

Grid'5000 Science Advisory Board Report 2018

The Grid'5000 Science Advisory Board (SAB) met on April 4th at the INRIA Sophia Antipolis. In attendance were Henri Bal (Vrije University, the Netherlands), André Brinkmann (Johannes Gutenberg-Universität Mainz, Germany), Vincent Breton (CNRS-IN2P3, France), Marco Danelutto (University of Pisa), Wu Feng (Virginia Tech, USA), Kate Keahey (Argonne National Laboratory, USA), Pierre Kuonen (HES-SO, Switzerland), and Manish Parashar (Rutgers University, USA). The advisory board was chaired by Kate Keahey.

The board meeting opened with presentations from two Grid'5000 scientific projects, the first on energy efficiency by Laurent Lefèvre and the second on a holistic framework to facilitate experimentation with OpenStack by Adrien Lebre. These scientific talks were then followed by a pair of general presentations about the Grid'5000 project and its future: (1) a status report on the health of the Grid'5000 project as well as a summary response to the Grid'5000 SAB recommendations by Lucas Nussbaum and (2) a proposed evolution of the FIT and Grid'5000 projects into a "Super Infrastructure for Large-Scale Experimental Computer Science" (SILECS) project by Frédéric Desprez. The presentations were followed by a closed discussion of the SAB, followed by a question-and-answer session with members of the Grid'5000 project. The SAB was asked to address two general issues: (1) a general evaluation of the Grid'5000 project and (2) advice on moving forward with the SILECS project. This report summarizes the discussion and recommendation on both issues.

Evaluation of the Grid'5000 project. As in prior meetings, the SAB was very favorably impressed by the continued achievements of Grid'5000 in building an experimental platform. The numbers of active users, usage numbers, and the recent publications all demonstrate that the project continues to make a significant impact and provides an essential platform for active research. Apart from the quantitative assessment, the qualitative synergy between an experimental platform and the user community that it fostered has been outstanding and unprecedented.

Relevant to this assessment, the SAB made the following recommendations:

- *Development of a proactive vision.* While the project excels in its response to the needs of its user base -- in other words developing a reactive vision of capabilities of an experimental testbed -- its ability to persuade its stakeholders of the importance of further investment in the instrument would be significantly strengthened if it could present a proactive vision of capabilities of the testbed. What are the emerging (rather than existing) research challenges in areas related to the mission of the testbed? What is the research agenda for the next 10 years? What kind of experimental platform is needed to serve that research agenda? Is this need fully served by the existing platforms (such as for example, commercial clouds) and, if not, how should it be supplemented? What important types of experiments – and

thus research – will be impossible to carry out without such a platform? The SAB recognizes that developing a proactive vision for an experimental platform, predicated as it is on developing a vision for the research itself first, is a challenging task. Thus, the SAB suggests that the project seeks the active participation and backing of the excellent user community it developed.

- *Dissemination of achievements.* The challenges of supporting experimental research for Computer Science are still not well – and certainly not broadly – understood. Grid’5000 is uniquely positioned to provide insights into operational models for Computer Science testbeds. A thoughtful analysis of information from the testbed usage logs, user surveys, projects, availability, and other qualitative and quantitative operational indicators would be important to not only make the achievements of the project better known and recognized, but would also form a valuable contribution to the general discussion on operational models for supercomputing. Venues such as the SC “State of the Practice” track or the LISA conference can provide good platforms for the dissemination of these insights.
- *Workforce.* A longstanding problem in the project is the low allocation of effort devoted to Grid’5000 for the permanent engineering staff: engineers working on short-term contracts create a situation where high turnover is the norm, causing inefficiencies due to significant ramp-up times required to obtain specialized knowledge and putting an unsustainable – and unfair – amount of pressure on the time of the few permanent staff members. In the long-term, this situation is detrimental to the project and poses a significant risk. The SAB recognizes that this situation is caused by factors that are difficult to change but encourages the project to work with the local institutions to allocate permanent staff to the project so that the “turnover effect” is limited. An ideal solution would be provided by recognition of the project’s value and “institutionalization” of its operation so that it obtains a permanent base support similar to the building of a datacenter, where a long-term commitment is made to the building and its operations.

A Path Forward: the SILECS project. The SAB was extremely excited about the new opportunities implied in SILECS: the ability to create a platform serving a broad range of emergent Internet of Things (IoT) research will ensure that a proven and exceptionally successful experimental research platform will evolve as the research frontier evolves. The evolution of such a platform is important to ensure that new ideas are capable of being worked on and validated – and entrusts this important task to the team that has the experience and proven track record of serving experimental needs.

To make this evolution smoother, we recommend the following:

- While it is important that support for new experimental models is developed, it is essential that support for the existing research is not neglected. The existing instrument has had an enormous impact on the feasibility of many branches of research as evidenced by the continuing interest of the research community and the numerous publications that resulted from it. This

support of the traditional Grid'5000 mission of research on HPC/cloud/datacenter should be continued and expanded to support research topics in data science, machine learning, and disaggregated hardware, to name just a few. In other words, while SILECS is an exciting direction for Grid'5000, care needs to be taken that the set of experiments currently supported by the platform continues to be supported and developed along the trajectory dictated by its own priorities.

- Research related to the emergent IoT opportunity is in relatively early stages; it is not always clear where the new opportunities and priorities lie. The SAB recommends that SILECS develops a well-defined set of application use cases and partners with application projects to better understand and define research challenges that need to be addressed, and then use those challenges to drive the design of the experimental platform. An example of such a use case is the Clermont-Ferrand environmental observatory [1] that combines relevant environmental indicators with infrastructure for data collection, storage, modeling, analysis and simulation. This combined platform is used to evaluate the role that environmental factors can play on the yield and quality of agricultural products – as well as the impact of these agricultural systems on the evolution and the quality of the environment. Combining wireless systems with datacenter management, this use case thus drives challenges of precisely the kind that will require a combined Grid'5000 and FIT capabilities. While this “use cases -> research agenda -> experimental platform” design approach puts together domain scientists with Computer Science researchers, it will be invaluable in developing a vision articulating the need for the SILECS platform and persuading funding institutions that a platform for Computer Science research is aligned with national and local priorities.

[1] Clermont-Ferrand environmental observatory:

<http://cap2025.fr/recherche/challenges-scientifiques/les-agroecosystemes-durables-dans-un-contexte-de-changement-global/challenge-1-agro-ecosystemes-durables-dans-un-contexte-de-changement-global-32027.kjsp?RH=3365383655036558>