Site monitoring and diagnosis tools

Andrea Sciabà

CERN

WLCG collaboration Workshop

1-2 September 2007, Victoria, BC

Status of the ARDA dashboard

- **Job Monitoring**
  - ➢ The "standard" view, useful to see how jobs are distributed, how they finish, to whom and to which activity they belong, etc.
- **Task Monitoring**
  - ➢ Allows CRAB users to see what's happening to their tasks
  - ➢ Support for MC production jobs will come later
- **Visualisation of SAM tests**
  - ➢ Will be THE web interface for the SAM tests and for the availability
  - ➢ Only a prototype for now
- **Monitoring of the I/O rate between WN and SE**
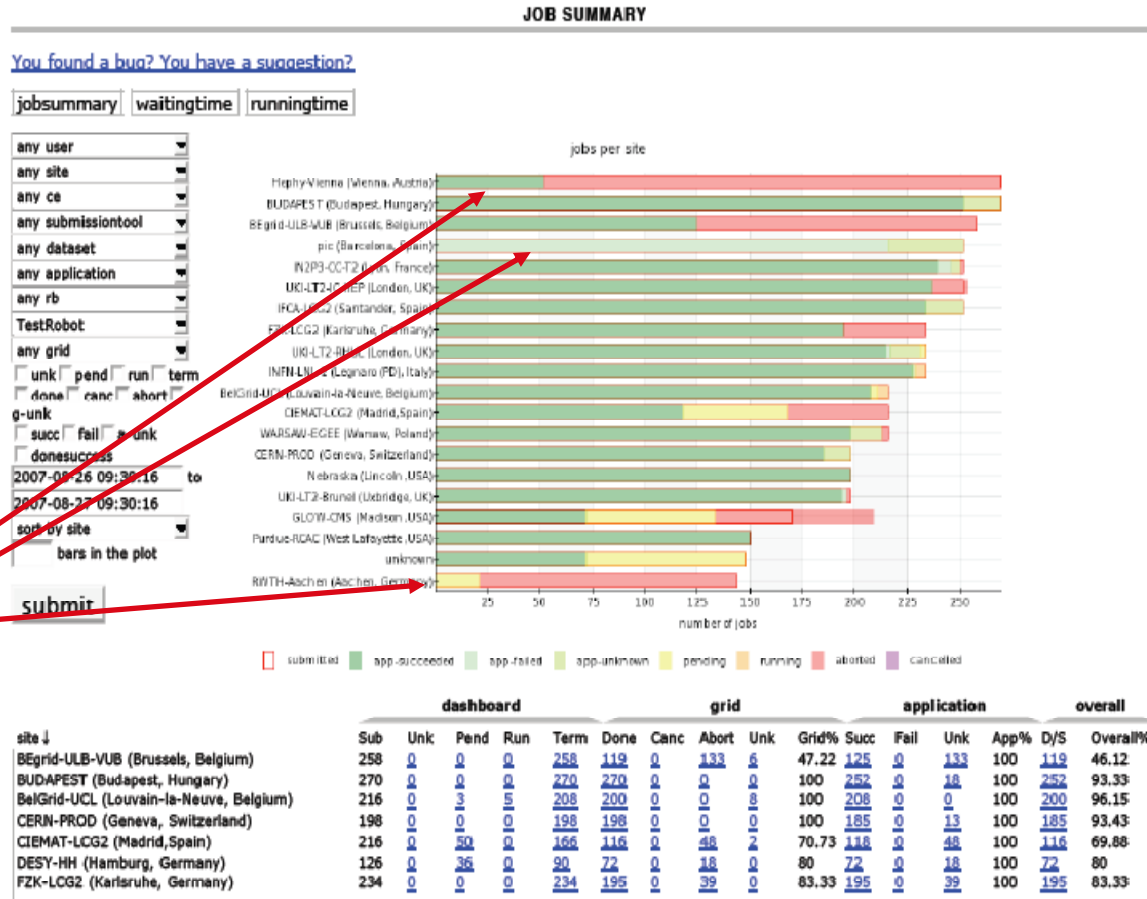 - ➢ Shows the I/O data rate generated by analysis jobs

