



**SEVENTH FRAMEWORK PROGRAMME
Research Infrastructure**

**FP7-INFRASTRUCTURES-2010-2 – INFRA-2010-1.2.3:
Virtual Research Communities**

**Combination of Collaborative Project and Coordination and Support
Actions (CPCSA)**



**LinkSCEEM-2
Linking Scientific Computing in Europe and the Eastern
Mediterranean – Phase 2**

Grant Agreement Number: RI-261600

**D5.7
Report on prospects for the sustainability of the LinkSCEEM-2 e-
Infrastructure**

Final

Version: 1.0
Author(s): Jens Wiegand, CaSToRC
Date: 18/03/2015
LinkSCEEM-2 - RI-261600

Project and Deliverable Information Sheet

LinkSCEEM Project	Project Ref. №: RI-261600	
	Project Title: LinkSCEEM-2	
	Project Web Site: http://www.linksceem.eu/ls2/	
	Deliverable ID: < D5.7 >	
	Deliverable Nature: <DOC_TYPE: Report>	
	Deliverable Level: PU *	Contractual Date of Delivery: 28 / 02 / 2015
		Actual Date of Delivery: 14 / 04 / 2015
EC Project Officer: Leonardo Flores Anover		

* - The dissemination level are indicated as follows: **PU** – Public, **PP** – Restricted to other participants (including the Commission Services), **RE** – Restricted to a group specified by the consortium (including the Commission Services). **CO** – Confidential, only for members of the consortium (including the Commission Services).

Document Control Sheet

Document	Title: < Report on prospects for the sustainability of the LinkSCEEM-2 e-Infrastructure >	
	ID: <D5.7>	
	Version: <1.0 >	Status: Final
	Available at: http://www.eniac.cyi.ac	
	Software Tool: Microsoft Word 2007	
	File(s): LinkSCEEM-2-D5.7 Report on prospects for the sustainability of the LinkSCEEM-2 e-Infrastructure.docx	
Authorship	Written by:	Jens Wiegand, CaSToRC
	Contributors:	WP Leaders and Participants
	Approved by:	Constantia Alexandrou, CaSToRC

Document Status Sheet

Version	Date	Status	Comments
0.1	18/02/2015	Draft	first version
1.0	18/03/2015	Final version	
1.0	14/04/2015	Approved	

Document Keywords

Keywords:	LinkSCEEM-2, Computational Science, HPC, e-Infrastructure, Eastern Mediterranean
------------------	--

© 2010 LinkSCEEM-2 Consortium Partners. All rights reserved.

Table of Contents

Project and Deliverable Information Sheet	ii
Document Control Sheet.....	ii
Document Status Sheet	iii
Document Keywords	iii
Table of Contents	iv
1 Introduction.....	5
2 Outlook on regional HPC resources	5
2.1 CaSToRC- Cy-Tera	5
2.2 BA.....	6
2.3 SESAME.....	6
2.4 Other LinkSCEEM relevant regional HPC centers.....	7
3 Outlook on regional cooperation	8
4 Outlook on access to resources.....	10
5 Outlook on training efforts	11
6 Outlook on private sector involvement	12
7 Conclusions.....	12

1 Introduction

A primary purpose, and indeed a major success, of the LinkSCEEM-2 project was the creation of the LinkSCEEM e-Infrastructure. It has always been the goal that the LinkSCEEM-2 project creates a sustainable structure that enables regional scientists to access HPC resources, thereby growing computational science in the region beyond the project duration. With this in mind, the project consortium began implementing a sustainability strategy during the third year of LinkSCEEM-2. This strategy is based on the uninterrupted provision of HPC resources to regional scientists to support the shift to scientific computing on parallel systems, the creation of a user support team, the provision of appropriate training and educational material, accessible beyond the project duration, and the development of regional cooperation to help stimulate government funding for hardware and attract further competitive research funding.

This report gives an overview of the HPC situation in the Eastern Mediterranean as the LinkSCEEM-2 project comes to an end and outlines the actions that have been taken or are planned to ensure the sustainability of the activities of LinkSCEEM that have been initiated. Information on High Performance Computing (HPC) resources, regional cooperation and planned training programs are given. The report concludes with a summary of the actions connected with the sustainability of LinkSCEEM.

