ROSPACEGRID

ANNUAL REPORT 2010

project: “Romanian GRID middleware repository for Space Science Applications”,
Contract No 98050

Ion-Sorin ZGURA-Principal Investigator
Bucharest •July, 2010

 Ion-Sorin ZGURA • szgura@spacescience.ro •

Technical objectives planned for the reporting period

WBS (work breakdown structure) of the WP3

M1 Finalization of the equipment acquisitions

A1.1 Auction

1.1.1 Elaboration of the requirements for equipments and the auction document

1.1.2 Realizing the auction

1.1.1 Evaluation of the equipment proposals

A1.2 Equipment procurement

1.2.1 Elaboration of the contract between Institute and the equiment supplyer

1.2.2 Reciving the equipment

M2 Instalation, configuration and deploying the RoSpaceGRID cluster

A 2.1 Hardware

2.1.1. Installation

2.1.1.1 Installing the cluster in a new rack

2.1.1.2 Installing the UPSs

2.1.1.3 Installing the switches

2.1.1.4 Installing the servers

2.1.2. Conections

2.1.2.1 Electrical

2.1.2.2 Network

A 2.2 Software

2.2.1 Installation

2.2.1.1 Installation of the operation system on the front-end server

2.2.1.2 Installation of the management sistem on front-end server

2.2.1.3 Installation of the worker nodes

2.2.2 Configuration

2.2.2.1 Configuration the network devices

2.2.2.2 Configuration the specific software

2.2.3 Testing

2.2.3.1 Testing the operation system
2.2.3.2 Testing the management system
2.2.3.3 Testing the specific software

A 2.3 Deploying the computing resources

A 2.4 Maintanance of the cluster in 24/24 regim

M3 Monitoring, Controlling and R&D parameters of the Data Center

A 3.1 Monitoring, Control and R&D of the environment parameters
3.1.1 Installation of the temperature sensors
3.1.2 Development of data collection programs
3.1.3 Data Storing in MySql databases

A 3.2 Monitoring, Control and R&D the cluster parameters
3.2.1 Moniroring CPU load
3.2.2 Monitoring CPU usage
3.2.3 Monitoring Storage usage
3.2.4 Monitoring Network usage

M4 Starting CAMSAD collaboration

A 4.1 Started contact with Prof. Juhani Huovelin, University of Helsinki (UH) Finland.

A 4.2 Kick-off Meeting in Helsinki

A 4.3 ISS milestone and tasks in the frame on CAMSAD proposal

A 4.4 Strengthening institutuional relationships with CAMSAD members

M 5 Dissemination

Snapshots on RoSpaceGRID activities

Instalation of tne new rack:

Installing the UPSs

2	**APC Smart-UPS RT 6000VA RM 230V**	
 	APC smart UPS RT, Rack mounting support 3U Output power capacity 4200Watts/ 6000 VA, Nominal voltaje 230 V, Efficiency at full load 92% Output connections 8 IEC 320 C13 backup + 2 IEC C320 C19, Typical recharge time: 2,5 hours Interface port DB-9 RS-232/ RJ-45 10/100 Base-T, Smart Slot Audible alarm distinctive low battery alarm Tipycal backup time: 15.8 minutes (2100 Watts), 5.3 minutes (4200 Watts) Filtering Full time multi-pole noise filtering 0,3% IEEE surge let-throught Web/SNMP Management card Regulatory approvals: C-tick, CE, EN 50091-1, EN 50091-2, EN 55022 Class A, EN 60650, en 61002, GOST, VDE Dimensions: 130x432x660 mm height x width x depth Net weight: 54.55kg	

Installing the switches

HP ProCurve Switch 2810 48G	
	HP ProCurve rackmounting Switch 1U , J9022A 44 10/100/1000 ports (IEEE 802.3 Type 10Base-T, IEEE 802.3u Type 100Base-TX, IEEE 802.3ab 1000Base-T Gigabit Ethernet); 1 RJ-45 serial console port; 4 dual-personality ports - each port can be used as either an RJ-45 10/100/1000 port (IEEE 802.3 Type 10Base-T, IEEE 802.3u Type 100Base-TX, IEEE 802.3ab 1000Base-T Gigabit Ethernet) or an open mini-GBIC slot (for use with mini-GBIC transceivers), Switching capacity 86 Gbps, Mac address table size 8000 entries Memory and processor MIPS 264 Mhz, 16 MB flash, 64 MB sdram, packet buffer size 1,5 MB Management HP ProCurve Manager Plus, command line interface, web browser, configuration menu: out-of-band management Network management IEEE 802.1AB link layer discovery protocol (LLDP) RFC 2819 four groups of RMON/ RFC 3176 sFlow/ SNMPv1/v2/v3 Voltaje 100-127 VAC/ 200-240VAC/ Current 1,5 A Maximum power rating 92W Regulatory approvals: C-tick, CE, EN 50091-1, EN 50091-2, EN 55022 Class A, EN 60650, en 61002, GOST, VDE Net weight: 3,9kg

