Enabling Grids for E-sciencE

CMS/ARDA activity within the CMS distributed system

Julia Andreeva, CERN
On behalf of ARDA group

CHEP06
Mumbai , February 2006

http://arda.cern.ch

Information Society

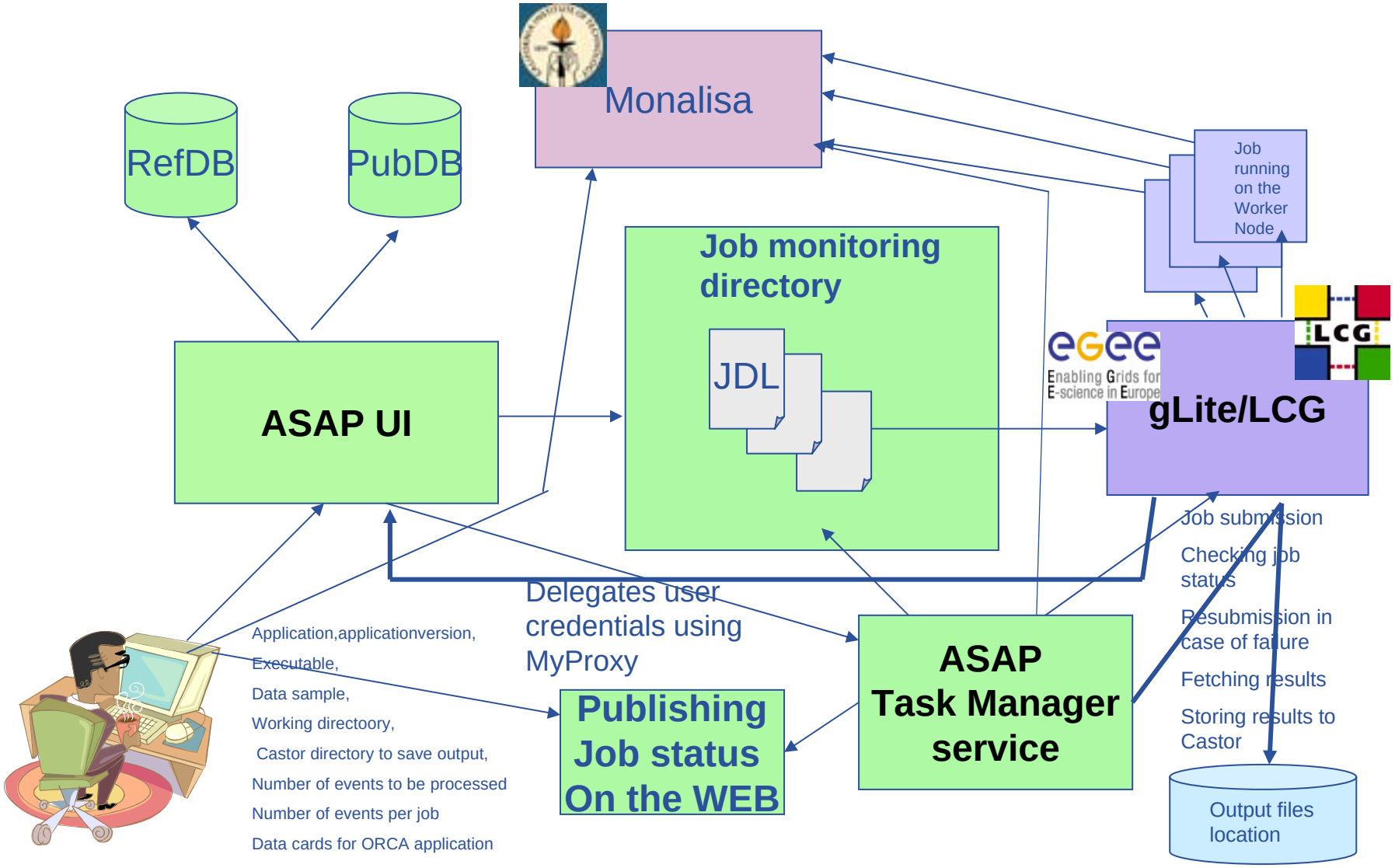
ARDA project

- **ARDA is an LCG project**
  - Main objective is to enable LHC analysis on the grid
  - Contribute to the distributed analysis activity of each LHC experiment
  - Build on top the new EGEE middleware (gLite)
    - ***Use the grid software as it matures (exposed to all pre-releases)***
    - ***Verify the components in an analysis environments***
    - ***Provide early and continuous feedback***
- **Two main directions of ARDA-CMS development**
  - Task Manager and Task Monitor for supporting user analysis on the Grid
  - Dashboard – monitoring system providing complete view of the Grid related and experiment specific information collected from the CMS distributed system

ASAP

- **ASAP stands for ARDA Support for Analysis Processing**
- **Started to develop end-to-end prototype enabling CMS analysis on the gLite middleware platform in May 2004**
- **Early phase of the gLite middleware development, very unstable and evolving environment**
- **Need feedback of the real users. How to attract them?**
- **Create the system which would make running analysis on the Grid for the user as simple as possible , requiring as little of time and attention as possible.**

ASAP work and information flow



Really less effort ?

- **Still user needs to**
  - *create a configuration file for a given task*
  - *run job generation command*
- **The rest can be done by the Task Manager**
  - *job submission*
  - *resubmission in case of Grid or application failure*
  - *generation of the web pages with the task monitoring information*
  - *saving of the output to the final location*
- **User wants to get a control over his task back and run in a simple mode without using a service**

  Whether it is possible?

  Yes, at any given moment
