



Monitoring the ALICE Grid with MonALISA

Costin Grigoras

2008-08-20

ALICE Workshop @ Sibiu

Monitoring the ALICE Grid with MonALISA

- MonALISA Framework
- ApMon library
- Data collection and storage in ALICE
- Visualization methods
- Process automation
- Tools
- Site monitoring
- Summary

MonALISA Framework

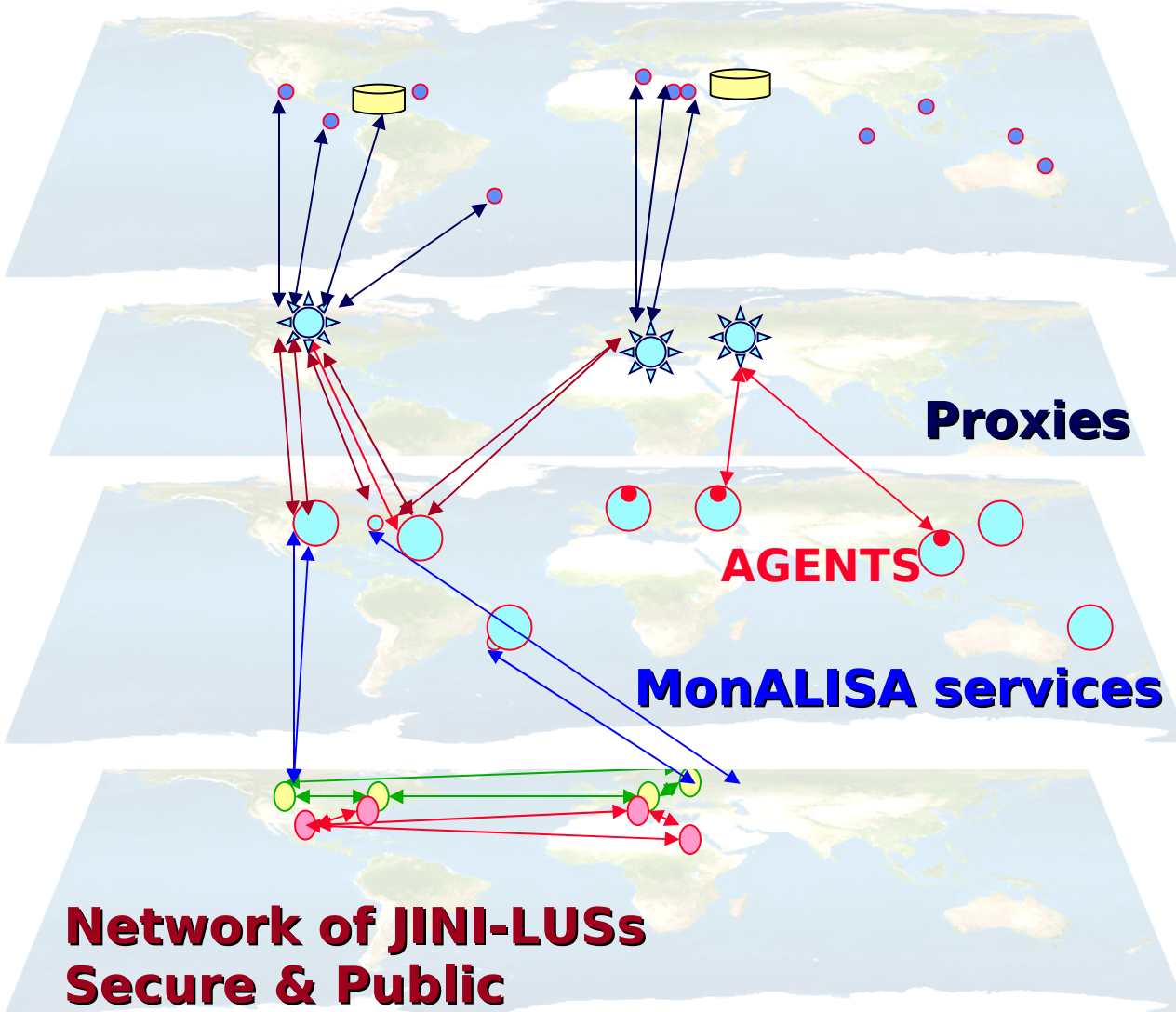


- MonALISA is a distributed service able to:
 - collect any type of information from different systems,
 - analyze this information in near real time,
 - take automated decisions,
 - optimize workflows in complex environments.
- The system is made of autonomous, self-describing agents that:
 - register themselves in global directories,
 - discover other agents and
 - cooperate with them to perform monitoring tasks.
- Project was started by Caltech and developed in collaboration with Politehnica University of Bucharest and CERN.

