Procedia Computer Science

Volume 51, 2015, Pages 1685–1693

ICCS 2015 International Conference On Computational Science

Computational Science Research Methods for Science Education at PG Level

Nia Alexandrov[1] and Vassil Alexandrov[2*]
[1]*Barcelona Supercomputing Centre, Barcelona, Spain*
[2]*ICREA-Barcelona Supercomputing Centre, Barcelona, Spain*
nia.alexandrov@bsc.es, vassil.alexandrov@bsc.es

Abstract
The role of Computational Science research methods teaching to science students at PG level is to enhance their research profile developing their abilities to investigate complex problems, analyze the resulting data and use adequately HPC environments and tools for computation and visualization. The paper analyses the current state and proposes a program that encompasses mathematical modelling, data science, advanced algorithms development, parallel programming and visualization tools. It also gives examples of specific scientific domains with explicitly taught and embedded Computational Science subjects.

Keywords: Computational Science Research Methods, Postgraduate Education, Science Subjects

1 The Need

Current expectations from university education are to prepare Postgraduate (PG) level students for careers both in academia and industry. This became now even more relevant in the sciences where the needs of knowledge based economy with increasing number of smart jobs require science graduates to be "research ready" for a dynamically changing and computationally heavy project environment using High Performance Computing machines, environments and tools. The professional training courses currently available to computational scientists, or those established just few years ago, were targeting mainly computer science graduates, and although they are able to provide some practical skills which can resolve short term programming deficiencies or introduce some simulation techniques, they are not designed to substitute university education.

*Distinguished Visiting Professor, ITESM-Monterrey Tech, Monterrey, Mexico

Selection and peer-review under responsibility of the Scientific Programme Committee of ICCS 2015

doi:10.1016/j.procs.2015.05.305

Large scale computing in science and industry has become an indispensable way to tackle societal and scientific Grand Challenges, and to address the needs of industry to innovate in terms of products and services (COM EC, 2012) and (Joseph, E. et al, 2010, September). The major societal drivers such as Energy, Climate Change, Urbanization, etc. (Luebkeman 2009), require advanced scientific methods and in particular Computational Science and HPC (High Performance Computing) are the key methods for helping address these challenges. The platforms allowing us to advance on such scale and magnitude are the Exascale Computing systems. Recently, Exascale Computing, an attempt to harness a thousand fold projected increase in computational power, has emerged as a Grand Challenge research area (Dongarra, J. et al, 2009). In addition, the development of novel mathematical methods and scalable algorithms for exascale in itself has been identified as a Grand Challenge (Dongarra et al, 2014).

Computational approaches to scientific Grand Challenge problems such as the detection and treatment of diseases like cancer, modelling of the human brain, and climate forecasting are beginning to bear fruit. Computational Science, an interdisciplinary field that melds basic sciences, mathematical modelling, quantitative analysis techniques, algorithms, parallel programming and tools and HPC techniques, is proving integral in addressing the big problems in industries ranging from manufacturing and aerospace, to drug design and risk management.

Recent US reports (Simon et al, 2007) and (NECCEA, 2011) focused on Grand Challenges in Climate Change and Environmental Modelling, Energy, Biology, Astrophysics and Socioeconomics outlined the key Grand Challenges in each of these areas. For example, in Climate Change and Environmental Modelling focus on improving "our understanding of the complex biogeochemical cycles underpinning the global ecosystem functions and control sustainability of life on Earth" (Simon et al, 2007) and (NECCEA, 2011). The urgent questions to be answered are: characterizing and bounding the coupled Earth system, are extreme weather events leading to significant shifts in regional climates, tropical rain forest sustainability, stability of polar caps and Arctic ice sheets, sustainability of sea life and agricultural ecosystems, link between socioeconomic and climate responses (Simon et al 2007). This in turn requires accurate models of coupled physical and biochemical systems, integration of models and observations of the carbon cycle, process level modelling of biochemical cycle (Simon et al 2007). All rely on developing of new mathematical models with high resolution able to assimilate variety of data including ones from satellites, detailed modeling of variety of ecosystems, process scale models for biochemical, ecological and aerosol processes. To address these challenges it is critical to have an increased simulation capacity and advanced mathematical methods and algorithms as well as new software tools and rapid development environments to enable tackling all these problems (Simon et al 2007) and (NECCEA, 2011).