- **Whether users prefer to dump their work to the Task Manager rather then following themselves their tasks ?**

  Yes, they do prefer to use Task Manager service

Example of monitoring user task with ASAP

```
** 07/02/2006 16:51:57 => Setting Attribute jobid := log cms https://gdrb01.cern.ch:9000/dTpS9lc8Dx7yDoN-4ouHLA
** 07/02/2006 16:51:57 => Setting Attribute status := SUBMITTED
** 07/02/2006 16:51:57 => Setting Attribute gridstatus := submitted
** 07/02/2006 17:14:15 => Setting Attribute runnr := 196800019
** 07/02/2006 17:14:15 => Setting Attribute first_event := 0
** 07/02/2006 19:32:26 => Setting Attribute reason := Cannot plan: BrokerHelper: no compatible resources
** 07/02/2006 19:32:26 => Setting Attribute gridstatus := Aborted
** 07/02/2006 19:32:28 => Setting Attribute status := FAILED
** 07/02/2006 19:32:29 => Setting Attribute jobid :=
** 07/02/2006 19:32:29 => Setting Attribute gridstatus :=
** 07/02/2006 19:32:29 => Setting Attribute reason :=
** 07/02/2006 19:32:29 => Setting Attribute status := RESUBMIT_REQ
** 07/02/2006 19:33:14 =>
Select
Connec
Loggin

******

 The
 Use edg-job-status command to check job current status. Your job identifier (edg_jobId) is:

 - https://gdrb01.cern.ch:9000/BZuqyVdTWmM_WFEGCFISDg

**********************************************************************************************

** 07/02/2006 19:33:14 => Setting Attribute jobid := log cms https://gdrb01.cern.ch:9000/BZuqyVdTWmM_WFEGCFISDg
** 07/02/2006 19:33:14 => Setting Attribute status := SUBMITTED
** 07/02/2006 19:33:14 => Setting Attribute gridstatus := submitted
** 07/02/2006 22:23:55 => Setting Attribute ml_ev_total := 715
** 07/02/2006 22:24:00 => Setting Attribute reason := Job successfully submitted to Globus
** 07/02/2006 22:24:00 => Setting Attribute destination := ce101.cern.ch:2119/jobmanager-lcglsf-grid_cms
** 07/02/2006 22:24:00 => Setting Attribute gridstatus := Running
** 07/02/2006 22:24:01 => Setting Attribute status := RUNNING
```

