AliEn site services and monitoring

Miguel Martinez Pedreira

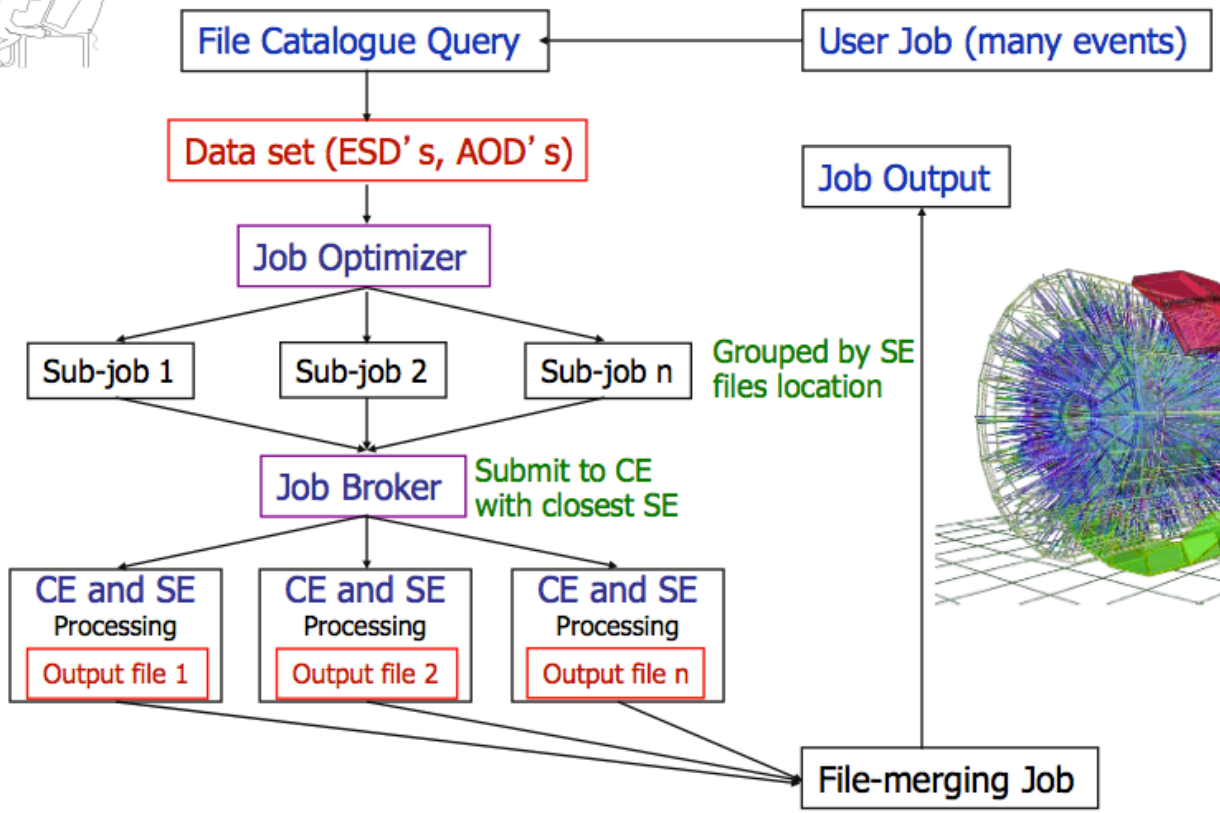
What is AliEn ?

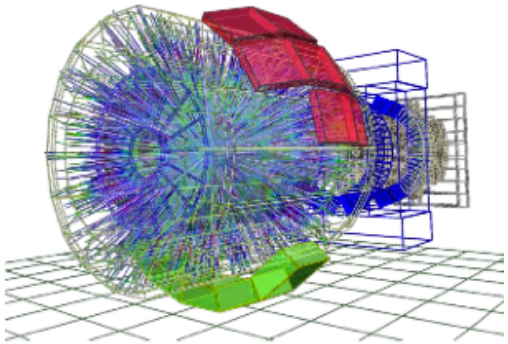
- AliEn (ALICE Environment) is a lightweight Open Source Grid Framework built around other Open Source components using the combination of a Web Service and Distributed Agent Model
  - designed to comply with the offline world of a HEP experiment
    - massive amounts of data implies distributing its storage and processing
- It started within the ALICE Off-line Project at CERN and constitutes the production environment for simulation, reconstruction, and analysis of physics data of the ALICE Experiment
- The current status of the ALICE grid operation can be found at the MonALISA Grid Monitoring

Virtual Organisations

- Users (jobs submission for data analysis) + Central management (the "brain" of the GRID) + Sites (the "muscle" of the GRID)

