

Maintenance and upgrade of the services of AliEn, the ALICE Grid Environment.

Narine Manukyan

ANSL/ALICE and ALICE offline teams,

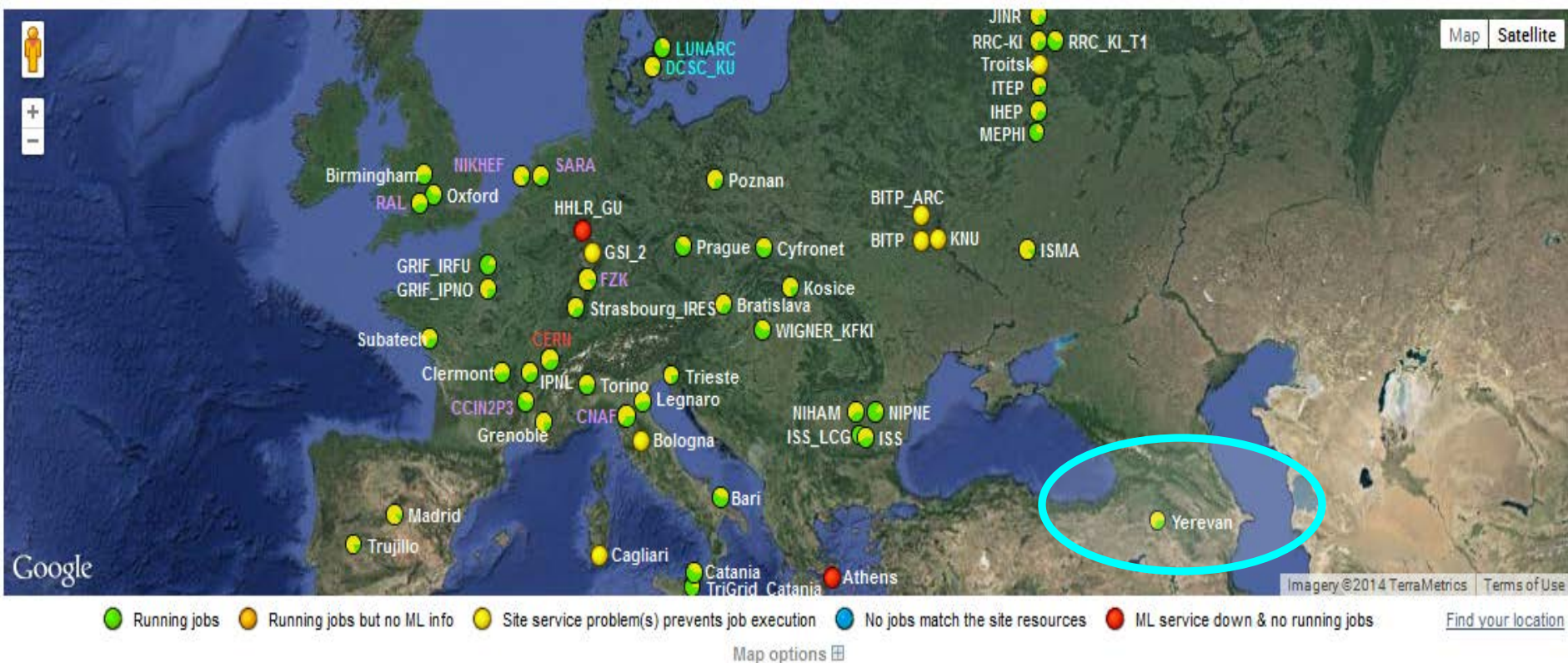
Master study student at SEUA

nmanukya@ { mail.yerphi.am, mail.cern.ch }

- **ANSL** site of **LHC** and **ALICE** Computing Grid
 - ❖ Deployment and 3.5 year of operation
- Work on the upgrade of **ALICE** computing Environment :
 - ❖ Upgrade of the **AliEn Package Manager Service**
 - ❖ Design and Development of the **AliEn File Access Monitoring Service**
- Publications and Presentations
- Participation in the **ALICE** shifts
- Current activity
- Student activity

Deployed at the end of 2010.

