

The LinkSCEEM FP7 Infrastructure Project:

Linking Scientific Computing in Europe and the Eastern Mediterranean

Constantia Alexandrou
Cyprus Institute

Linking Scientific Computing in Europe and the Eastern Mediterranean (LinkSCEEM) was an FP7 infrastructure project that established a high-

performance computing (HPC) ecosystem in the Eastern Mediterranean region by interlinking and coordinating regional compute resources to form an integrated e-Infrastructure. The overarching objective of LinkSCEEM was to enable scientific research in the region by engaging and supporting research communities, with an initial emphasis in the fields of climate research, digital cultural heritage, and synchrotron radiation applications.

The FP7 infrastructure project Linking Scientific Computing in Europe and the Eastern Mediterranean (LinkSCEEM) established a high-performance computing (HPC) ecosystem in the Eastern Mediterranean region by interlinking and coordinating regional compute resources to form an integrated e-Infrastructure. It was the follow-up project to a coordination and support action that preceded LinkSCEEM. The coordinating institution was the Computation-based Science and Technology Research Center (CaSToRC) of the Cyprus Institute (CyI), with partners from Cyprus (CYNET), Egypt (BA and NARSS), France (ESRF), Germany (FZJ/JSC and MPG), Israel (IUC), Jordan (SESAME and JUNET), and the US (NCSA).

LinkSCEEM's overarching objective was to enable scientific research in the region by engaging and supporting research communities, with an initial emphasis in the fields of climate research, digital cultural heritage, and synchrotron radiation applications. To achieve its mission, LinkSCEEM adopted the following three-fold approach:

- Optimally integrate computational resources contributed by HPC centers in the Eastern Mediterranean region while establishing links with leading HPC centers and thus developing and sharing best practices for managing these resources.
- Create user support and training programs, in parallel to an active networking process that engages and integrates research communities and enables scientists in the region to

utilize HPC for addressing complex scientific and engineering problems, thereby diminishing the digital divide.

- Develop a subset of HPC applications of particular relevance to the region—namely climate science, cultural heritage, and synchrotron radiation applications—and provide the links to leading groups in these fields to enable research of the highest standing. Thanks to the participation of world-leading HPC centers and research teams, core expertise of the highest level was introduced to the region both in the service provided and the spectrum of research enabled.

In summary, the main objectives of this project were to interlink and coordinate regional computational resources to form an integrated e-Infrastructure, as well as to provide the associated training activities and user support, engaging regional communities through networking activities such as workshops, exchange of visitors, organization of joint events, and outreach. This contributed to the creation of an HPC ecosystem from the individual group cluster to the Tera-flop scale machines provided by CaSToRC and Bibliotheca Alexandrina (BA), and promoted the coordination of practices and methodologies.

The specific objectives of the project were as follows:

- *Objective 1:* Promote the establishment of user communities in computational science in the Eastern Mediterranean.
- *Objective 2:* Establish a resource allocation mechanism to coordinate access to the integrated simulation platform and its usage for research.
- *Objective 3:* Provide adequate training programs.
- *Objective 4:* Develop a dissemination and outreach program to publicize the project to the HPC and computational science communities in Europe and the Eastern Mediterranean. The program included the organization of an international conference, a continuously updated website, and focused outreach programs for computational scientists.
- *Objective 5:* Provide adequate user support.
- *Objective 6:* Adapt and develop software to support the optimization of parallel applications, data management, and visualization tools: (a) provide advanced user support via crossdisciplinary research activities, (b) port data-management and scientific workflow codes to the regional computational resources and optimize to the regional connectivity bandwidth limitations, and (c) implement and deploy visualization software for developing collaborative virtual spaces.
- *Objective 7:* Optimize scalability of existing climate codes and develop new climate codes for the infrastructure provided by the project.
- *Objective 8:* Exploit the benefits of visualization in cultural heritage research: (a) prototype a small, portable imaging center at CaSToRC, and (b) produce software tools for tele-immersive collaborative environments.
- *Objective 9:* Develop an online cultural heritage digital library for the Eastern Mediterranean.
- *Objective 10:* Port existing common and new algorithms at synchrotrons to GPUs.

