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 (CP- CSA)

LinkSCEEM-2

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

Grant Agreement Number: RI-261600

D7.3

Report on synchrotron data transfer: Report on the monitored date transfers between SESAME and CyI

Version: 1.0

Author(s): Andreas Panteli (CaSToRC)

Date: 14/08/2014

Project and Deliverable Information Sheet

LinkSCEEM Project	**Project Ref. № RI-261600**	
	Project Title: LinkSCEEM-2	
	Project Web Site: http://www.linksceem.eu	
	Deliverable ID: < **D7.3** >	
	Deliverable Nature: <DOC_TYPE: Report>	
	Deliverable Level: PU *	**Contractual Date of Delivery:** 31 / 08 / 2014
		Actual Date of Delivery: 16 / 09 / 2014
	EC Project Officer: Sonia Spasova	

* - The dissemination levels 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**	
	ID: <D7.3>	
	Version: <1.0 >	**Status:** Final
	Available at: http://www.eniac.cyi.ac	
	Software Tool: Microsoft Word 2010	
	File(s): LinkSCEEM-2-D7_3.doc	
Authorship	**Written by:**	Andreas Panteli (CaSToRC)
	Contributors:	WP7 participants
	Reviewed by:	
	Approved by:	PMO

Document Status Sheet

Version	Date	Status	Comments
0.1	31/07/2014	first draft	
1.0	14/08/2014	final	

Document Keywords

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

Table of Contents

References and Applicable Documents

[R1] Project LinkSCEEM– RI-222904– Annex I – Description of Work

[R2] The Cyprus Institute www.cyi.ac.cy

[R3] Cyprus Research and Academic Network www.cynet.ac.cy

[R4] Synchrotron-Light for Experimental Science and Applications in the Middle East www.sesame.org.jo

[R5] SCP linux.die.net/man/1/scp

[R6] BBCP www.slac.stanford.edu/~abh/bbcp/

List of Acronyms and Abbreviations

CyI	The Cyprus Institute
CyNet	Cyprus Research and Academic Network
HPC	High Performance Computing; Computing at a high performance level at any

	given time; Often synonymous with Supercomputing
IP	Internet Protocol
RTT	Round Trip Time
SCP	Secure copy
SESAME	Synchrotron-Light for Experimental Science and Applications in the Middle East
TCP	Transport Control Protocol

1 EXECUTIVE SUMMARY

LinkSCEEM's mission to establish a High Performance Computing (HPC) eco-system in the Eastern Mediterranean region is clearly affected by the network connectivity between the cooperating scientific sites. This fact is the result of the innate need of HPC to produce and manage huge amounts of data, for instance Synchrotron-light radiation experiments produce big amounts of data that have to be analysed on HPC recourses. In order for the LinkSCEEM's mission to be accomplished, the quality of the network connectivity between The Cyprus Institute (CyI) and the Synchrotron-Light for Experimental Science and Applications in the Middle East (SESAME) is of great importance.

An evaluation of the network quality for big data transfers has been conducted between CyI and SESAME. The evaluation involved a series of data transfers, which were performed in a time frame of two weeks; in order to assess the transfer of synchrotron data using two data transfer protocols, namely SCP and BBCP. The data transfer tests between the two sites were repeated every half an hour, thus covering the evaluation of the network connectivity quality during both peak and off-peak times. Besides the two aforementioned data transfer protocols, ping and iperf were used to evaluate the network connection quality in terms of throughput and latency respectively.

The results derived from the assessment described above, indicated a poor quality network connection between CyI and SESAME. The average throughput between CyI and SESAME is relatively low. More specifically the measured data transfer rates are 3.38 Mbps from CyI to SESAME and 2.5 Mbps in the opposite direction, on average. The maximum average throughput is 6.44 Mbps and 5.17 Mbps and the minimum average throughput is 0.05 Mbps and 0.02 Mbps for the two aforementioned directions respectively. Round Trip Time (RTT) ranges from 125 ms to 926 ms, with an average of 237 ms. Additionally, the large standard deviation on both Throughput and RTT is indicating the unstable nature of the network and therefore a poor network connectivity quality.

In conclusion, the current network connectivity between CyI and SESAME is of insufficient quality for sustained data transfers of large data sets. The reasons for this low connectivity can be partially blamed on the very low internet connection of SESAME (10 Mbps), but also on the high latency due to connections through central Europe rather than directly from Jordan to Cyprus. In order for the two sites to have a fruitful collaboration in the near future, in terms of analysis of synchrotron radiation produced data, the upgrade of SESAME's internet connection bandwidth is inevitable.

2 INTRODUCTION

HPC, by its nature, produces and manages big amounts of data. In addition, Synchrotron-light radiation experiments produce a lot of data that need to be analysed on HPC recourses. The mission of the partnership between the two centres, namely CyI located in Cyprus and SESAME, located in Jordan, is for the latter to gather data from the synchrotron-light radiation experiments and for the former to analyse them on the Cy-Tera HPC resource hosted at CyI. Therefore, the data transfer rate between the two sites is of significant importance.