Installing the servers

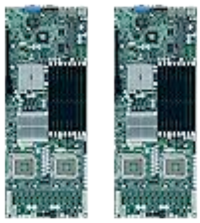
Nr.	Products Tip	Specifications **Supermicro SYS-6015TW-T TM** **SUPERMICRO PROPRIETARY TRADE MARK**
1	**Processors**	**4 x** Intel® Xeon™ Quad Core **E5430 2.66GHz** Processor with 12MB L2 Cache 1333MHz 80W
2	**Chassis**	**Form Factor: 1U**, rackmountable/ Supports 2 proprietary 16" x 6.5 " (40.6cm x 16.5cm) motherboards **Model: SC808T-980V** (B)
3	**Dimensions**	27.75" (705 mm) Depth 1.7" (43mm) Height/ 17.2" (437mm) Width
4	**Mounting Rails**	Compatibil rack 19'"CSE-PT51L (1U Mounting Rail Kit)
5	**Processor/Cache (per Node)**	Dual 771-pin LGA Sockets Supports up to two Intel® 64-bit Xeon®processor(s) of the same type below: Quad-Core Intel®Xeon®Processor 5400/5300 sequence (Harpetown/ Clovertown processor) up to 3.20 GHz System Bus: 1600/ 1333 / 1066 MHz system bus
7	**Memory Capacity (RAM)**	**32 GB 667 FB-DIMM ECC DDR2 SDRAM 72-bit, 240 pin gold-plated DIMMS** **16 x 2 GB** DDR2 module DDR2 667 Mhz ECC FB Memory 72-bit, 240-pin gold-plated DIMMs Corrects single-bit errors/ Detects double-bit errors (using ECC memory) / Supports Intel® x4 and x8 Single Device Data Correction (SDDC)
8	**Power Supply**	980W AC-DC power supply /AC Voltaj : 100-240 V, 60-50 Hz

9	**On-Board Devices (per Node)**	**Super X7DWT - Supermicro Proprietary** - Intel® 5400 (Seaburg) chipset / ESB2 6-Phase-switching voltage regulator with auto-sense from 0.8375V-1.60V 6x Cooling Fans -8 sloturi memorii/ DDR2 ECC FB-Dim up to 64 GB Memory - FSB 1600/ 1333/1066/667 SYSTEM BUS - 2 x Dual 771-pin LGA Sockets - 4x Hot-swap SATA 2 Drive Bays - 2Intel® (ESB2/Gilgal) 82563EB Dual-Port Gigabit Ethernet Controller - Enhanced Intel Speedstep Technology, - Intel Virtualization Technology, - RAID 0, 1, 5, 10 suport Windows/RAID 0, 1, 10 Linux, - 1 1 (x16) PCI-e Generation II slot - 4 USB 2.0 Compliant - clock generator CK410B - Intel® I/OAT support for fast, scaleable, and reliable networking/ Winbond 83627HF chip - 1 Fast UART 16550 serial port / 2 x RJ45 output - 8x fans with tachometer status monitoring - Status monitor for speed control - Monitoring for CPU and chassis environment - PSU I^2C temperature sensing logic
10	**Network Controller (per Node)**	Intel® (ESB2/Gilgal) 82563EB Dual-Port Gigabit Ethernet Controller Supports 10BASE-T, 100BASE-TX, and 1000BASE-T, RJ45 output Intel® I/OAT support for fast, scaleable, and reliable networking
11	**Graphics (per Node)**	ATI ES1000 controller with 32 MB of video memory
12	**Hard-disks (per Server)**	**2x500** GB Hdd WD 7200 rpm/ WD RE3 3GB/s 4 x 3,5” Hot-swap SATA Drive Bays
13	**USB ports (per Server)**	4x USB rear ports/ USB 2.0 Compliant 4x USB internal headers

14	**Software compatibility**	**Linux 32 bit / 64 bit** **Other operating system compatibility**: MS Win 2000 Advanced Server + SP4, MS Win XP Pro + SP2(x64), MS Win 2003 Server Enterprise Edition(x64), Redhat Linux ES 3 (Update 7/8), Redhat Linux ES 4 (Update 2/3/4), SuSE Linux 10/10.1, SLES 10, RedHat Linux 3 (Update 8) x64, RedHat Linux 4 (Update 2/3/4) x64, SuSE Linux 10 x64,
15	**Power Supply**	PWS-981-1S/ 980W 1U Cold Swap 12V high eficiency Power : 100-240 V, 60-50 Hz, 14-6 Amp
16	**Keyboard & mouse**	Not included
17	**Case**	**SC808T-980V** (B), **Supermicro Proprietary** 1U, RACKMOUNT, black

Physical Views of CE

Supermicro SYS-6015TW-TB	
Case SM 808T	

Motherboard Super X7DWT + Super X7DWT	

Maintenance of the PLANCK CLUSTER

The cluster comprises about 15 servers with 8 cores, 2.6GHz and 16 GB RAM.

CPU load for entire period of usage

At the end of this year, the ISS-PLANCK cluster has been created for performing parallel computing with MPI for Planck mission. Coordinator for this is the Planck spokesperson from ISS Dr. Lucia Popa.

Bellow there are giving the performances of this cluster monitoring with ganglia. In the Figure 1 is shown the Planck cluster load from the beginning to present.

