

ES

Experiment Support

CERN IT  
Department

# WLCG Workshop

# ALICE SESSION



- ❑ *2010 Data Taking: ALICE Report*
  - Patricia Mendez Lorenzo (CERN)
- ❑ *T2 feedback on setting-up and operating storage for the Alice VO*
  - Jean-Michel Barbet (Subatech)
- ❑ *The WLCG services required by ALICE at the T2 sites: evolution of the CREAM-system, VOBOXES and monitoring*
  - Dagmar Adamova (Prague)
- ❑ *Alice tier 1 testbed at KISTI-NSDC*
  - Christophe Bonnaud (KISTI)

**MANY THANKS TO GALINA SHABRATOVA (JINR) COAUTHOR OF THE 1st AND 4th PRESENTATIONS**

**Concurrent Grid session at CERN during the ALICE Offline Week**

# 2010 Data Taking: ALICE Report

Patricia Méndez Lorenzo on behalf of the  
ALICE Offline team

COAUTHOR: **Galina Shabratova** (JINR)

Thanks to **Andreas Morsch** (ALICE) for his  
contributions

This talk introduces the topics that the speakers will treat in detail in this session:

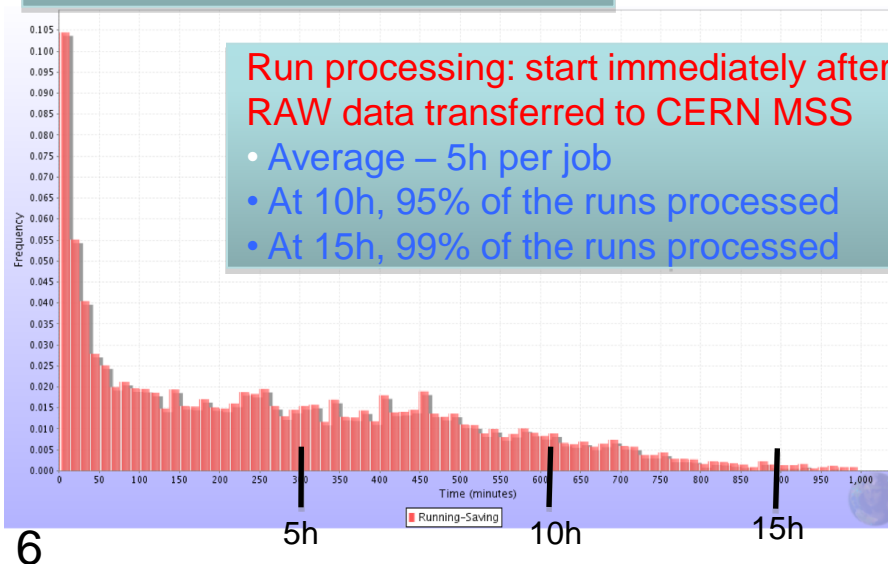
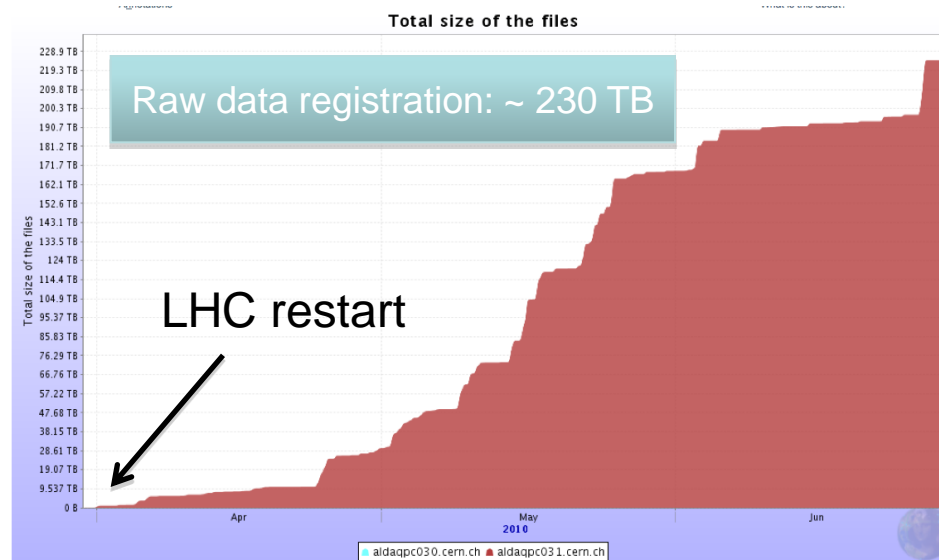
- 2010 data taking results
- Raw data transfers
- New AliEn2.18
- User Analysis
- WLCG services
- Monitoring and operations

