

# ISS Grid activities - status report and future plans

**Liviu IRIMIA**, Ionel STAN, Adrian SEVCENCO

Institute of Space Science  
P.O. Box: MG-23, RO 077125, Bucharest-Magurele ROMANIA  
<http://www.space-science.ro>

RO-LCG 2018 “Grid, Cloud and High-Performance Computing in Science”  
17-19 October 2018

# ISS Computing

2

Cluster	Number of servers	Cores
ISS-ALICE	49	1216
RO-13-ISS	8	128
PlanckGrid	16	144
RoSpaceGrid	50	784
Total	123	2272

RO-LCG 2018 “Grid, Cloud and High-Performance Computing in Science”  
17-19 October 2018

# New ISS Computing Infrastructure

3

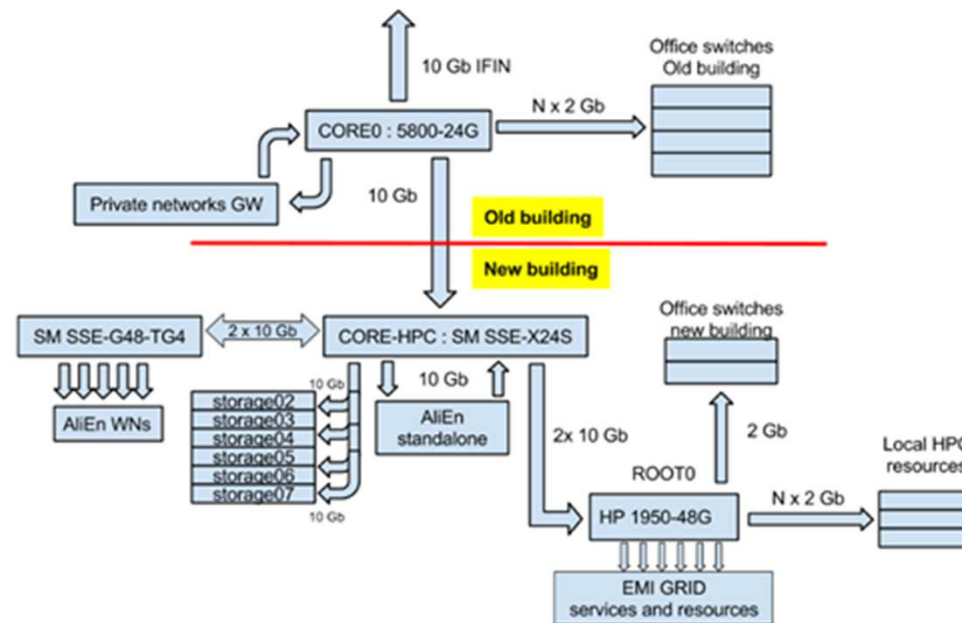


- Designed for high density computing (Hot Aisle, InRow cooling)
- Scalable solution for future investments
- UPS Power : 48 kVA (with N+1 redundancy power units)
- Cooling capacity : 80 kW installed (2N capacity redundancy)

# HARDWARE AND TOPOLOGY OF COMPUTING FACILITY

4

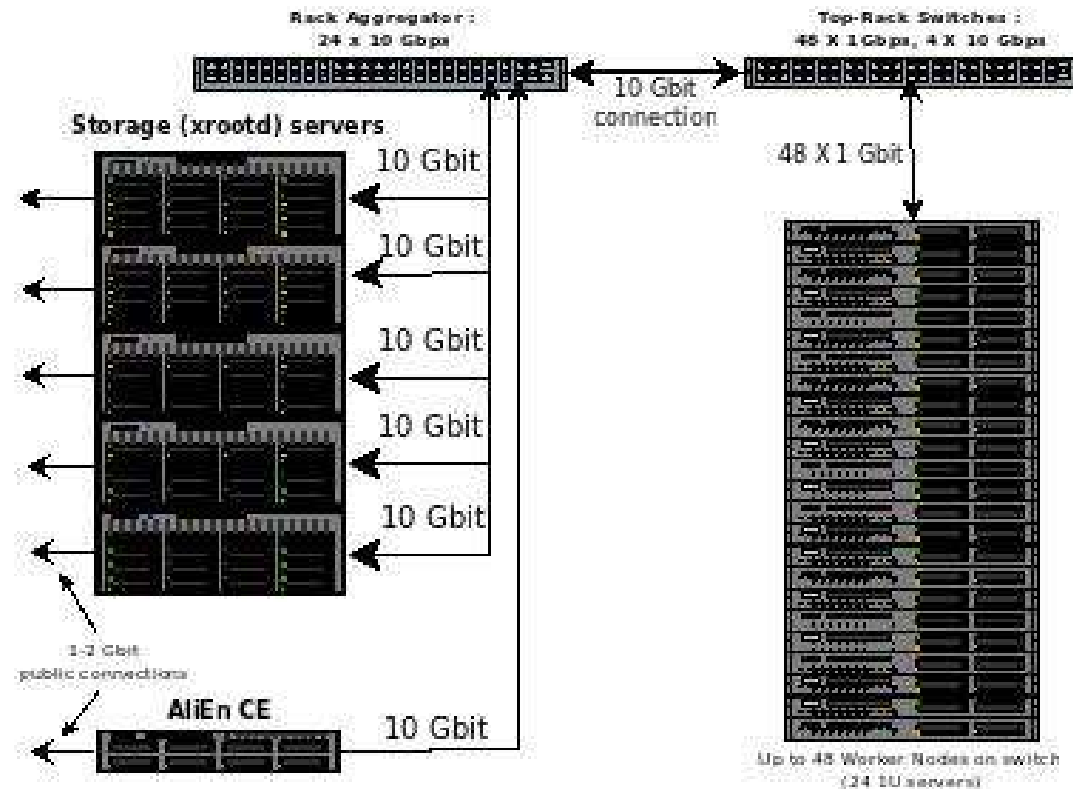
- Our hardware is mainly comprised of SuperMicro machines that were chosen for the great resource density/price ratio. For computing nodes we use Twin servers which give us densities of 4 sockets/1U and for the storage we use servers with 24, 36 drives and JBOD cases with 45 drives in 4U of rack space.
- Generic schematic of ISS computing facility :