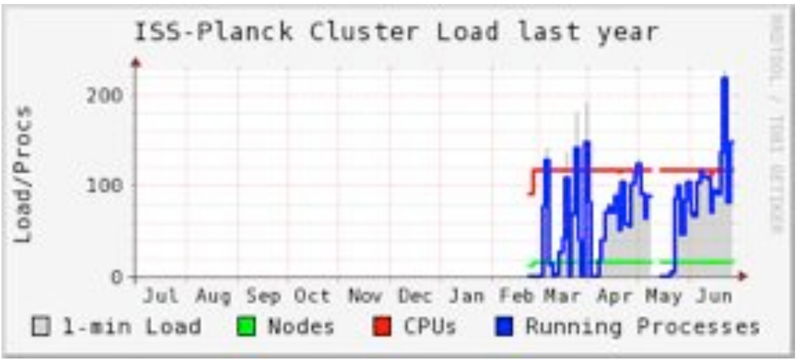


Figure 1. *Planck cluster load from beginning to present.*

The usage of Plank cluster is pretty good of about 35%. Figure 2 shows the usage percentage for entire cluster from beginning to present.

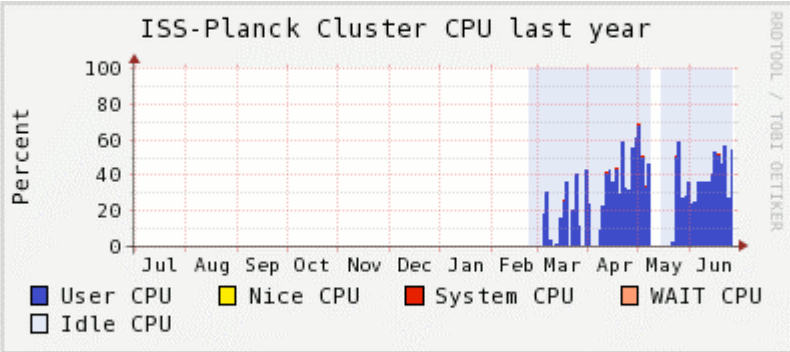


Figure 2. *Percentage load of the entire cluster from beginning to present*

Figure 3 shows that every node in the cluster are approximatively equally used with exception of the node 5 which is the front-end machine used to as muster for lunching the jobs.

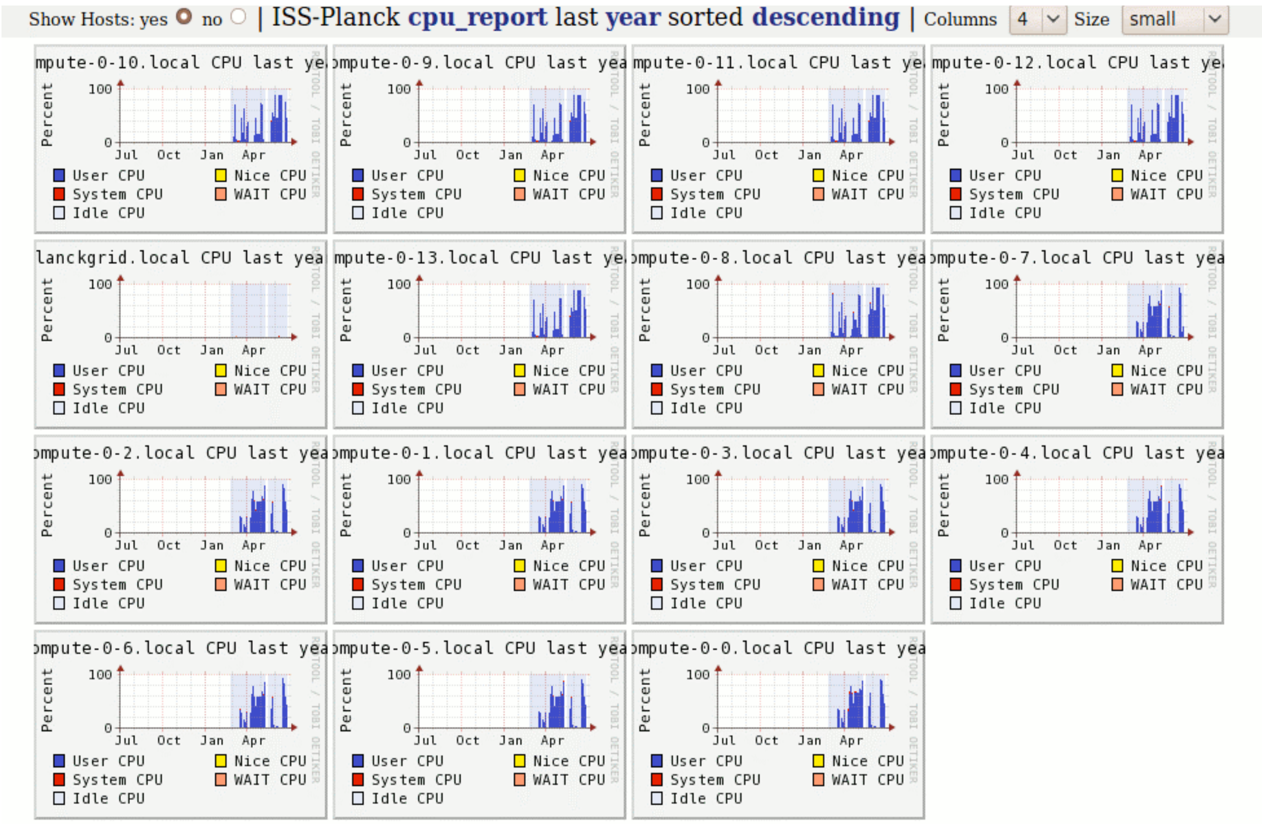


Figure 3. *The worker node load from the beginning to present*

CPU usage for last month (Figure 4)