- ❑ We finished 2009
  - All sites migrated to SL5 (VOBOX and WNs)
    - ✧ Only 4 sites were blacklisted
  - Lack of CREAM services at the T2 sites
- ❑ At this moment
  - All ALICE sites in production (no blacklisted sites)
  - All sites have a CREAM service
    - ✧ Basically the WMS service has been deprecated (only used at CERN)
  - New AliEn2.18 version
  - Xrootd used as the unique transfer protocol
  - Increasing relevance of the Analysis Facilities (CAF)
    - ✧ New TF and AF Meeting

- ❑ Since Feb. till end of Mar. cosmic-ray data taking
  - $\sim 10^5$  events
- ❑ pp run since March 30<sup>th</sup>
  - 7 TeV: 490Mio events
  - 0.9 TeV: 13Mio events

### MC production

Several production cycles for 0.9, 2.36 and 7TeV pp with various generators and conditions from real data taking



Run processing: start immediately after RAW data transferred to CERN MSS

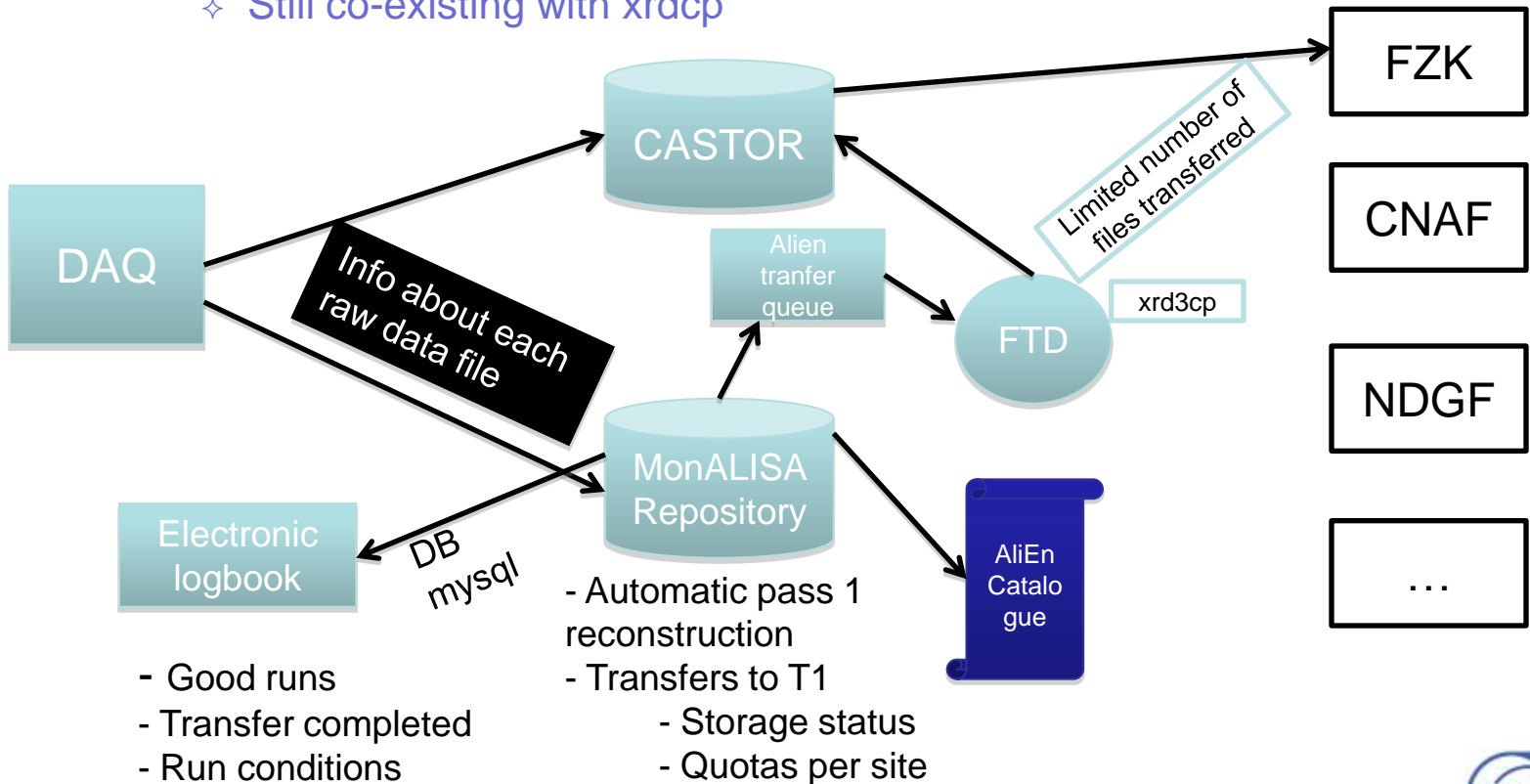
- Average – 5h per job
- At 10h, 95% of the runs processed
- At 15h, 99% of the runs processed