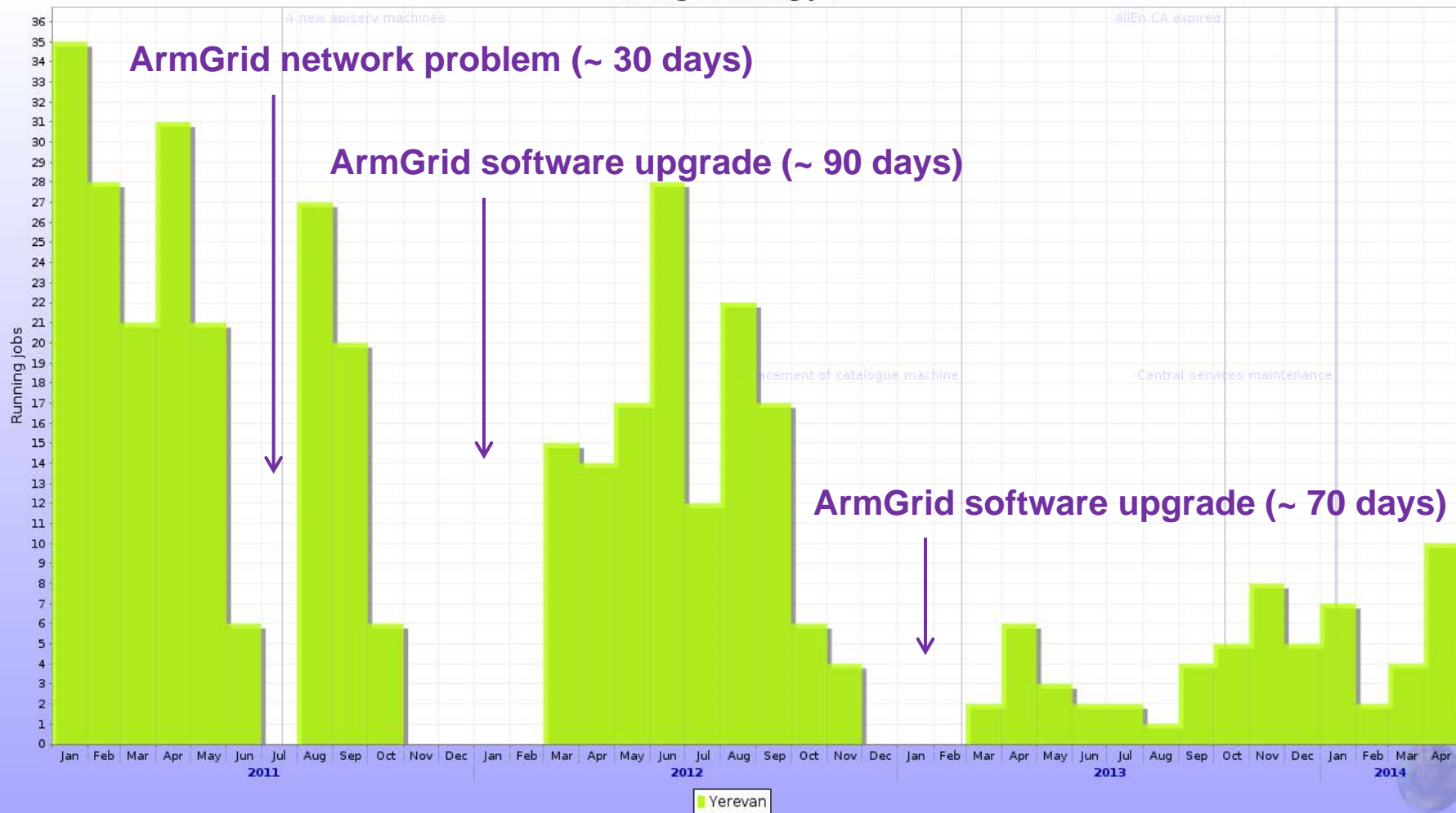
The site is used actively by **ALICE** members for the *Monte Carlo* simulations (the analysis jobs are not run on the site because the site has no storage capacity).