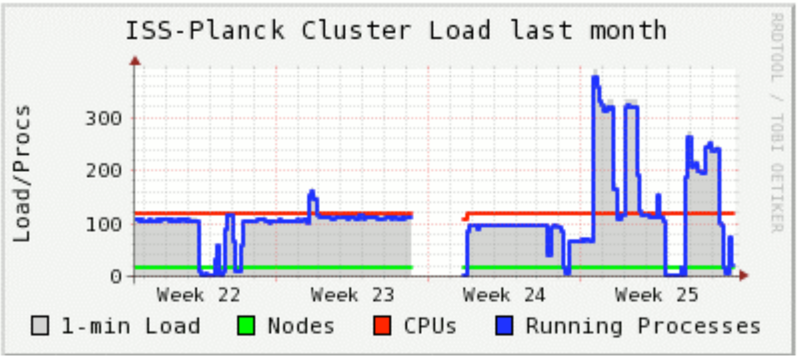


Figure 4. Overall *CPU load on the last month*

Figure 5 shows the CPU usage percentage on the last month.

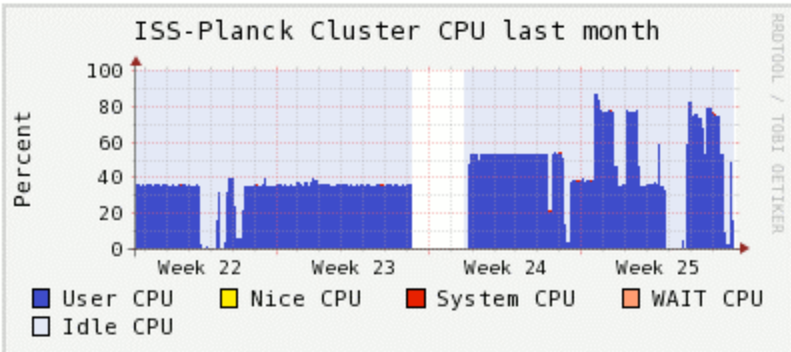


Figure 5. *The CPU percentage usage of entire cluster*

1. *Worker nodes load on the last week*

Figure 6 shows that jobs are lunches on the CPUs in order they are listed in the configuration files.

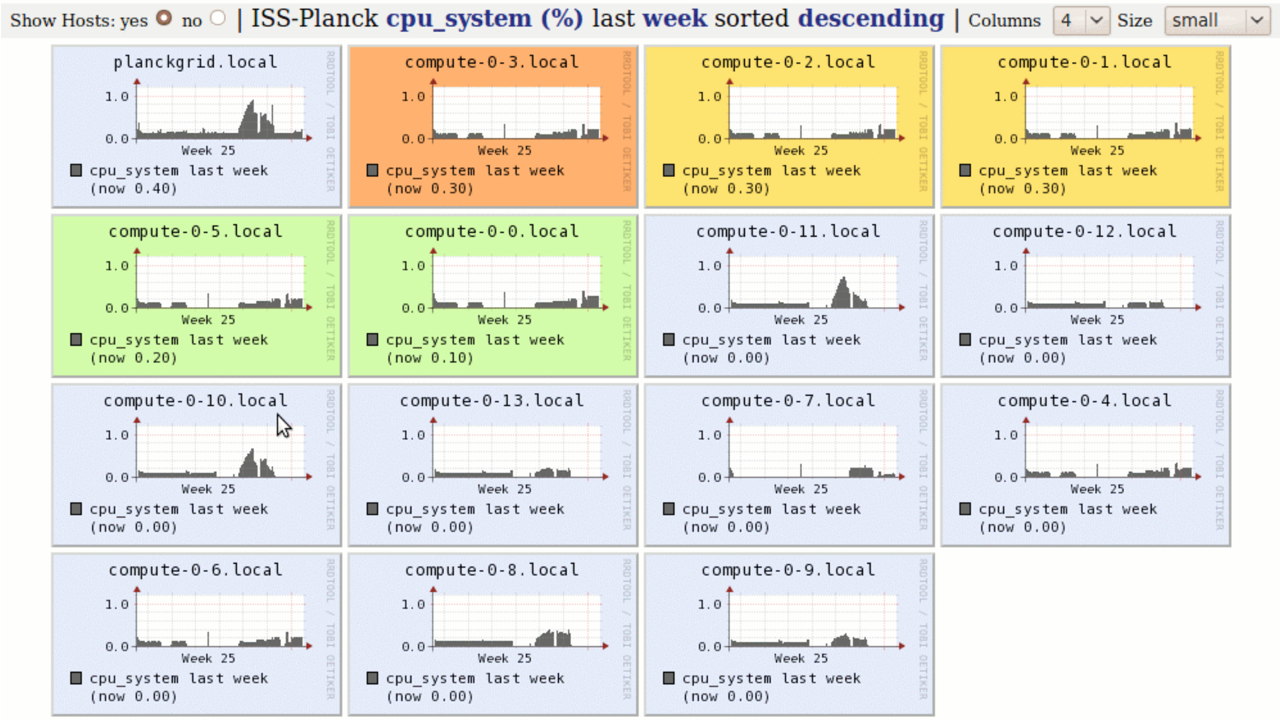


Figure 6. *Worker nodes CPU load for the last week*

2. *Memory and network*

In Figure 7 and 8 are given the overall memory usage and networking traffic on the last month.