In the Energy area, the focus is on combustion, nuclear fusion, clean energy which includes solar energy, etc. (Simon et al 2007) and (NECCEA, 2011). Al these require high-fidelity simulation of combustion phenomena, efficient particle and fluid simulation including fine-scale turbulence, for solar energy developing of models to model complex systems (specifically nano-systems with tens of thousands of atoms dynamically for long time (Simon et al 2007).

Biotechnology, bioinformatics and breakthroughs in healthcare technology are also a key focus areas both in Europe (Horizon 2020), USA (NECCEA, 2011) and Latin America (Alexandrov et al, 2014). The focus here, for example, is on advanced methods and technologies for preventive medicine as well as creating complete DNA sequencing for major diseases, (see, for example, Horizon 2020 and NECCEA, 2011).

In Socioeconomic modelling the focus is on integrated modeling of the social, economic and environmental systems coupling these elements and treating uncertainties and non-linearities (Simon et al 2007). For example, the challenges here are: how will different adaptation and mitigation strategies affect energy supply and demand, the overall economy, the environment , the public, public

health etc.; how the demographic change, economic growth etc is connected to the above etc .(Simon et al 2007).

2 Current State

The Recent SIAM (Society for Industrial and Applied Mathematics) working group report (SIAM Working Group on CSE Education, 2014) and the USA DoE Assessment of Workforce Development Needs in office of Science Research Disciplines (Chapman et al, 2014), as well as our own assessment have shown that despite increasing acceptance of Computational Science as interdisciplinary science there is still a lot of silos based approaches to university education and the compartmentalised culture towards curricula development and course design are preventing us from really truly bearing the fruit of implementing Computational Science research methods into variety of science degrees at time when we are at a critical junction when the research and industry are beginning to experience a stifling effect of lack of properly educated and trained specialists in Computational Science (Chapman et al, 2014) and (SIAM Working Group on CSE Education, 2014). We are at a crossroad where we need to decide what the way forward would be: poaching the few available ones from each other if we can, or re-design science degrees at PG level where the cycle is shorter and more flexible to update and thus provide the right calibre graduates.

Firstly, there are simply too few Master's and PhD programs in Computational Science and related areas (Scientific Computing, High Performance Computing, and Supercomputing). As noted previously (Joseph et al, 2010, July) and (Joseph, E. et al, 2010, September), the number of such programs (MScs) has been actually dropping. In EU, for example, there are established MSc programs in Computational Science and related areas in ETH (Switzerland), University of Amsterdam (Netherlands), KTH (Sweden). The remaining programs outside the EU are principally in the USA and Canada, in institutions such as the George Mason University, Georgia Tech, University of California San Diego, Stanford University, Universities of Illinois and Rice, University of Toronto and the University of Waterloo. Outside the EU and North America, there are several scattered programs in Singapore, South Korea and Australia.

Secondly, many of the existing Computational Science programs (sometime despite what their promotional literature says) tend to be highly focused on one part of the "Computational Science pipeline". They produce highly skilled personnel in mathematical modelling or HPC computing techniques but they don't expose them to the pipeline from basic science, through modelling and simulation, HPC application design, implementation and evaluation – all in combination with data-intensive computing. While technical depth is important, we believe the leadership in interdisciplinary areas like Computational Science, requires a breadth of experience and knowledge.

Thirdly, existing programs often expose students to a narrow range of application domains. For example, they may have great depth in matrix based and/or fluid-flow problems but do not expose students to applications based on string or geometric data (bioinformatics) or large scale graphs (web analytics). This narrowness not only limits an individual's employment opportunities, it also impedes cross-fertilization between Computational Science application domains. This is also linked with the silos culture and compartmentalization dominant in many academic environments.

Fourthly, most existing programs in Computational Science don not have the opportunity to expose students to the practice of science in a multi-country/multi-cultural setting. This is a wasted opportunity. The practice of science, especially in the context of Grand Challenge problems, is becoming steadily more global. These problems are simply too big for single countries to tackle without global collaborations. Even industrial R&D teams now typically span multiple sites in multiple countries and time zones. We believe that learning to work effectively outside one's own cultural home is an essential skill in being a productive team member or leader of large scale R&D efforts.

Finally with the emergence of Data Science focused on dealing with Big Data an additional dimension emerges that requires universities and research institutions to establish a common Data Science competencies profile and a common modularized (component-based) curriculum for education and training to the required job profiles (Demchenko et al, 2014), (High Level Expert Group on Scientific Data, 2010) , (Big Data Analytics, 2014) and (Manyika, J. et. al, 2011).