Some of LinkSCEEM's accomplishments are briefly outlined below.

COMPUTATIONAL ECOSYSTEM FOR THE EASTERN MEDITERRANEAN

LinkSCEEM successfully developed an operational model for an integrated HPC e-Infrastructure. Two regional HPC facilities have contributed resources and services in this project. The Cyprus Institute (CyI) through CaSToRC contributed Cy-Tera, an HPC facility of 35 Tflops peak performance, with an associated research and educational program devoted to computational science.

The BA installed a Sun cluster of peak performance of 11.8 Tflop/s, and also maintained an archival system. BA had storage systems of 3.7 Pb, of which 1.6 Pb were previously unused. The BA was engaged in joint activities and partnerships in visualization projects with universities in

Egypt, the region, and internationally. The consortium subsequently administered this e-Infrastructure and various access calls for regional computational scientists.

Access to the systems was given through frequent calls for proposals. Users could apply for three types of access:

- *Production*. A large number of computational resources were granted over a 12-month access period for scientific computing.
- *Preparatory*. A small amount of resources were allocated over a 6-month period for code development and testing.
- *Educational*. Resources were allocated for educational purposes (for example, to lecturers and students of university courses).

The access calls for preparatory and educational access were continuously open. Production access calls were issued twice yearly.

Figure 1 identifies the scientific field and principal investigator home country for the computational projects carried out on the infrastructure.

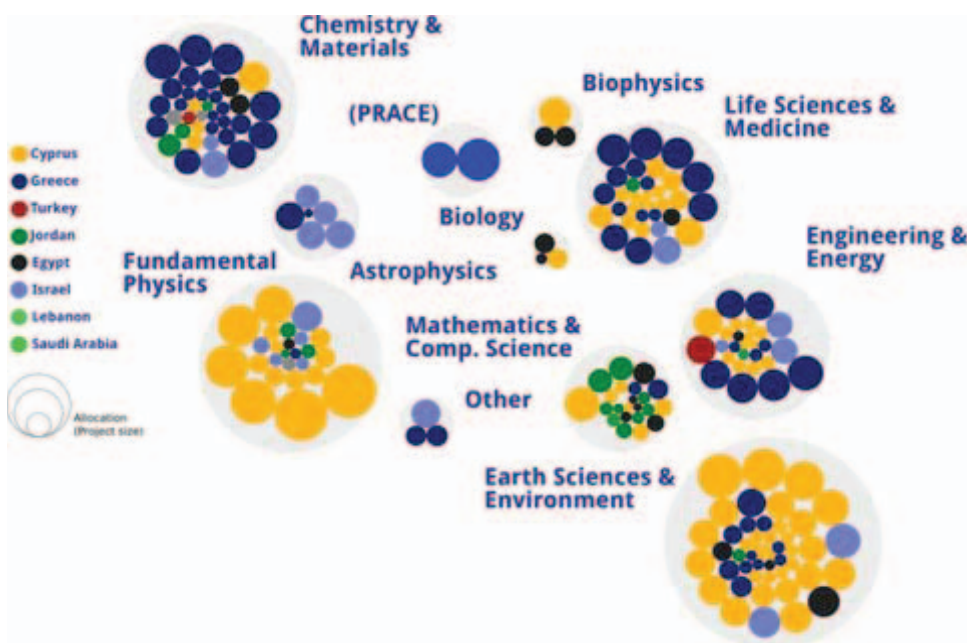


Figure 1. Scientific fields and countries represented in the computational projects carried out as part of LinkSCEEM's integrated HPC e-Infrastructure.