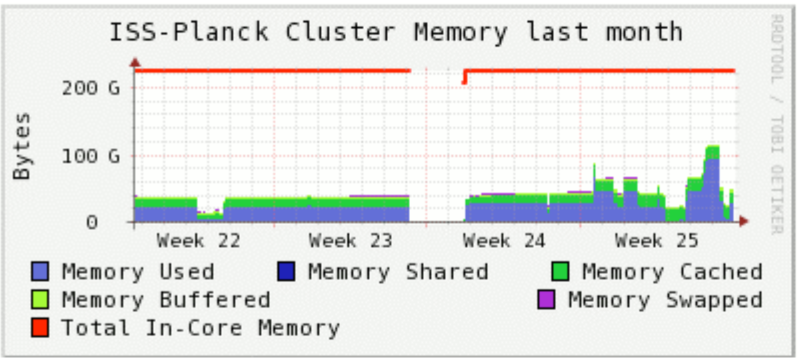


Figure 7. *Memory usage on last month*

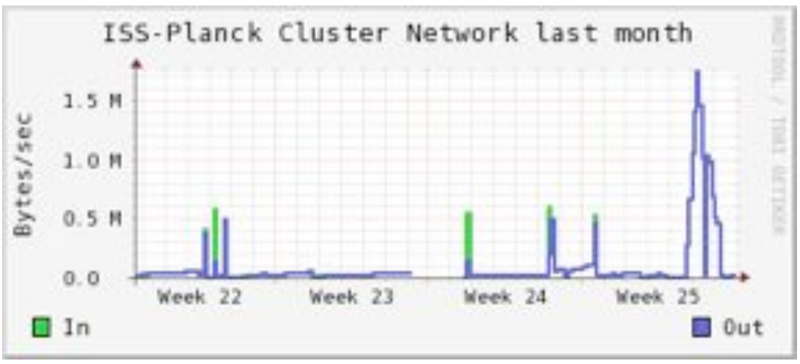


Figure 8. *Network transfer on last month*

Maintenace of the ISS-ROSPACEGRID CLUSTER

The cluster comprises about 27 servers with 8 cores, 2.6 GHz and 16 GB RAM.

ROSPACEGRID CPU load for entire period of usage

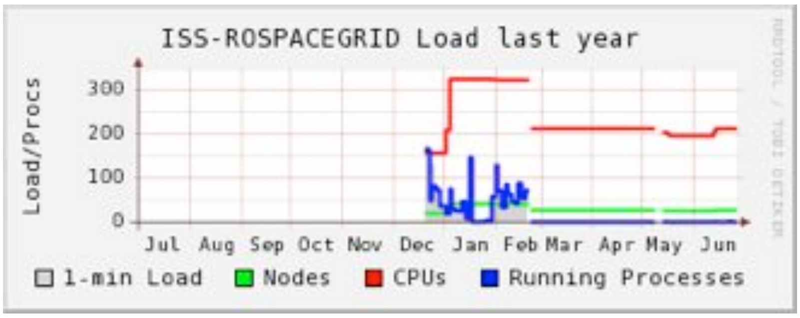


Figure 9. *CPU load for Rospacegrid from beginning to present*

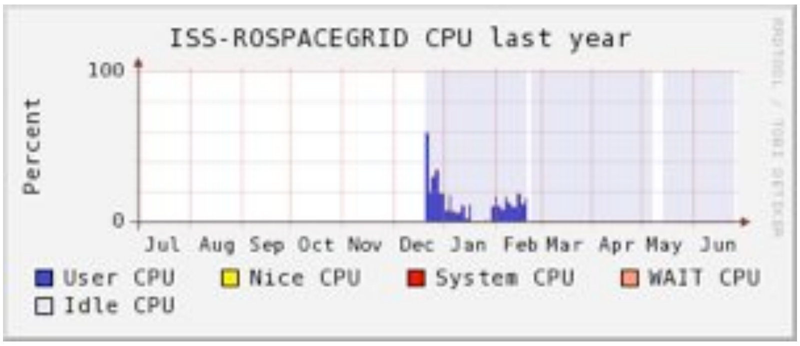


Figure 10. *Percentage CPU for Rospacegrid from beginning to present*

The load of the worker nodes is given in Figure 11. for entire period of working.

Figure 11 Working nodes load

Starting the CAMSAD collaboration

CAMSAD collaboration started arround the need to analyze the data from multiple satellites.