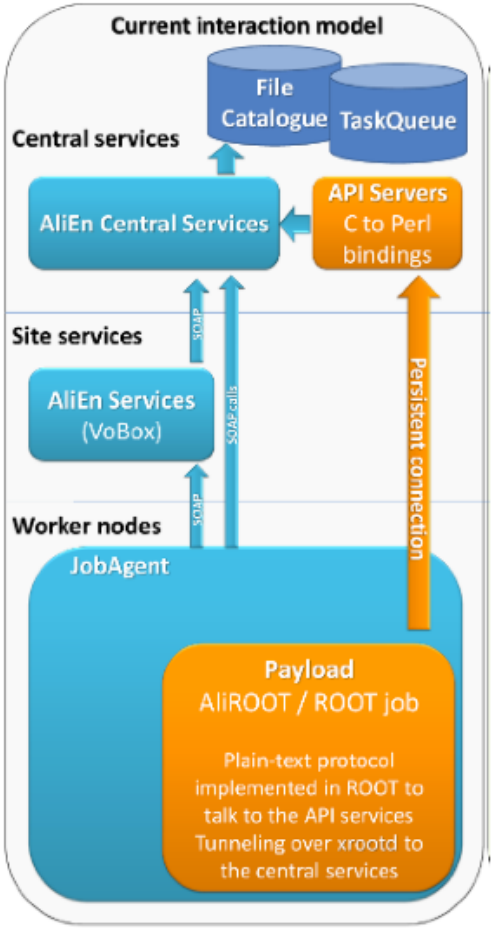
Distributed Analysis





AliEn summary

- 3-layer system that leverages the deployed resources of the underlying WLCG infrastructures and services
- Interfaces to AliRoot via ROOT plugin (TAlien) that implements AliEn API
- Complex workflows including distributed analysis built on top of AliEn API
- Used by ALICE and PANDA



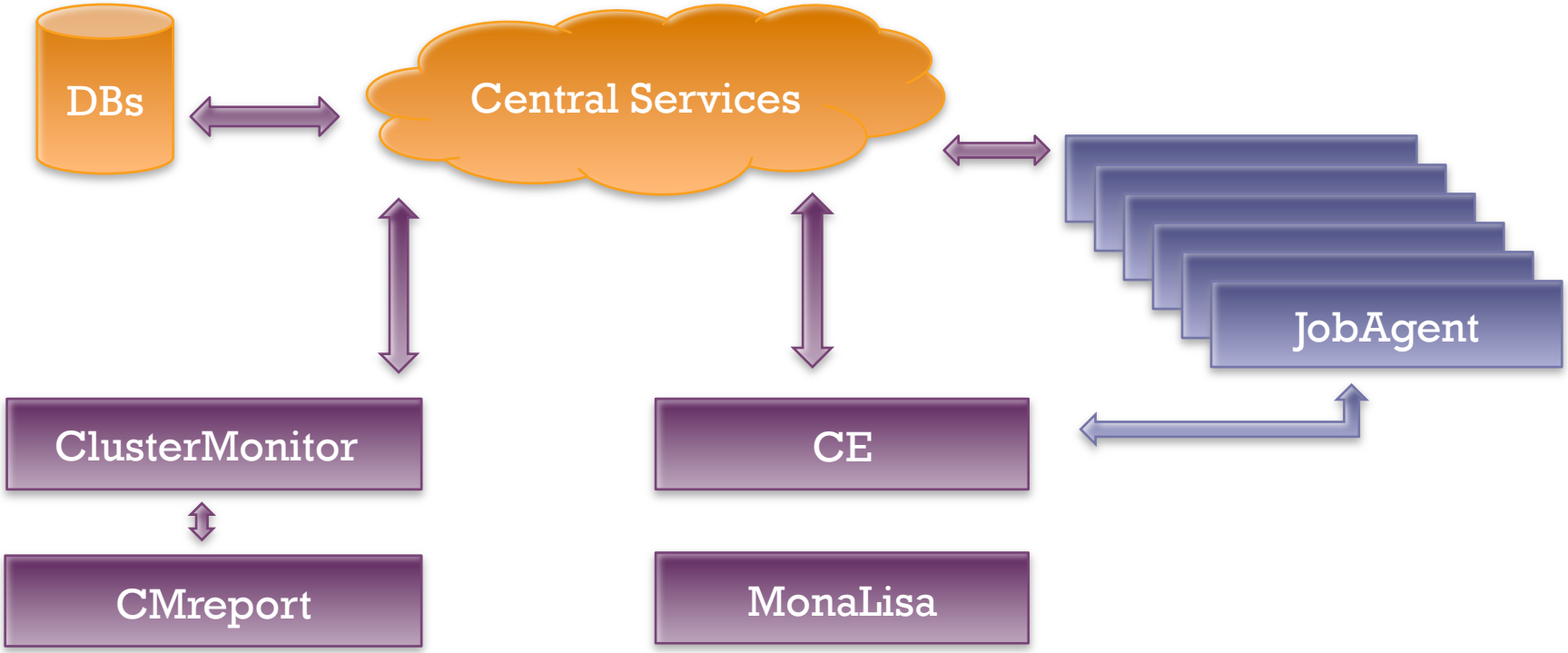
Site services summary

- CE
  - Component that stays in a loop, the one submitting jobs to batch system
  - Comunicates with the JobBroker to see if the parameters of the cluster and the jobs waiting fit
- ClusterMonitor (and CMreport)
  - Proxies several kinds of communications, like process information from jobs, mainly connects nodes to VoBox, gets job slots for CE, other external calls…
    - CMreport parses and sends messages stored by the CM to the CS (joblogs)
- MonaLisa
  - Monitoring component. Controls services, and sends all kind of information to the ML central repository
- ~~PackMan~~
  - See CVMFS later, used to manage which software packages (AliRoot, etc) are installed in the nodes, triggers installations and so on. Previous technology used for this was torrent