Similar is the situation at doctoral level. The problem here is particularly acute since in many cases the PhD students are usually selected from the corresponding MScs and since these are clearly much underrepresented, here is even increasing shortage. The latest results of the IDC surveys (Joseph et al, 2010, July) and (Joseph, E. et al, 2010, September), as well as RISC project survey results (Alexandrov et al, 2014), show that we observe the same gaps in developed economies such as EU, USA and in BRICS including Latin America.

Analyzing overall situation it appears that the major problem still remaining is compartmentalization, Chapman (July 2014). The current system of such programs produces number of graduates in USA, for example, (Chapman et al, 2014) and (SIAM Working Group on CSE Education, 2014), which still remain very low as a number to satisfy the needs. Similar is the situation in Europe.

The major drawbacks of such programs have been identified, at least in USA, that they do not provide exposure to real-world applications, e.g. students are not able to grasp the complexities in the field (Chapman et al, 2014). Also the US NSF taskforce in 2011 (NSF-ACCI Report, 2011) has concluded that in general the universities are not adequately preparing students with the right skills to become tomorrow's computational scientists and engineers, the study also stated that "the current programs do not teach students the skills essential to apply Computational Science and Engineering in modern scientific and technological enterprises." It was also pointed out that there was "almost no university that have or are likely to develop curriculum focused on topics associated with petascale and exascale science."

Additionally, with the advent of Big Data we observe the emergence of Data Science (Chapman et al, 2014), (High Level Expert Group on Scientific Data, 2010) and (EC Report, 2012) and the need for relevant skills for data-intensive science and also Big Data needs of industry (EC Report, 2012), (Big Data White Paper, 2013), (Hey et al, 2009) and (ABSE, 2014). The applications from the exemplar areas above in addition of being compute intensive are predominantly data intensive as well. Recent USA DoE report in 2013 has identified as a result also the challenges in Data Intensive Science and Exascale Computing (DOE ASCAC Data Subcommittee, 2013).

It is time to accept that a way forward could be an introduction of Computational Science into the core subject area of PG science degrees, not just computer science and/or mathematics graduates but as well those in life, earth and natural sciences are inevitably going to use HPC methods, tools and environments in their working life and a part of their education should prepare them to at least understand and implement, if not design, parallel algorithms and be familiar with the HPC machines and HPC tools and environments.

3 Computational Science Research Methods

The graduates require set of skills to tackle both compute and data intensive applications independently if they will follow a career in academia or industry. In particular, the Grand Challenges outlined above require knowledge and expertise in mathematical modelling (multi-model, multi-scale, modeling continuous processes (PDEs, ODEs, etc), modelling discrete events), discretization techniques, advanced algorithms development (parallel algorithms, scalability, numerical and non-numerical algorithms), software implementations on variety of advanced architectures, program execution, tools for analysis, visualization etc, data analysis, visualization and validation of the results.

Data Scientists or data science teams on the other hand focus on solving complex data problems by employing deep expertise in one or more of these disciplines, as well as business strategy and domain knowledge. Personal skills in communication, presentation and inquisitiveness are also very important. In terms of skills a Data Scientist is seen as "a practitioner who has sufficient knowledge in the overlapping regimes of expertise in business needs, domain knowledge, analytical skills, and programming and systems engineering expertise to manage the end-to-end scientific method process through each stage in the big data lifecycle" (Manyika, J. et. al, 2011), (EC Report, 2012), (Big Data White Paper, 2013), (Hey et al, 2009) and (ABSE , 2014).

From point of view of computational scientists working in Supercomputing area, the body of knowledge the students need to acquire to be able to tackle the above Grand Challenges and to be able to successfully tackle project work is expected to cover:

- Mathematical modelling and algorithms: requiring systematic approach to modelling (complex systems - introduction, overview, etc.), modelling continues events/systems (Ordinary Differential Equations (ODEs), Partial Differential Equations (PDEs) and systems of PDEs, multi-level, multi-scale methods) and discrete events modelling, numerical analysis (Linear Algebra, Optimization, etc.), discretization, stochastic numerical methods and stochastic modelling, advanced algorithms (numerical and non-numerical including parallel algorithms and parallelization techniques);
- Programming Environments and tools: programming languages, including parallel programming approaches, advanced programing models and tools, variety of performance tools for parallel programs/computing including for HPC;
- Data Analysis: stochastic and optimization methods for data analysis, data visualization techniques (scientific visualization, where appropriate Virtual Reality approaches, etc.).