# HARDWARE AND TOPOLOGY OF COMPUTING FACILITY

5

- The AliEn cluster has at his core a 10 Gbps aggregating switch which is connected to the top-of-rack switch of the computing nodes. In the aggregating switch are connected the private/public interfaces of the storage node, a topology which give a high bandwidth connection between worker nodes and storage with very little oversubscribing.



# CLUSTERING TECHNOLOGIES IN ISS DATA CENTER

6

## Rocks Clusters

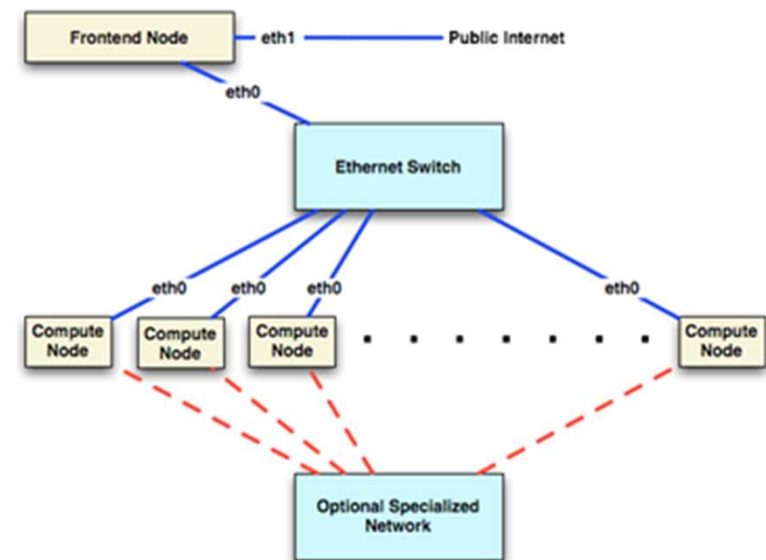
- was chosen due to ease of deployment, management, maintenance and flexibility , facile extensibility of software packages via software entities known as “Rolls”.
- is more than a middleware. It is a system composed of open source operating system that is based on a Red Hat flavor - CentOS (which comes with a modified interface for Anaconda for setup of cluster parameters and services) and a database with a corresponding set of scripts that simplify the process of node deployment, authorization of services and cluster monitoring.
- Installation can be customized by using so-called software Rolls. These Rolls expand system capacity by integrating effortlessly and automatically software packages and management mechanisms used by the Rocks administration tools, simplifying installation and configuration over a large number of computing nodes. Rocks Clusters brings a large number of rolls, such as: Torque, SGE, Condor, Ganglia, Java, HPC, etc.

# CLUSTERING TECHNOLOGIES IN ISS DATA CENTER

7

## Rocks Clusters

- used in more than 1800 clusters worldwide
- the model used is the server (frontend) - client model in which the central server will contain all necessary central clustering services that will serve a number of clients types like the computing nodes (main computing resources) and the login nodes – the nodes for user interaction with the cluster.
- On the computing nodes the first Ethernet interface will be used for the private network but Rocks have the possibility of setting up over a large cluster of nodes of other various private network used for dedicated usage like high speed networking or special storage sharing (10 Gbps, Infiniband).
- Rocks use the well-known and established tools for clustering, “Maui” and “TORQUE”.
- “Maui” - We use a priority based node allocating policy that would give priority to the most unoccupied servers
- “TORQUE” - We use several queues for the jobs to be submitted with restrictions in place for memory consumption (3584 MB resident memory and 4096B virtual memory) and wall time (36 hours).