Monitoring of the JobRobot jobs using Dashboard interactive UI
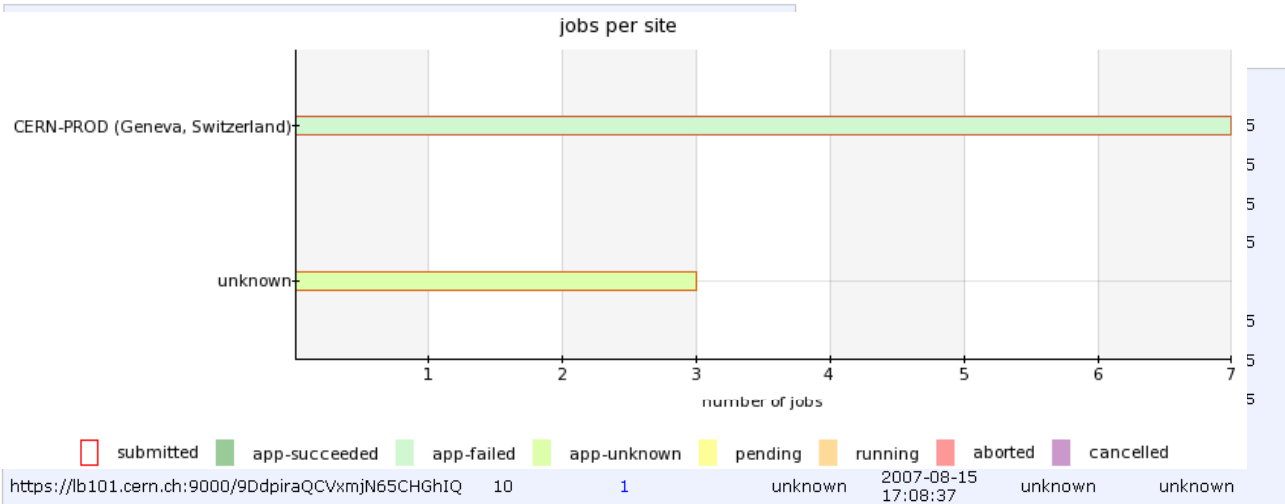
Sites having troubles