Job was resubmitted by the Task Manager multiple times due to the Grid failures

Currently is running and had processed 715 events

196800022_0 DONE/OK/STORED Cleared #events: 1000 (1 -> 1000) CARF Total Errors 0 0 0

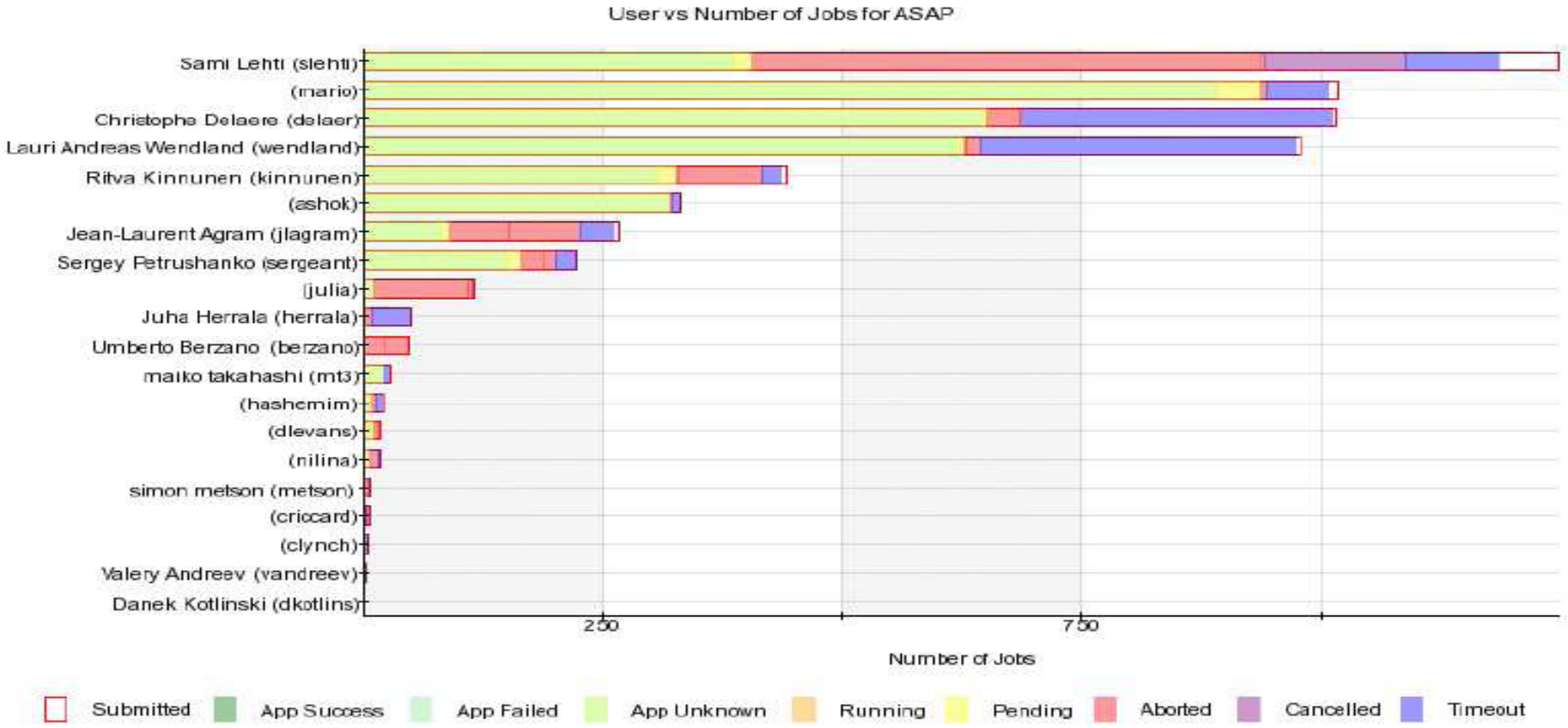
Use of ASAP

Submitted After 2006-01-25 10:50:30 , Submitted Before 2006-01-27 22:03:52 , Status At 2006-02-07 22:29:11 latest

Submit

Note: time format is yyyy-mm-dd or yyyy-mm-dd hh:ii:ss
Your current timezone is the same as UTC+1 (CERN), click here if you want to change it.