To be able to bridge the skills gap there is a need to change the current education programs at University PG level as well as the professional training. It requires paradigm shift in terms of introducing advanced mathematical modelling methods and scalable algorithms (stochastic, deterministic and hybrid ones) allowing change of thinking in terms of high levels of parallelism for exascale and beyond; advanced programming models and tools allowing advanced modelling and simulation (including simulation and modelling at scale); scientific visualization of data and Big Data processing.

The challenges are to integrate these key components in the curricula as separate subjects as well as embed components in domain specific subjects for maximum impact.

4 Integration at Curricula Level

The standard taught MSc program in EU is usually 120 ECTS (European Credit Transfer System) which comprises four semesters consisting usually of three main blocks:

- one to two semesters of core subject modules with a few or in some cases no electives,
- at least one semester of specialization where the elective subjects are clustered,
- between half and one semester consisting of a final project and dissertation.

It is common practice to include some research skills modules, e.g. presentation, writing, subject specific project management. The available credits for taught modules are about 90 ECTS. Depending on the structure and the subject area, that usually means teaching between nine and twelve separate modules with electives rarely being more than two to three modules (between 1/3 and ¼ of the modules).

In EU there are several established Computational Science MSc degrees with the possibility to specialise in a certain scientific discipline such as ETH Zurich ”Diploma in Computational Science and Engineering”, KTH Sweden “International Master Program in Scientific Computing”, University of Amsterdam “MSc in Computational Science”. On the other hand the US experience shows two currently predominant models of Computational Science programs at PG level: one delivered at computer science or mathematics departments and another in scientific disciplines departments with the participation of computer science and engineering where the computational science subjects are given as a specialisation (Chapman et. al, 2014). The issue both in EU and US, seems to be the low numbers of graduates which are by no means adequate to answer the need of required specialists.

If the skills gap is to be bridged, the way forward should be introducing Computational Science research skills as a part of the core subject block on the majority of the MSc science degrees. Our proposal is to embed several taught modules delivering the core theoretical knowledge and introduce at the dissertation module project specific practical skills training.

The taught modules could be in a block of five or as a “lighter version” in a block of three. Significant Computational Science element, a combined degree in its essence, with five modules (at least 25 ECTS) would include: **introduction to mathematical modelling** (introduction to PDEs and ODEs, discretization approaches, linear algebra, introduction to parallel algorithms); **advanced mathematical modelling** (complex systems and systems approach to modelling, multi-level, multi-scale methods, elements of optimization, advanced linear algebra, advanced parallel algorithms and parallelization techniques); **introduction to programing environments and tools** (introduction to parallel programing, MPI and OpenMP, introduction to performance tools needed for parallel codes debugging and optimization); **advanced programming environments and tools** (programing models for petascale and concepts of exascale, high performance tools, experimenting with approaches to scale up exemplar application); **data analysis** (stochastic and optimization methods for data analysis, scientific visualization and elements of VR).

A more introductory level approach, building understanding rather than professional capacity, would include three modules (15 ECTS): **introduction to mathematical modelling** (introduction to PDEs and ODEs, discretization approaches, linear algebra, introduction to parallel algorithms); **introduction to programing environments and tools** (introduction to parallel programing, MPI and OpenMP, introduction to performance tools needed for parallel codes debugging and optimization); **data analysis** (stochastic and optimization methods for data analysis, scientific visualization and elements of VR).

In order to ensure that students gain set of practical skills that can be applied in their further studies or work a set of project specific training segments can be introduced during the dissertation stage. This should not be lecture heavy and will have the aim of providing the practical skills needed to tackle a problem. The whole block should not exceed 5 ECTS and could consist of 1 or 2-3 day segments training in the use of particular HPC tools or programming environments, the use of visualisation suits or data analytics techniques facilitating the progress of an MSc level dissertation in the said science area with sufficient Computational Science element. These practical skills segments constitute in essence professional training which is currently provided by the Supercomputing centres. In Europe an example of such training is the developed and delivered in the past three years training program by PRACE – The Partnership for Advanced Computing in Europe (PRACE, 2015). The six PRACE Advanced Training Centres (PATC) had collaborated to produce professional training curricula which aims at delivering practical skills training to computer scientists, domain scientists and HPC administrators (PATC OMB, 2013). Some of the PATC such as BSC and EPCC, already have started collaborating with local universities on their MSc degrees. In particular, in Barcelona, FIB (Faculty of Informatics) is crediting BSC PATC courses on their MIRI MSc program.