### Raw data processing

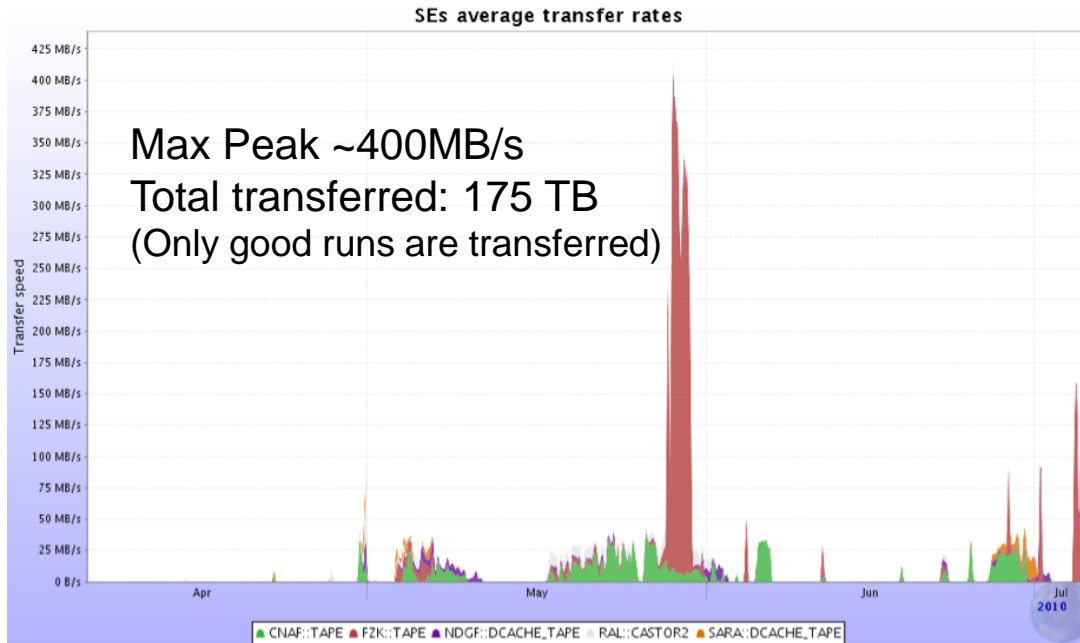
- Pass1-6 completed for 0.9 and 2.36 TeV data
- Pass1@T0 for 7TeV data follows data taking
- Analysis train running weekly: QA, physics working groups organized analysis



- T0-T1 transfers performed via xrootd
  - Latests version (<https://savannah.cern.ch/projects/xrootd>) enables the 3<sup>rd</sup> party copy (xrd3cp)
    - ✦ Implemented and already in production for FZK and CNAF
    - ✦ Still co-existing with xrdcp



- Good runs
- Transfer completed
- Run conditions



- Full runs transferred to each T1 site
- SE choice based on the ML tests at transfer time

- First approach: Equal conditions, SE taken randomly
- Procedure changed to chose the SE based on the number of resources provided by the site
- Distribution already defined in SC3

Transfer requests ([see new issues](#))

ID	Path	Target SE	Status	Progress	Files	Total size	Started	Ended
325.	/alice/data/2010/LHC10b/000115393/collection	ALICE::LBL::SE	Error	<div style="width: 100%; background-color: red;"></div>	366	963.8 GB	07 May 2010 16:23	today 06:22
324.	/alice/data/2010/OCDB	ALICE::LEGNARO::SE	Running	<div style="width: 100%; background-color: gray;"></div>	77074	10.36 GB	03 May 2010 18:24	
323.	/alice/data/2010/LHC10b/000115399/collection	ALICE::RAL::CASTOR2	Done	<div style="width: 100%; background-color: green;"></div>	133	333 GB	30 Apr 2010 09:59	06 May 2010 02:58
322.	/alice/data/2010/LHC10b/000114930/collection	ALICE::NDGF::DCACHE_TAPE	Running	<div style="width: 10%; background-color: green;"></div>	19	35.51 GB	30 Apr 2010 09:01	
321.	/alice/data/2010/LHC10b/000114931/collection	ALICE::FZK::TAPE	Done	<div style="width: 100%; background-color: green;"></div>	95	248.1 GB	30 Apr 2010 09:01	07 May 2010 05:24
320.	/alice/data/2010/LHC10b/000115056/collection	ALICE::NDGF::DCACHE_TAPE	Running	<div style="width: 10%; background-color: green;"></div>	2	86.03 MB	30 Apr 2010 09:01	
319.	/alice/data/2010/LHC10b/000115165/collection	ALICE::SARA::DCACHE_TAPE	Done	<div style="width: 100%; background-color: green;"></div>	10	1.237 GB	30 Apr 2010 09:01	09 May 2010 19:47
318.	/alice/data/2010/LHC10b/000115173/collection	ALICE::RAL::CASTOR2	Done	<div style="width: 100%; background-color: green;"></div>	10	1.071 GB	30 Apr 2010 09:01	06 May 2010 02:33
317.	/alice/data/2010/LHC10b/000115186/collection	ALICE::NDGF::DCACHE_TAPE	Running	<div style="width: 20%; background-color: green;"></div>	40	107.1 GB	30 Apr 2010 09:01	
316.	/alice/data/2010/LHC10b/000115193/collection	ALICE::NDGF::DCACHE_TAPE	Running	<div style="width: 30%; background-color: green;"></div>	90	236 GB	30 Apr 2010 09:01	