Fully distributed system with no single point of failure



Clients, repositories, other HL services



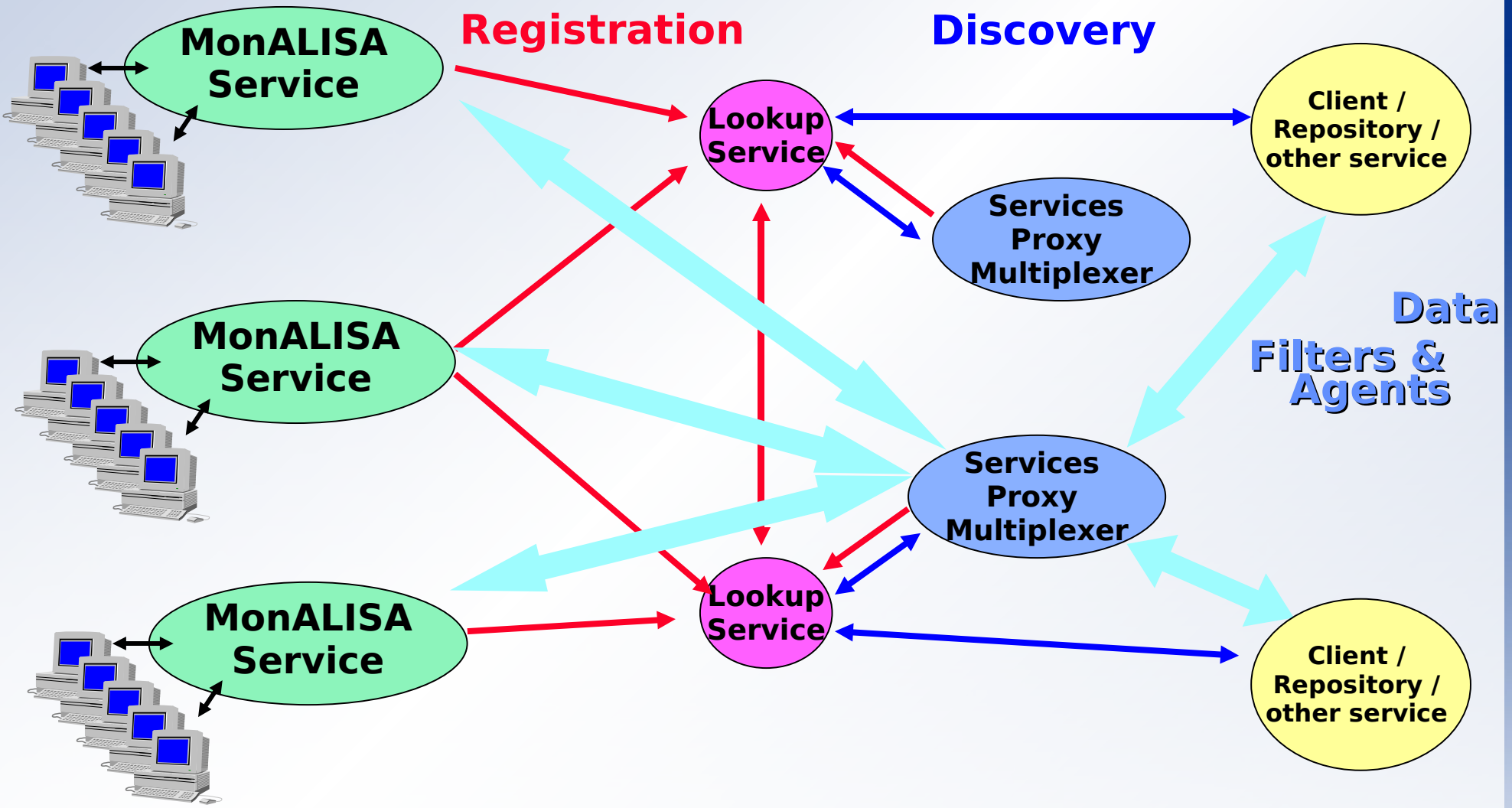
Global Services or Clients

**Dynamic load balancing
Scalability & Replication
Security AAA for Clients**

**Distributed System
for gathering and
Analyzing Information.**

**Distributed Dynamic
Discovery based on a
lease Mechanism**

MonALISA discovery mechanism

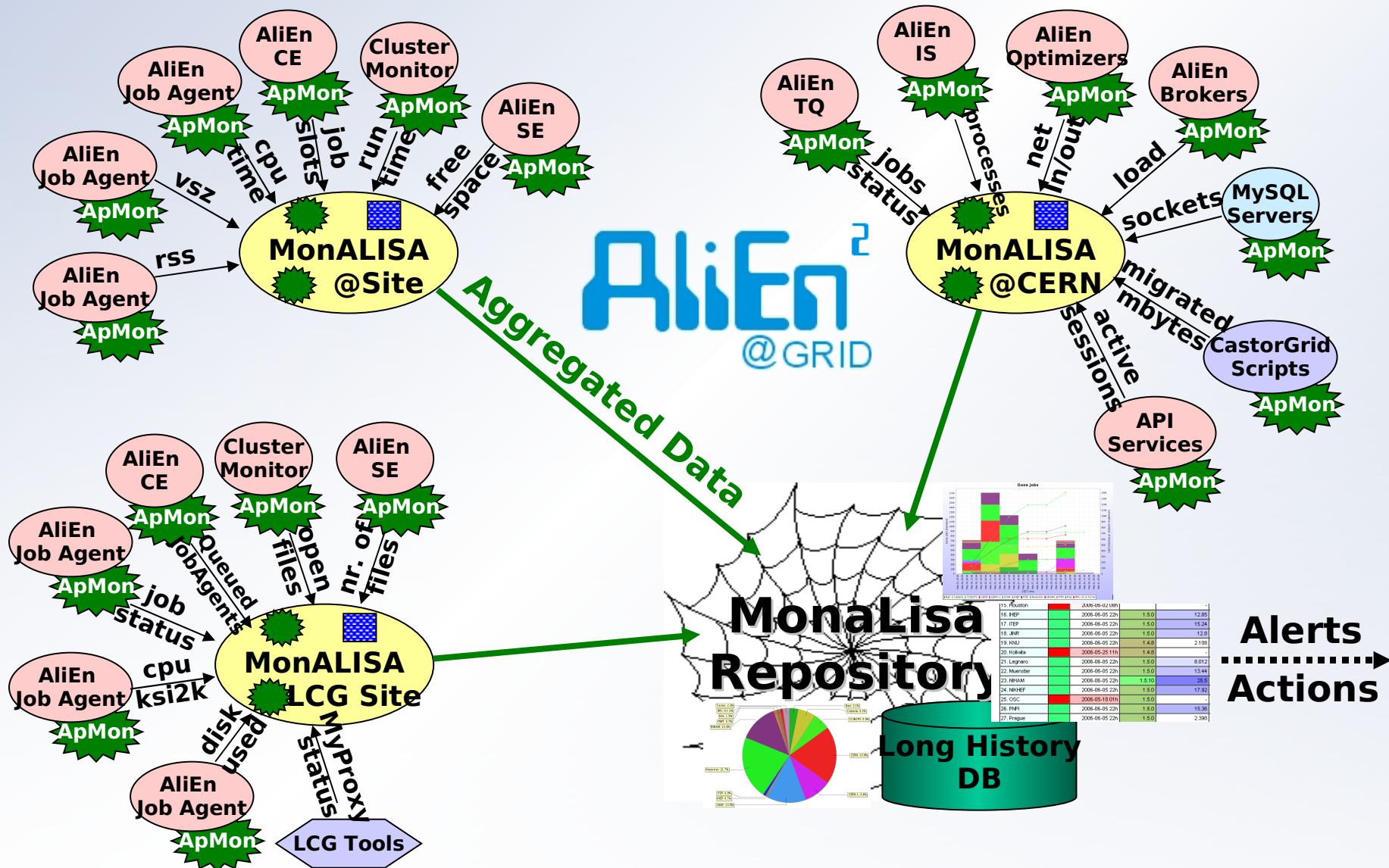


ApMon library



- ApMon is a lightweight library of APIs (C, C++, Java, Perl, Python) that can be used to send any monitoring information to MonALISA services.
- Uses UDP packets in XDR format for low footprint on application and network traffic.
- Provides by default:
 - system-level monitoring for the host where it runs,
 - application monitoring and accounting.
- In ALICE it is embedded in AliEn services, Job Agents and xrdcp processes.
- Can also be run in stand-alone mode, making cluster monitoring very simple (CAF, GSI, Muenster).
- http://monalisa.cern.ch/monalisa__Download__ApMon.html

Data collection and storage in ALICE



Data collection and storage in ALICE



- MonALISA services and the repository share the same storage API:
 - in-memory buffer of the recent history, continuously adjusting its size depending on how much memory is allocated to the JVM and how it is used,
