

LHCb Analysis Test at Tier1's

Or “Running a user Application over large amounts of input data”

By R. Graciani, A. Puig
for the LHCb Production Team

Overview

- Original Idea:
 - Understand current limitations for Data access by User Analysis Jobs
- Procedure
- Running the test
- Some Improvements
- Results
- Summary

Procedure

- Select a user with a working analysis code and need to run over a large data sample:
 - Albert Puig, preparing the code for the calorimeter calibration with first data.
- Select Large data samples:
 - Inclusive B samples, replicated to Disk storage at all Tier1's.
- Job submission and analysis of output.
- Repeat the submission with different samples periodically (2 – 3 times a week).

Running the Test

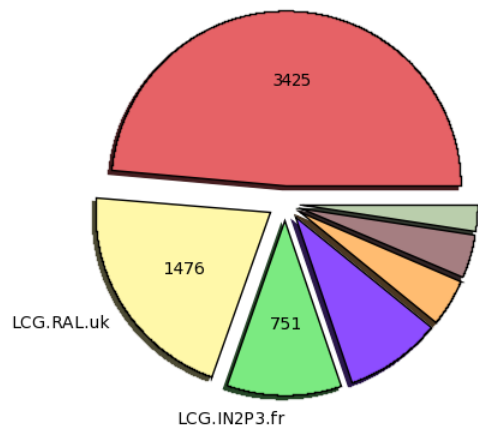
- Each submission consist of:
 - ~ 600 jobs, reading 100 different files with 500 events per file.
 - Average file size 200 MB
 - **1 Job = 100 Files = 50.000 events ~ 20 GB ~ 1-2 hours~ 2-3 MB/s**
- InputData:
 - User submits list of LFNs
 - DIRAC checks LFC for replicas and check availability at the sites.
 - Job receives SURLS.
 - **SURLS are converted to tURLs by asking to SRM (getturls)**
- Application:
 - Root base application open the files in sequence, reads and analyzes all events **directly from the SE (without copy to WN)**.

Some Improvements

- After the initial iterations we noticed that dCache sites (GRIDKA, IN2P3, NIKHEF, PIC) performance was significantly worse than for other sites (CERN, CNAF, RAL).
 - Work to tune:
 - DCACHE_RAHEAD (finally set to 1)
 - DCACHE_RA_BUFFER (finally set to 102400)
 - (thanks PIC for feed back and investigation).

Results (Global)

Total Number of Jobs by Site (Sum: 7046)
LCG.CERN.ch

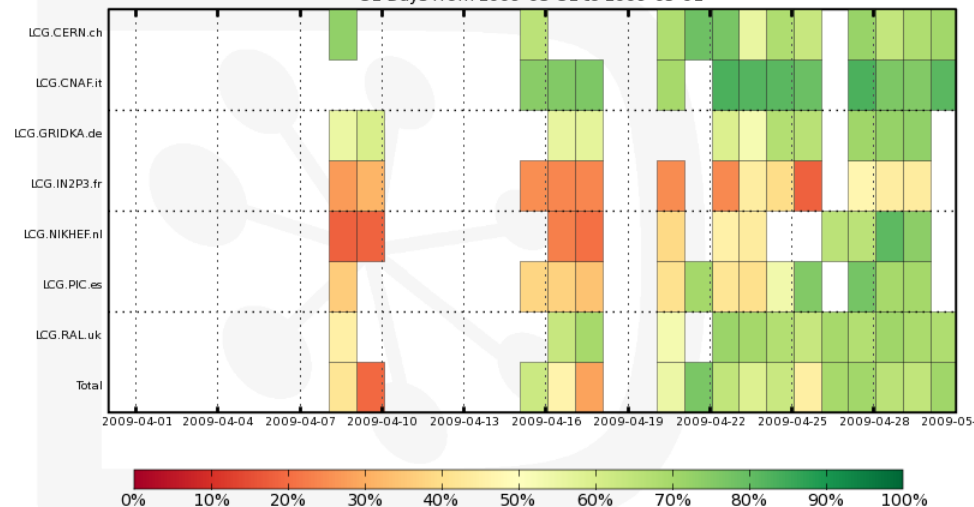


■ LCG.CERN.ch (3425)
 ■ LCG.RAL.uk (1476)
 ■ LCG.IN2P3.fr (751)
 ■ LCG.CNAF.it (630)
 ■ LCG.NIKHEF.nl (308)