All applications underwent a rigorous and transparent review process, which ensured that the expensive computational resources were put to good use. Preparatory and educational access requests only underwent a brief technical review to ensure that the proposed applications were suitable for the available systems. These requests were usually processed within a few days. Production access applications underwent a more elaborate process as the allocated computer time was significantly larger. These applications underwent a peer-review process of international standard, which ensured the fair and transparent allocation of resources to users based on scientific excellence. The process included a technical review and a scientific assessment by international experts. It is a significant success of LinkSCEEM to have been consistently able to identify and motivate international experts to scientifically review proposals submitted by regional scientists. A total of 87 production projects and 183 preparatory access projects were accommodated on the supercomputers, resulting in 76 scientific publications from 53 completed projects and more than 80 presentations in conferences and scientific meetings.

One of the success stories of LinkSCEEM has been the availability of a training cluster to the Eastern Mediterranean region. This cluster, called Euclid, was hosted at the CyI and was a hybrid CPU/GPU Linux cluster composed of six 8-core computer nodes with two NVIDIA Tesla T10 processors attached to each node. Academic institutions were given access for teaching purposes. Euclid was also used for training events and user meetings, where trainee user accounts could be set up for hands-on training. Individual users who were starting with parallel programming and required a cluster for their educational and testing needs were also given access.

REGIONAL TRAINING PROGRAM

LinkSCEEM implemented an ambitious training workshop program. The program consisted of an annual series of five large 3-day workshops that were held in alternating locations across the Eastern Mediterranean. The program included one large cross-sectional workshop (usually in winter) and a General User meeting (usually in the summer) on general HPC-related topics such as parallel programming, performance analysis, and HPC applications. This was supplemented with three workshops in the project thematic areas (climate modeling, cultural heritage, and synchrotron radiation). The thematic workshops focused on community-specific lectures regarding methods and applications. In total, the project organized 20 workshops as part of this training program.

Complementing the intensive multiday workshop program was a series of relatively brief basic training sessions cast into the format of an HPC roadshow. The aim of this event series was to promote the available HPC facilities and provide basic training in how to access and use the supercomputers. In addition, very detailed presentations were given regarding available training events in general and online training materials in particular. All participants were offered access to Euclid, the LinkSCEEM training system and connected during the event, when possible. This was a very effective way of tempting new users into a supercomputing environment. The format of the roadshow was designed to maximize impact and outreach while minimizing organizational overhead. All roadshow events followed a predefined program that was also made available in video and audio. This meant that local trainers could be relatively easily recruited and many events were organized locally without having to pay for relatively expensive regional travel. The format of the roadshow proved very successful. More than 800 people participated in 37 single events in the framework of 3 integrated roadshows.

Figure 2 shows the number of LinkSCEEM events carried out throughout the region.



Figure 2. LinkSCEEM events.

One major highlight of the LinkSCEEM project was the organization of the Conference on Scientific Computing (CSC 2013), which took place in Cyprus in December 2013. Leading international researchers in computational science gathered to present highlights of their scientific work performed using Partnership for Advanced Computing in Europe (PRACE) and LinkSCEEM computational resources (see Figure 3).

CSC 2013 brought together 120 international computational scientists, amongst them the directors of the largest HPC facilities in Europe and the US. The European research community, local academics, and scientists from the Eastern Mediterranean region and US had the opportunity to exchange ideas, develop synergies, and set out a vision for future development on various aspects of HPC.



Figure 3. Professor Bothina A. Hamad from the University of Jordan at the 2013 Conference on Scientific Computing. She was assisted through LinkSCEEM to install the first cluster enabling research in computational solid state, projects, and access to Cy-Tera.

LinkSCEEM also had other dissemination activities. In principle, three dissemination channels were used to actively promote the project. The first channel was the events organized by the project, the second was the participation in regional and international conferences and meetings, and the third was the project website. In addition, the project was also acknowledged in the publications of users. In total, LinkSCEEM was presented in 109 dissemination activities.