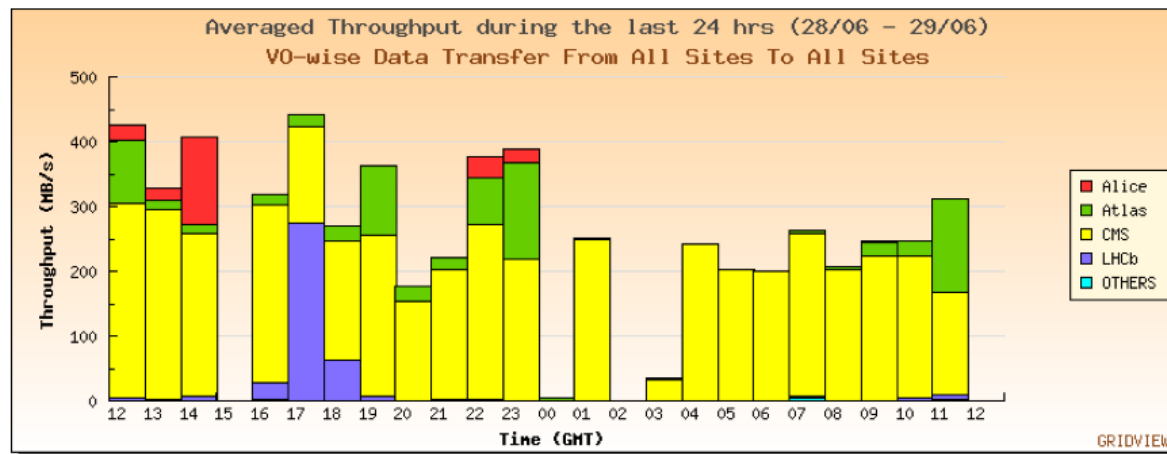


- Status of the transfers daily followed in Dashboard

Transfers done on: Wed 05 May 2010

Site (click on any site)	Successful transfers	Failed transfers	Efficiency
ALICE::LEGNARO::SE	90	4997	1.77 %
ALICE::SARA::DCACHE_TAPE	72	509	12.39 %
<b>Error message</b>			<b>Counter</b>
ALICE::CERN::CENTRAL7:			291
ALICE::CERN::CENTRAL8: contacting the Broker/Transfer			192
ALICE::CERN::CENTRAL8: getting the file root://voalice09.cern.ch:1094//castor/cern.ch/alice/raw/global/2010/04/05/13/10000115369018.20.root			19
ALICE::CERN::CENTRAL1: getting the file root://voalice09.cern.ch:1094//castor/cern.ch/alice/raw/global/2010/04/10/11/10000115888014.100.root			7
ALICE::NDGF::DCACHE_TAPE	194	41	82.55 %
<b>Error message</b>			<b>Counter</b>
ALICE::CERN::CENTRAL7:			29
ALICE::CERN::CENTRAL8: contacting the Broker/Transfer			12
ALICE::RAL::CASTOR2	339	18	94.96 %
ALICE::FZK::TAPE	331	16	95.39 %
ALICE::CNAF::TAPE	726	0	100.00 %