<http://alimonitor.cern.ch/map.jsp>

Regrettably, many (and long-lasting) outages happen that do not allow to speak about a high reliability performance of the site:

Average running jobs



Done jobs in Yerevan



Administration of the site consists of the following routine:

- Regular monitoring of **AliEn** services using **AliEn MonALISA** monitoring tool;
- Updates of the **AliEn** software installed on *VO-BOX* server;
- Maintenance of the server hardware (changing broken hardware parts, cleaning from dust, etc, etc);
- Solving with the **AliEn** administrators the *VO-BOX* – related problems

The Package Manager (**PackMan**) service of **AliEn**: Automatic installation, removal, upgrade and configuration of application software packages.

2011, 2012 –Upgrade of **PackMan**: An essential speed-up of the execution time of the commands of this service has been achieved.

The upgraded code has been implemented in the version 2.20 of AliEn software

File Access Monitoring Service –FAMoS (with A. Abramyan)

The **FAMoS** is to monitor the frequency of the accesses to the data files in the **AliEn** storage nodes which is one of the crucial parameters that we need to know in order to optimize the use of available storages.

The FAMoS has been deployed and put into exploitation on the central servers of AliEn on August 2013.

The detailed talk about these works has been presented by Armenuhi Abramyan (ANSL ExpDep seminar, 16th of April 2014)

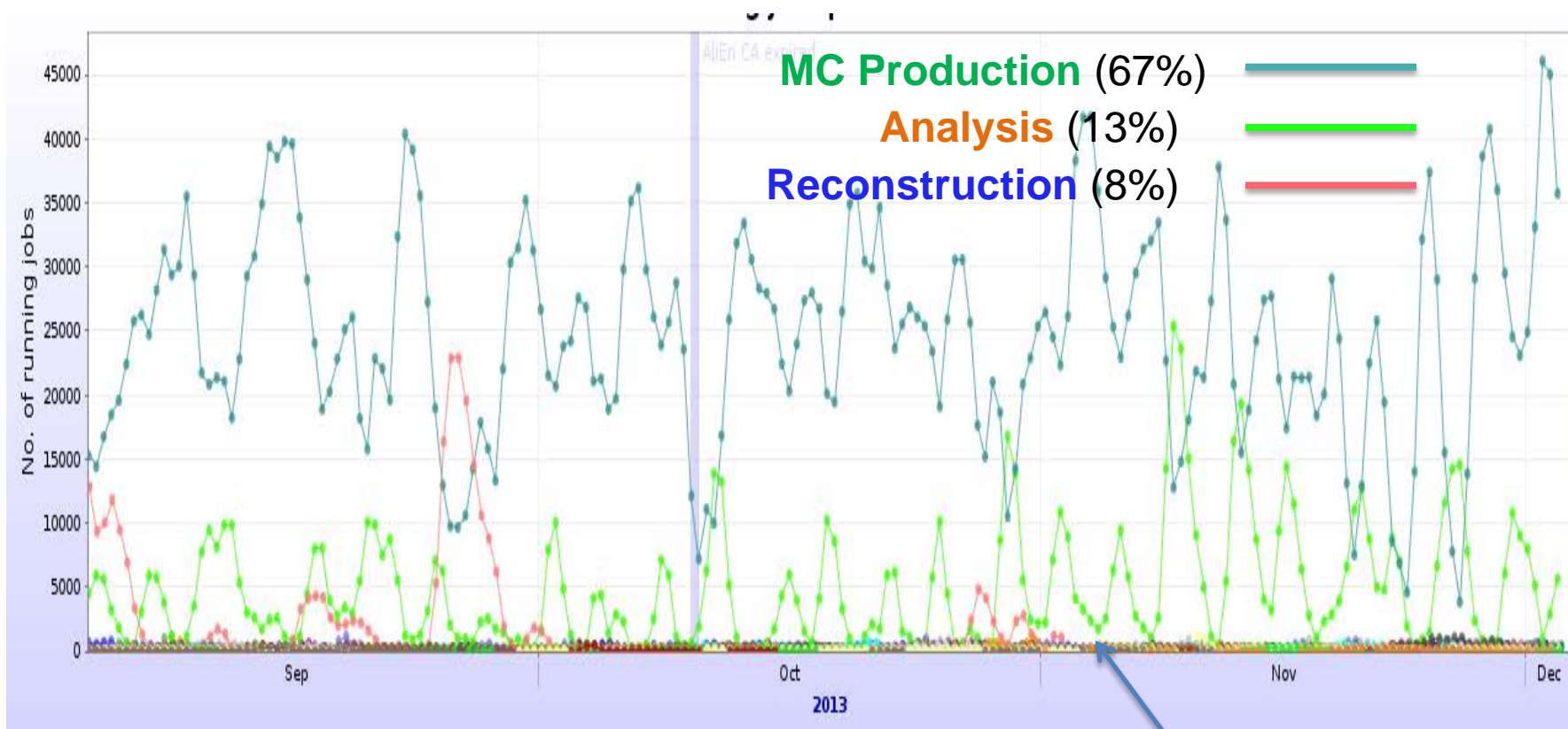
Regular presentation of the results of our studies and work by me and my colleague **Armenuhi Abramyan** at **ALICE Offline meetings**.