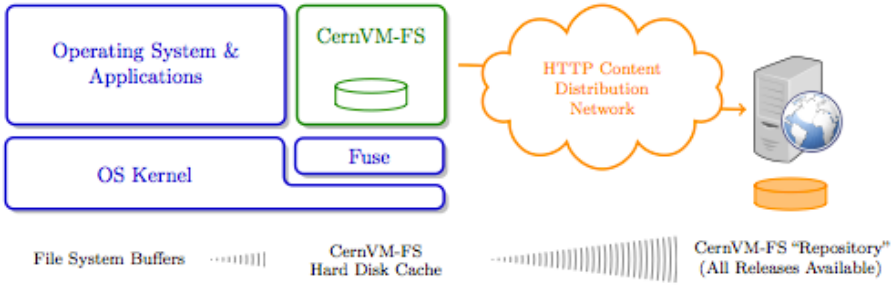
JobAgent

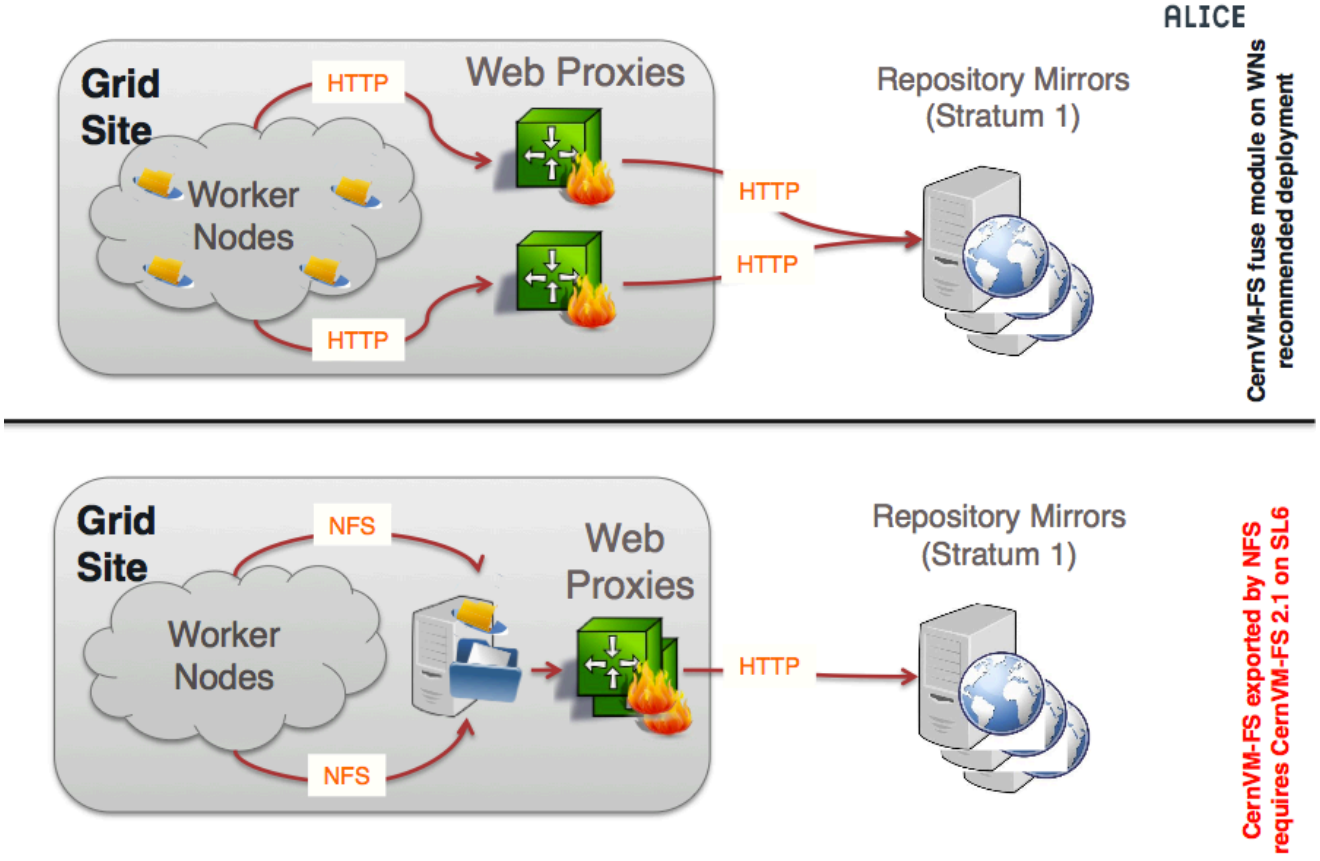
- Also commonly called ‘Pilot’
- Runs the job itself
  - Starts from a script sent to batch system from the CE
  - Gets a job from the TaskQueue
    - Initialization, sandbox, proxy, environment…
  - Several processes: 1 communication daemon, 1 control process, 1 worker process
  - Measures cpu, memory, disk usages, sends heartbeats…
  - Uploads output files to the storage elements
- Continues running other jobs if possible
  - It has its own TTL