To this end, a series of data transfer tests have been performed between CyI and SESAME, with the participation of Cyprus Research and Academic Network (CyNet), in order to evaluate the data transfer quality between the two sites.

3 SYNCHROTRON DATA TRANSFER MONITORING

The evaluation of the data transfer between the two sites was performed using real synchrotron-light data provided by ESRF, with a series of data transfer experiments. At first, two attempts have been done to transfer 349MB of real synchrotron-light data, provided by SESAME. For this first transmission, two data transfer protocols have been used, namely SCP and BBCP. SCP is a very well known protocol, which supports data transfers between hosts on a network. It uses SSH for authentication and encryption of data during transfer. Due to its popularity, SCP has been used as the baseline for the evaluation of another data transfer protocol, namely the BBCP. BBCP is a point-to-point file copy application, which is capable of transferring files at approaching line speeds in the network. The performance improvement of BBCP derives from the following two main characteristics of the protocol: a) BBCP uses multiple TCP/IP streams (default is 4) and the data to be transferred is shared between those streams. b) The data is not encrypted prior to the transfer by default.

Figure 1 summarizes the results of the two data transfer tests, comparing the two protocols mentioned above. As it can be seen from the figure, the two protocols have nearly of the same maximum performance, at around 5 Mbps. It should be noted that the theoretical internet connection bandwidth of the two sites is 100Mbps for CyI and 10Mbps for SESAME. Therefore, the theoretical bandwidth of the connection between the two sites is 10Mbps, as it is limited by SESAME's internet connectivity. The fact that the two protocols have around the same maximum performance is caused by the low bandwidth of the connection. In high performance networks, the computation overhead that is needed by SCP in order to encrypt the data becomes the bottleneck of the transfer. For the evaluated network this is not the case, as the low bandwidth of the connection becomes the bottleneck of the data transfer. Additionally, due to the low available bandwidth of the connection, BBCP cannot exploit the multiple TCP/IP streams to improve the transfer performance. Thus, the theoretical eminence of BBCP in comparison to the SCP was not revealed for the network connection under evaluation.

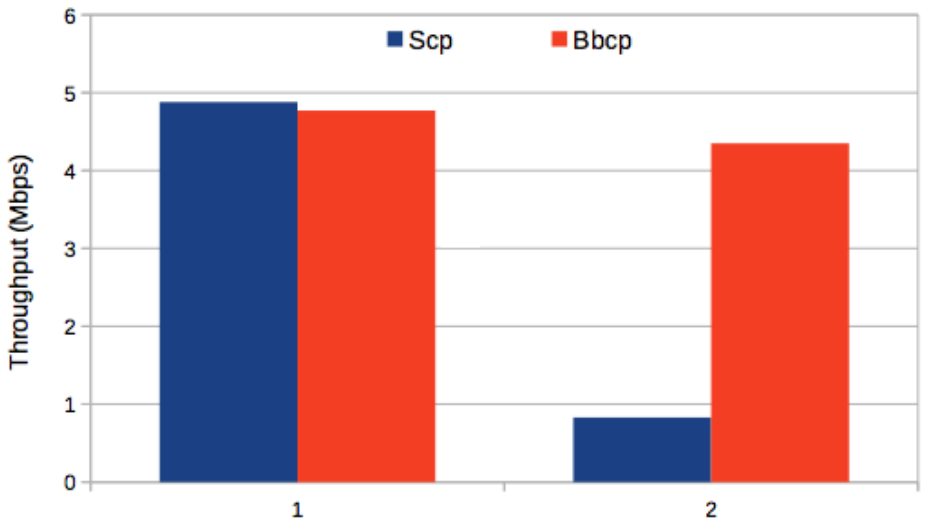

Figure 1: SCP Vs BBCP Throughput in Mbps

Furthermore, as seen in Figure 1, the results we got when we performed the data transfer test at two different time points are not consistent. This observation let us to perform further tests, in order to examine the behaviour of the network, during a longer period of time. We therefore performed a number of tests using iperf and ping, to evaluate the network connection in terms

of throughput, using iperf and latency, using ping. The data transfer tests were performed at peak and off-peak times and were continuously monitored. More specifically, the tests have being performed every half an hour, for 15 consecutive days in the period between 26/11/2013 and 10/12/2013, as it is shown in Figure 2 and Figure 3.

Figure 2 shows the bi-directional throughput between the two sites measured with iperf. It is evident that the main drawback of the connection between the two sites is the variation of the throughput over time. This confirms the conclusions drawn from the first test, where the results regarding the throughput were not consistent for the two transfer tests. Additionally, the variation of the network throughput does not show any specific pattern, e.g. having lower performance during business hours, or having higher performance during the weekends. This adds another drawback for the data transfer between Cy-Tera and SESAME, as the performance of the network is not predictable. Finally, it can be seen from the figure that the throughput from Cy-Tera to SESAME is almost always higher than the throughput in the opposite direction.