# GRID SERVICES AND MIDDLEWARE DEPLOYED IN DATA CENTER

## AliEn – Alice Environment

- is a lightweight Grid framework which is built around Open Source components using the Web Services model. It has been initially developed by the ALICE collaboration (ALICE Offline Project) as a production environment for the simulation, reconstruction, and analysis of physics data in a distributed way.

## UMD/EMI –Unified Middleware Distribution/European Middleware Initiative

- is a Service Oriented Grid middleware providing services for managing distributed computing and storage resources and the required security, auditing and information services. Target server platform is Red Hat Linux or any binary compatible distribution, such as SL and CentOS
- is a close collaboration of three major middleware providers, ARC, gLite and UNICORE, and other specialised software providers like dCache.
- The products, managed in the past by these separate providers, and now developed, built and tested in collaboration, are for deployment in EGI as part of the Unified Middleware Distribution - UMD



# UMDUI - User interactive tool

9

Umdui - ISS Proxy LCG Info Manage Jobs File Management

## Proxy Management

Grid-proxy-init

Grid-proxy-info

Grid-voms-info

Grid-voms-init

Grid-proxy-destroy

### Command output:

```
subject : /DC=RO/DC=RomanianGRID/O=ISS/CN=Liviu IRIMIA/CN=1977948870
issuer  : /DC=RO/DC=RomanianGRID/O=ISS/CN=Liviu IRIMIA
identity : /DC=RO/DC=RomanianGRID/O=ISS/CN=Liviu IRIMIA
type   : RFC 3820 compliant impersonation proxy
strength : 1024 bits
path   : /tmp/x509up_u501
timeleft : 0:00:00
```

```
subject : /DC=RO/DC=RomanianGRID/O=ISS/CN=Liviu IRIMIA/CN=1977948870
issuer  : /DC=RO/DC=RomanianGRID/O=ISS/CN=Liviu IRIMIA
identity : /DC=RO/DC=RomanianGRID/O=ISS/CN=Liviu IRIMIA
type   : RFC3820 compliant impersonation proxy
strength : 1024
path   : /tmp/x509up_u501
timeleft : 00:00:00
key usage : Digital Signature, Key Encipherment, Data Encipherment
```

### Grid-proxy-init dialog

Certificate Password:

Close

Execute

# UMDUI - User interactive tool

10

Umdui - ISS Proxy LCG Info Manage Jobs File Management

## General informations

LCG-info

LCG-infosites

### Command output:

Avail Space(kB)	Used Space(kB)	Type	SE
1968384708	628175	SRM	grid02.spacescience.ro
88709068147	30812943865	SRM	seau.spacescience.ro

## Infosites Dialog

VO: dteam

Filter: Filter(Optional)...

Site

Verbose

CE

SE

Space

closeSE

tag

lfc

wms

site names

Close

Execute

# Future applications - Monitoring tools

11

	ISS-CORE-0	ISSAF	ProCurve2810-48G_-_CORE1	ProCurve2810-48G_-_CORE2	ProCu		
	Description	Admin status	Operation status	In discards	In errors	Out discards	Out errors
1		up	up	0	6	164866	0
2		up	down	0	0	0	0
3		up	down	0	0	0	0
4		up	down	0	0	0	0
5		up	down	0	0	160	0
6		up	down	0	0	0	0
7		up	up	0	1	6739	0
8		up	down	0	0	0	0
9		up	down	0	9	54287	0
10		up	down	0	0	0	0
11		up	down	0	0	0	0
12		up	up	0	0	160305	0
13		up	down	0	0	0	0
14		up	down	0	0	0	0
15		up	down	0	0	0	0
16		up	down	0	1	22251	0
17		up	up	0	98	170637	0
18		up	down	0	0	0	0
19		up	up	0	1	170646	0
20		up	up	0	0	159117	0
21		up	up	0	0	167051	0

# Future applications - Monitoring tools

12

Number	Name	Disks	State	Total Capacity	Free capacity	HDD capacity	HDD channels
ARC-1280							
0	"raid0"	12	"Normal"	12000000	0	1000000	"1.2.3.4.5.6.7.8.9.10.11.12"
1	"raid1"	12	"Normal"	12000000	0	1000000	"13.14.15.16.17.18.19.20.21.22.23.24"