Site services



CVMFS


Operating System & Applications
CernVM-FS
HTTP Content Distribution Network
OS Kernel
Fuse
File System Buffers
CernVM-FS Hard Disk Cache
CernVM-FS "Repository" (All Releases Available)


ALICE
Grid Site
HTTP
Web Proxies
Worker Nodes
HTTP
HTTP
HTTP
Repository Mirrors (Stratum 1)
CernVM-FS fuse module on WNs recommended deployment
Grid Site
NFS
Web Proxies
Worker Nodes
NFS
HTTP
Repository Mirrors (Stratum 1)
CernVM-FS exported by NFS requires CernVM-FS 2.1 on SL6

New (WLCG) VoBox pseudo-script

- Put host certificates under /etc/grid-security
  - mkdir –p /etc/grid-security
  - cp </path_certs/host*> /etc/grid-security
  - chmod 400 /etc/grid-security/hostkey.pem
- Install WLCG software
  - yum install yum-priorities yum-protectbase
  - rpm -Uvh http://repository.egi.eu/sw/production/umd/3/sl6/x86_64/updates/umd-release-3.0.1-1.el6.noarch.rpm
  - rpm -Uvh http://linuxsoft.cern.ch/wlcg/sl6/x86_64/wlcg-repo-1.0.0-1.el6.noarch.rpm
  - yum install wlcg-vobox
- Configure VoBox
  - mkdir yaim
  - Put yaim files in the created folder
  - /opt/glite/yaim/bin/yaim -c -s /root/yaim/voboxalice-site-info.def -n VOBOX
- Open ports [22, 1975, 1093, 8084, 9000, 7001, 8884, 9000, 9930]
- Install meta-package for all library dependencies
  - yum install HEP_OSlibs_SL6

New (WLCG) VoBox pseudo-script

- Install CVMFS (squid servers?)
  - rpm -Uvh https://ecsft.cern.ch/dist/cvmfs/cvmfs-release/cvmfs-release-2-4.el6.noarch.rpm
  - yum install cvmfs cvmfs-keys cvmfs-init-scripts
  - CVMFS configuration file: cp cvmfs-default.local /etc/cvmfs/default.local
  - cvmfs_config setup
  - cvmfs_config chksetup
  - cvmfs_config probe
  - Make sure you can see it:
    - ls -l /cvmfs/alice.cern.ch/bin/

New (WLCG) VoBox pseudo-script

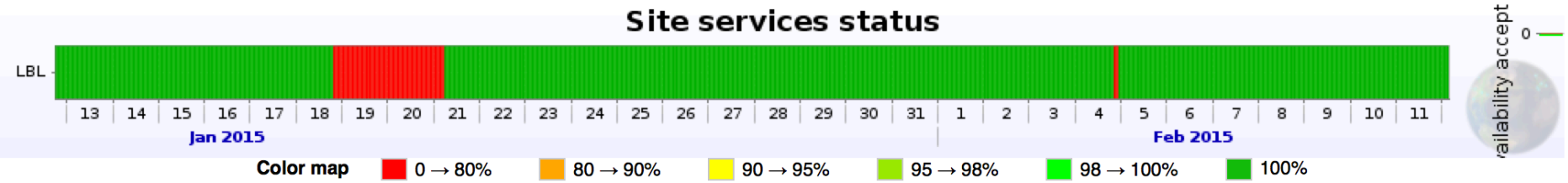
- Links in $HOME/bin to AliEn bins
- Proxy steps (http://alien2.cern.ch/index.php?option=com_content&view=article&id=30&Itemid=77)
- Some files for AliEn (under $HOME/.alien)

```
[alicesgm@voboxalice2 .alien]$ ls -l
total 12
-rw-r--r--. 1 alicesgm alice  132 Jan 27 11:32 alice.conf
-rw-r--r--. 1 alicesgm alice  213 Jan 27 11:38 Environment
drwxr-xr-x. 3 alicesgm alice 4096 Jan 23 17:02 etc
[alicesgm@voboxalice2 .alien]$ cat alice.conf
CLUSTERMONITOR_ADDRESS http://voboxalice2.cern.ch
CLUSTERMONITOR_SOAPTYPE httpd
JOB_MANAGER_ADDRESS  https://aliendb8.cern.ch:8083

[alicesgm@voboxalice2 .alien]$ cat Environment
export ALIEN_USER=mmmartin
export ALICE_LDAP_DN=aliendb06a.cern.ch:8389/o=alice,dc=cern,dc=ch
export ALIEN_DOMAIN=cern.ch
export CREAM_CEREQ='"LSFResource==\"wigner && amd && !fullsmt && !smt && type==SLC6_64\""'
[alicesgm@voboxalice2 .alien]$ ls -l etc/aliend/ALICE/
total 4
-rw-r--r--. 1 alicesgm alice 112 Jan 27 11:42 startup.conf
[alicesgm@voboxalice2 .alien]$ cat etc/aliend/ALICE/startup.conf
AliEnUser=alicesgm
AliEnServices="Monitor CE CMreport MonaLisa"
AliEnCommand="/cvmfs/alice.cern.ch/bin/alien"

[alicesgm@voboxalice2 .alien]$ cat etc/aliend/startup.conf
ALIEN_ORGANISATIONS="ALICE"
```