 - persistent storage in a database, PostgreSQL or MySQL (PG is embedded in the ML installation kit).
- In ALICE persistent storage is disabled for the site services, to minimize the impact on the VoBox operations:
 - the default 64MB allocated to JVM usually allow keeping ~200K values in memory (2h history on a small site / a few minutes on a big site),
 - memory size can be tuned at instance level.

Data collection and storage in ALICE



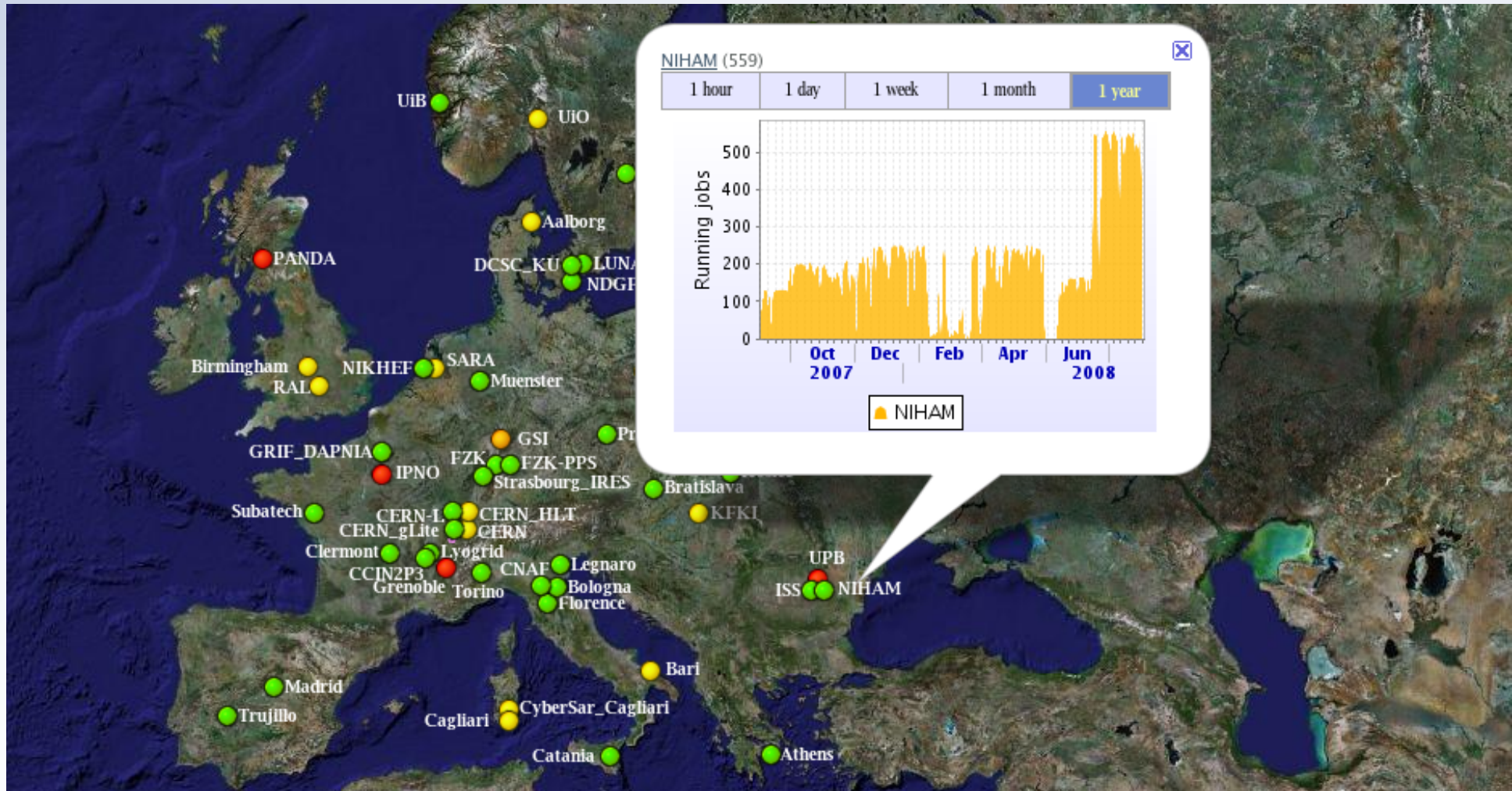
- 100 MonALISA services publish some 800K unique parameters with a total rate of 1KHz.
- Out of these $\sim 70K$ (raw and aggregated) time series are stored in the repository DB with a rate of 30Hz.
- New series can be defined on the fly, changes to the collection filters are applied right away without any service restart.
- We have accumulated 155GB of monitoring history data for more than 2 years of operation, using the following data reduction schema:
 - 2 minutes bins for the last 2 months
 - 30 minutes bins for the last 6 months
 - 2.5 hours bins for forever
- On average users are calling dynamic charts every 2-5 seconds.
- One machine is handling both the history database and the web interface, with very low load (0.4 avg on 24h, 91% CPU Idle time).