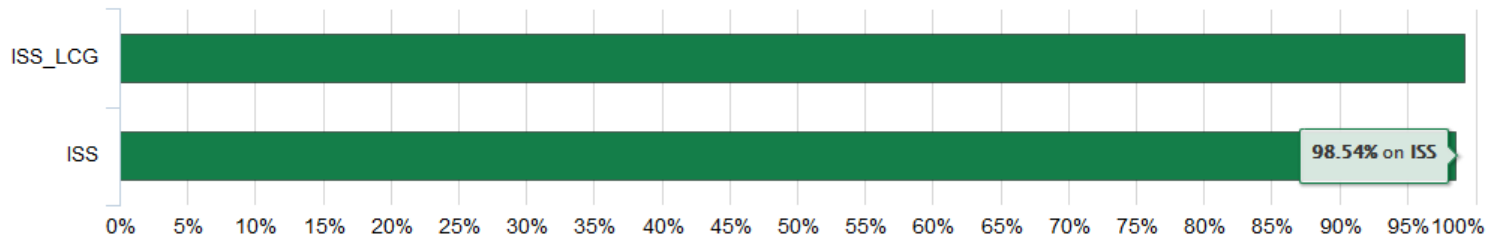
# Global Performance

13

## SITE AVAILABILITY

### Site Availability using ALICE\_MON\_CRITICAL

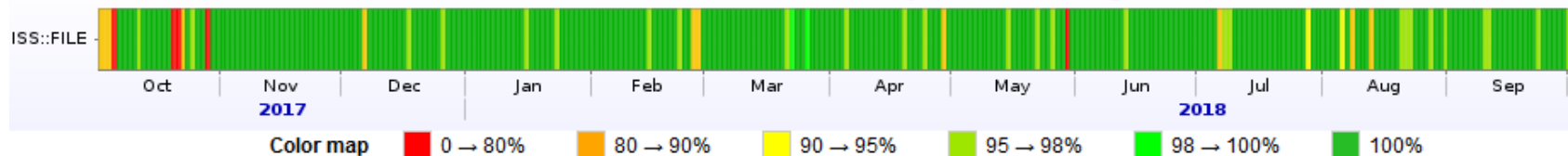
From 2017/10/01 to 2018/10/01



<http://wlcg-sam-alice.cern.ch>

## STORAGE AVAILABILITY

### AliEn SEs availability for reading



Statistics						
Link name	Data		Individual results of reading tests			Overall
	Starts	Ends	Successful	Failed	Success ratio	Availability
ISS::FILE	01 Oct 2017 23:09	01 Oct 2018 22:23	8592	159	98.18%	98.22%

<http://alimonitor.cern.ch>

# Statistics per site - RO-13-ISS

14

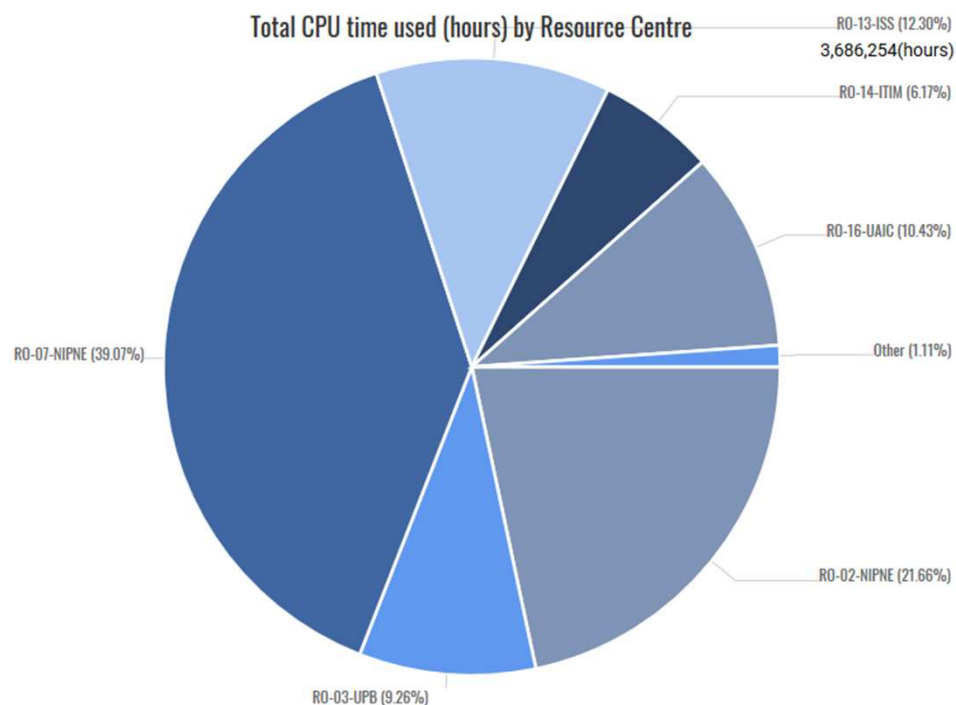
## Done jobs

Resource Centre RO-13-ISS — Total number of jobs by VO and Resource Centre (LHC VOs)