■ LCG.GRIDKA.de (281)
 ■ LCG.PIC.es (175)

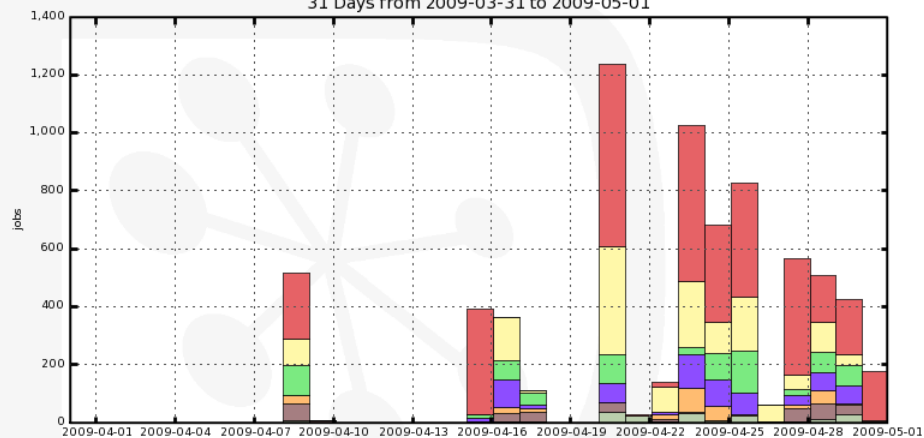
Total Number of Successful Jobs

Job CPU efficiency by Site
31 Days from 2009-03-31 to 2009-05-01



CPU/WallClock for Successful Jobs

Jobs by Site
31 Days from 2009-03-31 to 2009-05-01



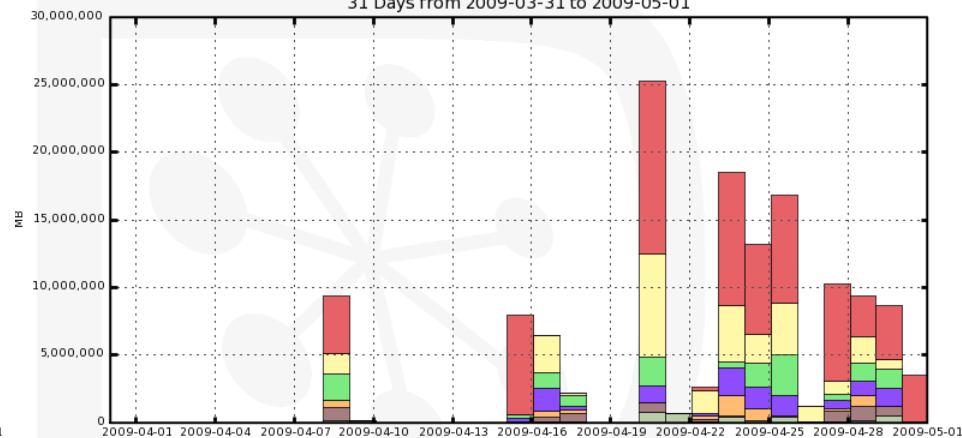
■ LCG.CERN.ch
 ■ LCG.RAL.uk
 ■ LCG.IN2P3.fr
 ■ LCG.CNAF.it
 ■ LCG.GRIDKA.de