Jobs Status
show table: Yes plot: Bar Plot Note: How job status and success rate are calculated



Goal of the Dashboard project

- The CMS Dashboard project aims to provide a single entry point to the monitoring data collected from the CMS distributed system.
- Dashboard development is a part of the CMS integration project, a joined effort of ARDA and MonAlisa teams
- There is a plan to follow Dashboard development as a joined project for CMS and ATLAS

Currently, main activities are concentrated on the job monitoring part

The objective is provide a complete view of how the jobs of CMS VO are doing on the Grid

(both for LCG/gLite and OSG middleware platforms)

Use cases

- -Quantities (how many jobs per site, per user, per submission tool, per data collection…), distribution over time
- -Usage of the resources (CPU, memory, IO rates) aggregated on different levels
- -Sharing of the resources between production and analysis, different analysis groups, individual users
- -How Grid is behaving, success rate, failure reasons…
- -How CMS application is behaving , success rate , failure reasons…
- -How CMS data is distributed over the sites
- -Indication of the problems related to

Pure Grid troubles(lack of stability, scalability, performance)

Site configuration

Data publishing

Data access

Software distribution

…….

Sources of information

- **Currently two main sources of information are used:**

RGMA for getting Grid related info (LCG/gLite logging and book-keeping)

Investigating possibility to use GridIce and LCG2 Real Time Monitor Resource Broker XML Files to complement data got via RGMA

MonAlisa for CMS related data and system information

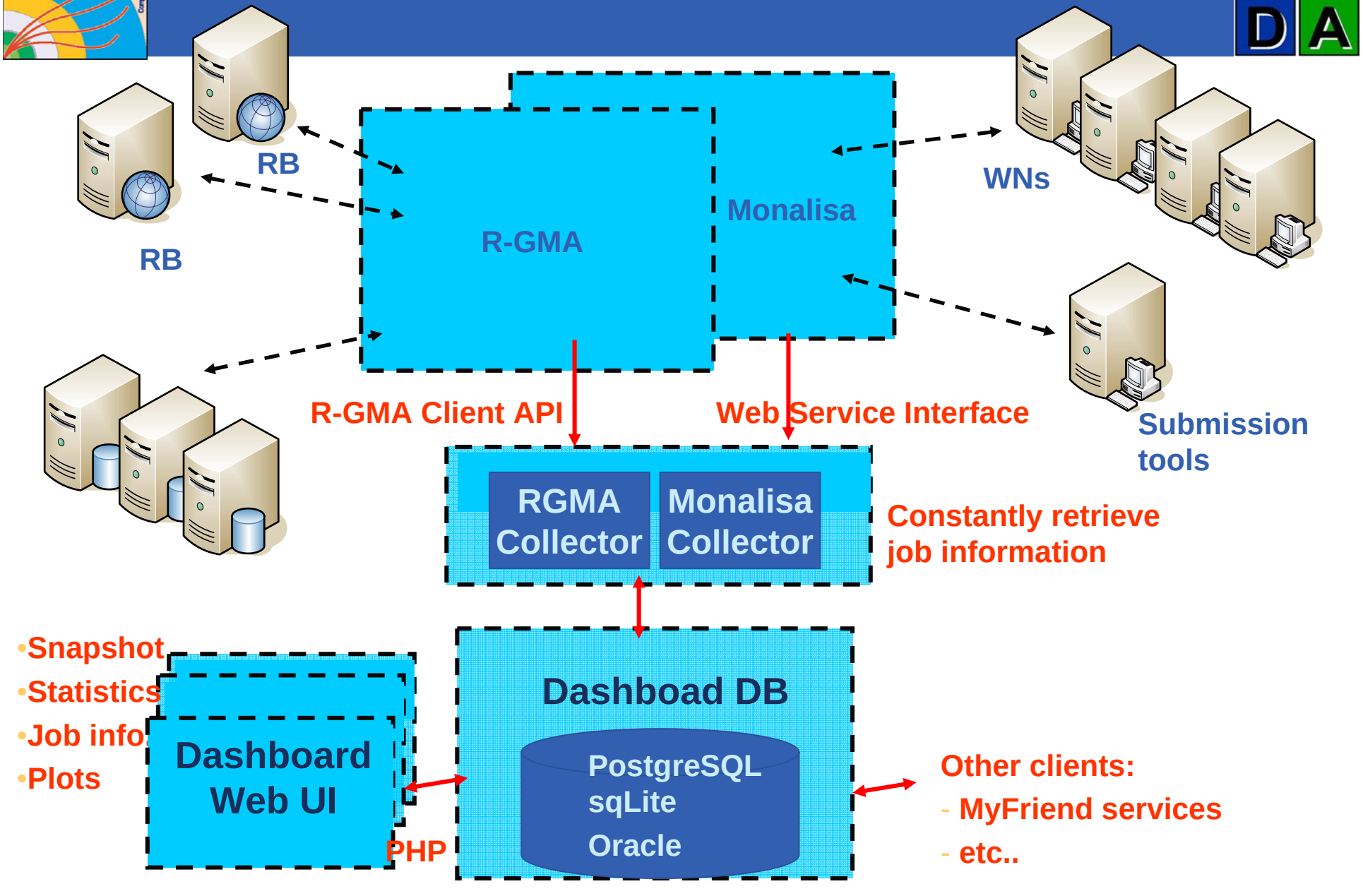
CMS has a very positive experience for using MonAlisa for system monitoring.