Authors	#	Reference
N. Manukyan	1	"ANSL site of LHC and ALICE Computing Grids. Deployment and Operation" . Proc. Int. Conf. "Computer Science and Information Technologies", Yerevan, pp. 374-376, 2011.
N. Manukyan, A. Abramyan	1	"AliEn File Access Monitoring Service – FAMoS" . Proc. 9th Int. Conf. "Computer Science and Information Technologies", Yerevan, Armenia, pp. 352-355.
N. Manukyan, A. Abramyan and 11 members of ALICE Offline software team	3	"AliEn Extreme JobBrokering" . Proceedings of Computing in High Energy and Nuclear Physics (CHEP) conference, New York City, NY, USA (2012)
		"ALICE Environment on the GRID" . Proceedings of Computing in High Energy and Nuclear Physics (CHEP) conference, New York City, NY, USA (2012)
		"Creating a simplified global unique file catalogue" . 20th International Conference on Computing in High Energy and Nuclear Physics (CHEP2013), Amsterdam, Beurs van Berlage (2013)
N. Manukyan, et al. (ALICE Collaboration)	2	...

Participation in ALICE shifts

(2011) - 6 (night) shifts have been served for the **Data Acquisition System (DAQ)**

Current activity – a very hot problem in AliEn



Individual user (12%)



The reduction of execution time of ALICE analysis jobs

~ **30%** of analysis jobs read experimental data and ~ **70%** read MC data

The execution time for **analysis** jobs ~ **4.5 hours**, from half of which is time spent on the waiting in the queue.

Waiting time for analysis jobs is too large!

A sensitive reduction of the whole time of the analysis jobs are possible only through the reduction of the time spent on the waiting in the queue