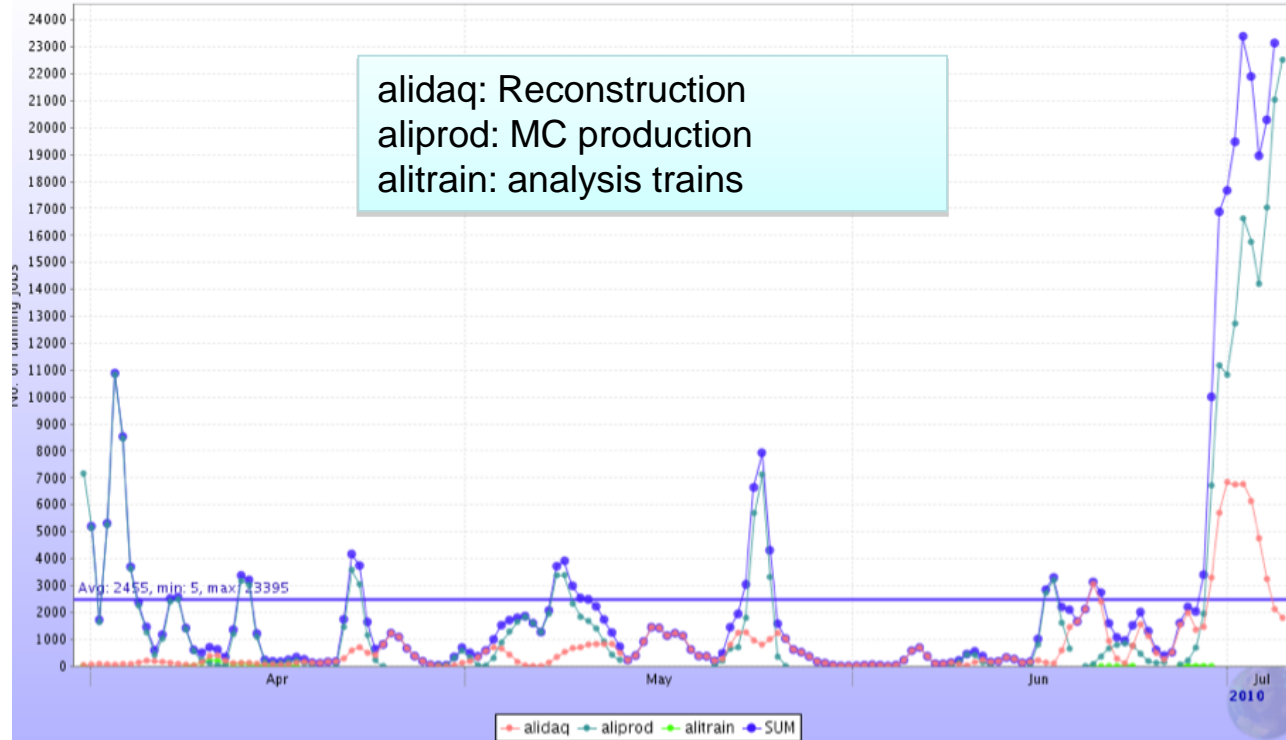
- Since few weeks also available in Gridview



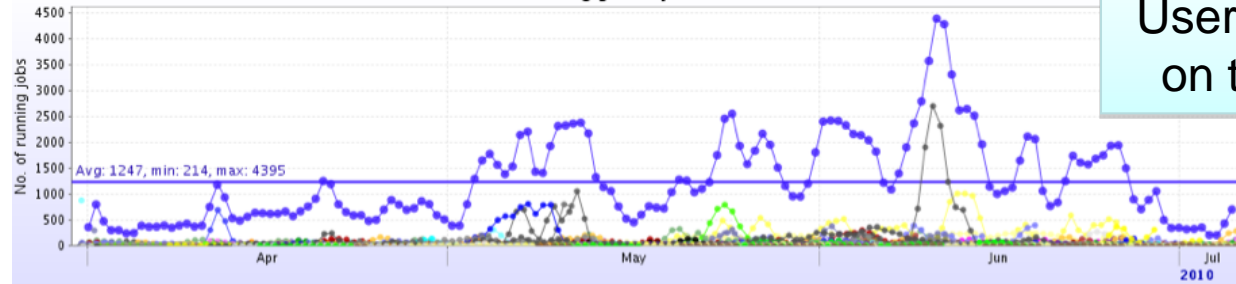
- ❑ Registration of raw data in CASTOR2 and AliEn file catalogue
  - Routine operation, high redundancy of servers and software
  - Problem-free operation
  - All updates carried out transparently by IT/FIO and ALICE Offline
- ❑ Replication of raw data to T1 sites
  - All T1 sites being currently used
    - ✧ Transfers to Lyon still pending
      - The xrootd with tape backend storage at Lyon under tuning
      - Entering in production in a short time
  - Data rates not fully maintained
    - ✧ No escalation issues found



