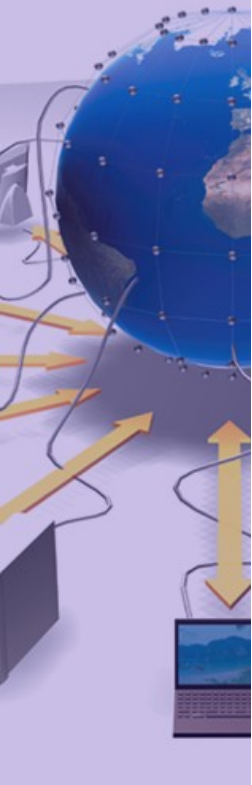


Monitoring: Tracking your tasks with Task Monitoring

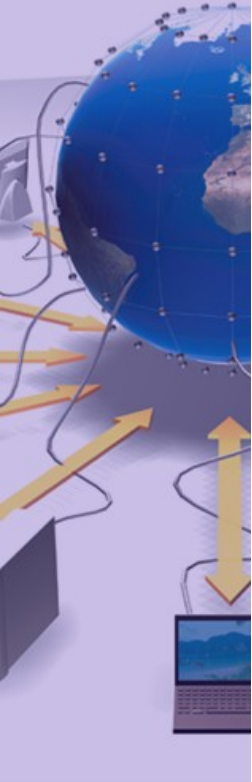
PAT eLearning – Module 11

Edward Karavakis

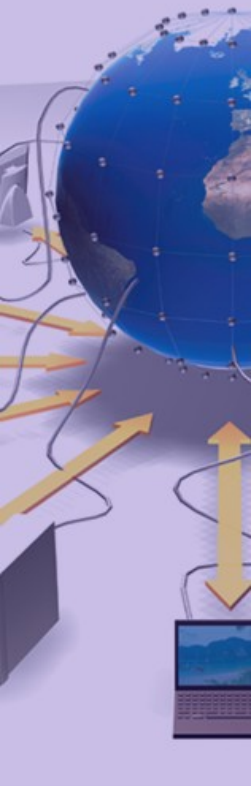
On behalf of the Dashboard Team



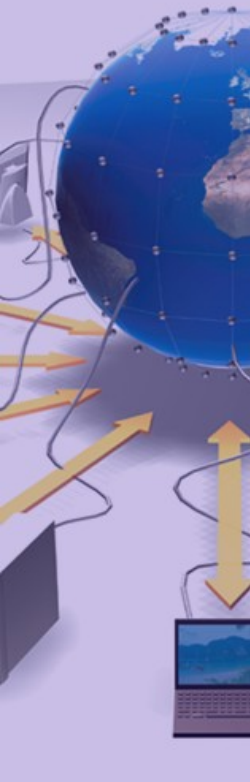
- You should be able to monitor your analysis tasks – without any hassle!
- Task Monitoring collects and exposes a user-centric set of information to the user regarding submitted tasks.
- Part of the Dashboard Framework:
 - <http://dashboard.cern.ch>
- Uses the job monitoring information from the Dashboard database.
- Available at: <http://dashboard.cern.ch/cms>



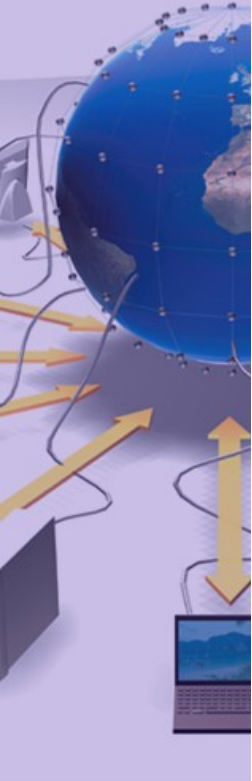
- Focused on the user's perspective.
- Easy to use and to navigate.
- Intuitive in layout.
- Fast with very low latency.
- Updates in 'real time'.
- Bookmark your favourite tasks.
- Offers a wide selection of graphical plots.
- User-driven development.