Monitoring

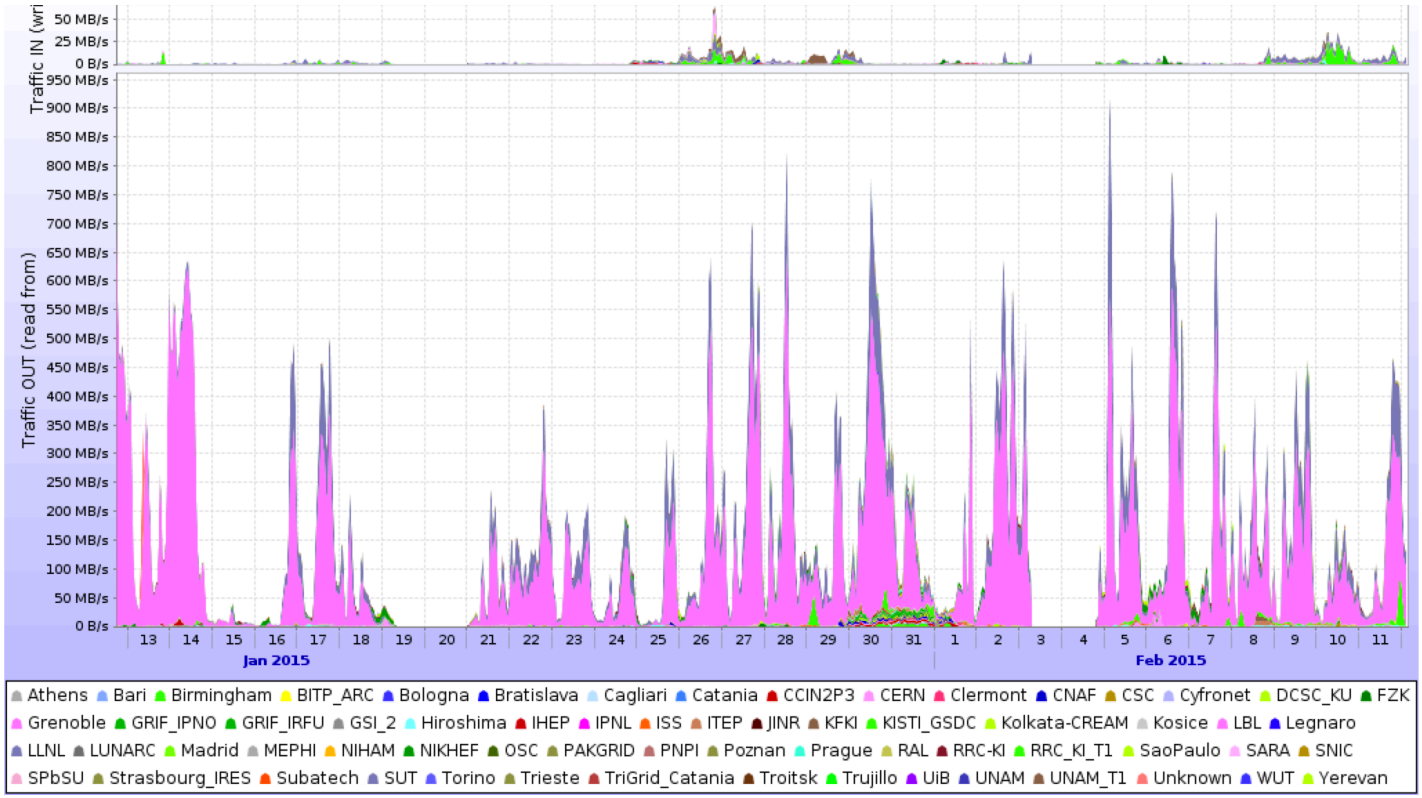


Statistics

Link name	Data		Overall	
	Starts	Ends	Reliability	Availability
LBL	12 Jan 2015 18:09	12 Feb 2015 03:26	98.19%	92.33%

- Reliability = yes / (yes + no)
- Availability = yes / (yes + no + unknown)
- What do you want to see?