Example of above proposed approach is given below:

- Scientific Domain: Environmental Sciences
  - Mathematical modelling: Complex Systems; Systems approach to modelling; Multi-level multi-scale methods; PDEs and ODEs; Discretization;
  - Programming Environments and tools: Parallel programming; MPI and OpenMP; Performance Analysis;
  - Data Analysis: Stochastic and Optimisation Methods for data analysis; Visualisation;
  - Embedded in the Dissertation Project Module: Introduction to Earth Science Simulation Environments, Performance tools (BSC ETT, 2015).

- Scientific Domain: Life Sciences
  - Mathematical modelling: Complex Systems Modelling and Simulation; Stochastic Methods; Parallel Numerical and Non-numerical Algorithms;
  - Programming Environments: Parallel Programming, MPI, OpenMP, CUDA;
  - Data Analysis: 3D Visualisation of Data, Bid Data Analysis;
  - Embedded in the Dissertation Project Module: Simulation Environments for Life Sciences; Parallel and Distributed Programing Models, Virtual/ Augmented Reality.

- Scientific Domain: Material Science
  - Mathematical modelling: Complex Systems Modelling and Simulations; Stochastic and Deterministic Methods; Parallel Algorithms;
  - Programming Environments: Parallel Programming, OpenMP;
  - Data Analysis: 3D Visualisation of Data, Stochastic and Optimisation Methods of Data Analysis;
  - Embedded in the Dissertation Project Module: Parallel and Distributed Programming Models, Virtual/Augmented Reality.

5 Conclusion

Computational Science research skills gap is evident from variety of reports and studies, for example, (Chapman et al, 2014) and (SIAM Working Group on CSE Education, 2014). IDC has also identified the skills which are most difficult to find to tackle the inflection points, for example: Scientists with HPC capabilities ("Combined scientific background and HPC programming skills", "Computational scientists"), Parallel Programmers ("Experience in parallel software development", "Engineers and scientists that can program in HPC/parallel Fortran" , "Parallel code porting/optimization"), Algorithm Developers ("For computational science people who can help researchers develop and implement new algorithms"), System Administrators with high-end computing experience "Scientific computing system management experience", "System administrators with HPC expertise"), see (Joseph et al, 2010, July) and (Joseph, E. et al, 2010, September).

The graduates require set of skills to tackle both compute and data intensive applications independently if they will follow a career in academia or industry. In particular, the scientific Grand Challenges outlined, require knowledge and expertise in mathematical modelling, advanced algorithms development, software implementation on variety of advanced architectures, program execution, data analysis, visualization and validation of the results.

Apart from the Computational Science degrees that are taught in Computer Science and Mathematics departments, the graduates of MSc courses taught at domain science departments also

must be addressed. It is necessary to increase the general understanding in Computational Science research methods. This can be achieved by introducing explicitly taught modules on mathematical modelling, HPC methods, tools and programming as well as provide opportunities to train graduates in practical skills as part of their dissertation projects.

References

ABSE (2014, August), Advisory Board to the Secretary of Energy, (Report of the Task Force on High Performance Computing, Department of Energy U.S. August 18, 2014). [Online] Available from: http://energy.gov/sites/prod/files/2014/10/f18/SEAB%20HPC%20Task%20Force%20Final %20Report.pdf [Accessed: February 2015].

Alexandrov, V. et al (2014), RISC Roadmap, [Online] Available from: http://www.risc-project.eu/publication_type/green-papers/ [Accessed: February 2015].

BSC ETT, BSC Education and training team. (2015) PATC training timetable 2014-15 [Online] Available from: http://www.bsc.es/education [Accessed: February 2015].

Chapman, B. et. al, (July 2014), *DOE: Assessment of Workforce Development Needs in office of Science Research Disciplines*, [Online] Available from: http://science.energy.gov/ ~/media/ascr/ascac/pdf/charges/ASCAC_Workforce_Letter_Report.pdf [Accessed: February 2015].

Big Data White Paper (2013) Unleashing the potential of Big Data. A Whitepaper "World Summit on Big Data and Organization Design, 2013 [Online] http://www.essi.upc.edu/~aabello/publications/ 13.Whitepaper.pdf [Accessed: February 2015].

Big Data Analytics (2014), Assessment of demand for Labour and Skills 2013-2020, [Online] https://www.thetechpartnership.com/globalassets/pdfs/bigdata_report_nov14.pdf [Accessed: February 2015].