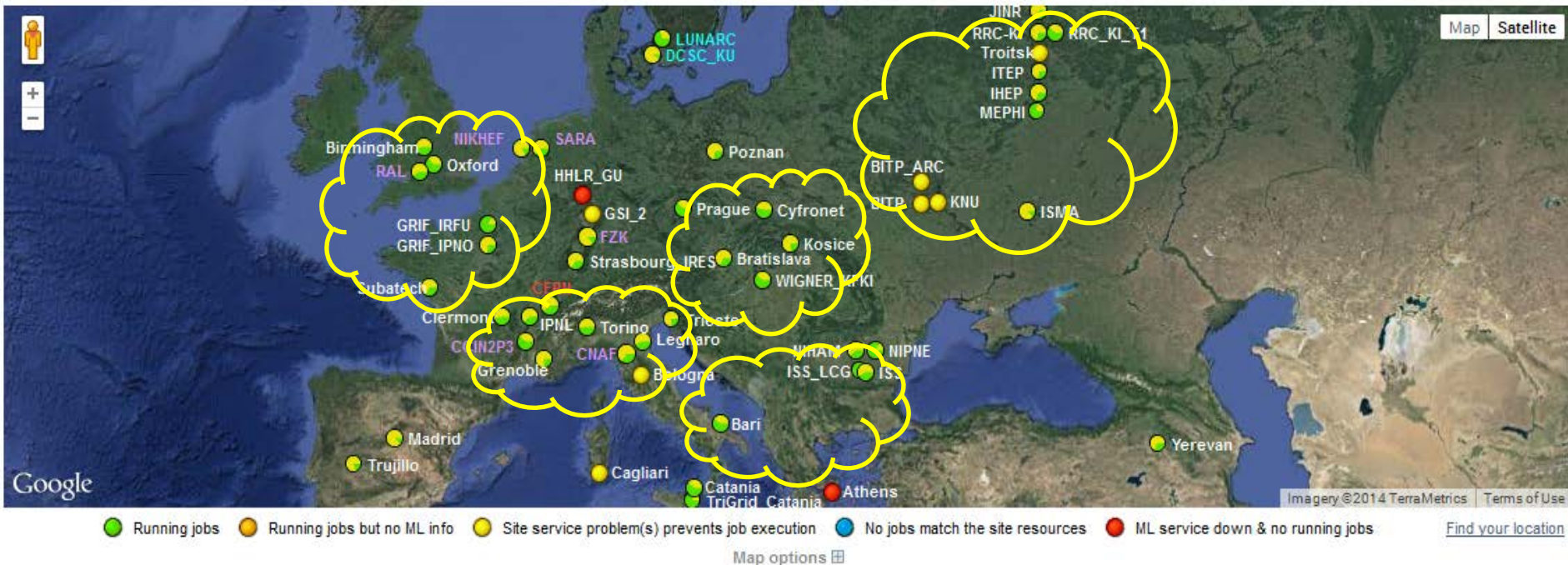
Why waiting time is large

- Almost all jobs are executed on the sites which contain the *InputData* required by these jobs;
- **Jobs are waiting in the queue for the site where the input data for these jobs are located (if that site is not available the jobs can wait for very long time)**

New strategies have to be considered

Reduce the number of sites on which a given simulation and analysis jobs are going to be scheduled for execution.

By creating the clouds of the sites from actual set of AliEn sites



Idea

- Schedule the running of the simulation jobs on the sites of the one of these clouds;
- Keep the output of the simulation result in the storages of the sites where simulation jobs were running
- Schedule the execution of the Analysis jobs on the aforementioned clouds of the site (where the input data are located).

Expectation

This strategy is supposed to assure the minimization of data flow and to constraint the subsequent steps of processing the analysis jobs to a fewer number of sites while maintaining the sufficient number of file replicas that guarantee overall dataset availability even if certain number of sites become unavailable.

Work to be done

The cloud creation algorithm should be elaborated. Simulations within **Discrete Event Simulation** approach are in progress

*Bachelor diploma thesis “**Regulation of the flow of files in the Grids**” has been defended **magna cum laude** at the end of **May 2013***

Master thesis devoted to ALICE computing environment is in progress

THANK YOU



Backup Slides

Job's execution average time

User	Job execution total time (hour)
alidaq	7.94
alitrain	4.375
aliproduct	24.357
individual user	11.397

alitrain job's status transition average times

State transition	alitrain hour
<i>Inserting - Waiting-Assigned</i>	1.71
<i>Assigned-Started</i>	0.73
<i>Started-Running</i>	0.01
<i>Running-Saving</i>	1.84
<i>Saving-Saved</i>	0.08
<i>Saved-Done</i>	0.005

Waiting time is too large for alitrain jobs!