Visualization methods



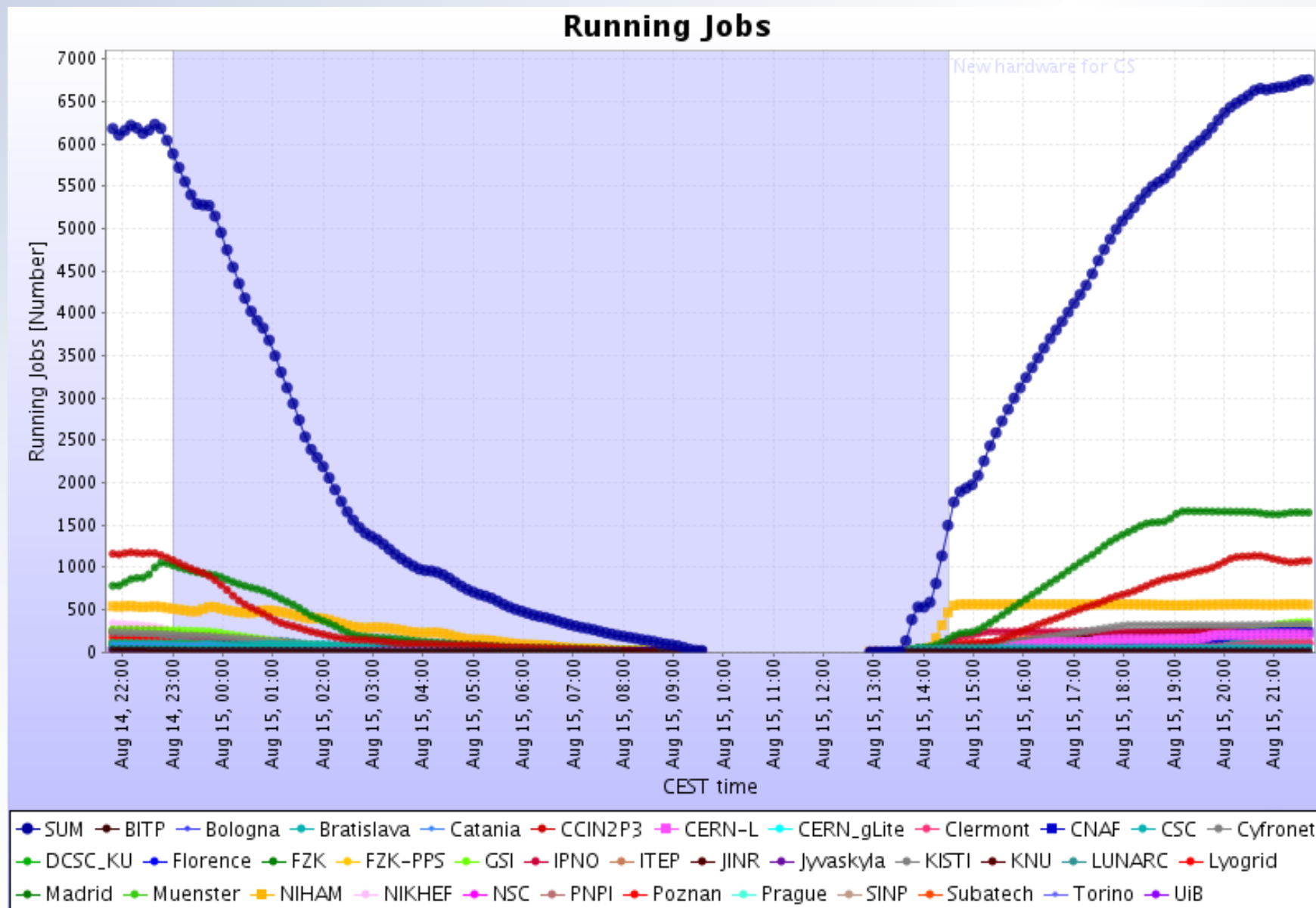
- Various types of charts to present history data and the current state of the Grid:
 - interactive map of all the sites,
 - general interest widgets in every displayed page,
 - generic charts based on configuration files to present time series as history of points, areas or bars, pie/bar/spider charts, scatter plots etc,
 - daily/weekly/monthly reports with resources' usage,
 - specialized pages for particular purposes:
 - overview of sites' status,
 - RAW data registration and analysis requests,
 - AliEn jobs status and history,
 - PWGs production requests,
 - other administrative pages.
- Users can dynamically define new charts that display any time series in the DB in either history or scatter charts.