■ LCG.NIKHEF.nl
 ■ LCG.PIC.es

Maximum: 1,235 , Minimum: 0.00 , Average: 220.31 , Current: 174.06

Daily Number of Jobs by Site

Input data by Site
31 Days from 2009-03-31 to 2009-05-01



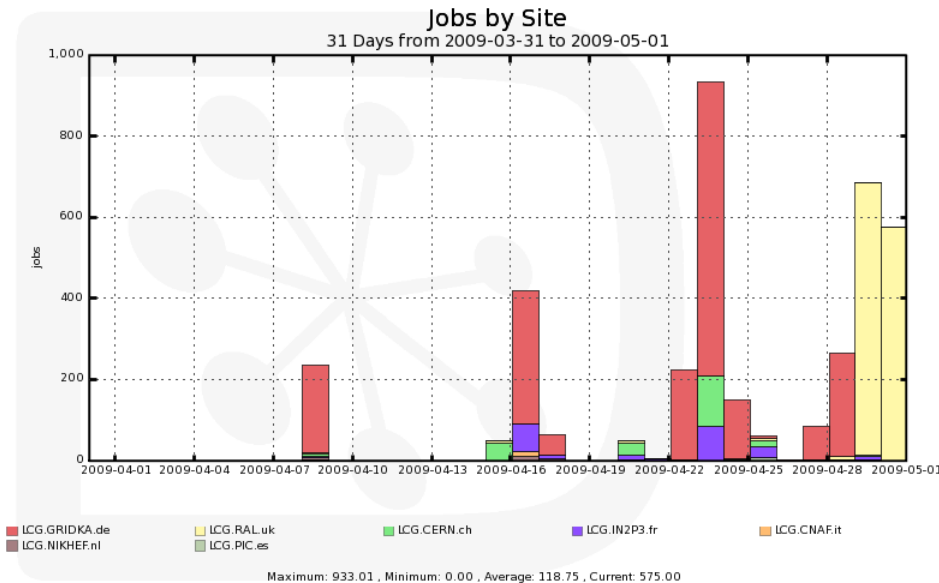
■ LCG.CERN.ch
 ■ LCG.RAL.uk
 ■ LCG.IN2P3.fr
 ■ LCG.CNAF.it
 ■ LCG.GRIDKA.de

■ LCG.NIKHEF.nl
 ■ LCG.PIC.es

Maximum: 25,285,244 , Minimum: 0.00 , Average: 4,246,034 , Current: 3,486,462

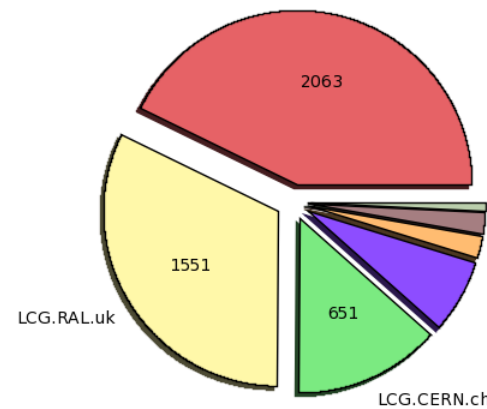
Daily Amount of Data read by Site

Results (Failures)



Daily Number of Failed Jobs by Site

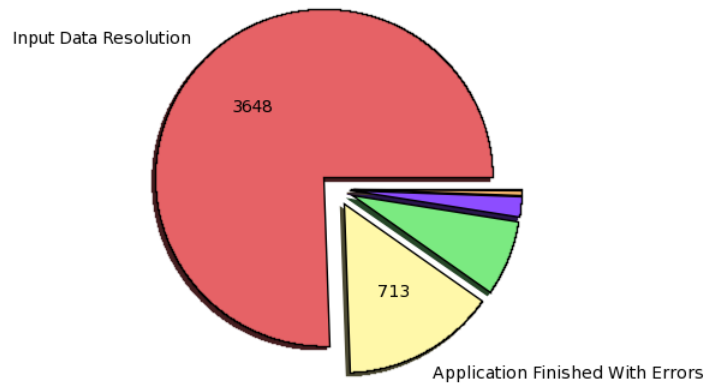
Total Number of Jobs by Site (Sum: 4827)
LCG.GRIDKA.de