Proposal full title: **Combined Analysis of Multi-Source Astronomical Data**
Proposal acronym: **CAMSAD**
Type of funding scheme: **Collaborative Project (CP)**
Work programme topics addressed: SPA.2010.2.1-03 **Exploitation of space science and exploration data**

Name of the coordinating person: **Juhani Huovelin**

List of participants:
Participant organisation name Country
1 (Co) University of Helsinki (UH) Finland
2 P1. University of Southampton (SOTON) United Kingdom
3 P2. Danish Technical University (DTU) Denmark
4 P3. Institute for Space Sciences (ISS/INFLPR) Romania

Concept and objectives

1.1.1 Introduction

Astronomical observations, and remote sensing data in general, naturally include a combination of information from all phenomena that are active in the observed target. Also phenomena in the medium between the target and the observing instrument, background sneaking in the instrument aperture, and radiative interactions in the instrument and satellite around the detector add their dedicated signatures to the signal that is finally transformed in digital form called scientific data. The most feasible approach in the analysis would thus be to include a complete set of phenomena relevant in causing signal in the scientific instrument in the physical model that is then fitted to the data. Such models are usually developed for the analysis of radiation in specific wave-length or energy interval that coincides with the sensitivity range of an instrument. In some cases a model is in a very good agreement with the observations, but much too often there are discrepancies and conflicts, many of which remain unexplained. A significantly bigger challenge is faced while combining the analyses of two or more instruments, since the number of instrument-related and other local environmental parameters is multiplied, and thus the models be-come more complex. Explanations to observed anomalies and differences, if any, are usually found after tedious and lengthy handwork, and the interpretations are often processed further by discussions in the scientific community.

1.1.2 Scientific and Technical Objectives

1.1.2.1 Core objective

The core objective of the CAMSAD project is to improve our understanding of astronomical data from space missions by exploiting as fully as possible all useful data in the scientific analysis. For this we apply data mining methods and modern workflow engines, and develop software algorithms, which will speed up the scientific analysis of data from several instruments and satellites at a time. This will also enable and speed up resolving previously unknown effects that cause anomalies in the data and can mislead the interpretation of the results, and can also cause mismatch in the results from different instruments.
The core objective of the project directly contributes to the exploitation of data from three ESA satellites

(Integral, XMM-Newton, SMART-1), and also contributes to the national space programmes in Finland, Denmark and UK, which are members of the core science teams and in-vest a significant amount of resources to the utilization of the data from the three satellites. The project aims also at strengthening the collaboration between South-East European countries with Western European countries via partnership of ISS/INFLPR from Romania in this project.

1.1.2.2 Scientific objectives

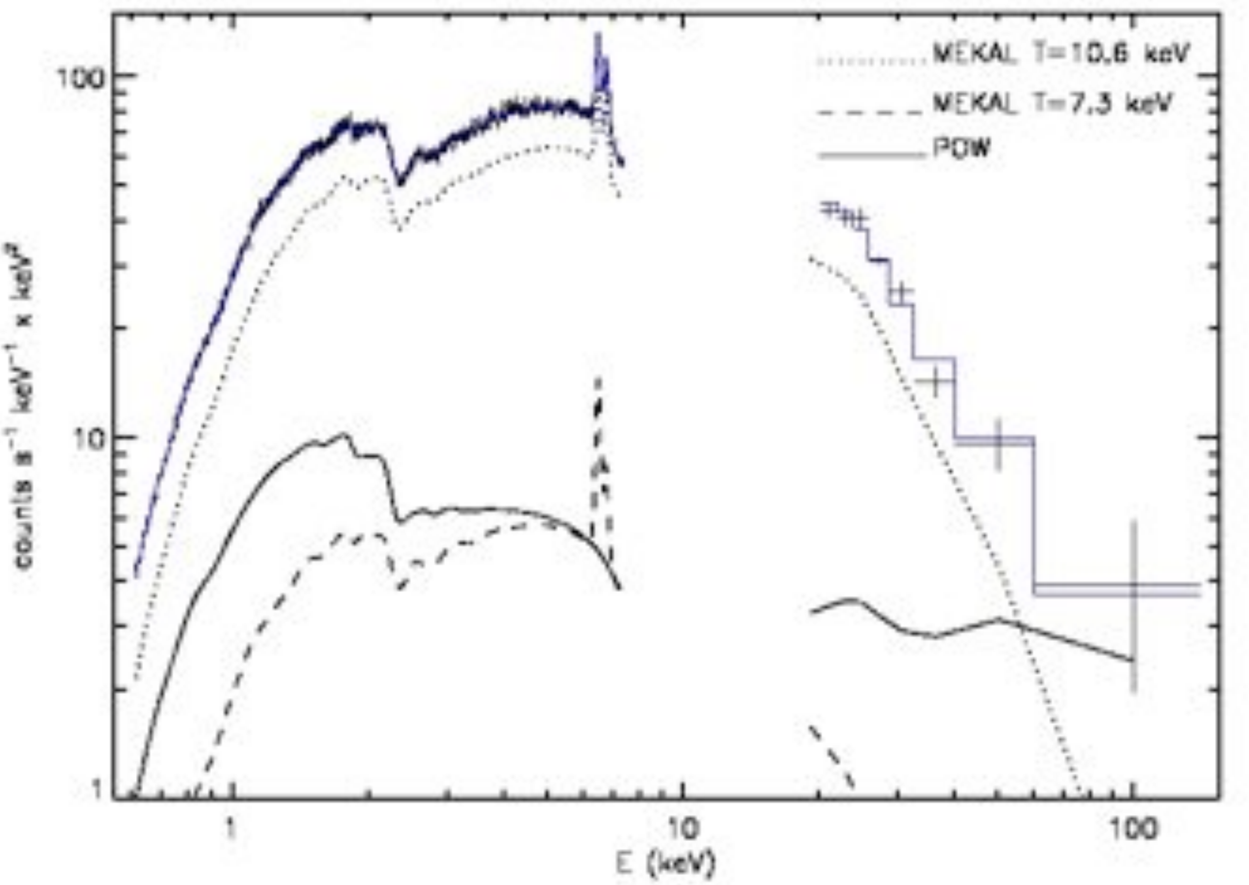
The scientific objectives of this research are, 1) an improved understanding of the nature and structure of astronomical sources by broadband spectral modeling with simultaneous use of data from two or more instruments i.e. applying cross-mission and cross-instrument analysis, 2) improved calibration by simultaneous use of scientific, housekeeping, and auxiliary data from two or more satellites and/or instruments, 3) improved background determination by using all connected data, 4) an improved X-ray source database by simultaneous use of data from two or more instru-ments.