Map view



<http://pcalimonitor.cern.ch/>

Default history view



Other views and chart types



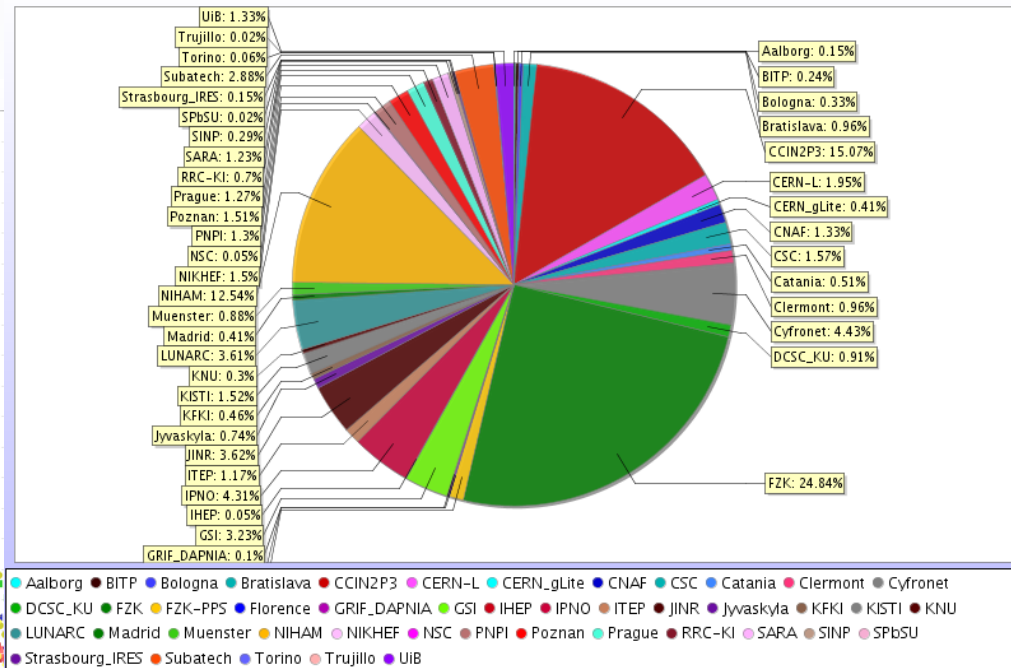
Proxies

Service	AliEn Tests		LCG Tests					
	AliEn proxy		Delegated proxy		Proxy Server		Proxy of the machine	
	Status	Time left	Status	Time left	Status	Time left	Status	Time left
1. Aalborg		1d 5:41	-	-	-	-	-	-
2. aliendb5.cern.ch		0min	-	-	-	-	-	-
3. Athens	Fail...	-	Cann...	0min	Fail...	0min		20:46
4. Bari		0min	Un...			0min		21:45
5. Birmingham		0min	Un...			0min		22:57

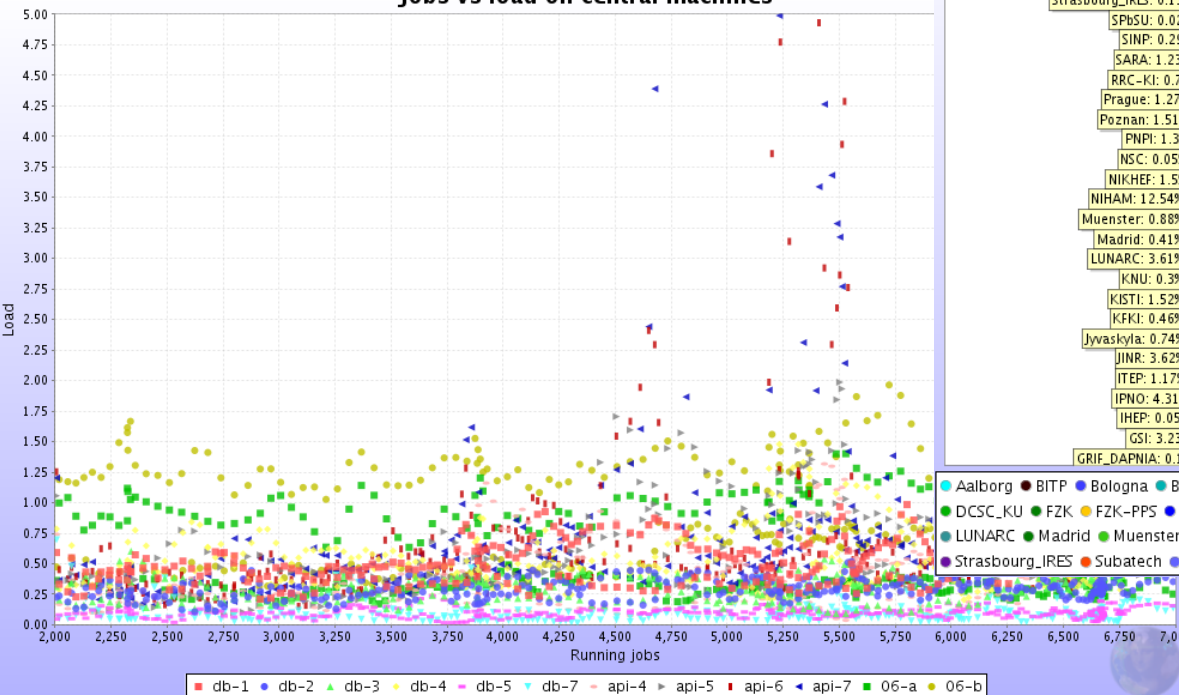
Click for more details

Cannot read USER_PROXY

Total CPU time for ALICE jobs [hours]



Jobs vs load on central machines



Specialized view: site summary

<http://pcalimonitor.cern.ch/siteinfo/>



Select site: »

MonALISA information Version: 1.8.6 (JDK 1.6.0_06)

Running on: hgate.nipne.ro

Administrator: Claudiu Schiaua <Claudiu.Schiaua@cern.ch>

Service health NTP: SYNC, offset: 0.013s

Services status

AliEn: v2-15.41

CE: OK

SE: OK

ClusterMonitor: OK

PackMan: OK

FTD: OK

Proxies status AliEn proxy: OK (28 days, 23:55)

Delegated proxy: n/a (n/a)

Proxy server: n/a (n/a)

Proxy of the machine: n/a (n/a)

SAM tests

Delegated proxy duration: ERROR

Proxy of the machine: OK

Proxy renewal: OK

Proxy server registration: OK

RB status: n/a

Software area: ERROR

User proxy registration: ERROR

WMS stats: WARNING

Make everything OK

Current jobs status

Assigned: 2

Running: 561

Saving: 1

Accounting
(last 24h)

Success jobs: 1395 (profile)

Failed jobs: 0

Error jobs: 2122

kSI2k units: 471 / 330 pledged

Site averages
(last 24h)

Active nodes: 77.39

Average kSI2k/node: 1.512

Storages status

Name	Status	Size	Used	Free	Usage	No of files	Type	ADD test
ALICE::NIHAM::File	OK	117.2 TB	8.166%	107.6 TB	9.57 TB	5.326 M	File	OK

VoBox health

CPUs: 4x 2992MHz
Mem usage: 36.28% of 1.979 GB
Processes: 150
Sockets: 371 TCP / 27 UDP
Uptime: 39 days, 10:14

CPU usage
(last 1h avg)

Load: 0.37
User: 2.252%
System: 0.71%
IOWait: 3.291%
Idle: 93.4%

Int: 0.004%
Soft int: 0.302%
Nice: 0%
Steal: 0%

AliEn LDAP var	VoBox path	Size	Used	Free	Use%
TMP	/scratch/alice32/tmp	49.24 GB	17.91 GB	28.79 GB	39%
LOG	/scratch/alice32/log	49.24 GB	17.91 GB	28.79 GB	39%
CACHE	/scratch/alice32/cache	49.24 GB	17.91 GB	28.79 GB	39%