In ARDA we started to use MonAlisa for job monitoring.

There is a plan to use MonAlisa to get batch system related information for the OSG sites

RGMA and MonAlisa are complementary regarding the type of information they provide. However using two sources of information make possible crosschecking between the two . Our experience shows that it allows to increase the reliability of the system.

Architecture



Main principles

- **Provide a necessary level of interactivity**

  Not just precooked views , but possibility to dig in in case of troubles, possibility to ask detailed questions.

- **Foresee in future the possibility not only to collect and to expose information, but to analyze it and to generate alarms in case of evident problems**

- **Serve users with different level of requirements**

  - Managers (computing projects managers, site managers, production managers) - CMS global views, site views

  - Users running their tasks on the grid - task views

- **Follow input and feedback of the CMS community**

  - What kind of information to collect

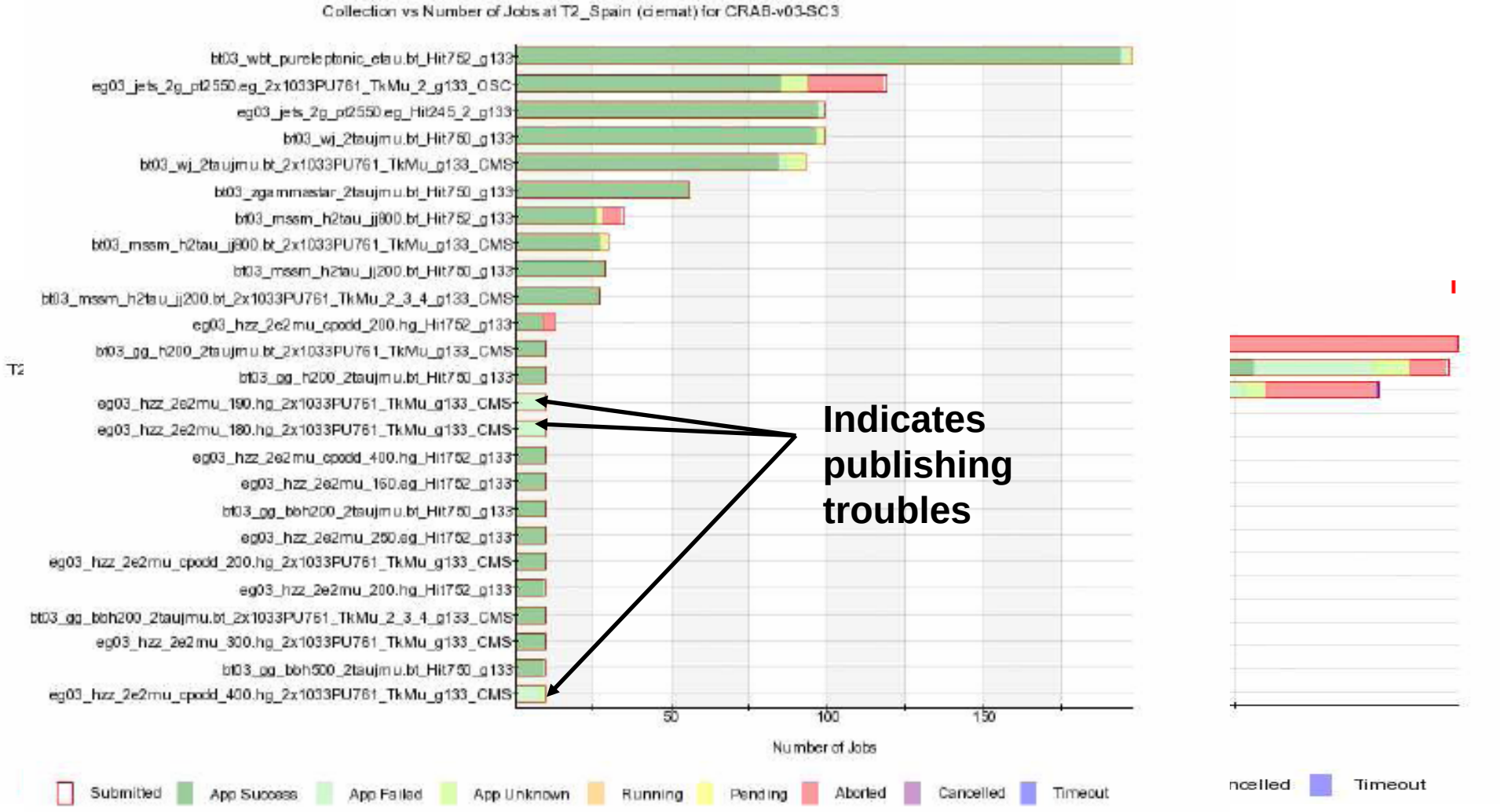
  - At what level to aggregate information

  - How to present collected information