A potential bonus, which may be of high scientific value itself, is finding new object classes and structures as a result from applying data mining methods to large astronomical databases for the distinguished 4 objectives above.

In the following, we explain in more detail the objectives of our plan.

1) Cross-mission and cross-instrument analysis

Any given astronomical instrument only works in a certain subspace of observational parameters. For example, some detectors (like XMM-Newton PN) only detect photons with energy in the 0.1-10 keV range, whereas some detectors (like INTEGRAL ISGRI) are sensitive only to photons with energy higher than 10 keV. The usage of only one instrument may limit the derived information on the astrophysical object, since different processes may yield emission that peaks at different wavelengths. For example, the thermal emission in clusters of galaxies dominates the 0.1-10 keV energy band, while the non-thermal inverse Compton component is better visible at higher energies. Thus, combining the XMM-Newton PN and INTEGRAL ISGRI data we are able to separate the non-thermal component in Ophiuchus cluster of galaxies and derive its properties (Nevalainen et al., 2009, see Figure 1.) which would have been impossible by using one of the two instruments alone.



The black crosses show the data and the statistical uncertainties. The blue line shows the total best-fit convolved model. The non-thermal and thermal components are also show separately by solid and dotted lines, respectively.

Table 1.3 a: Work package list

Work package No1	Work package title	Type of activity2	Lead participant No3	Lead participant short name	Person months4	Start month5	End month
WP1	Project management	MGT	1	UH	8	M1	M48
WP2	Cross-mission calibration	RTD	1	UH	32	M1	M42
WP3	Smart Data Processing	RTD	1	UH	20	M1	M24
WP4	Knowledge Management	RTD	1	UH	32	M7	M48
WP5	Knowledge Warehouse Management	RTD	4	ISS/ INFLPR	57	1	42
WP6	Data Quality and Machine Learning	RTD	2	SOTON	47	M1	M48
WP7	Statistical Multi-source and Multi-set Analysis	RTD	3	DTU	67	M1	M42
WP8	Portal & Visualisation	RTD	1	UH	26	M1	M42
WP9	Dissemination and Exploitation	OTHER	1	UH	26	M31	M48
				Total:	**315**		

Description on WP5

Work package number	WP 5	Start date or starting event:					
Work package title	Knowledge Warehouse Management						
Activity Type:	RTD						
Participant number	1	2	3	4			
Participant short name	UH	SOTON	DTU	**INFLPR**			
Person-months per participant:	2			55			

Objectives

We plan to use state-of-the-art GPU grid infrastructure for execution of intensive and time con-suming processes, for data mining in particular.

GPU grid will also be needed for data fusion and data set combination in the scientific analysis. GPU's are particularly efficient in matrix calculus, for instance, processing of digital images, which basically are always matrices. The distribution of processes into grid will be coordinated by certain specialized tasks within the workflows, while development of new workflow tasks may be required for the implementation of the GPU grid into our system.

- Setup of Warehouse infrastructure and databases
- Developing algorithms for data mining
- Monte-carlo simulations
- Delivery of Knowledge Gluon

Description of work

WP 5.1 Setup of Warehouse and Databases
Setup of warehouse infrastructure for data storing and data processing including persis-tent databases.
WP 5.2 Developing of algorithms
Developing algorithms for data mining and data image processing using CUDA and OpenCL tech-nologies
WP 5.3 Monte Carlo simulations with GEANT 3 & Fluka
WP 5.4 HPC system implementation in ISS
Data storing and data processing to produce Knowledge Gluon with delivery mechanism.

Deliverables

D5.1 Warehouse infrastructure with databases - M12

D5.2 Algorithms for data mining and data image processing -M24

D5.3 Results from Monte Carlo simulations -M30

D5.4 HPC system in ISS for data storing and data processing -M42

Unfortunately the project did not pass. But we are going to "polish" the proposal for a new call.