2 Outlook on regional HPC resources

2.1 CaSToRC- Cy-Tera

CaSToRC started a national computational facility with the installation of the Cy-Tera computer, its main computational resource, along with a number of smaller prototype machines based on innovative computer technologies. The Cy-Tera machine was funded through the Cy-Tera project, whose consortium consists of The University of Cyprus (UCY), the Juelich Supercomputing Center (JSC), the National Center for Supercomputing Applications (NCSA) at the University of Illinois at Urbana Champaign and the Synchrotron Light Source Facility SESAME. The aim of the facility is to provide all Cypriot scientists and the researchers at its partner institutions in particular as well as the wider region in general, an internationally competitive infrastructure for large computation. An integral component of CaSToRC's strategy was the establishment of a user support team and the development of training and educational programs for scientists and HPC administrators in the usage of supercomputers. Such a structure on the national scale is crucial for supporting research and innovation. LinkSCEEM provided the resources that helped establish the support, training and education necessary for the full exploitation of Cy-Tera.

Thus the Cy-Tera and LinkSCEEM projects have led to the creation of a National Computational Center with an infrastructure of regional significance, which is a landmark not only for CaSToRC and the Cyprus Institute, but also for the whole of Cyprus.

The Cy-Tera supercomputer, provides a peak performance of 35 TFlops to national and regional researchers. A total of 90% of Cy-Tera's capacity is offered for national use (60%) and regional use (30%). Cy-Tera also participates in the DECI program of PRACE by contributing 10% of its capacity to European users. The Cyprus Institute has already pledged to continue offering 30% of its capacity to regional users beyond the duration of the

D5.7 Report on prospects for the sustainability of the LinkSCEEM-2 e-Infrastructure

LinkSCEEM-2 project. The associated e-Infrastructure that was created through the LinkSCEEM and LinkSCEEM-2 projects will continue to operate beyond the end of LinkSCEEM-2.

CaSToRC computational facility hosts a number of small scale experimental prototypes (for test-driving emerging hardware architectures) and also Euclid, the training facility offered for educational purposes. The knowhow that activities associated with innovative computer technologies bring will be central in pursuing funding opportunities but also with designing the next production systems contributing to the sustainability. Euclid was utilised intensively for educational and training purposes throughout LinkSCEEM-2. Numerous University lecturers use it for their courses and LinkSCEEM-2 training events relied on Euclid for practical hands-on sessions. Free and open access to Euclid will also continue beyond the LinkSCEEM-2 project.

Cy-Tera was inaugurated in January 2012 and will be operational for at least another 2 years. CaSToRC is actively pursuing upgrade options of the system. The Institute is committed to enabling the upgrade of Cy-Tera during the next two years ensuring the sustainability of National Computing facility. Options of funding include structural funds as was the case for the initial establishment of Cy-Tera and utilizing core funds of the institute in a lease agreement with vendors.

2.2 BA

The Bibliotheca Alexandrina installed its High-Performance Computing (HPC) cluster in 2009 with the aim of providing researchers nationally and regionally with the necessary tools for conducting computational research. Upon inauguration of the cluster, the BA announced that the primary purpose of the cluster will be to serve researchers affiliated with public Egyptian universities and research institutions. Indeed, over the ensuing years, the BA has hosted computational research projects in diverse fields for researchers at various public research institutions in Egypt, including the Faculty of Science and the Faculty of Engineering at Alexandria University, the Faculty of Computers and Informatics and the Faculty of Engineering at Cairo University, and the Egyptian Nanotechnology Research Center. As LinkSCEEM-2 project partner, the BA has also hosted computational research projects affiliated with research institutions in the larger Eastern Mediterranean region. Beyond LinkSCEEM-2, the BA plans to continue to provide access to researchers in Egypt and possibly also to researchers in the region through future joint efforts. The BA is already participating in a joint call for proposals with the Cyprus Institute ("7th CfP" jointly organized by the CyI and the BA).

BA's plans for the HPC cluster also include providing computational resources for researchers working in a variety of fields such as, the Regional Action on Climate Change (RACC), an initiative by the BA for setting up a forum for researchers working in the field of climate change to collaborate with one another.