RAW data registration



RAW Data Registration, Transferring and Processing

Login

Run#	Partition	Chunks	Total size	First seen	Last seen	Transfer status	Processing status	Staging status	
	- All -			- All -		Completed	Started	- All -	»
42371	LHC08b	620	566.5 GB	29 Jun 2008 23:42	30 Jun 2008 00:41	ALICE::GSI::SE	615 / 620 (99%)		
42365	LHC08b	96	85.78 GB	29 Jun 2008 23:38	30 Jun 2008 00:07	ALICE::GSI::SE	⚠ 92 / 96 (95%)		
42351	LHC08b	149	136 GB	29 Jun 2008 23:38	30 Jun 2008 00:00	ALICE::GSI::SE	⚠ 3 / 149 (2%)		
42275	LHC08b	11	5.947 GB	29 Jun 2008 05:37	29 Jun 2008 14:21	ALICE::GSI::SE	⚠ 9 / 11 (81%)		
42211	LHC08b	4	1.188 GB	29 Jun 2008 03:05	29 Jun 2008 03:05	ALICE::GSI::SE	⚠ 1 / 4 (25%)		
42207	LHC08b	4	498.5 MB	29 Jun 2008 03:05	29 Jun 2008 03:05	ALICE::GSI::SE	4 / 4 (100%)		
42192	LHC08b	4	1.943 GB	29 Jun 2008 03:04	29 Jun 2008 03:05	ALICE::GSI::SE	0 / 4 (0%)		
42188	LHC08b	15	5.383 GB	28 Jun 2008 01:44	29 Jun 2008 03:05	ALICE::GSI::SE	15 / 15 (100%)		
42184	LHC08b	5	582.3 MB	28 Jun 2008 01:44	28 Jun 2008 01:45	ALICE::GSI::SE	4 / 5 (80%)		
42081	LHC08b	40	36.02 GB	27 Jun 2008 13:38	27 Jun 2008 16:36	ALICE::GSI::SE	33 / 40 (82%)		
10 runs		948 files	839.8 GB			0 0 10	0 10 0	1 9 0	

2 subjobs have failed with ERROR_V

<http://pcalimonitor.cern.ch/DAQ/>

Administrative pages



RAW Data Processing Requests

Welcome grigoras! [Add new request »](#)

Run # (chunks)	Partition	Pass	Requested	Status	Options
	- Any -	- Any -		- Any -	Filter
49424 (27)	LHC08c	1	31 Jul 2008	Completed	Delete Reprocess
49424 (27)	LHC08c	2	03 Aug 2008	Completed	Delete Reprocess
48868 (2)	LHC08c	2	03 Aug 2008	Requested by mmeoni	Reprocess
48868 (2)	LHC08c	1	31 Jul 2008	Completed	Delete Reprocess
48860 (2)					
48860 (2)					
48589 (3)					
48589 (3)					
48584 (1)					
48584 (1)					
47747 (24)					

New request

Add new request to the list

Run number(s) (comma separated list):

Pass:

Analysis requests

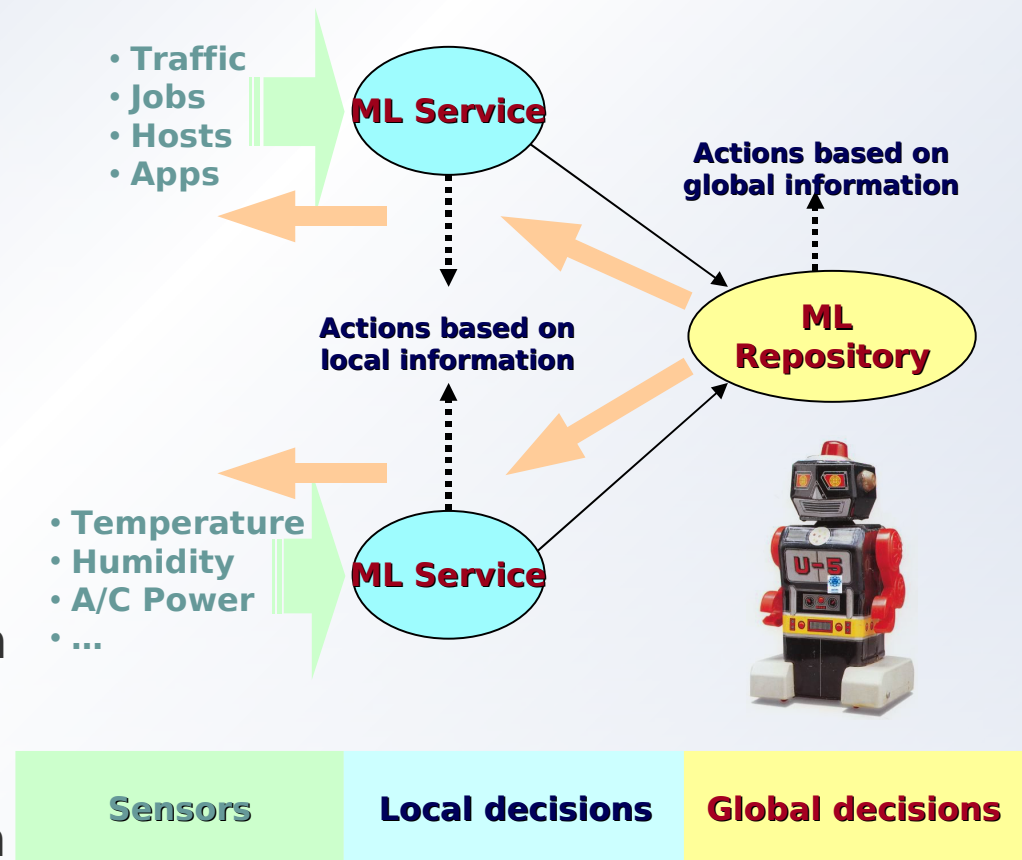
Software packages management

Install packages							Remove installed packages	
Package name	Platforms					Install all	Package name	Remove
	Linux-i686	Linux-x86_64	Linux-ia64	Darwin-i386	Darwin-x86_64			
APISCONFIG V2.4	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	VO_ALICE@AliRoot::v4-11-Rev-05	<input type="checkbox"/>
AliRoot v4-09-Rev-01	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	VO_ALICE@AliRoot::v4-12-Rev-02	<input type="checkbox"/>
AliRoot v4-10-Rev-01	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	VO_ALICE@AliRoot::v4-13-Rev-02	<input type="checkbox"/>
AliRoot v4-10-Rev-02	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	VO_ALICE@AliRoot::v4-13-Rev-03	<input type="checkbox"/>
AliRoot v4-11-Rev-02	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	VO_ALICE@AliRoot::v4-13-Rev-06	<input type="checkbox"/>
AliRoot v4-11-Rev-03	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	VO_ALICE@AliRoot::v4-14-Rev-01	<input type="checkbox"/>
AliRoot v4-11-Rev-04	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	VO_ALICE@AliRoot::v4-14-Rev-02	<input type="checkbox"/>
AliRoot v4-11-Rev-05	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	VO_ALICE@APISCONFIG::V2.4	<input type="checkbox"/>
AliRoot v4-12-Rev-02	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	VO_ALICE@GEANT3::v1-9-2	<input type="checkbox"/>
							VO_ALICE@GEANT3::v1-9-3	<input type="checkbox"/>

Process automation



- Two levels of decisions:
 - local (autonomous),
 - global (correlations).
- Actions triggered by:
 - values above/below given thresholds,
 - absence/presence of values,
 - correlations between any values.
- Possible action types:
 - alerts (emails/instant msg/atom feeds),
 - running an external command,
 - automatic charts annotations in the repository,
 - running custom code, like securely ordering a ML service to (re)start a site service.



Automatic actions in ALICE



- ALICE is using the monitoring information to automatically:
 - resubmit error jobs until a target completion percentage is reached,
 - submit new jobs when necessary (watching the task queue size for each service account)
 - production jobs,
 - RAW data reconstruction jobs, for each pass,
 - restart site services, whenever tests of VoBox services fail but the central services are OK,
 - send email notifications / add chart annotations when a problem was not solved by a restart,
 - dynamically modify the DNS aliases of central services for an efficient load-balancing.
- Most of the actions are defined by few lines configuration files.

Automatic job management



Subject: LPM in action for aliproduct

From: monalisa@mail.cern.ch

Date: 09:43 PM

To: Costin.Grigoras@cern.ch, lbetev@mail.cern.ch

LPM settings: https://pcalimonitor.cern.ch:8443/lpm/lpm_manager.jsp

Account: aliproduct

PackMan seems ok

resubmitting error jobs

```
pid 19295764 had 12 error jobs (job is 73% done - 739 out of 1000)
pid 19297368 had 15 error jobs (job is 68% done - 681 out of 1000)
pid 19298390 had 0 error jobs (job is 22% done - 247 out of 1100)
pid 19298550 had 26 error jobs (job is 54% done - 544 out of 1000)
pid 19298739 had 55 error jobs (job is 56% done - 564 out of 1000)
pid 19298930 had 49 error jobs (job is 53% done - 539 out of 1000)
pid 19302697 had 35 error jobs (job is 35% done - 351 out of 1000)
pid 19304987 had 83 error jobs (job is 27% done - 271 out of 1000)
pid 19304988 had 27 error jobs (job is 10% done - 102 out of 1000)
pid 19307971 had 28 error jobs (job is 0% done - 0 out of 1000)
pid 19309315 had 18 error jobs (job is 0% done - 0 out of 1000)
pid 19310455 had 35 error jobs (job is 0% done - 0 out of 1000)
pid 19311619 had 28 error jobs (job is 0% done - 0 out of 1000)
pid 19312795 had 0 error jobs (job is 0% done - 0 out of 1000)
pid 19314234 had 0 error jobs (job is 0% done - 0 out of 1000)
total resubmitted : 411
```

there are 3691 jobs waiting in queue for user aliproduct

target queue size is 4000

submitting 1 new job(s):

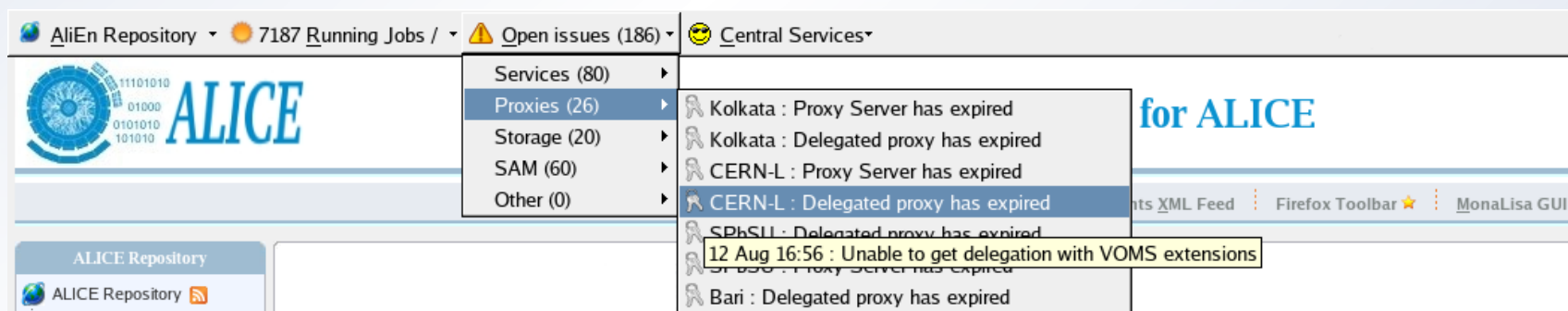
submitting run 401815 of /alice/cern.ch/user/a/aliproduct/PDC_08/LHC08b_test/JDL with arguments: '401815'

```
Aug 15 21:43:19 info Submitting job '/alice/bin/aliproduct_new '...
Aug 15 21:43:19 info There is no price defined for this job in the jdl. Putting the default '1.0'
Aug 15 21:43:19 info Calling directly getListPackages (list -silent -all)
Aug 15 21:43:19 info Job is going to be splitted for production, running from 1 to 1000
Aug 15 21:43:19 info Input Box: {CheckESD.C Config.C CreateAODfromESD.C rec.C sim.C simrun.C tag.C}
Aug 15 21:43:19 info Command submitted (job 19315658)!!
Job ID is 19315658 - 0
AliEn exit code: 0
```

Tools



- Subscription-based system to notify system administrators in case of problems with various components of the system:
 - central or site services problems,
 - failure of storages,
 - proxies' expiration,
 - general announcements.
- You can opt for either email or RSS feed at:
<http://pcalimonitor.cern.ch/xml.jsp>
- The same information is available in the Firefox toolbar:



- Certificate-/role-based administrative pages for day-to-day operations of the Grid (interaction with the site services, software packages deployment, job definitions and so on).