MIDDLE EASTERN WOMEN IN SCIENCE PROGRAM

LinkSCEEM also ran a Middle Eastern Women in Science program, which supported women in various ways as they engaged with computational science. As an example, a European contribution through the LinkSCEEM project provided funding for fellow Hadeer EL-Habshy (see Figure 4) to attend a workshop program at the Jülich Supercomputing Centre in Germany.

RESEARCH ACTIVITIES

LinkSCEEM's crossdisciplinary activities are aimed at research on simulation techniques across scientific fields, and at providing sophisticated support in the optimization of algorithms used in HPC applications. The crossdisciplinary research team worked in close coordination with the user support group and also significantly contributed to training activities. Their approach toward performance analysis and optimization was based on existing efforts and experiences

within the Virtual Institute for High Productivity Supercomputing (VI-HPS). A number of tools were installed on all systems to target specific aspects of performance, such as single-node performance, parallel performance, debugging, instrumentation, measurement, and visualization. Training events showcasing some of these tools were provided throughout the project.

Climate Modeling Research

LinkSCEEM's climate modeling research successfully brought expertise to and subsequently enabled research at the forefront of climate studies of the Eastern Mediterranean region. The parallel scalability of the atmospheric general circulation model that is used to predict regional climate change was improved as part of the research done in LinkSCEEM.

Cultural Heritage Research

The objectives of LinkSCEEM's cultural heritage research were twofold. The first task was to establish a novel small-object imaging center at CaSToRC that would make use of multiple high-resolution imaging techniques for the documentation of a wide range of ancient artefacts. The center was designed to focus in particular on two innovative approaches to imaging: reflectance transformation imaging and 360-degree imaging of cylindrical objects, especially cylinder seals. A second goal was to collaborate with the InscriptiFact Digital Image Library to facilitate the worldwide distribution of the center's images. This objective was broadened throughout the project duration as it became possible to extend InscriptiFact capabilities by developing a tailor-made content management system that was tuned for regional artifacts by the use of the open source MEDICI software. InscriptiFact is designed to host a very limited number of digital formats, mainly linked to inscriptions. MEDICI however, allows hosting of a wide range of digital formats, including 3D models.

In addition, a tele-immersive visualization system that would create a portable, virtual collaborative space where scholars could jointly examine artefacts in a 3D environment was developed and tested across the region.



Figure 4. Fellow Hadeer EL-Habashy (middle row, far right): "When I got accepted to the Guest Student Program on Scientific Computing, Jülich Supercomputing Centre (JSC), Germany, I realized that it marked a turning point in my career." EL-Habshy received support to attend JSC as part of LinkSCEEM's Middle Eastern Women in Science program.

Synchrotron Radiation

Synchrotron radiation facilities such as the Synchrotron-light for Experimental Science and Applications in the Middle East (SESAME) and the European Synchrotron Radiation Facility (ESRF) enable scientific research in a broad variety of fields: atomic and molecular physics, biology, material science, spectroscopy, archaeology, and imaging. LinkSCEEM contributed significantly to the efforts to support SESAME in developing the ecosystem around the emerging synchrotron, which became operational in 2016. The research conducted under LinkSCEEM sought to

- port existing synchrotron applications to GPUs,
- develop new synchrotron applications for GPUs, and
- support training programs for future SESAME users.

The research was conducted by experienced researchers at the ESRF jointly with staff at SESAME. Many exchanges and visits ensured that the generated knowledge was also transferred to SESAME so that development could continue on the ground beyond the duration of LinkSCEEM. The research on synchrotron applications mainly focused on the use of accelerators, in particular graphic processing units (GPU), in synchrotron data processing within LinkSCEEM.

Beyond the provision of computing resources and training to computational scientists, LinkSCEEM also developed data-management systems for regional use and provided performance analysis tools and expertise for application support.