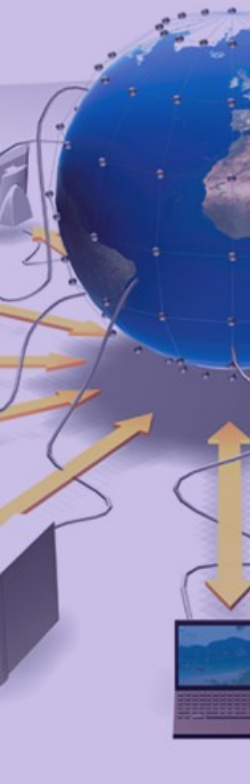
- We don't have your grid certificate – cannot directly query the Grid Logging and Bookkeeping System.
- We rely on the job status sent to the Dashboard either from the jobs themselves from the CRAB UI via MonALISA or on the job status information on RGMA and ICRTM.
- Provides monitoring functionalities regardless of the submission method or the middleware.



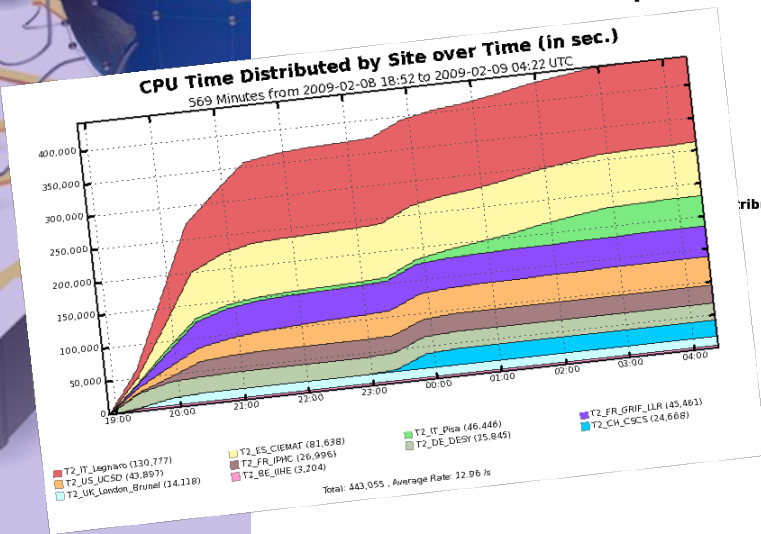
- What happens if you see an inconsistency between Dashboard Task Monitoring and crab -status ?
- As soon as the job is finished at the worker node it is reported to Dashboard as finished. CRAB reports a finished job only when it is considered DONE by the Grid and normally, a small delay is introduced by the Grid Services.
- But, if you notice that some data is missing regarding your jobs, you should believe crab -status.



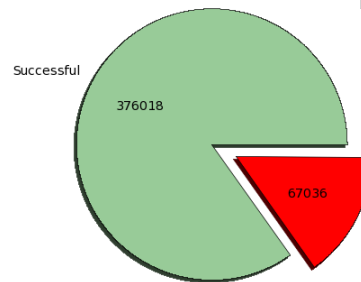
- Many UI improvements.
- URLs to every task, user and particular views of a given task.
- Consumed Time information:
 - Task-specific: *Average Efficiency, total & average CPU and Wall Clock time usage, average CPU time per event.*
 - Job-specific: *Efficiency for a specific job.*
- Extended the selection of the plots.