The BA HPC cluster is capable of a theoretical peak performance of approximately 11.8 Tflops, and it provides 1 GB of memory per each of its 1040 CPU cores. The cluster currently lacks GPU accelerators. Taking note of the memory constrain that limited the number of projects BA was able to accommodate during LinkSCEEM-2 Calls for Proposals, it was decided to implement an upgrade of its cluster.

D5.7 Report on prospects for the sustainability of the LinkSCEEM-2 e-Infrastructure

2.3 SESAME

Jordan has initiated the establishment of a national supercomputing facility during the LinkSCEEM-2 project. The IMAN1 supercomputer (<http://www.iman1.jo/>) was built using 2260 PS3 devices with IBM Cell Processors connected together on a very fast fibre based network; basically, turning a video gaming console into a supercomputer powerhouse that is capable of performing up to 25 trillion mathematical operations per second (i.e. 25 TFlops). The main use of this initial system was planned to be for educational and scientific applications.

The IMAN1 project is now operated and funded by the Jordan Atomic Energy Commission (JAEC) and SESAME, supporting their very crucial missions and fulfilling their High-end computing needs. While it is still available to all the universities and scientific institutions in Jordan, an upgrade of the now aging hardware is being planned to specifically cater for the needs that arise from SESAME. Building on the existing infrastructure, new hardware will be gradually purchased to meet the computational demands from future SESAME users and the Jordanian scientific user community. The planned upgrade will include the following hardware:

- 45 Two Intel based Xeon Proc E5-2698 each 16 x 8 128GB TruDDR4 RAM Memory
- 10 Two Intel Xeon Proc E5-2698 + 10 NVIDIA Tesla K20, each 16 x 8 128GB TruDDR4 RAM Memory
- 9 Two Intel based Xeon Proc E5-2698 + 9 Intel Xeon Phi 5110P
- InfiniBand interconnect
- 32x 10 TB operational storage

LinkSCEEM-2 was actively involved in providing advice on running and expanding the system, but also through the provision of training to the scientific community in Jordan. One example for these activities is the first IMAN1-LinkSCEEM User's Meeting that was held on the 12th of March 2014 on the IMAN1 premises in Amman. A user's survey and feedback collection were performed during the meeting, which showed the high impact of LinkSCEEM project on the Jordanian community. On the other hand, users also exposed the need for continuation of HPC-related support and resource upgrades to fulfil the rapidly growing research demands.



The Iman1 supercomputer and participants of the IMAN1-LinkSCEEM user meeting 2014

The cooperation between CyI and SESAME will continue beyond the duration of LinkSCEEM. It is envisaged that IMAN1 will join the HPC access calls that were established through LinkSCEEM-2 once the system is upgraded.

2.4 Other LinkSCEEM relevant regional HPC centers

Greece

D5.7 Report on prospects for the sustainability of the LinkSCEEM-2 e-Infrastructure

The Greek Research and Technology Network (GRNET) has signed a contract for the development of a high-performance computing system (HPC) to support large-scale scientific applications. The new system, which is currently being installed, marks a return to supercomputing for the country which hasn't had a TOP500-level system since 2000.

The 180-teraflops machine will be based on IBM's NeXtScale platform, which according to the vendor provides a "building-block approach to hyperscale computing," enabling users to start small and scale rapidly as needed. The 426-node system will be powered by Intel Xeon E5 v2 processors and will comprise 8,500 processor cores, interconnected via FDR InfiniBand technology. It will come with one petabyte of storage, based on the IBM General Parallel File System (GPFS). The national supercomputing resource is being developed under the "PRACE-GR – Developing National Supercomputing Infrastructure and Related Services for the Greek Research and Academic Community" project, which is co-funded by the Operational Program "Attica" and the European Regional Development Fund (ERDF).

LinkSCEEM-2 has attracted many users from Greece and also held a number of joint training events with GRNET, which is hosting the new system. The arrival of such large scale system will significantly enhance supercomputing resources in the region. Some access to the system is planned in the context of the VI-SEEM proposal that was submitted jointly by GRNET and CyI to the Horizon 2020 program under the Research Infrastructures work program.

Lebanon

The American University of Beirut has very recently commissioned a hybrid CPU/GPU cluster with a peak performance of 4.2 TFlops. The cluster is equipped with an Infiniband interconnect and 64 GB of memory per node. The cluster will be open to AUB students. LinkSCEEM-2 has engaged the systems administrators group to share best practices. Two system administrators joined for the System administrators' workshop that was held in January 2015.