COM EC (2012). Communication from the Commission to the European Parliament, The Council, The European Economic and Social Committee and the Committee of the Regions. *High-Performance Computing: Europe's place in a Global Race*. European Commission. Brussels, 15.2.2012 COM(2012) 45 final.

Demchenko, Yuri, Emanuel Gruengard, Sander Klous (2014), *Instructional Model for Building effective Big Data Curricula for Online and Campus Education.* 1st IEEE STC CC and RDA Workshop on Curricula and Teaching Methods in Cloud Computing, Big Data, and Data Science, in Proc. The 6th IEEE International Conference and Workshops on Cloud Computing Technology and Science (CloudCom2014), 15-18 December 2014, Singapore.

DOE ASCAC Data Subcommittee (2013, March) Report, *"Synergistic Challenges in Data-Intensive Science and Exascale Computing,* [Online] Available from: http://science.energy.gov/ ~/media/40749FD92B58438594256267425C4AD1.ashx [Accessed: February 2015].

Dongarra, J. et al, (2009) *The international exascale computing project*, [Online] Available from: http://hpc.sagepub.com/content/23/4/309.short [Accessed: February 2015].

Dongarra et al, (2014) *Applied Mathematics Research for Exascale Computing, US – DOE Report, March 2014*. [Online] Available from: http://www.netlib.org/utk/people/JackDongarra/ PAPERS/doe-exascale-math-report.pdf [Accessed: February 2015].

EC Report, (2012, May), Skills and Human Resources for e-Infrastructures within Horizon 2020, The Report on the Consultation Workshop, May 2012 [Online] http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/report_human_skills.pdf [Accessed: February 2015].

High Level Expert Group on Scientific Data, (2010, October). Riding the wave: How Europe can gain from the rising tide of scientific data. [Online]. Available from: http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf [Accessed: February 2015].

Horizon 2020, The EU Framework Programme for Research and Innovation, [Online] Available from: http://ec.europa.eu/programmes/horizon2020/ [Accessed: February 2015].

Joseph, E., Conway, S., and Wu, J. (2010, July) *A Study of the Talent And Skill Set Issues Impacting HPC Data Centers Conducted On Behalf Of The US Department Of Energy*. IDC.

Joseph, E. et al. (2010, September), *A Strategic Agenda for European Leadership in Supercomputing: HPC2020 – IDC Final Report of the HPC Study for the DG Information Society of the European Commission*. IDC. Hey,T., Tansley, S. and Tolle, K. (2009). The Fourth Paradigm: Data-Intensive Scientific Discovery. Microsoft Corporation, October 2009. ISBN 978-0-9825442-0-4 [Online]. Available: http://research.microsoft.com/en-us/collaboration/fourthparadigm/ [Accessed: February 2015].

Luebkeman, C. (2009) and the Foresight team at ARUP (Jennifer Greitschus eds.), *Drivers of Change,* Prestel Verlag and Arup, 2009, ISBN 978-3-7913-4224-5.

Manyika, J. et. al (2011) Big data: The next frontier for innovation, competition and productivity, Report, McKinsey Global Institute, [Online] www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation [Accessed: February 2015].

NECCEA, (2011), National Economic Council, Council of Economic Advisers, and Office of Science and Technology Policy, *A Strategy for American Innovation: securing our economic growth and prosperity, February 2011*. [Online] Available from: http://www.whitehouse.gov/sites/default/files/uploads/InnovationStrategy.pdf [Accessed: February 2015].

NSF-ACCI Report, (2011, March) NSF Advisory Committee for Cyberinfrastructure Task Force on Grand Challenges, Final Report, NSF, [Online] https://www.nsf.gov/cise/aci/taskforces/TaskForceReport_Software.pdf. [Accessed: February 2015].

PATC OMB (2013), *PRACE Advanced Training Centres HPC Curriculum 2013*-2014, PRACE IP3, April 2014.

PRACE, *Partnership for Advanced Computing in Europe website* [Online] Available from: http://www.prace-ri.eu [Accessed: February 2015].

SIAM Working Group on CSE Education (2014),*SIAM: Graduate Education for Computational Science and Engineering* [Online] Available from: http://www.siam.org/students/resources/report.php, [Accessed: February 2015].

Simon, H., Zcharia, T. and Stevens R. (2007) *Modelling and Simulation at the Exascale for Energy and the Environment, US Office of Science.* [Online] Available from: http://science.energy.gov/~/media/ascr/pdf/program-documents/docs/Townhall.pdf. [Accessed: February 2015].