Results achieved

The main results of the WP3 are:

- Install, configure and deploying the new cluster
- training of the team to use the hardware and software (middleware) and to offer support to interested physics groups(Plack-ESA Mission) . We obtained a near 100% uptime.
- Start up the CAMSAD collaboration and lanch fist proposal to FP7
- Maintenace of the RoSpaceGRID cluster

As a by-product, team's activity contributed to the results:

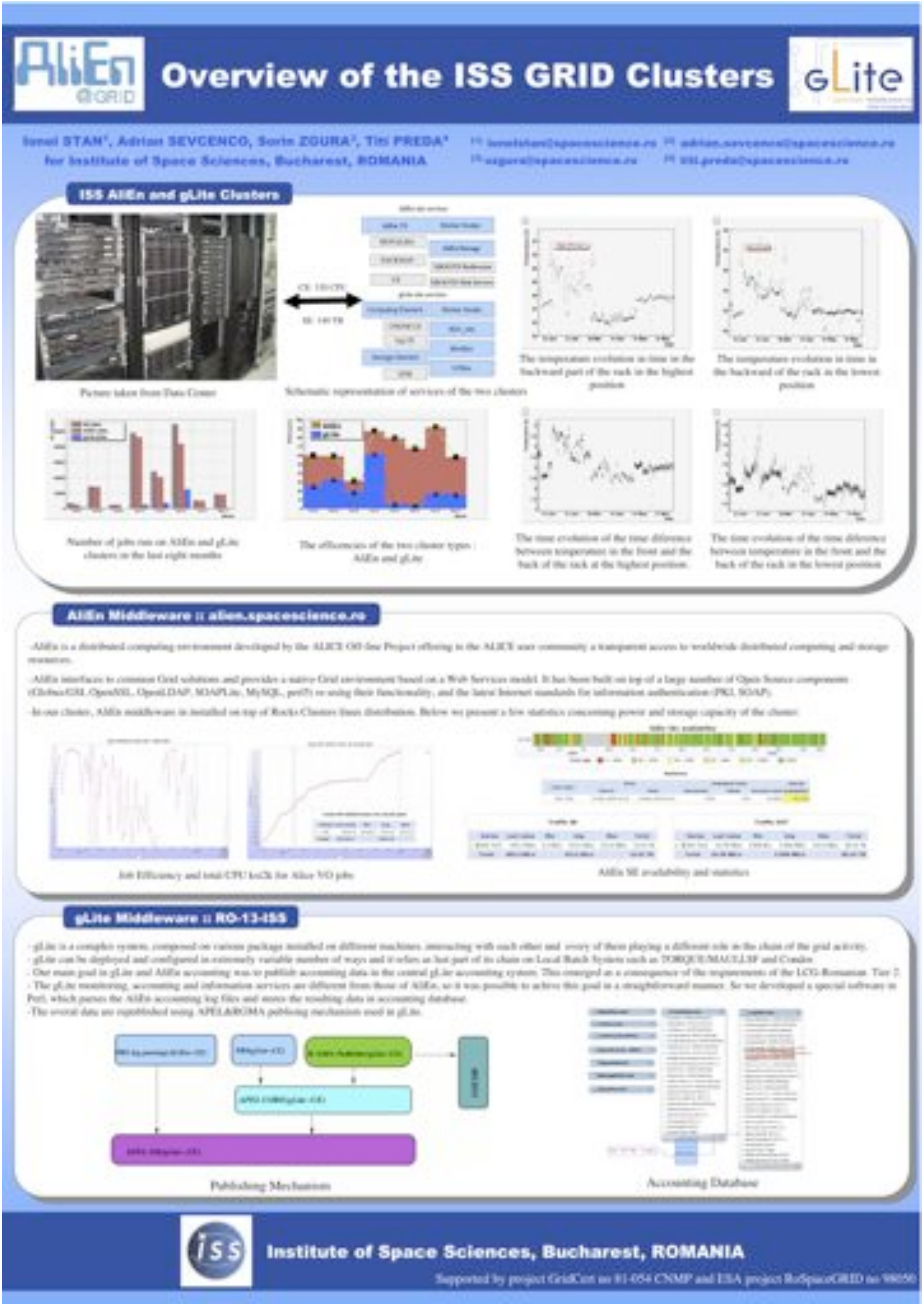
- The 4th International Conference "Distributed Computing and Grid-technologies in Science and Education", LIT-JINR, Dubna, Russia, "Overview of the ISSS GRID Clusters"
- 9th RoEduNet International Conference, Sibiu 2010
- 1st Workshop of Ro-eIRG, Timisoara 2010

Changes proposed for the next period

At this time the is no need to change the initial master plan of the WP4.
We have in plan to organize in the near future an workshop with topics (GRID, HTC and dataminnig for data from satellites)

Financial balance of the project against planned budget

	Total project (EURO)	Actual 2007-2010 (EURO)
Salaries	510000	388,167
Consumables	62,000	48,001
Travel	43,000	13,494
Overhead	85,800	89,874
Subtotal	700,800	539,536



AliEn @GRID
Overview of the ISS GRID Clusters
gLite
Ionel STAN, Adrian SEVCENCO, Sorin ZGURA, Titi PREDA
for Institute of Space Sciences, Bucharest, ROMANIA
ISS AliEn and gLite Clusters
Picture taken from Data Center
Schematic representation of services of the two clusters
The temperature evolution in time in the backward part of the rack in the highest position
The temperature evolution in time in the backward of the rack in the lowest position
Number of jobs run on AliEn and gLite clusters in the last eight months
The efficiencies of the two cluster types : AliEn and gLite
The time evolution of the time difference between temperature in the front and the back of the rack at the highest position
The time evolution of the time difference between temperature in the front and the back of the rack in the lowest position
AliEn Middleware :: alien.spacescience.ro
Job Efficiency and total CPU to Clk for Alice VO jobs
AliEn SE availability and statistics
gLite Middleware :: RO-13-ISS
Publishing Mechanism
Accounting Database
Institute of Space Sciences, Bucharest, ROMANIA