3 Outlook on regional cooperation

LinkSCEEM-2 was a constant presence in the region over the past four and a half years. Counting all workshops and networking events, a total of 62 events were organised over the duration of the project. A number of regional cooperation initiatives have emerged from the effort.

Cooperation with the South East Europe (SEE) region

The SEE region suffers from similar HPC related issues as the Eastern Mediterranean. There is a generally relative low availability of HPC resources that are accessible by researchers and a relatively high need for training and community building. Parallel to LinkSCEEM-2 the HP-SEE project built joint access mechanisms to HPC resources across the SEE region. This was also complemented with a training and outreach program. HP-SEE was coordinated by GRNET, which also participated in the first LinkSCEEM project. HP-SEE and LinkSCEEM-2 collaborated strongly by co-organising training and networking events (e.g. HP-SEE/LinkSCEEM Athens Summer Training in 2011 and various LinkSCEEM roadshow events in Greece). Also, some inter-project access trials were implemented. For instance, LinkSCEEM-2 production projects were given access to one machine in Bulgaria through the HP-SEE project.

D5.7 Report on prospects for the sustainability of the LinkSCEEM-2 e-Infrastructure

The shared challenges and similar objectives clearly call for closer cooperation in future activities. In order to solidify and structure the cooperation, the concept of the South-East-Europe and Eastern Mediterranean (SEEM) region was established.



The focus of this collaboration is to develop joint strategies for the provision of computing and data resources to computational users across the entire region, develop training programs to target regional needs, support connectivity development in the region and give a strong voice for the region within PRACE. Driven by GRNET and CaSToRC, funding opportunities to support HPC related activities were pursued during the first round of H2020 calls for proposals. This resulted in the submission of a joint proposal under the H2020 call EINFRA-9-2015 *e-Infrastructures for virtual research environments (VRE)*. The proposal aims at establishing a joint regional e-Infrastructure that provides access to computation (HPC, Grid) and data resources. Focusing on key regional user communities (Climate science, Life science, Cultural heritage), the proposal will also provide integrated access to related datasets, data management tools and specific user developed software stacks as well as a supporting training program. The evaluation of the proposal is pending.

In addition, PRACE members of the region also support efforts to re-evaluate the training program in PRACE. Within the framework of the PRACE-4IP project, that just started, the potential of a new, perhaps regional distributed PRACE Advanced Training Center (PATC) is being promoted.

Cooperation on synchrotron applications

Despite suffering multiple delays, SESAME is firmly on track to start operation in 2016. This will give scientists from many different research fields access to a synchrotron facility and result in many computational needs and challenges.

Through the framework of the LinkSCEEM-2 project, key players for the success of SESAME were brought together to embed SESAME into European efforts related to Synchrotron light and HPC. ESRF provided competence in synchrotron applications and is the link to the European Synchrotron community, while CaSToRC established the HPC link to Europe. While LinkSCEEM-2 only provided assistance in the preparatory phase of SESAME due to the delays in starting operation at SESAME, the cooperation of all three institutions will continue beyond the project duration. CaSToRC will provide further assistance in developing HPC capacity at SESAME and access to computational resources, while ESRF is helping to link SESAME better to European efforts in the field of Synchrotron radiation. A first step in this direction was the PANDAAS proposal that was submitted under the H2020 call INFRADEV-4-2014-2015 *Implementation and operation of cross-cutting services and solutions for clusters of ESFRI and other relevant research infrastructure initiatives*. SESAME and CaSToRC joined the proposal consortium that was coordinated by

D5.7 Report on prospects for the sustainability of the LinkSCEEM-2 e-Infrastructure

ESRF. The proposal aimed at developing a joint platform for Synchrotrons and Neutron sources to enable users to remotely analyse, visualise and manage their experimental data. This type of activity will have crucial impact for SESAME users since the capacity to transport and store data is often very low. The proposal has been evaluated favourably, but a final funding decision is still pending. In case no funding can be secured, a reduced scale activity with similar scope will be pursued regardless.