Use of the Dashboard during SC3

ASAP

Arda Support for cms Analysis Process



Indicates publishing troubles

Use of the Dashboard during SC3

Conditions:
User: ALL, Site: ALL, Job Submission Tool: CRAB-v03-SC3, task: ALL, collection: ALL
Submitted Time: 2005-12-05 ~ 2005-12-08 , Status at 2006-02-01 17:11:04

Close

Scheduler JobId	site	Status Reason	JobMonitorId	TaskName	Events PerRun	Events Processed	Application Exit Status	Submitted at	Finished at
https://gdrb08.cern.ch:9000/U_qqUTSp4wl-zbAenq3pew	desy.de	user retrieved output san...	000046_131.169.223.107	SC3.051205.DESY.ExSi...	-	-	137	2005-12-05 15:59:33	2005-12-06 02:23:20
https://gdrb01.cern.ch:9000/_BKI5QBW_9ZFvOKobLR2CA	desy.de	user retrieved output san...	000077_131.169.223.117	SC3.051205.DESY.ExRo...	-	55	137	2005-12-05 16:14:37	2005-12-06 15:14:06
https://gdrb01.cern.ch:9000/jLDP-srA7nZU27pJSCfQ4Q	desy.de	user retrieved output san...	000047_131.169.223.114	SC3.051205.DESY.ExSi...	-	-	137	2005-12-05 16:34:55	2005-12-06 14:44:41
https://gdrb01.cern.ch:9000/nCnCW9NKdrxJEScfSBoMdQ	desy.de	user retrieved output san...	000074_131.169.223.92	SC3.051205.DESY.ExSi...	-	-	8	2005-12-05 16:44:44	2005-12-06 14:51:12