LCG.GRIDKA.de (2063) LCG.RAL.uk (1551) LCG.CERN.ch (651) LCG.IN2P3.fr (329) LCG.CNAF.it (101)
LCG.NIKHEF.nl (94) LCG.PIC.es (38)

Total Number of Failed Jobs by Site

Total Number of Jobs by FinalMinorStatus (Sum: 4829)

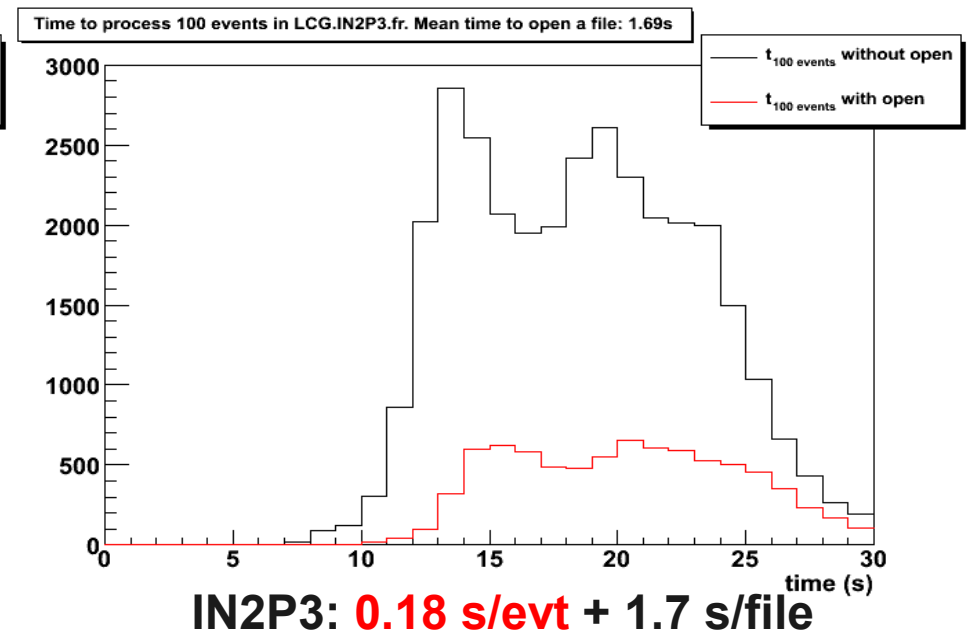
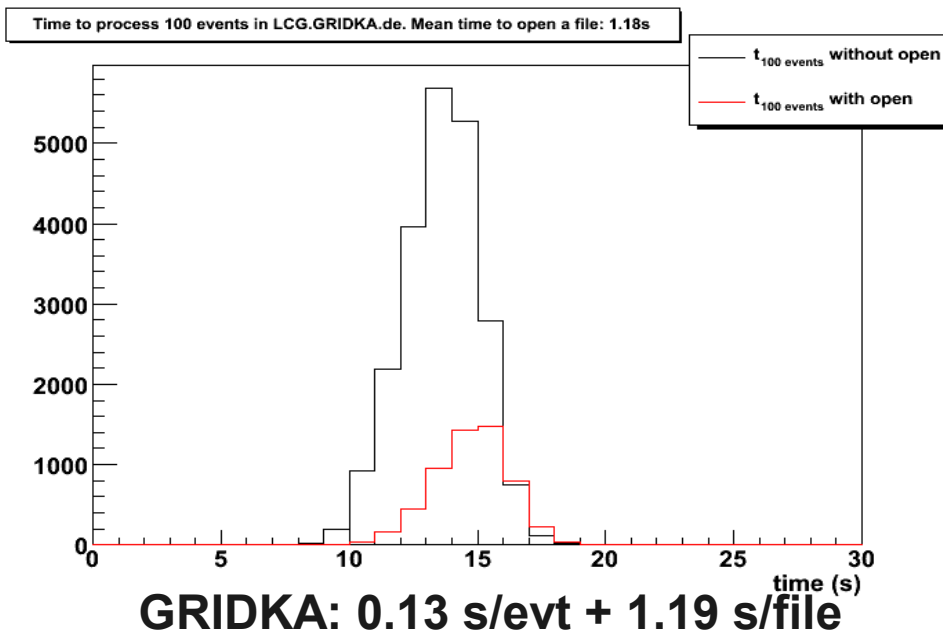
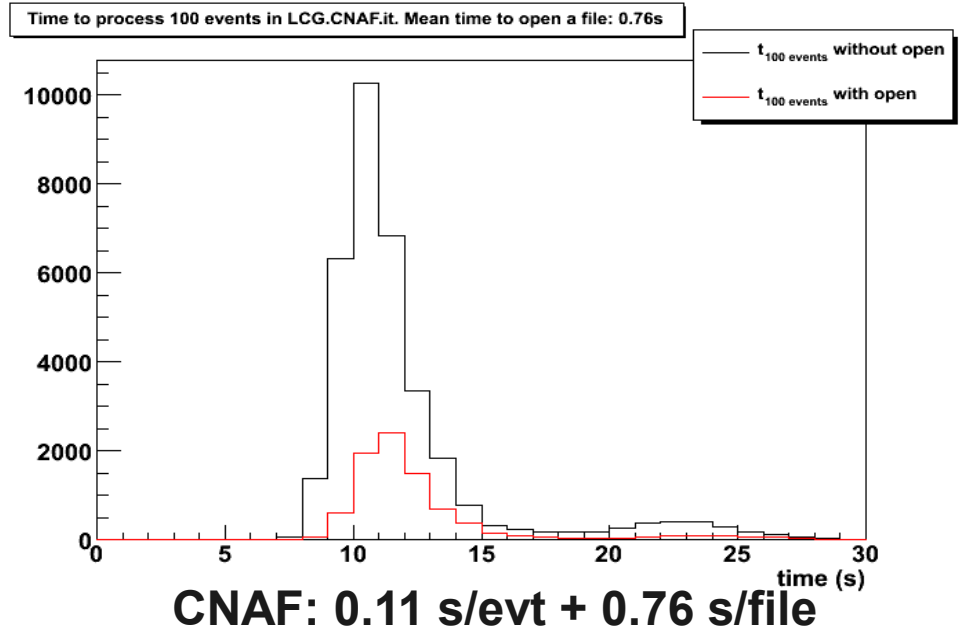