Monitoring



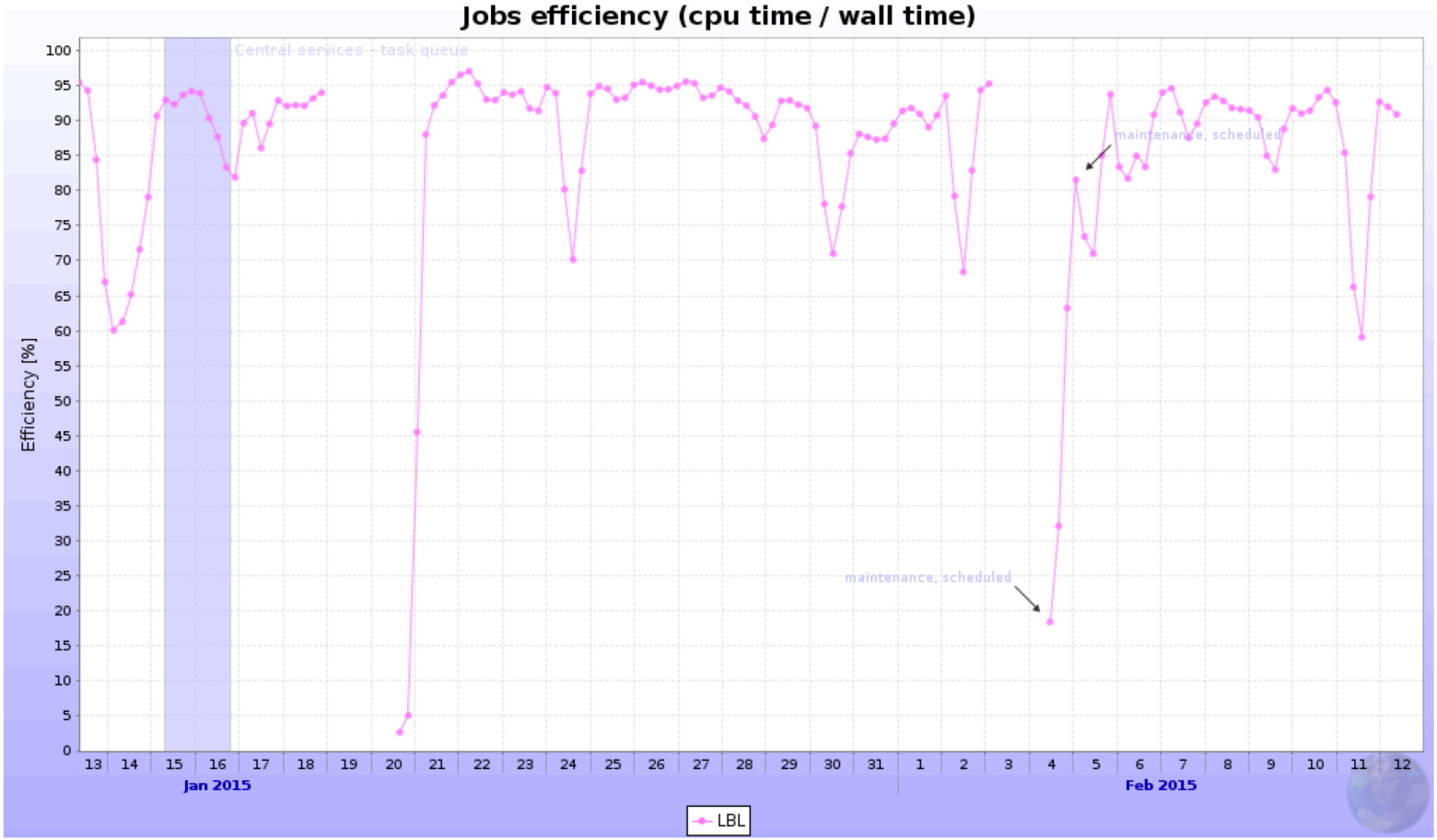
Traffic IN

	Series	Last value ▲	Min	Avg	Max	Total
33.	LLNL	8.314 MB/s	0 B/s	1.278 MB/s	122.4 MB/s	3.199 TB
31.	LBL	0.655 MB/s	0 B/s	0.243 MB/s	16.44 MB/s	623.4 GB
61.	UNAM_T1	0.185 MB/s	0 B/s	0.485 MB/s	22.47 MB/s	1.214 TB
40.	PAKGRID	0.138 MB/s	0 B/s	0.27 MB/s	6.621 MB/s	673.9 GB
28.	KISTI_GSDC	52.89 KB/s	0 B/s	1.01 MB/s	136.6 MB/s	2.528 TB