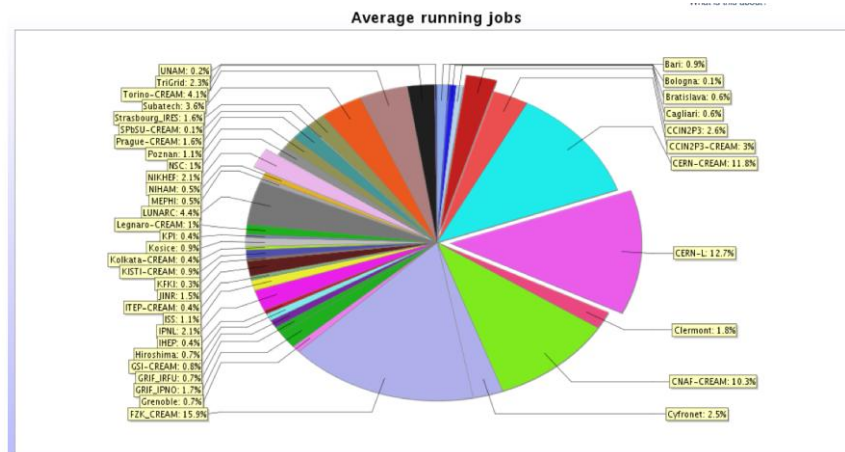
Running jobs per user



Running jobs per user



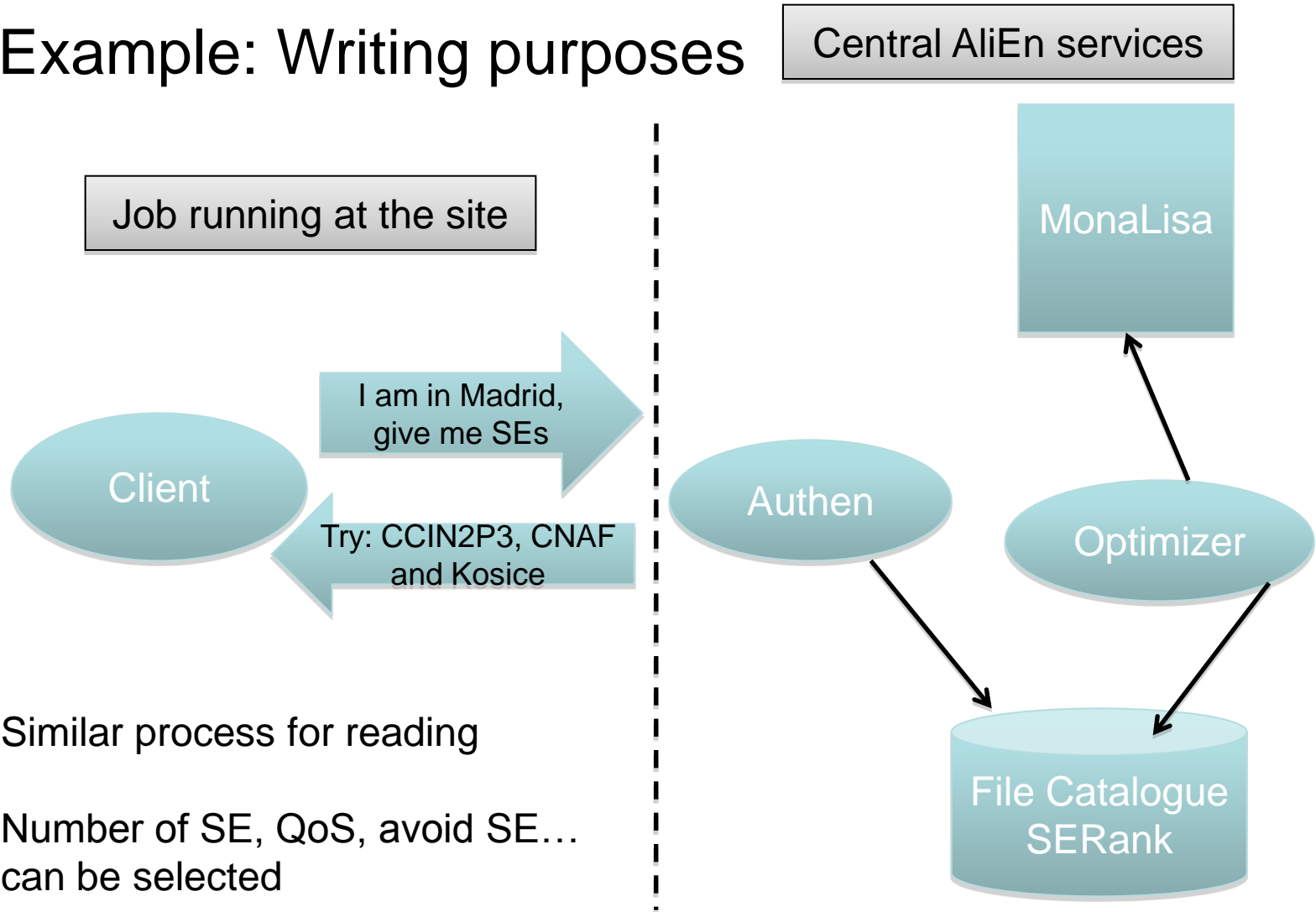
User analysis on the Grid



- ❑ Many new features included in AliEnv2.18 solving quite a lot of previous challenges
  - Deployment of this version transparently from central services
  - Simultaneously to the startup of data taking
- ❑ We mention here two important improvements:
  - Implementation of Job and File Quotas
    - ✧ Limit on the available resources per user
    - ✧ Jobs: # jobs, cpuCost, runtime
    - ✧ Files: # files, total size (including replicas)
  - Improved SE discovery
    - ✧ Finding the closest working SEs of a QoS to the worker or the user once the file has been registered in the catalogue
      - For reading and writing and taking into account ML tests and the network and geographical topology it discovers
    - ✧ Simplifying the selection of SE
      - Giving more options in case of special needs



