver Fuhrer and Christoph Schär, ETH Zurich).

Introduction of questions and issues from the review panel (Jeff Nichols, chair of the review panel).

General discussion

13.00 – 14.00 Standing Lunch / poster session

14.00 – 16.00 Parallel tracks

Track A: Geophysics (Auditorium)

Review Board members: Lippert and Tang.

Fichtner	HPC Application Support for the Solid Earth Dynamics Community
Fichtner	Geoscale: A framework for multiscale seismic modelling and inversion
Tackley	GeoPC: Infrastructure development for hybrid parallel smoothers for multigrid preconditioners
Chopard	Optimal deployment of multiscale applications on a HPC infrastructure

Track B. Material sciences (Aula magna)

Review Board members: Harmon and Nichols.

Hutter	Application-support specialists for materials simulations
Goedecker	ENVIRON: A Library for Complex Electrostatic Environments in Electronic-structure Simulations
Hutter	Reduced scaling electronic structure calculations
Luisier	ANSWERS: Accelerating nano-device simulations with extreme-scale algorithms and software co-integration

16.00 – 16.30 Coffee break / poster session

16.30 – 18.00 Parallel tracks

Track A: Physics (Auditorium)

Review Board members: Harmon and Tang.

Mayer	DIAPHANE: A common platform for application-independent Radiative Transport in astrophysical simulations
Villard	Particles and Fields

Track B: generic tools (Aula magna)

Review Board members: Lippert and Nichols.

Fuhrer	Grid Tools: Towards a library for hardware oblivious implementation of stencil based codes
Hoefler	A Heterogeneous Compiler Platform for Advanced Scientific Codes
Goedecker	A highly efficient implementation of direct GPU to GPU communication in a library for direct exchange calculations

19.30 – 22.00 Dinner at Canvetto Luganese

Morning 4th December 2015

9.00 – 10.30 Parallel tracks

Track A: life sciences (Auditorium)

Review Board members: Harmon and Nichols.

Koumoutsakos	Angiogenesis in Health and Disease: In-vivo and in-silico
Xenarios	Portable Scalable Concurrency for Genomic Data Processing
Krause	HPC Framework for Coupled Cardiac Simulations
Obrist	AV-Flow
Caflish	Scalable advanced sampling in molecular dynamics using standalone tools for data mining (CHARMMing PIGS)

Track B: social sciences (Aula magna)

Review Board members: Lippert and Tang.

Zollikofer	High-performance computing tools for agent-based general ecosystems models (HPC-ABGEM)
Horenko	Towards the HPC-inference of causality networks from multiscale economical data
Lomi	Snowball sampling and conditional estimation for exponential random graph models for large networks in high performance computing
Kuebler	Tackling large dynamic stochastic equilibrium models with occasionally binding constraints

10.30 – 11.00 Coffee break / poster session **(Aula magna)**

11.00 – 12.00 Plenary session: feedback of the review panel **(Aula magna)**

12.00 – 13.00 Standing Lunch (open) **(Aula magna)**

12.00 – 13.30 Lunch between SC and PRB, debriefing of the review