The project conducted research in three thematic areas of high regional relevance. The consortium contributed to developments by porting and accelerating HPC applications in the fields of climate modeling and synchrotron data analysis. Furthermore, an advanced imaging facility for cultural heritage applications was installed and operated by the project. Examples of the science enabled by LinkSCEEM are shown in Figure 5.

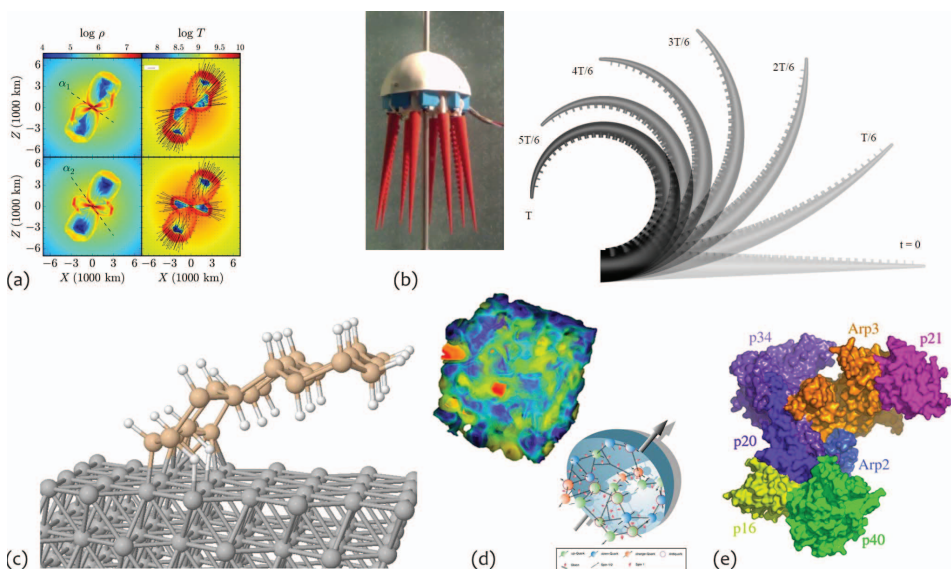


Figure 5. Examples of the science enabled by LinkSCEEM. (a) Modeling of supernovae explosions. (b) Modeling of aquatic swimming robot tentacles for the development of bioinspired robotic devices. (c) Modeling of layered materials (such as graphene). (d) Probing new physics beyond the standard model. (e) Modelling proteins to enhance the understanding of cancer development.

CONCLUSION

LinkSCEEM has had a strong impact on the scientific communities of the wider region. Significant computational resources were available across the region through a transparent excellence-based allocation process for the first time. This introduced regional scientists to international best practice and increased the competitiveness on a global level. Thanks to LinkSCEEM, a number of scientists were able to scale up their research work to fully use the opportunities of computational science similar to their counterparts in central Europe and the US. Furthermore, by being familiar with the access mechanisms introduced by LinkSCEEM, regional scientists will find it easier to compete for access to very large systems, as provided through PRACE. The impact of these scientists spreading their know-how to colleagues and local institutions should be emphasized, because this is an important instrument in establishing an HPC tradition.

The provision of computational resources to computational scientists has had a direct, profound scientific impact—this is apparent by the number and quality of the publications resulting from the scientific work done on the supercomputers available through LinkSCEEM.

Given that, at the end of LinkSCEEM, it was apparent that the regional scientific community required access to supercomputing systems to carry on, The CyI's and BA's provision of computational resources has continued. As of September 2017, two and a half years after LinkSCEEM ended, regional preparatory applications were still being received and applications were being accepted for the 12th Call for Production.

ABOUT THE AUTHOR

Constantia Alexandrou is an Institute Professor at the Cyprus Institute, and a professor of physics at the University of Cyprus, where she leads the Lattice Quantum Chromodynamics Computational Lab. Alexandrou received a PhD in theoretical nuclear physics from MIT. Contact her at alexandrou@cyi.ac.cy.