## Example: Writing purposes



Similar process for reading

Number of SE, QoS, avoid SE...  
can be selected



## □ Analysis trains

- Grouping many analysis tasks in a common data set
  - ✧ Allows for better CPU/Wall and reduces load on the storage servers
  - ✧ Pass 1 reconstruction is automatically followed by specific Pilot trains
    - Assess the quality of the run (detector by detector, global tracking and vertex position and stability)
  - ✧ Specific analysis activities required by the physics groups

113.	<code>#!pass1.sh</code>	<a href="#">/alice/cern.ch/user/a/alidaq/</a>		100 %	alidaq
203.	<code>/alice/cern.ch/user/a/akisiel/PDC09/TAGS/MergeOfficial/analysisMerge.jdl (edit)</code>	<code>#OUTPUTDIR#</code>	99%	100 %	alidaq
180.	<code>#.alien.lpm.RawQA 9</code>	<a href="#">/alice/cern.ch/user/a/alidaq/</a>	99%	100 %	alidaq
181.	<code>#.alien.lpm.RawQAMerge</code>		99%	100 %	alidaq
182.	<code>#.alien.lpm.CleanupAfterMerge</code>		100%	100 %	alidaq

- Reconstruction is started as soon as data is registered in CASTOR
- Analysis trains automatically triggered at the end of the reconstruction activities
- At the end of the analysis, merge and cleanup procedures are executed

- Chaotic analysis
  - User analysis on the Grid
  - Internal ALICE prioritization within the common task queue works well
    - ✧ Production user is demoted in favor of users in the queue
    - ✧ Generic pilot jobs assure fast execution of user jobs at the sites

- Fast and interactive analysis in PROOF clusters
  - Processing of reconstructed data, calibration and alignment
    - ✧ Set of pre-stage files already available at the WNs of the PROOF clusters
    - ✧ Possibility to pre-stage any kind of file required by the users and registered in the AliEn file catalogue
    - ✧ Limitations:
      - Available space in WNs
      - Memory consumption affecting the proof master
  - Very popular system for the users
  - CAF@CERN
    - ✧ For the moment unique stable system
    - ✧ 1 proof master and 26 nodes distributed in two sets (3 and 4 disk nodes)
  - GSI AF@GSI, SAF@Subatech, LAF@CCIN2P3
  - Status and development weekly followed during the TF&AF meeting

- ❑ gLite3.2 VOBOX
- ❑ CE front-end directly submitting to the local batch system
  - gLite sites have to provide a CREAM-CE (1.6 version)
  - Specific ALICE queue behind the CREAM-CE
  - Resource BDII providing information about CREAM-CE
- ❑ Xrootd based storage system
- ❑ CPU requirements
  - 2GB of memory per job for MC and reconstruction jobs
  - Optimization needed for some user analysis jobs

- ❑ 2010 goal: Deprecation of the gLite-WMS
  - Status: DONE. All ALICE sites are providing (at least) a CREAM service
    - ✧ AliEn2.18 foresees multiple CREAM-CE submission
  - Support to gLite-WMS still maintained in AliEn2.18
    - ✧ Still used at CERN (exclusively)
    - ✧ With the increase of the ALICE load in the last months the stability of the system has been confirmed
- ❑ Latest requirement: Service migration to CREAM1.6
  - Available for gLite3.2/sl5\_x86\_64 (version highly recommended by ALICE)
    - ✧ <https://savannah.cern.ch/patch/?3959>
  - A large number of bugs reported by ALICE site admins have been solved in this version
    - ✧ No new bugs have been reported by any site admin nor the experiment
  - It allows a lighter distribution of the gLite3.2 VOBOX

- ❑ CREAM1.6 distribution enables a lighter VOBOX
  - Direct job submission approach via CREAM requires the specification of a gridftp server to retrieve the OSB
    - ✧ ALICE solved it requiring a gridftp server with the VOBOX distribution (distributed with gLite3.2)
    - ✧ Retrieval of the OSB from CREAM disk exposed in CREAM1.6
- ❑ ALICE approach for the 2010 VOBOX
  - A single VOBOX is needed
  - What to do with the 2<sup>nd</sup> VOBOX (provided in 2009 by all those sites providing CREAM)?
    - ✧ Rescue it: FAILOVER MECHANISM
      - Both VOBOXEs run the same (redundant) services
    - ✧ This approach has been included in AliEnv2.18 to take advantage of the 2<sup>nd</sup> VOBOX deployed in several ALICE sites
      - ~25 sites currently providing  $\geq 2$  VOBOXES



- ❑ Support team priority for this year
- ❑ ALICE specifications:
  - Specific tests for the VOBOX service
  - Tests for the CE and the CREAM-CE sensors also running with ALICE credentials
    - ✧ Test suite belongs to the ops VO
    - ✧ Site availability calculation performed on the CE sensor results basis ONLY
      - OBSOLETE. This calculation has to be redefined in terms of the CREAM-CE only
  - The important point was to migrate the VOBOX specific tests
    - ✧ STATUS: DONE!