Task monitoring

SAM tests

- **Site and service availability based on the results of SAM tests**
- **Results of SAM tests are imported in real time to the dashboard DB**
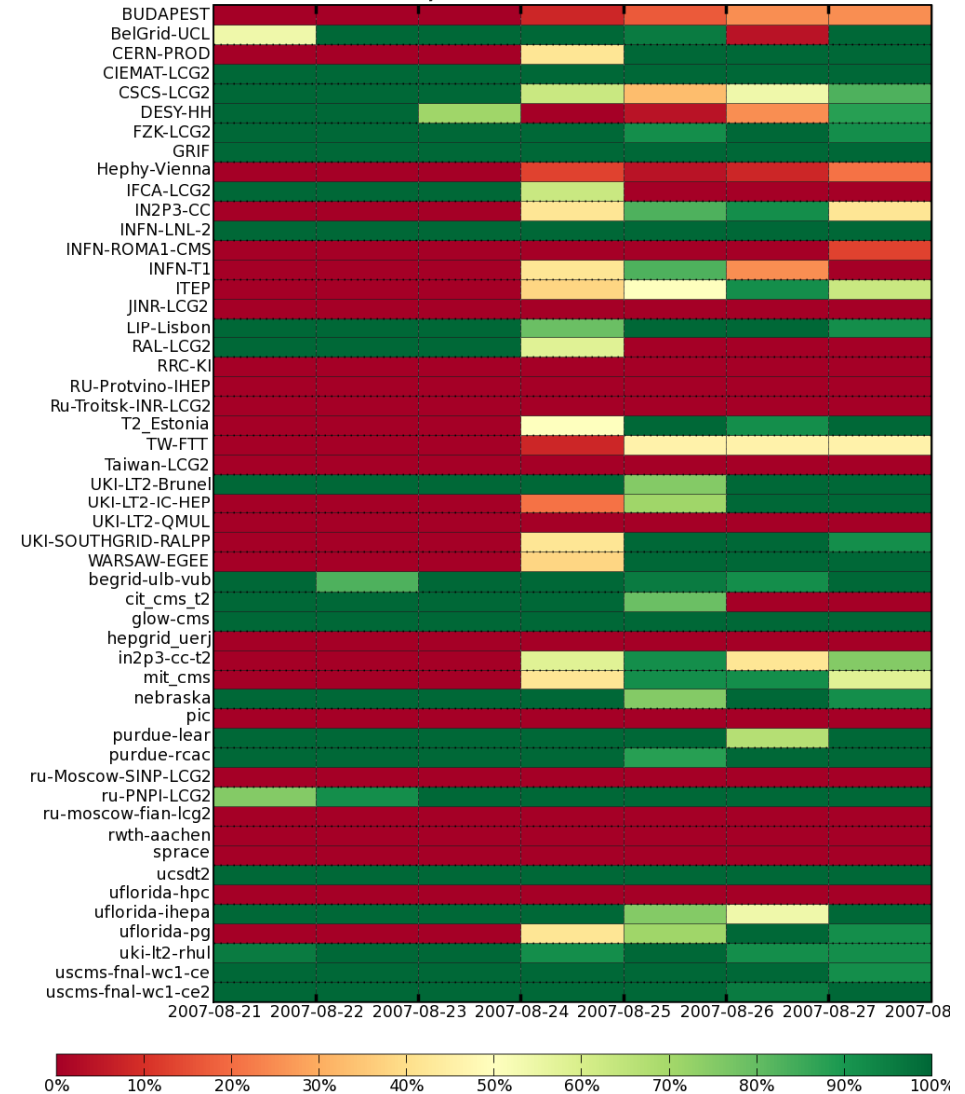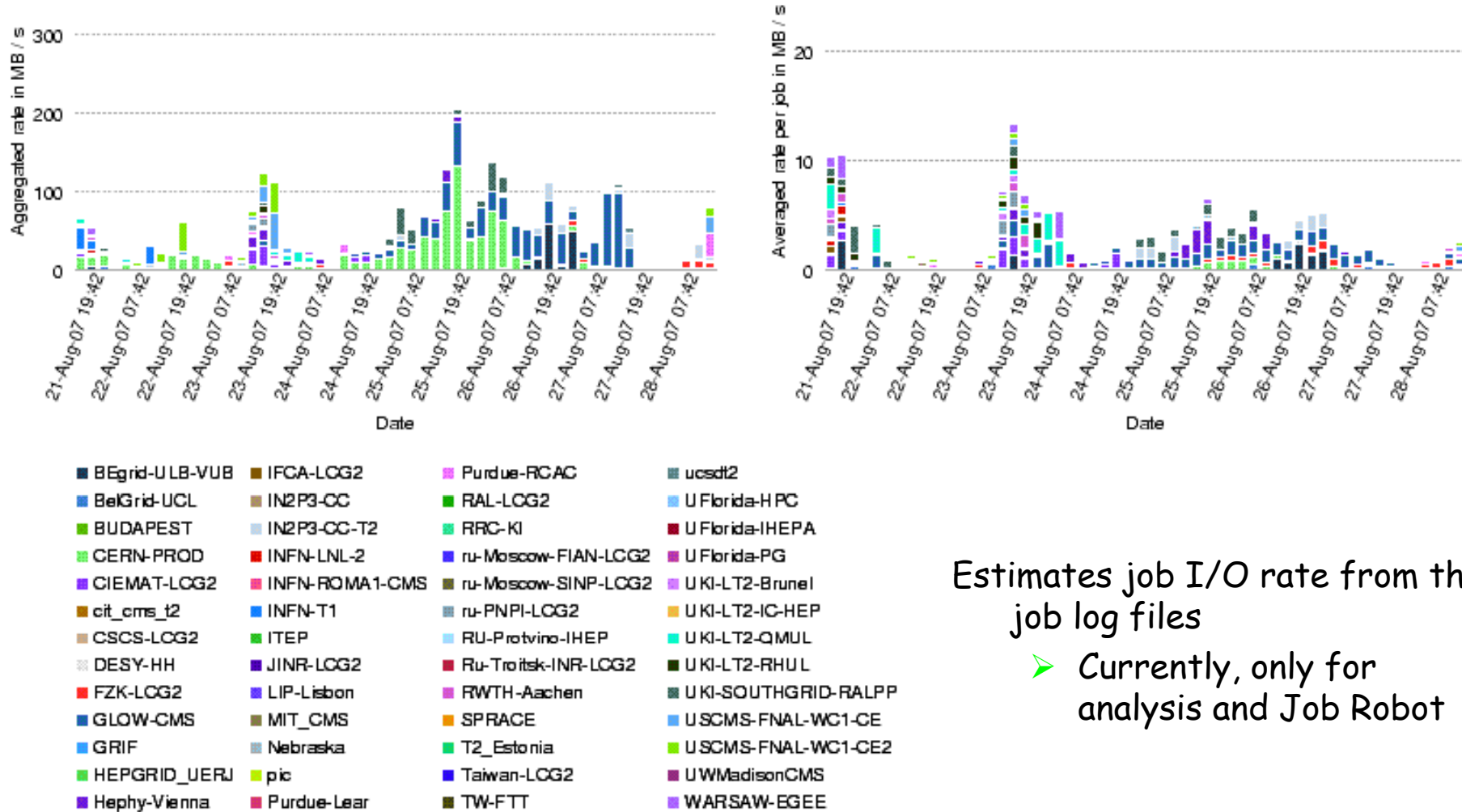
Estimates job I/O rate from the job log files