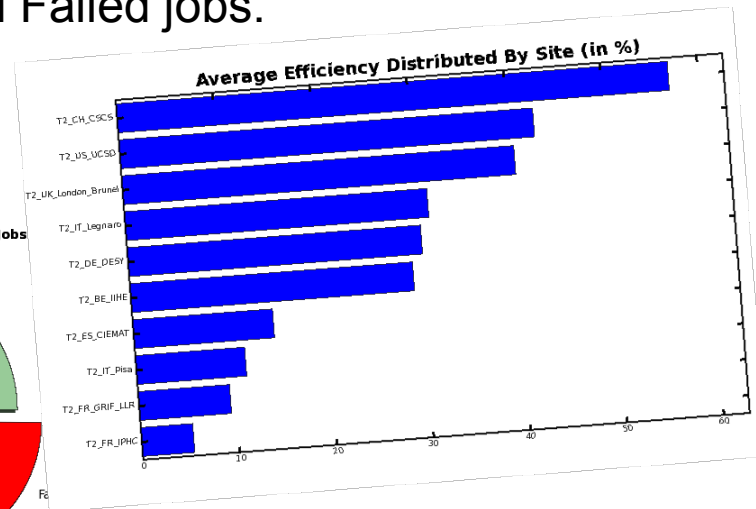
- Various Plots available, including:
 - Distribution by Site (successful, failed, running and pending, processed events).
 - Terminated jobs (in terms of Success/ Failures and over time).
 - Application-failed and Grid-aborted jobs by Reason of Failure.
 - Timing plots: Average Efficiency distributed by Site, CPU & Wall Clock time spent on Successful and Failed jobs.



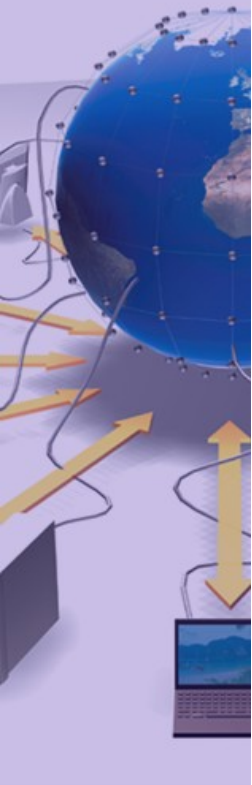
Distribution of CPU Time spent on Successful and Failed Jobs



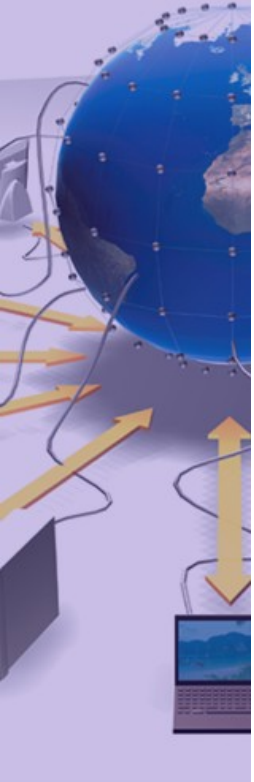
■ Successful (376019) ■ Failed (67036)



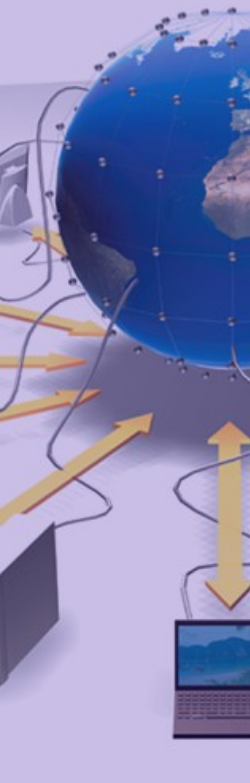
- Fully configurable time interval.
- Filters allowing you to search for detailed information about a specific task or job.
- Automatic generation of commands for: resubmissions / killings / getting logging info, retrieving output, ..



- The weakest point is failure diagnostics for both Grid and Application failures.
- It would be extremely useful to get not only the exit-code, which sometimes can be misleading, but a detailed reason of failure as well. *i.e.:* *'Could not save output file A on the storage element B'*.
- Requires modifications of the CRAB Wrapper.
- A user shouldn't have to search the log files to understand what went wrong.



- The Monitoring tool is available at:
<http://dashboard.cern.ch/cms>
- You can make it better!
- Please send us your suggestions and feedback at: dashboard-support@cern.ch



- Short demonstration of the application.