Cooperation of regional system administrators

An activity that gained significant momentum during the second half of the LinkSCEEM-2 project is the collaboration/networking between system administrators across the region. This activity emerged from the technical training program that was undertaken as part of WP4. The idea was to open these training activities to technical personnel outside of the consortium. The huge interest and overwhelmingly positive feedback encouraged the consortium to further expand this activity and resulted in a number of small technical workshops and culminated in a very large workshop with key note speakers from NCSA and JSC carried out in January 2015.

Keeping active links between regional system administrators is very important as it helps to establish some regional standards on how to handle hardware and software and can assist in establishing best practice. It also forms a knowledge pool, enables people to provide mutual assistance and allows monitoring the development of HPC resources (that are not always public knowledge). Following the final LinkSCEEM-2 workshop (held in January 2015) that was dedicated to system administrators, an on-line forum will be established to keep the communication going. The lack of funding for face to face meeting is compensated by the fact that remote assistance in respect of system administration is very well established.

4 Outlook on access to resources

A central pillar of the LinkSCEEM efforts was on the provision of computational resources free of charge to regional scientists. Through LinkSCEEM-2, resources from multiple sites were pooled and distributed to scientists through competitive calls. This access mechanism provided the users with a uniform HPC environment and software stack independent from the individual machine. It also introduced international best practices in HPC into the region. This is a major achievement of LinkSCEEM and a legacy that CaSToRC, together with BA and SESAME intend to preserve.

During the 3rd year of the project an agreement was made to issue calls that give access beyond the initial project duration. The consortium has further expanded this pledge to continue issuing competitive access calls beyond LinkSCEEM-2 (a 7th call for proposals has already been issued). During the final Steering Committee meeting that was held in December 2014, both CaSToRC and BA HPC sites agreed to commit similar amounts of resources to future regional access. This also includes any competitive proposals that would need resource commitments (e.g. CaSToRC and BA pledged some resources to the VI-SEEM proposal mentioned above). This is crucial as uninterrupted access is essential for computational scientists that want to build their research based on large scale computation.

The access mechanism established during LinkSCEEM-2 through the Resource Allocation Committee (RAC) has functioned very well and Norbert Attig from JSC, the chair of the RAC, kindly agreed to continue this role over the next three calls. A new RAC is already in place consisting of users of the LinkSCEEM project from Cyprus, Israel and Greece. The evaluation procedure will therefore continue unchanged. In the long run, it is planned to

D5.7 Report on prospects for the sustainability of the LinkSCEEM-2 e-Infrastructure

continue with a RAC along the one recently constituted containing regional and international computational scientists. This is also in line with international practice.

Apart from the continuation of access to resources, there are also plans to expand the available resources where possible. SESAME has already pledged to join the pool of HPC resources once their hardware upgrade is complete. There are currently no active plans of extension beyond SESAME, however, the consortium is open to include systems that show interest. Encouraging regional collaboration through the systems administrators group also helps to establish similar tools across regional systems, which also lowers the technical hurdle for joining the LinkSCEEM e-Infrastructure.

5 Outlook on training efforts

LinkSCEEM-2 implemented a very ambitious training program consisting of 2 major annual workshops on cross disciplinary HPC topics and three annual workshops focused on the thematic topics pursued in the research work packages (Climate, Cultural heritage, Synchrotron). This workshop program was complemented with numerous training events such as regional roadshows or technical training for system administrators. In total, LinkSCEEM-2 organised 24 workshops and 37 other training events. The aim of these training events was 1) to bring new users in touch with HPC technologies, 2) to provide application training for HPC community codes and 3) to educate scientists in basic and advanced parallel code development.

It is clear that this intense level of face-to-face training activity cannot be sustained beyond the duration of the project. It is nevertheless imperative to continue the provision of HPC training content in order to support the emerging HPC community. Therefore, WP2 and WP4 collected all training material from the workshops from the beginning of LinkSCEEM-2 and is making it permanently available online to ensure scientists have access to training resources. A major amendment of the LinkSCEEM-2 project, due to the cancellation of the connectivity upgrade between Cyprus and Jordan, allowed for resources to be allocated to WP8 for the development of a full-fledged training portal. The portal was developed by JSC throughout the 4th year and extension period of the project. The portal allows video footage of training sessions to be combined with the presented slides for a more effective user friendly presentation of the training material. More importantly, the portal also contains a highly innovative embedded browser based HPC training environment where users can do the tutorials described in the slides. For this, the training portal is linked to EUCLID, the HPC training cluster at CaSToRC, and provides actual HPC access to users.