Input Data Resolution (3648) Application Finished With Errors (713)
Uploading Job Outputs (350) Watchdog identified this job as stalled (93)
Input Sandbox Download (23) Exception During Execution (2)

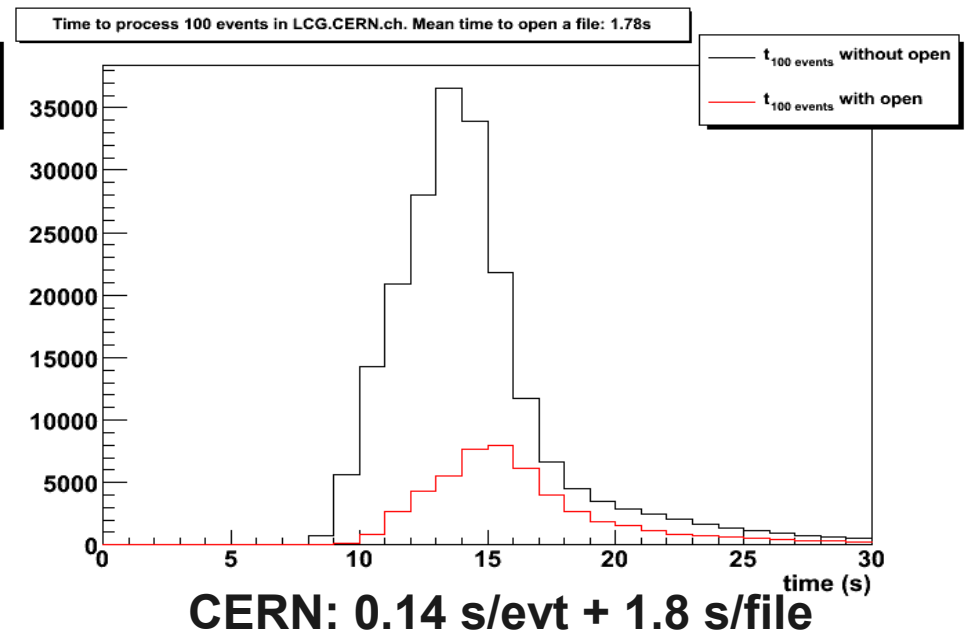
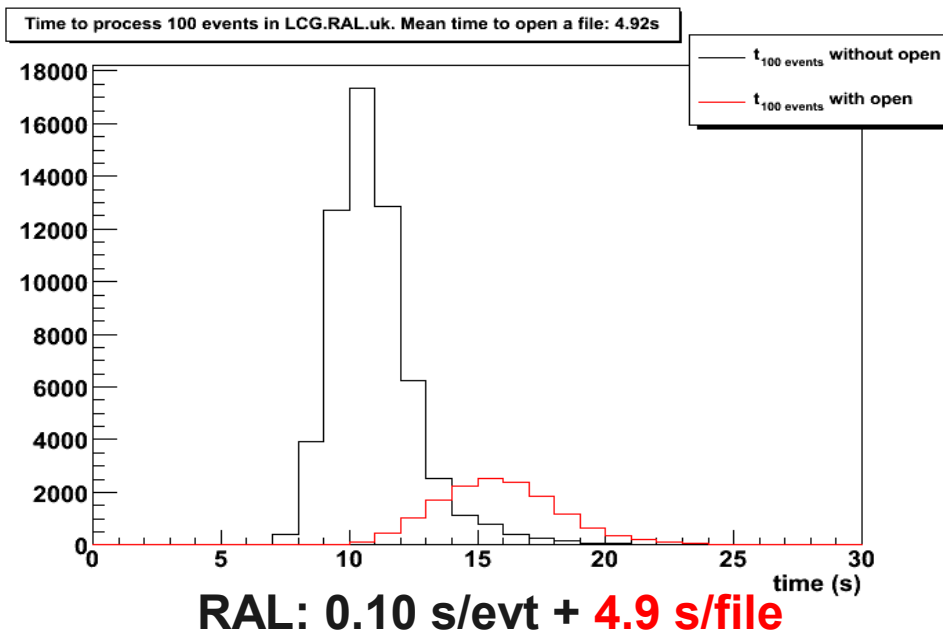
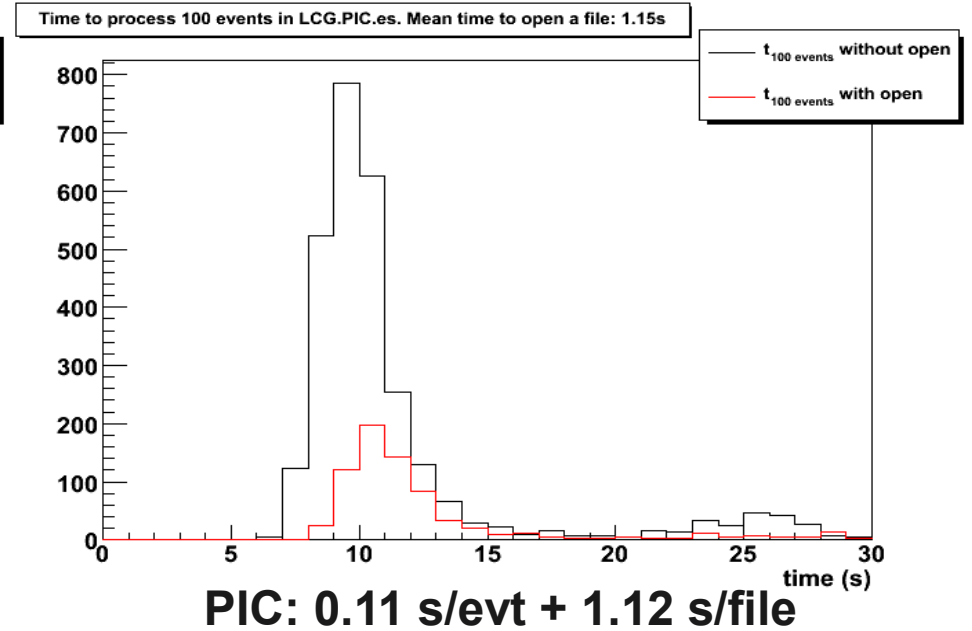
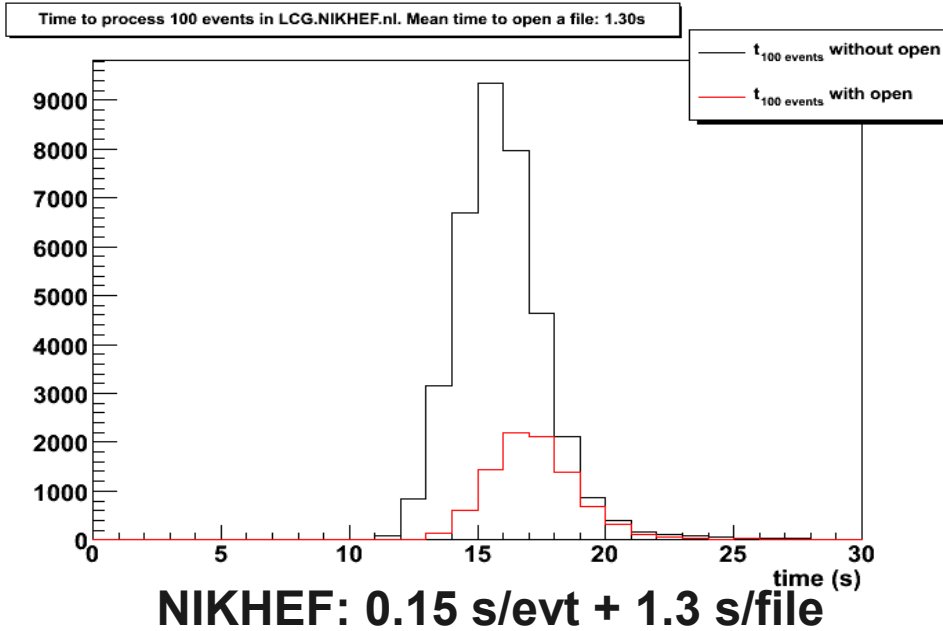
Total Number of Failed Jobs By Reason