VO	RO-13-ISS	Total	Percent
alice	1,047,437	1,047,437	100%
<b>Total</b>	<b>1,047,437</b>	<b>1,047,437</b>	
<b>Percent</b>	<b>100.00%</b>		

1 - 1 of 1 results < 1 > Number of rows per page 30

## Total CPU time



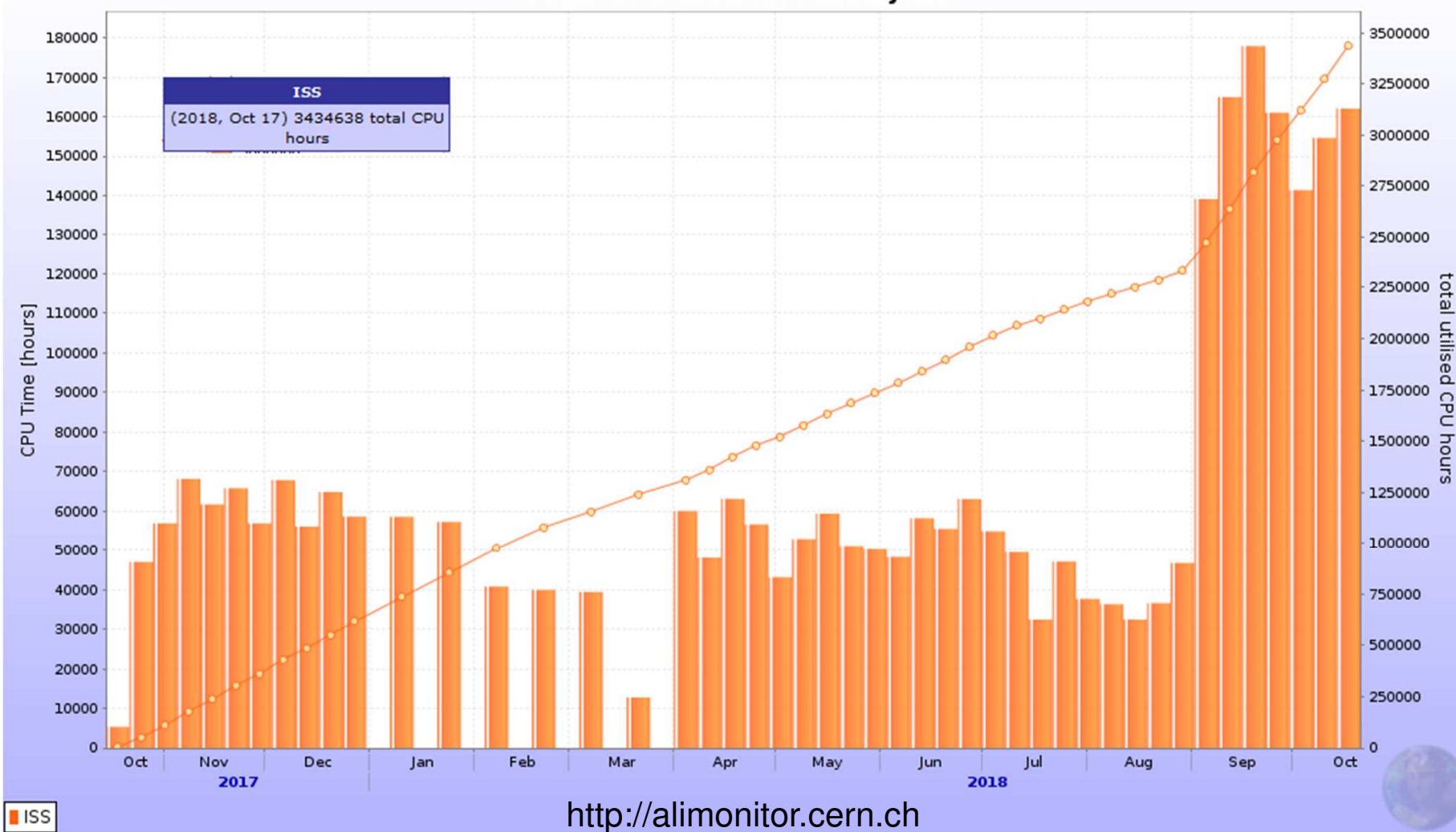
<http://accounting-next.egi.eu>



# Statistics per site - ISS-Alice

16

### Total CPU time for ALICE jobs

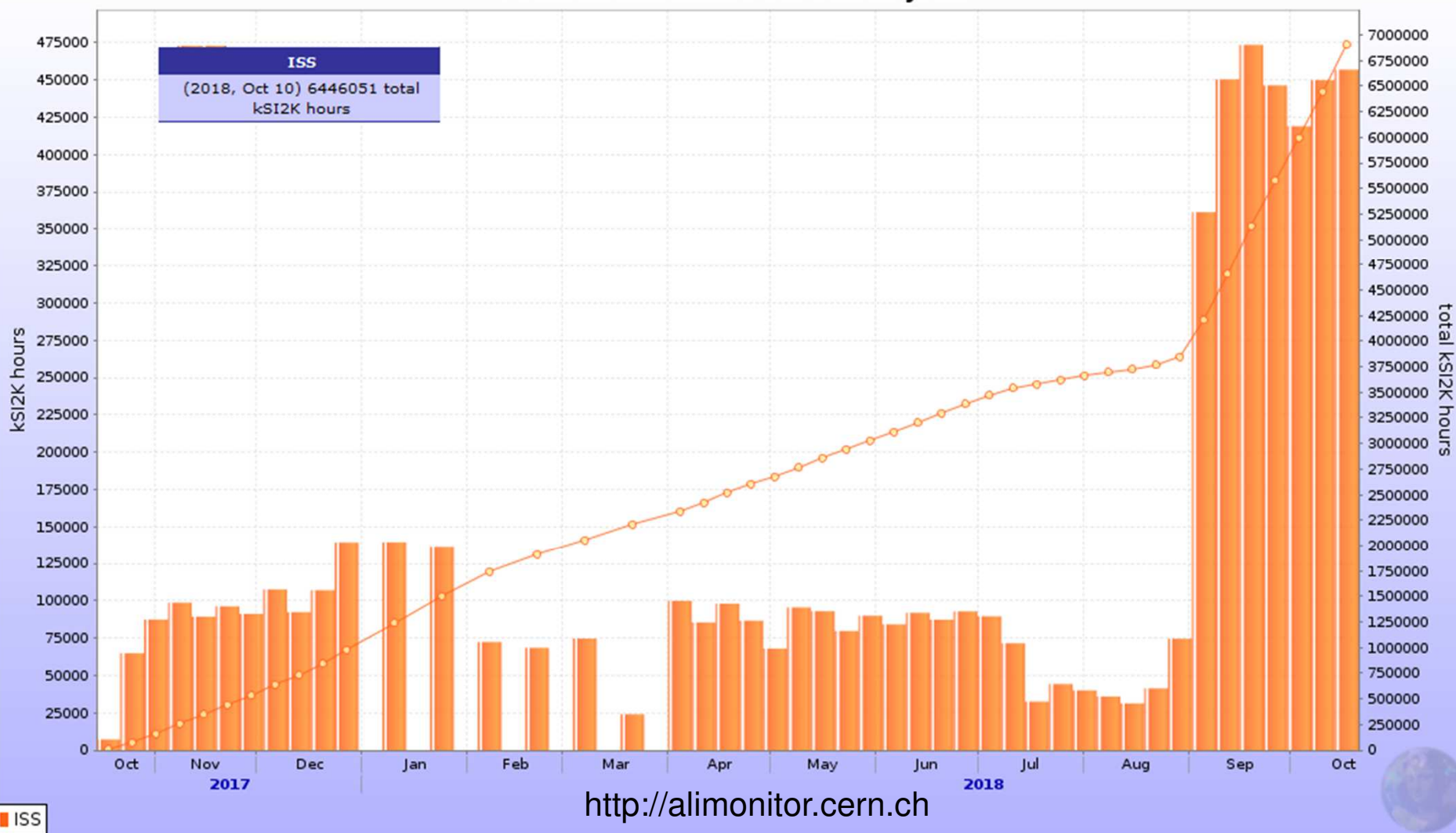




# Statistics per site - ISS-Alice

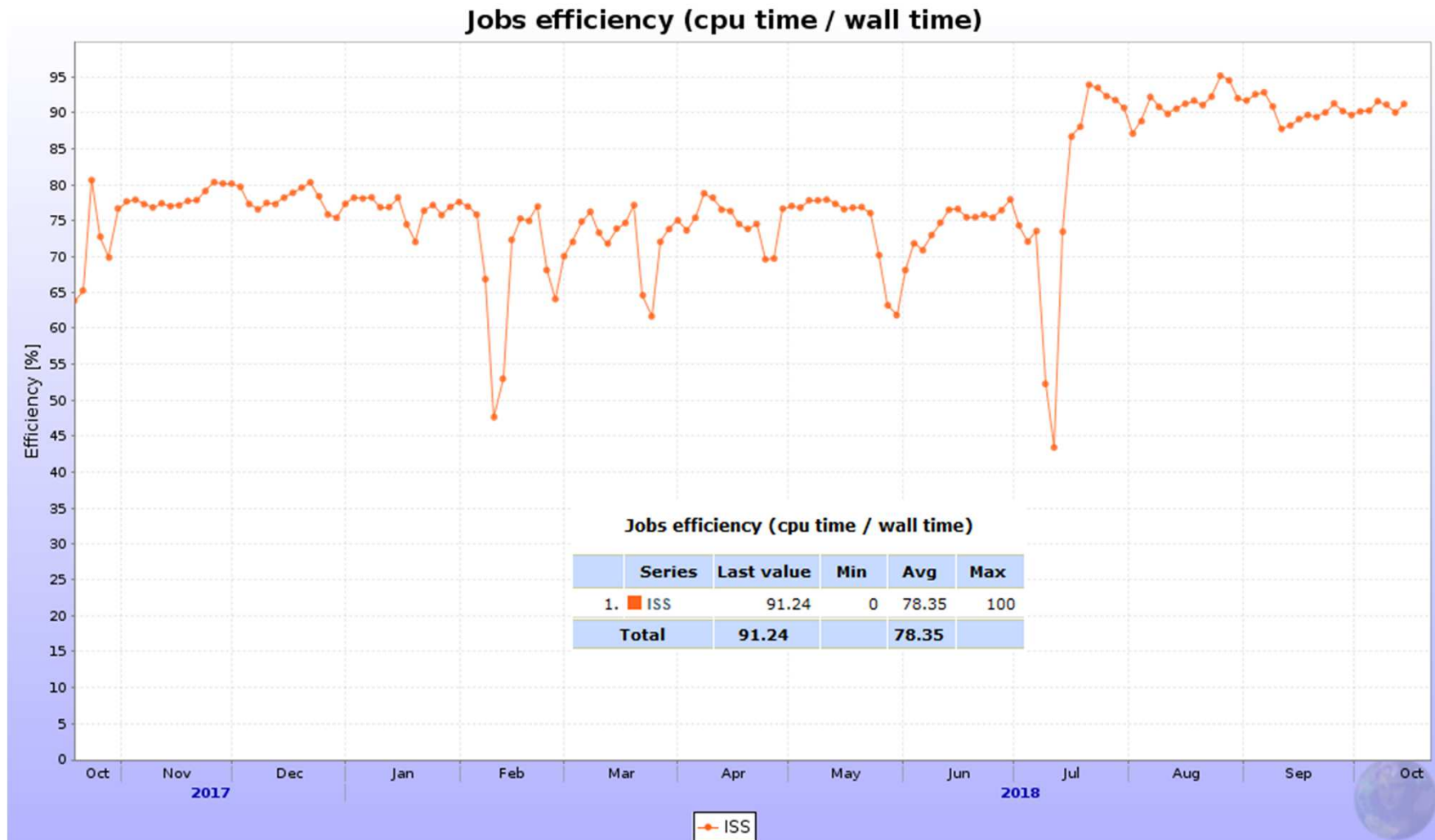
17

## Total kSI2K hours for ALICE jobs



# Statistics per site - ISS-Alice

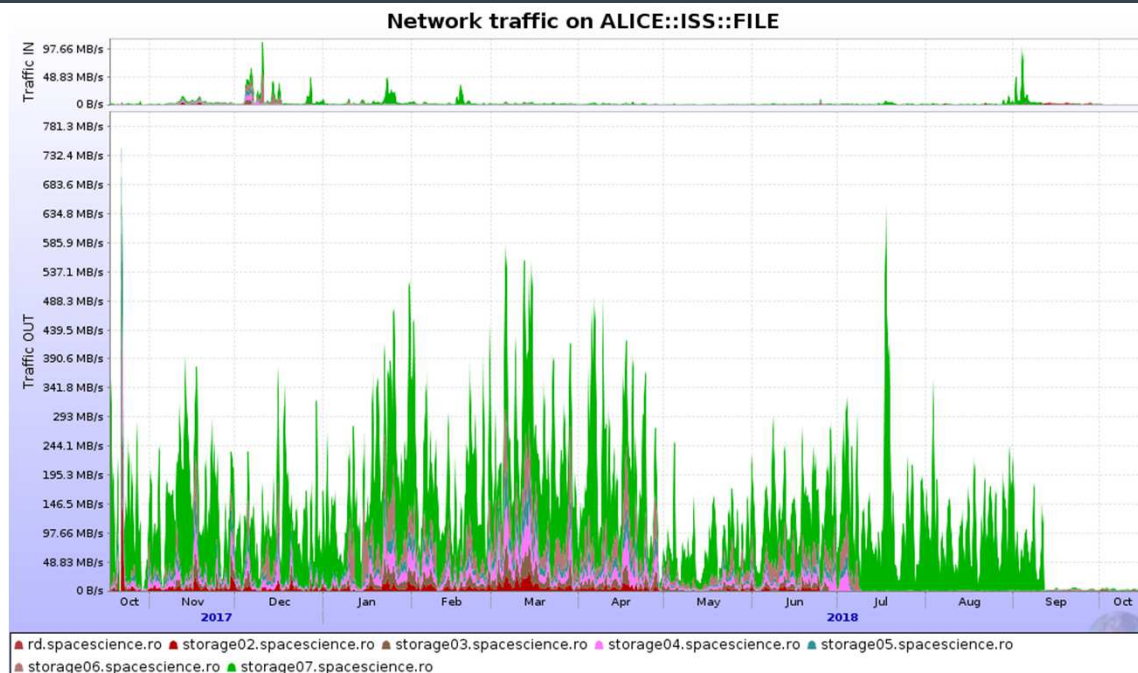
18



<http://alimonitor.cern.ch>

# Storage performance

19



Traffic IN

	Series	Last value	Min	Avg	Max	Total
1.	rd.spacescience.ro	15.49 KB/s	1.152 KB/s	16.93 KB/s	4.909 MB/s	508.2 GB
2.	storage02.spacescience.ro	3.682 KB/s	5.146 B/s	182.5 KB/s	62.87 MB/s	5.348 TB