- ✓ Migration of the test suite
- ✓ Upload of the ALICE software to the specific Nagios repository
- ✓ Build up of the software in the devel and testing Nagios infrastructure
- ✓ Rpm installed in the Nagios ALICE UI
- ✓ Population of the configuration files with the ALICE ROCs
  - ✓ CERN, Central Europe, Italy, NGI\_FRANCE, UKI, Russia
- ✓ Start up of the service
- ✓ ALICE is publishing the results in Nagios web interface since June the 8<sup>th</sup>

**Excellent support of the SAM/Nagios experts**

SERVICE\_VO-box (SERVICE\_VO-box)

Host	Status	Services	Actions
<a href="#">alibox.to.infn.it</a>	DOWN	5 OK 1 CRITICAL	
<a href="#">alice02.spbu.ru</a>	DOWN	6 OK	
<a href="#">alice11.spbu.ru</a>	DOWN	6 OK	
<a href="#">alicebox1.farm.particle.cz</a>	UP	6 OK	
<a href="#">alicegrid6.ba.infn.it</a>	UP	6 OK	
<a href="#">bovobox.bo.infn.it</a>	UP	6 OK	
<a href="#">ccicgalice01.in2p3.fr</a>	UP	6 WARNING	
<a href="#">ccicgalice02.in2p3.fr</a>	UP	6 WARNING	
<a href="#">clrvoboxalice1.in2p3.fr</a>	UP	6 OK	
<a href="#">clrvoboxalice.in2p3.fr</a>	UP	6 WARNING	
<a href="#">epgr03.ph.bham.ac.uk</a>	DOWN	6 OK	
<a href="#">gloci.itcp.ru</a>	UP	6 OK	
<a href="#">grid156.kfki.hu</a>	DOWN	2 OK 4 CRITICAL	
<a href="#">grinr004.inr.troitsk.ru</a>	UP	6 OK	
<a href="#">grinr07.inr.troitsk.ru</a>	UP	6 OK	
<a href="#">qt4.pnpi.nw.ru</a>	UP	6 OK	
<a href="#">house_grid.kiae.ru</a>	DOWN	6 WARNING	
<a href="#">ipnvobox.in2p3.fr</a>	UP	6 OK	
<a href="#">lcg53.sinp.msu.ru</a>	UP	6 OK	

<https://sam-alice.cern.ch/nagios/>

## FUTURE STEPS:

- Discrepancies found between SAM and Nagios in terms of results
- Specific CREAM-CE test suite
- Availability calculations
- Two additional tests for the VOBOX test suite required by ALICE

Host	Service	Status	Last Check	Duration	Attempt	Status Information
<a href="#">alibox.to.infn.it</a>	org.alice.vobox-PR-alice	CRITICAL	06-10-2010 11:24:33	0d 13h 0m 8s	1/2	Failed to execute /etc/init.d/alice-box-proxyrenewal status

- ❑ Issues daily reported at the ops. Meeting
- ❑ The weekly ALICE TF meeting includes now analysis items (TF & AF meeting)
  - Moved to 16:30 to contact with the American sites
- ❑ Latest issues at the sites
  - CAF nodes (T0)
    - ✧ Instabilities in some nodes have been observed in the last weeks
    - ✧ Thanks to the experts at the IT for the prompt answers and actions
  - AFS space (T0)
    - ✧ Replication of afs ALICE volumes
    - ✧ Separation in readable and writable volumes
  - CREAM-CE (all sites)
    - ✧ Bparser service down, connection refuses...
      - Once reported immediately solved, but daily observed

- ❑ Very smooth operations of all sites and services during the 2010 data taking
  - Fast reactions of site admins and experts in case of problems
- ❑ A new AliEnv2.18 has been deployed and will be the responsible of the data taking infrastructure of ALICE in the next months
  - Transparent deployment of the new version in parallel to the start up of the data taking
- ❑ Implementation of glexec in AliEn ongoing
- ❑ No remarkable problems found by the sites while migrating to CREAM1.6
  - However, in the last time we are seeing some instabilities (blparser service down, connection refused...)
    - ✧ Immediately solved when reported
    - ✧ Mandatory to change the availability calculation of the ALICE sites based on the CREAM-CE results
- ❑ New AliEn2.19 version foreseen for November 2010
  - 1<sup>st</sup> AliEn Jamboree in preparation (September 2010 at CERN)
- ❑ Ion data taking: end of November 2010