Results (Detail Performance)

- Distribution of Wall Clock time to process 100 consecutive events.
- **Distribution of Wall Clock time to process 100 consecutive events including a new file opening.**
- The difference of the means represents the file opening time.

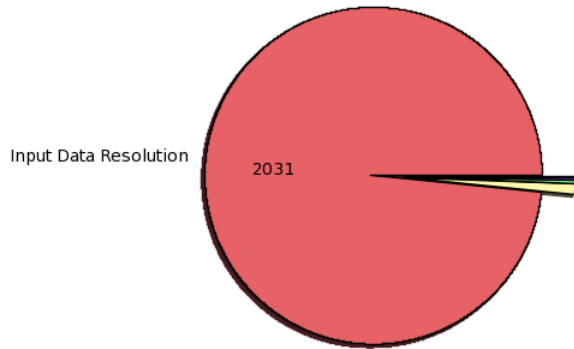


Results (Detail Performance)



Results (Failure details)

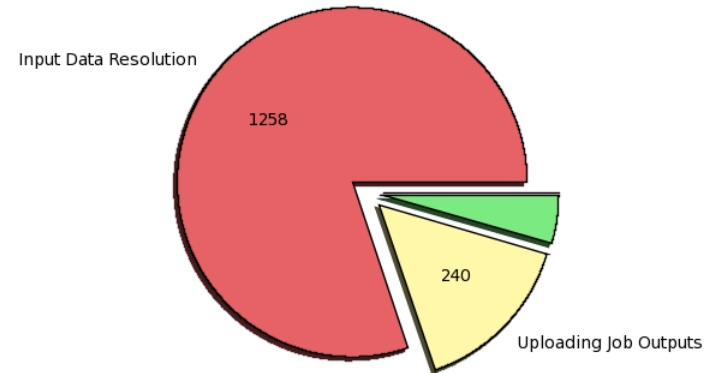
Total Number of Jobs by FinalMinorStatus (Sum: 2063)



■ Input Data Resolution (2031)
 ■ Application Finished With Errors (20)
 ■ Uploading Job Outputs (6)
 ■ Watchdog identified this job as stalled (5)
 ■ Input Sandbox Download (1)

Failure Reasons at GRIDKA

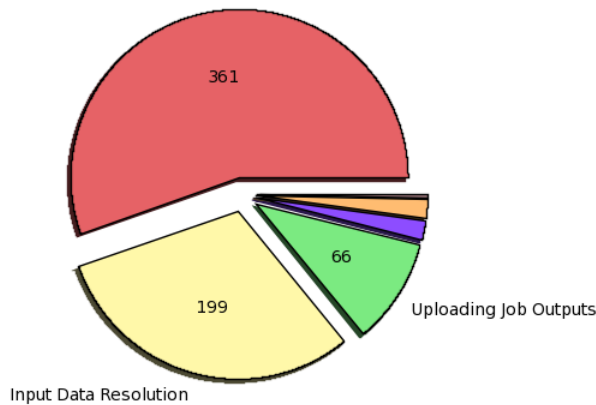
Total Number of Jobs by FinalMinorStatus (Sum: 1569)



■ Input Data Resolution (1258)
 ■ Uploading Job Outputs (240)
 ■ Application Finished With Errors (70)
 ■ Input Sandbox Download (1)

Failure Reasons at RAL

Total Number of Jobs by FinalMinorStatus (Sum: 652)
Application Finished With Errors



■ Application Finished With Errors (361)
 ■ Input Data Resolution (199)
 ■ Uploading Job Outputs (66)
 ■ Watchdog identified this job as stalled (12)
 ■ Exception During Execution (2)
 ■ Input Sandbox Download (12)

Failure Reasons at CERN

- Only most problematic sites shown.
- Dominated by SRM problems at GRIDKA and (recently) at RAL.
- Application Errors are typically due to errors while reading the input file.

Summary

- First “controlled” running of user analysis jobs.
- Promising results:
 - After tuning dCache average CPU efficiency in the 70-80 % range.
- Manage to Read up to 25 TB of data/day.
- Pending Issues:
 - Poor Performance of SRM at GRIDKA
 - Poor Performance of SE at IN2P3
 - Large File opening time at RAL
 - **Would like to test xrootd (volunteers)**
- Excellent performance of CNAF for file opening
- Any Question? Comments?

IN2P3 observations

- Poor CPU Efficiency at IN2P3 (< 50%)
- Large variation in time to process events
- Seems to be Job to Job variations
- Needs further investigations