Traffic OUT

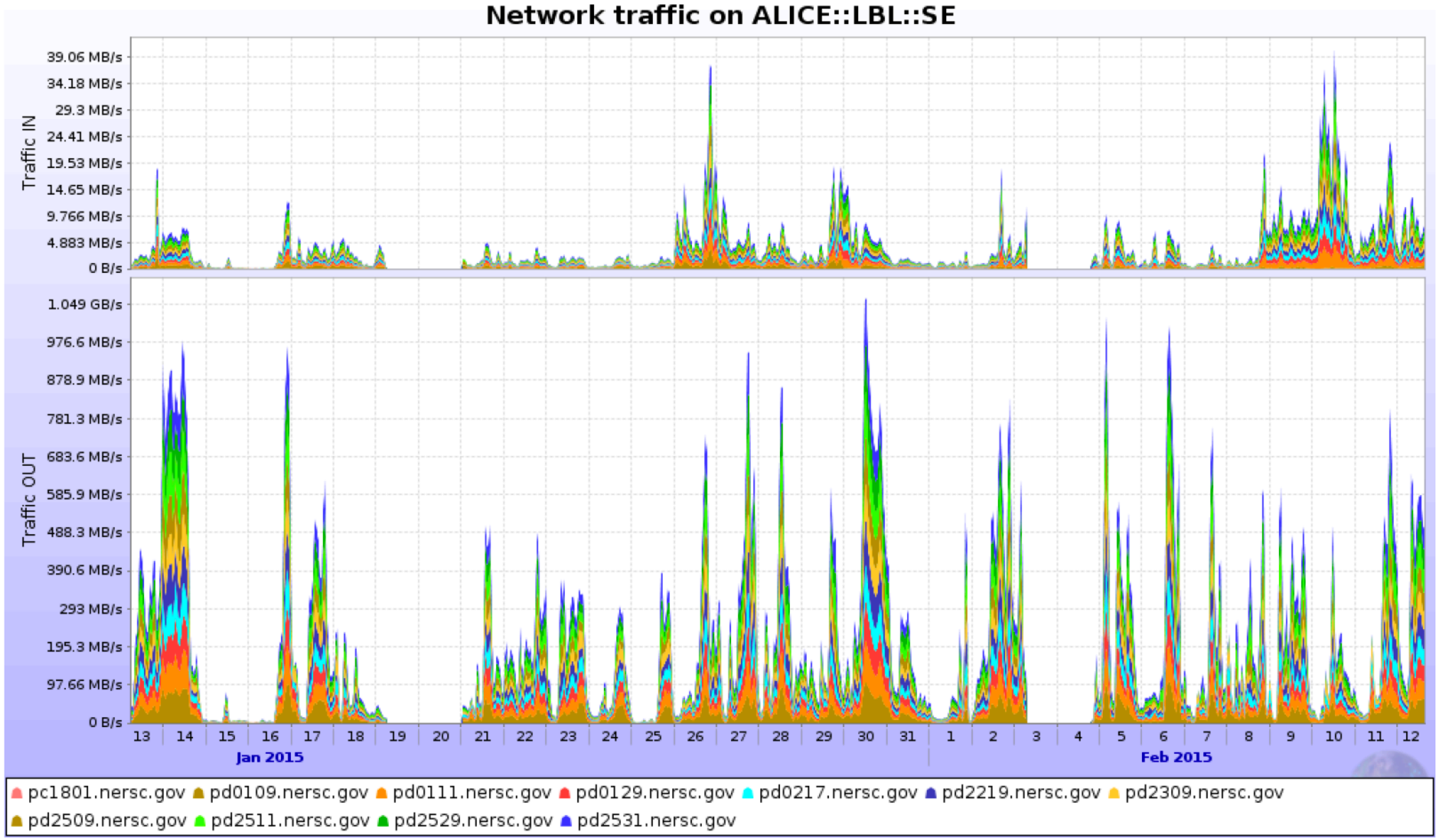
	Series	Last value	Min	Avg	Max	Total
31.	LBL	104.5 MB/s	0 B/s	124.8 MB/s	1.337 GB/s	312.1 TB
33.	LLNL	20.66 MB/s	0 B/s	33.17 MB/s	569.7 MB/s	83.04 TB
29.	Kolkata-CREAM	4.631 MB/s	0 B/s	0.231 MB/s	6.989 MB/s	588.5 GB
55.	Trieste	1.251 MB/s	2.458 B/s	0.262 MB/s	8.251 MB/s	639.5 GB
21.	Hiroshima	1.11 MB/s	0 B/s	0.438 MB/s	13.85 MB/s	1.094 TB

Monitoring



Jobs efficiency (cpu time / wall time)

	Series	Last value	Min	Avg	Max
1.	LBL	90.86	0	85.85	100
Total		90.86		85.85	



Traffic IN

	Series	Last value	Min	Avg	Max	Total
1.	pc1801.nersc.gov	38 KB/s	0.635 KB/s	10.3 KB/s	254.4 KB/s	25.78 GB
2.	pd0109.nersc.gov	452.5 KB/s	0.374 KB/s	271.5 KB/s	14.3 MB/s	679.6 GB
3.	pd0111.nersc.gov	891.9 KB/s	0.36 KB/s	462.8 KB/s	21.93 MB/s	1.131 TB
4.	pd0129.nersc.gov	991 KB/s	0.362 KB/s	456.4 KB/s	21.43 MB/s	1.116 TB
5.	pd0217.nersc.gov	431.3 KB/s	0.372 KB/s	452.3 KB/s	18.62 MB/s	1.106 TB
6.	pd2219.nersc.gov	1018 KB/s	0.352 KB/s	477.7 KB/s	32.01 MB/s	1.168 TB
7.	pd2309.nersc.gov	975.8 KB/s	0.358 KB/s	457.3 KB/s	23.66 MB/s	1.118 TB
8.	pd2509.nersc.gov	787.2 KB/s	0.358 KB/s	470.8 KB/s	24.73 MB/s	1.151 TB
9.	pd2511.nersc.gov	725.3 KB/s	0.368 KB/s	467.9 KB/s	27.45 MB/s	1.144 TB
10.	pd2529.nersc.gov	961 KB/s	0.362 KB/s	445.8 KB/s	25.33 MB/s	1.09 TB
11.	pd2531.nersc.gov	774.4 KB/s	0.354 KB/s	482.2 KB/s	25.84 MB/s	1.179 TB
	Total	**7.858 MB/s**		**4.351 MB/s**		**10.89 TB**