Site monitoring

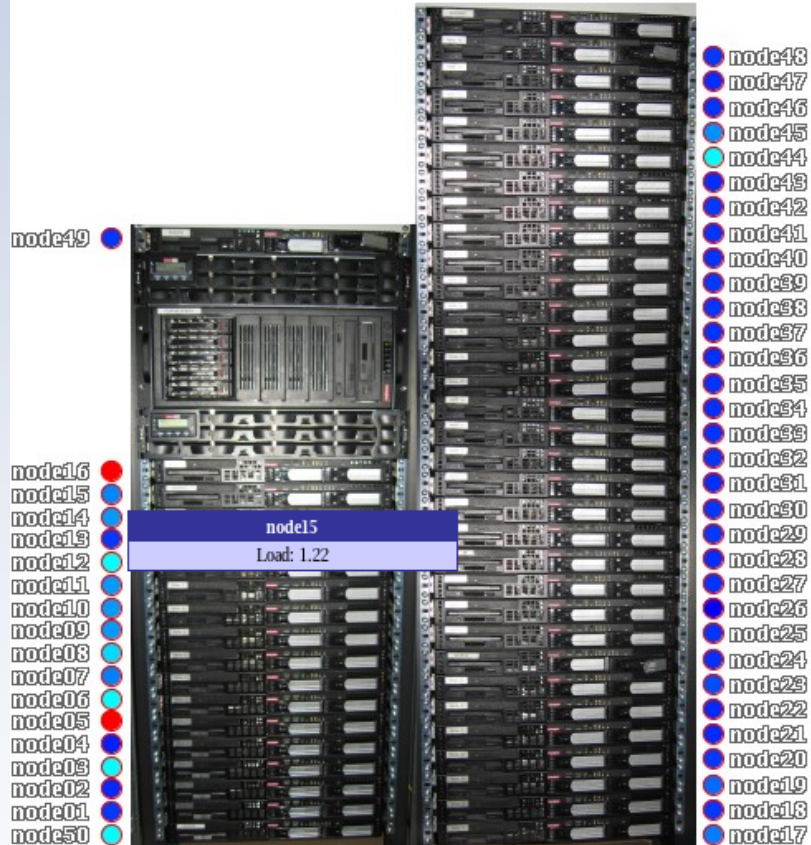


- Central ALICE ML repository stores a small portion of all the monitoring information only.
- Sites can run their own clients that store some other portion of the data.
- Examples:
 - CAF, GSI and Muenster use ApMon-based WN monitoring sending data to the VoBox ML instance,
 - NIHAM monitors the cluster with Ganglia but also uses the VoBox ML to export the values,
 - FZK used a simple client to store a few job accounting parameters to correlate with local BQ accounting.
- PANDA also uses AliEn and implicitly ML to monitor its Grid.

Site monitoring



Muenster Cluster Layout - Machines load status



GSI Grid Cluster

What is this about?

Machines status

Machine	Machine status			LSF	CPU					Memory			Swap		Networking		Processes
	Online	Ping	lustre		jobs	load	usr	sys	idle	Total	Cached	Free	Total	Free	IN	OUT	
lxb255.gsi.de	●	0.237ms	●		4	4.04	0.072	17.25	49.88	7.812 GB	2.96 GB	5.51 GB	1.953 GB	1.953 GB	5.625 KB/s	1.155 KB/s	217
lxb256.gsi.de	●	0.159ms	●		4	4.15	0.12	2.43	49.99	7.812 GB	1.676 GB	5.78 GB	15.27 GB	15.27 GB	6.826 KB/s	1.115 KB/s	221
lxb257.gsi.de	●	0.228ms	●		4	4.26	0.105	6.149	51.81	7.812 GB	2.185 GB	5.079 GB	15.27 GB	15.27 GB	6.856 KB/s	2.337 KB/s	240
lxb258.gsi.de	●	0.372ms	●		3	3.04	0.115	2.378	62.19	7.812 GB	2.408 GB	5.5 GB	14.91 GB	14.91 GB	6.256 KB/s	0.872 KB/s	214
lxb259.gsi.de	●	0.157ms	●		3	3.16	0.132	6.535	66.16	7.812 GB	2.524 GB	6.389 GB	14.91 GB	14.91 GB	6.956 KB/s	2.194 KB/s	213
lxb260.gsi.de	●	0.249ms	●		4	4.07	0.128	19.18	49.88	7.812 GB	2.896 GB	5.562 GB	14.91 GB	14.9 GB	6.217 KB/s	1.324 KB/s	217
lxb261.gsi.de	●	0.201ms	●		4	4.3	0.142	9.325	47.68	7.812 GB	2.33 GB	5.055 GB	14.91 GB	14.91 GB	6.818 KB/s	1.941 KB/s	237

Summary



- Monitoring is vital in a large distributed system, especially in high complexity environments such as ALICE.
- MonALISA provides 24/7 monitoring of all Grid components, from RAW data registration and central services monitoring to jobs status, resources usage accounting and data transfers.
- We've moved beyond monitoring by adding automatic decisions and end-user interfaces that aim at hiding the complexity of the system.
- Future plans:
 - move from current situation where we detect job and services failures to analysis of failure causes,
 - expansion of user-oriented section, more storage-related details and extended networking details.

The End :)



Questions ?



- <http://pcalimonitor.cern.ch/>
- <http://monalisa.cern.ch/>