➢ Currently, only for analysis and Job Robot

Dashboard plans

- UI to the production monitoring data, first prototype should be available end of September – middle of October

- Site/service availability interface based on the results of SAM tests. First prototype exists, still lot of improvements required to increase the level of interactivity, add possibility to calculate availability for different use cases, "custom" availability (end of September – middle of October)

- On the job interactive interface and task monitoring interface add the possibility to correlate job failures with the results of SAM tests

- Improvements to troubleshooting in case of application failures

- Continuation of work started with Job Robot (implemented by Olga Kodolova and Marcelo Baquero). In case of application failures more information from stderr and stdout is sent from the analysis job, which allows to improve diagnostics on the dashboard UI

- Enabling of better monitoring of jobs submitted via Condor-G due to instrumentation of Condor-G submitter for job status information reporting (first prototype of modified Condor-G version by end of October)

Status of the SAM tests

- Critical tests
  - Definition: a test is critical when its failure causes the resource to "disappear" from the BDII from the VO's point of view (unless it has been "whitelisted" in FCR)
  - A test should be critical only if its failure means that the resource is unusable
  - They are by definition the only ones taken into account by the **WLCG availability** calculation
  - Now all the CMS critical tests are ops tests, but there is nothing fundamental with that and CMS can choose any SAM test as critical
- FCR
  - All SEs are whitelisted in FCR
  - The CEs depend on the critical tests
    - ☞ Exception: the OSG CEs are whitelisted
- CE custom CMS tests
  - Test software area, CMSSW installations, Frontier/squid server, running from the WN
  - Typically not critical because their failure affects only specific types of jobs
- SRM custom CMS tests
  - Test the LFN→PFN mapping from TMDB, then `put`, `get metadata`, `get`, `advisory delete` with *srmcp*

CMS critical tests in FCR	
Test	**VO**
Computing Element	
CA certs	ops
CSH test	ops
Job submission	ops
VO tag management	ops

CMS site availability

- It is calculated only based on the CMS custom tests for the CE
- The current definition is very different from the WLCG availability
  - It is just the daily fraction of successful CMS custom tests run at the site!
  - This estimate is optimistic compared to the WLCG algorithm
  - Even if the same algorithm is used, remember that the tests considered are different ones!
- Different meanings:
  - WLCG avail.: says if the resource was usable at all
  - CMS avail.: says if the resource was usable by specific CMS tasks