The training portal does not only feature LinkSCEEM-2 generated content, but also contains links to relevant on-line training material from other sources. This will help users to navigate through the wealth of on-line training material by helping to focus on the relevant content. The online training portal was widely disseminated during the final LinkSCEEM-2 roadshows to ensure high awareness in the regional user community.

The educational access program of LinkSCEEM-2 was very successful. Six University lecturers used LinkSCEEM-2 facilities for their teaching courses. It is very important to make HPC part of undergraduate and graduate programs of study. The provision of access to training hardware supports lecturers across the region to include parallel applications into their curriculum. CaSToRC pledges to continue the free access to EUCLID as part of the educational access program beyond LinkSCEEM-2.

D5.7 Report on prospects for the sustainability of the LinkSCEEM-2 e-Infrastructure

For extended information on the content and capabilities of the Training Portal, we refer you to deliverable D4.6 “Report on Available Training Portal Content”. For a more detailed description of the training program of LinkSCEEM and a description of how to sustainably implement a HPC training program in the region, we refer you to deliverable D4.5 “Report on Adapting Needs and Emerging Trends in User Training Needs”.

Educational efforts have already attracted additional funding. CaSToRC is coordinating the European Joint Doctorate project HPC-LEAP that is funded under the Horizon 2020 Marie Skłodowska-Curie work program. HPC-LEAP is a highly interdisciplinary joint doctorate program realised by bringing together world-leading experts in applied mathematics, high performance computing technologies, particle and nuclear physics, fluid dynamics and life sciences to appropriately train researchers in Europe to exploit high performance computing, advance science and promote innovation.

6 Outlook on private sector involvement

Engaging the private sector with LinkSCEEM-2 was one of the great challenges that will need further attention. Significant economic uncertainty still exists in the target region. Most of the economies in the LinkSCEEM area were declining over the project duration with additional uncertainty caused by the turmoil that followed the Arab Spring. For example, a major target for private sector involvement in Cyprus was initially the financial sector, which almost collapsed in 2013.

Nevertheless, there is potential, in particular in the field of data management. For example, some cultural foundations have shown keen interest into the data management tools that were developed for the cultural heritage community. The digitisation of government services is also set to continue over the next years. This will increase not only the demand in tools, but also demand in knowledge. Another field of expected growth is the biomedical sector that is facing challenges due to the huge amount of data produced by latest generation of genetic sequencers.

Despite the difficult external circumstances, LinkSCEEM-2 tried to open up towards the private sector. The e-Infrastructure access was opened to allow for companies to apply without an academic partner (on the condition that results would be public). These efforts need to continue and perhaps be expanded to also offer professional training to systems administrators and programmers.

7 Conclusions

The LinkSCEEM-2 consortium has ensured that the LinkSCEEM e-Infrastructure continues to operate beyond the project duration. First access calls were already issued and it has been ensured that the structures created by the project remain functional for the next few years.

The greatest challenge for the long term prospects of LinkSCEEM is the timely upgrade of regional HPC systems. While there are adequate HPC resources available over the next couple of years, some major investment will be required to ensure activities born in LinkSCEEM can continue. Jordan will start contributing more processing power as SESAME readies itself for starting up its beam lines. There are also indications that systems will be upgraded in Egypt and a major investment was made already in Greece. Plans are currently being drawn for an upgrade of Cy-Tera that will be needed within the next two years.

D5.7 Report on prospects for the sustainability of the LinkSCEEM-2 e- Infrastructure

LinkSCEEM-2 has provided the technical framework to give adequate access to HPC training for regional scientists to continue the training efforts. The consortium will continue to seek competitive funding to boost training efforts. Several proposals for competitive funding were already submitted to Horizon 2020.

Ensuring timely investment by governments and providing the essential training to young scientists, even outside of externally funded projects, requires determination. The commitment to continuing the LinkSCEEM e-Infrastructure process is already visible with upgrades being planned at The Cyprus Institute and Bibliotheca Alexandrina as well as SESAME, which is preparing to join the e-Infrastructure.