3.	storage03.spacescience.ro	4.009 KB/s	5.406 B/s	218.2 KB/s	169.6 MB/s	6.395 TB
4.	storage04.spacescience.ro	12.07 B/s	5.152 B/s	216.1 KB/s	187.1 MB/s	6.334 TB
5.	storage05.spacescience.ro	2.471 KB/s	5.413 B/s	190.5 KB/s	170.9 MB/s	5.582 TB
6.	storage06.spacescience.ro	1.358 KB/s	2.978 B/s	608 KB/s	769 MB/s	17.82 TB
7.	storage07.spacescience.ro	6.2 KB/s	35.06 B/s	2.554 MB/s	1.129 GB/s	76.64 TB
<b>Total</b>		<b>33.22 KB/s</b>		<b>3.952 MB/s</b>		<b>118.6 TB</b>

Traffic OUT

	Series	Last value	Min	Avg	Max	Total
1.	rd.spacescience.ro	5.611 KB/s	1.164 KB/s	6.538 KB/s	349.1 KB/s	196.2 GB
2.	storage02.spacescience.ro	473.3 KB/s	1.062 B/s	4.222 MB/s	299.1 MB/s	126.7 TB
3.	storage03.spacescience.ro	615.8 KB/s	1.063 B/s	6.541 MB/s	291.7 MB/s	196.3 TB
4.	storage04.spacescience.ro	26.99 B/s	1.074 B/s	9.464 MB/s	380.1 MB/s	284 TB
5.	storage05.spacescience.ro	476.8 KB/s	1.074 B/s	8.484 MB/s	423.4 MB/s	254.6 TB