SAM Site Quality
30 Days from 2007-07-08 to 2007-08-07 UTC
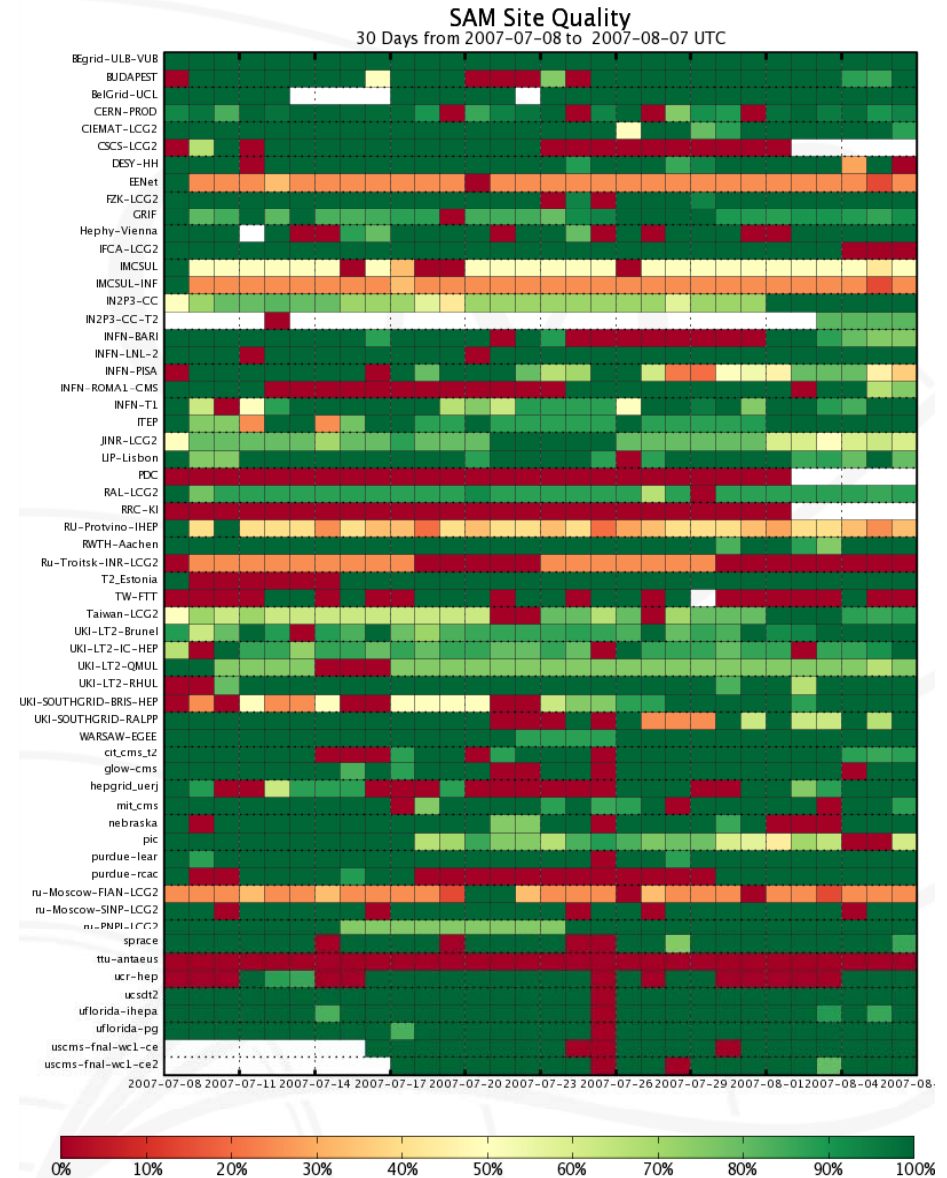
Issues with SAM

- None from the framework's point of view
- CE tests
  - The current tests are rather stable
  - Missing a test to "simulate" an analysis job (e.g. a test that tries to read few events from a dataset)
  - Need to identify all the "real life" problems which are not exposed by SAM and add new tests
- SE/SRM tests
  - The "standard" ops tests for the SE are of limited use for CMS, as they use the lcg_util commands and depend on LFC (and on the BDII)
  - Is it still useful to have an SE sensor at all? Probably not
  - The SRM tests work nicely, but they have a problem: they all depend on the success of the LFN $\rightarrow$ PFN test, which depends on what is written in the TMDB
    - ☞ CNAF had a very low WLCG availability for CMS in July because SRM always failed the TMDB test!
- FCR
  - Ok, apart from the fact that OSG sites have disappeared from it (but they will be back soon)

Sites

- **What the sites should be looking at every day**
  - From the page https://twiki.cern.ch/twiki/bin/view/CMS/SAMForCMS
    - ☞ Status of SAM CMS tests for all EGEE sites/OSG sites
      - To see how the CE tests are going
    - ☞ Status of the SRM tests on all EGEE/OSG sites
  - Is the "job submission" test failing?
    - ☞ Look at the logging info of the job from the SAM page
    - ☞ Check the GOC Wiki troubleshooting guide
      - http://goc.grid.sinica.edu.tw/gocwiki/SiteProblemsFollowUpFaq
  - Are the "real" CMS tests failing (basic, swinst, MC, squid, …)?
    - ☞ Look at the test outputs in SAM
    - ☞ Look in the SAMForCMS wiki
  - And again, don't forget the SRM tests!
  - And what about the availability?
    - ☞ Unfortunately the plot is not yet updated on a regular basis, but anyway check it to see how your site looks
- **Finally, write to hn-cms-comp-prep@cern.ch if you don't know how to solve a problem**

CMS

- Things to be improved
  - Visualization…
  - Documentation
    - ☞Build a knowledge base on the SAM errors
  - Whatever else the sites think is missing or unsatisfactory