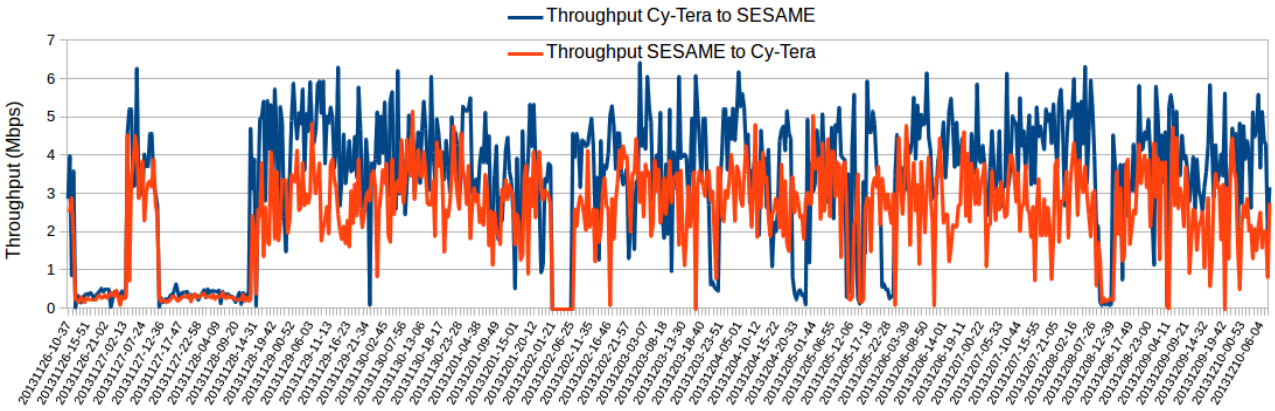

Figure 2: Throughput between Cy-Tera and SESAME

Figure 3 presents the correlation of the throughput, between the two sites, with the Round Trip Time (RTT). As it can be seen, when the RTT rises, the throughput of the data transfer decreases. There are some points in time that the RTT is approaching the value of 1 second, which indicates congestion in the network. The variations of the RTT and throughput, reveal the unstable nature of the network connection between the two sites and affect the data transfers directly.

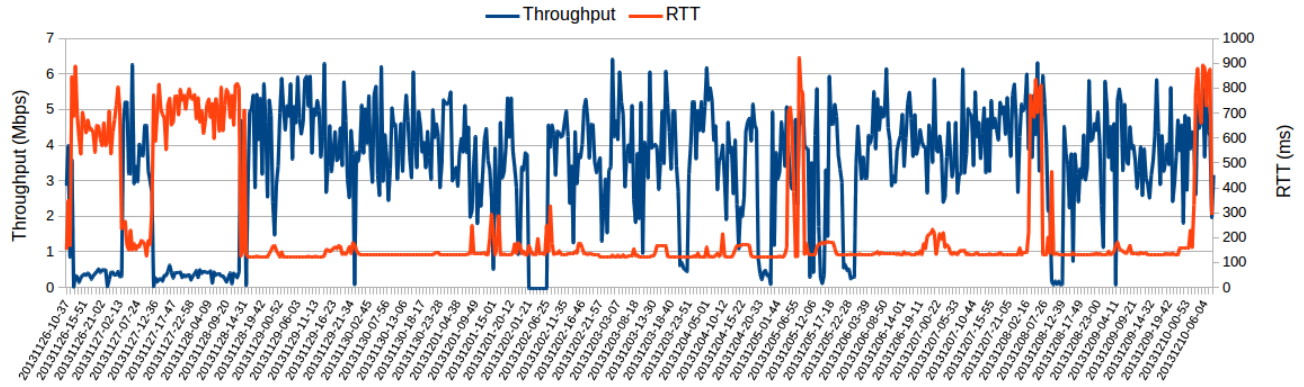

Figure 3: Throughput and RTT between Cy-Tera and SESAME

A statistical analysis of the results is given in table 1. One can see that the average throughput between Cy-Tera and SESAME is relatively low. More specifically we have measured data transfer rates of 3.38 Mbps from Cy-Tera to SESAME and 2.5 Mbps in the opposite direction, on average. Additionally, there is large standard deviation on both Throughput and RTT, indicating a lot of variations in the network and therefore a poor network connectivity quality.

Table 1: Statistics

	Cy-Tera to SESAME (Mbps)	SESAME to Cy-Tera (Mbps)	RTT (ms)
MAX	6.44	5.17	926.22
MIN	0.05	0.02	125.17
AVERAGE	3.38	2.50	236.76
STDEV	1.71	1.23	212.62

4 CONCLUSIONS

In conclusion, based on the results we have presented, derived from the experiments that have been performed in order to evaluate the data transfer quality between Cy-Tera and SESAME, there are many weaknesses in the network connection between the two sites. The low available bandwidth at SESAME site, in conjunction with the big variations in terms of throughput and RTT in the site-to-site connection, makes the data transfer between the two sites to be relatively slow and unreliable.