Traffic OUT

	Series	Last value	Min	Avg	Max	Total
1.	pc1801.nersc.gov	11.06 KB/s	0.296 KB/s	2.961 KB/s	63.02 KB/s	7.413 GB
2.	pd0109.nersc.gov	61.6 MB/s	0.138 KB/s	22.93 MB/s	165.9 MB/s	57.41 TB
3.	pd0111.nersc.gov	44.16 MB/s	0.224 KB/s	21.51 MB/s	190.8 MB/s	53.86 TB
4.	pd0129.nersc.gov	43.14 MB/s	0.224 KB/s	20.41 MB/s	145.4 MB/s	51.11 TB
5.	pd0217.nersc.gov	39.73 MB/s	0.216 KB/s	21.14 MB/s	185.6 MB/s	52.92 TB
6.	pd2219.nersc.gov	50.8 MB/s	0.192 KB/s	23.53 MB/s	160.3 MB/s	58.9 TB
7.	pd2309.nersc.gov	46.09 MB/s	0.223 KB/s	21.84 MB/s	187.6 MB/s	54.68 TB
8.	pd2509.nersc.gov	47.04 MB/s	0.148 KB/s	24.01 MB/s	207.8 MB/s	60.11 TB
9.	pd2511.nersc.gov	58.13 MB/s	0.204 KB/s	25.22 MB/s	183.3 MB/s	63.13 TB
10.	pd2529.nersc.gov	53.3 MB/s	0.225 KB/s	23.24 MB/s	196.5 MB/s	58.19 TB
11.	pd2531.nersc.gov	54.82 MB/s	0.183 KB/s	25.26 MB/s	179 MB/s	63.23 TB
	Total	**498.8 MB/s**		**229.1 MB/s**		**573.5 TB**

Monitoring

CERN	Networking on CERN-AURORA (**188.184.2.28,2001:1458:201:22:0:0:100:16**): low buffer size kernel parameters: tcp_wmem_max=4 MB only Networking on CERN-CVMFS (**voalice13.cern.ch**): low buffer size kernel parameters: tcp_wmem_max=4 MB only Networking on CERN-CVMFS (**voalice13.cern.ch**): No IPv6 public address Networking on CERN-SHA2 (**voalice11.cern.ch**): low buffer size kernel parameters: tcp_wmem_max=4 MB only Networking on CERN-SHA2 (**voalice11.cern.ch**): No IPv6 public address Networking on CERN-TEST (**voalice10.cern.ch**): low buffer size kernel parameters: tcp_wmem_max=4 MB only Networking on CERN-TEST (**voalice10.cern.ch**): No IP Please start deploying IPv6 in order to prepare for IPv6-only WNs. For more details see: http://ipv6.web.cern.ch/	Low long-term average job efficiency on CERN-TEST (**voalice10.cern.ch**)
CNAF ui01-alice.cr.cnaf.infn.it	ALICE::CNAF::TAPE redirector **ds-207.cr.cnaf.infn.it** doesn't see all nodes, only 170.9 TB out of 232.9 TB total space is connected ALICE::CNAF::TAPE redirector **ds-208.cr.cnaf.infn.it** doesn't see all nodes, only 170.9 TB out of 232.9 TB total space is connected Networking: No IPv6 public address ALICE::CNAF::TAPE: xrootd data server **ds-207.cr.cnaf.infn.it** fails GET test. ALICE::CNAF::TAPE: xrootd data server **ds-208.cr.cnaf.infn.it** fails GET test.	
FZK alice-kit.gridka.de	Networking: low buffer size kernel parameters: tcp_wmem_max=4 MB only Networking: No IPv6 public address ALICE::FZK::TAPE: xrootd data server **f01-070-130-e.gridka.de** fails DEL test. ALICE::FZK::TAPE: xrootd data server **f01-070-128-e.gridka.de** fails DEL test. ALICE::FZK::TAPE: xrootd data server **f01-070-122-e.gridka.de** fails DEL test. ALICE::FZK::TAPE: xrootd data server **f01-070-122-e.gridka.de** fails DEL test.	Low long-term average job efficiency

More monitoring

- LAN / WAN traffic
- Bandwidth tests

IN from

No.	ID	Site	Speed (Mbps)	Hops	RTT (ms)	Streams
1.	1992381	NERSC	1,006.69	2	0.40	1
2.	2022541	ORNL	671.12			1
3.	1266178	PDC	243.04	24	198.44	1
4.	2082762	Strasbourg_IRES	226.50	22	157.59	1
5.	1976089	KISTI-CREAM	192.95	11	144.74	1
6.	2080819	TriGrid_Catania	184.37	20	184.45	1

OUT to

No.	ID	Site	Speed (Mbps)	Hops	RTT (ms)	Streams
1.	1991741	NERSC	989.91	2	0.41	1
2.	2027623	ORNL	679.51			1
3.	2082660	GRIF_IPNO	335.56			1
4.	2083847	CNAF	285.23	21	168.89	1
5.	2080828	TriGrid_Catania	268.45	22	188.83	1
6.	2080017	NIHAM	260.06			1

- What do you want to have in ML that is useful for you ?