6.	storage06.spacescience.ro	277.4 KB/s	1.296 B/s	16.92 MB/s	980.4 MB/s	507.6 TB
7.	storage07.spacescience.ro	1003 KB/s	2.268 B/s	105.4 MB/s	1.509 GB/s	3.089 PB
<b>Total</b>		<b>2.786 MB/s</b>		<b>151 MB/s</b>		<b>4.426 PB</b>

<http://alimonitor.cern.ch>

# Statistics - summary

20

- High computing contribution/cost ratio
- High utilization by ALICE user analyses (over 633k done jobs)
- 78.35% Jobs efficiency (cpu time/wall time) over last year (76.68% average of ALICE GRID)
- High availability 98.4%, High success ratio 98.4% of our storages
- Over 4.5 PB data transfer in the last year, 93% storage space occupied

# Problems encountered

- Because of the aging hardware we encounter many stuck nodes due to insufficient memory relative to the actual needs of current software for analysis and reconstruction.
- The Information Systems and monitoring tools are fragmented and poorly documented and as such is difficult to decouple and interface with the said modules for providing complete and exact hardware and jobs status information.

# Future plans

- Implementation of IPv6 (until the end of this year)
- Increase network connection from 10 to minimum 20 Gb/s
- With the new data taking regime of LHC, the storage demands have greatly increased and starting since last year we replaced, within the limits of available funds, our old storage hardware with new resources, in order to fulfill our pledges and to offer stable computing services to the experiments in which we participate.
- For the next years we are planning to replace/upgrade our old hardware with new computing and storage resources to meet the demands of the new run periods of LHC experiments (RUN2, RUN3).

Thank you for your attention!