IO rates monitoring

ASAP

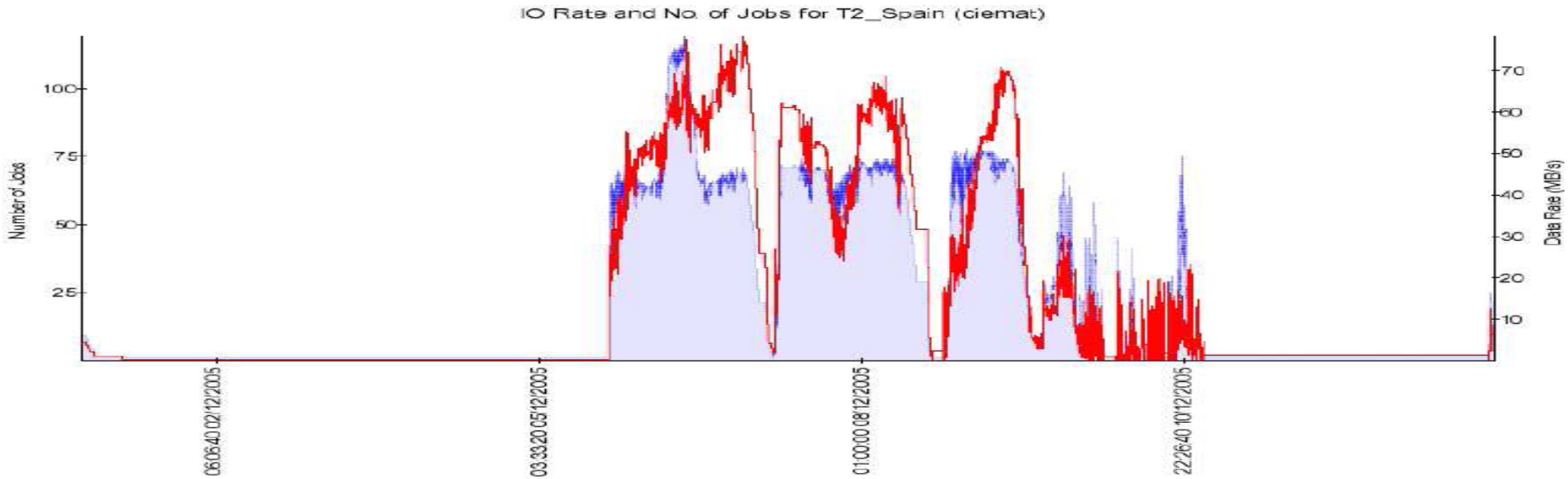
Arda Support for cms Analysis Process

NOTE: This may involve long queries, please be patient!

Dynamic IO Rate

User	All
Site	T2_Spain (ciemat)
Job Submission Tool	CRAB-v03-SC3
Task	All
Collection	All
Start Date (d/m/y HH:MM:SS):	01 | 12 | 2005 | 00 | 50 | 00
End Date (d/m/y HH:MM:SS):	15 | 12 | 2005 | 18 | 50 | 00

Submit Reset



Conclusions

- ARDA analysis system is used by the CMS physicists for running their analysis tasks on the Grid. Task Manager service aims to minimize the effort of the user and to increase the amount of successfully processed data. It got a positive feed back from the users and will be integrated in the new CMS analysis system.
- First prototype of the Dashboard and its use during CMS service challenge demonstrated need of the experiment in the common monitoring tool combining Grid and experiment specific information. LHC experiments are running on several middleware platforms, so the single entry point of monitoring information has an additional value.
- Still a lot of work has to be done to pass from the prototype to a production quality system providing a needed level of reliability, scalability and performance.

How to instrument jobs for MonAlisa reporting (backup slide)

- **At the submission time**

  - send meta information related to the task (application version, name of the executable, name of the input data collections…)

  - job related information (event range or other identifier of the job in the task scope, grid job ID)

- **From the worker node job wrapper should send**

  - results of the sanity checks

  - start and end flags of the job processing steps (staging in, running executable, staging out)

  - exit code of the execution of job processing steps

- **When job status is requested or job output is retrieved via job submission tool send to ML**

  - status and status reason string