

Distributed Analysis Status Report

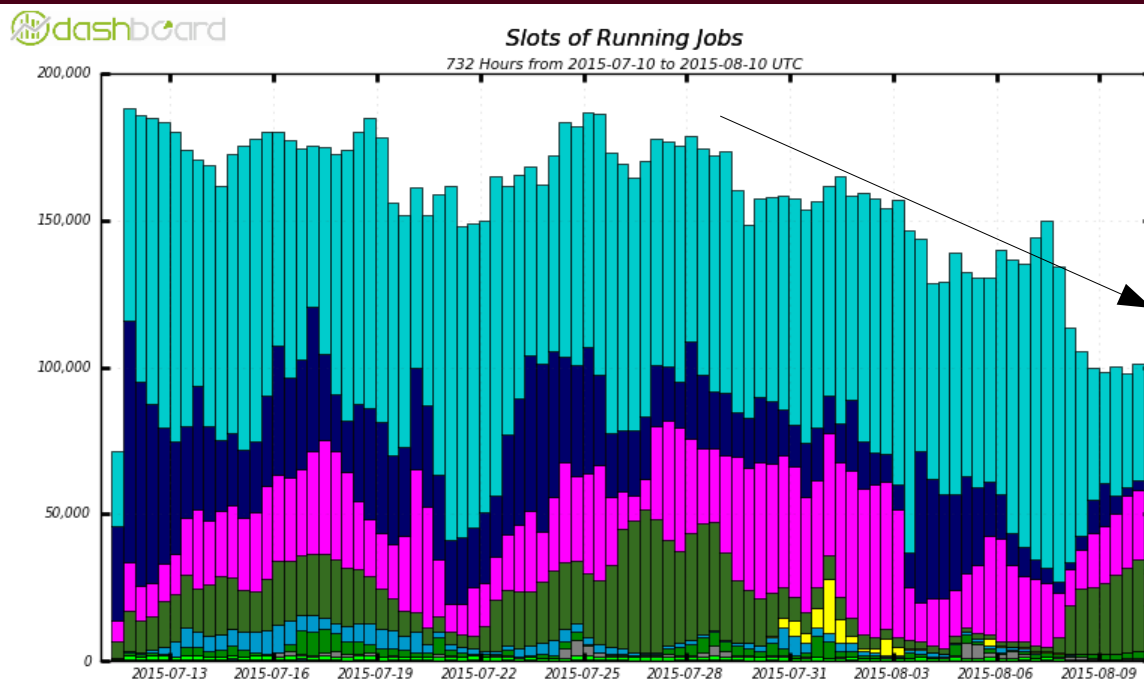
(+ some HC stuff)

A. Dewhurst (RAL), F. Legger (LMU)

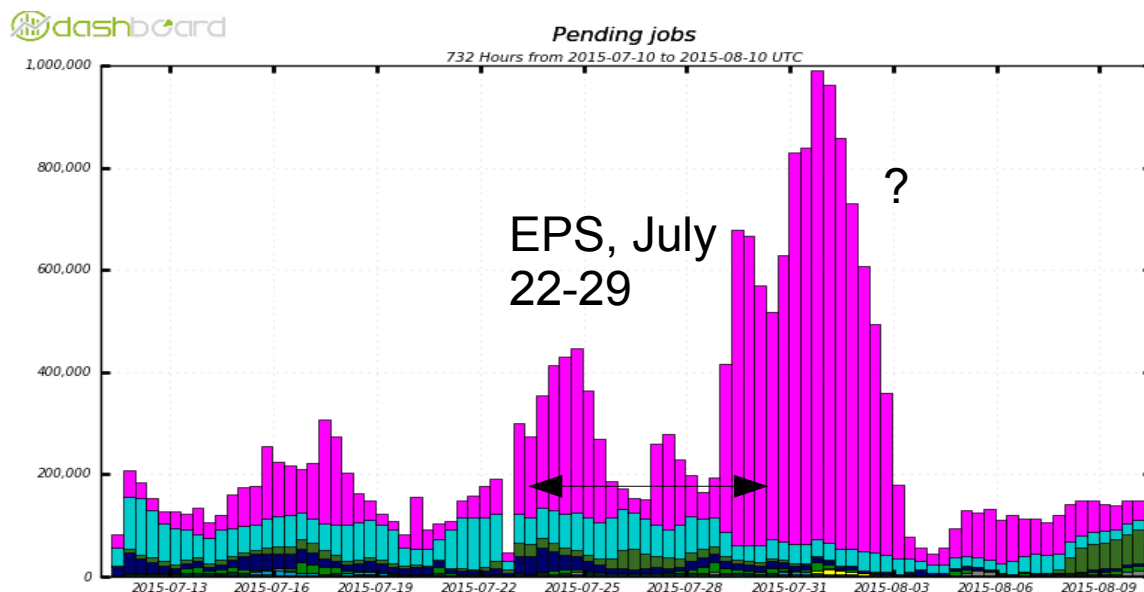
(+ HC team)



All jobs – last month – all T1s and T2s



- Most resources used for MC simulation
- Group production still small
- General decrease of slots used last week
- Usual picture: analysis usage in bursts, corresponding peaks of pending jobs



Maximum: 990,875 , Minimum: 0.00 , Average: 252,918 , Current: 71,906

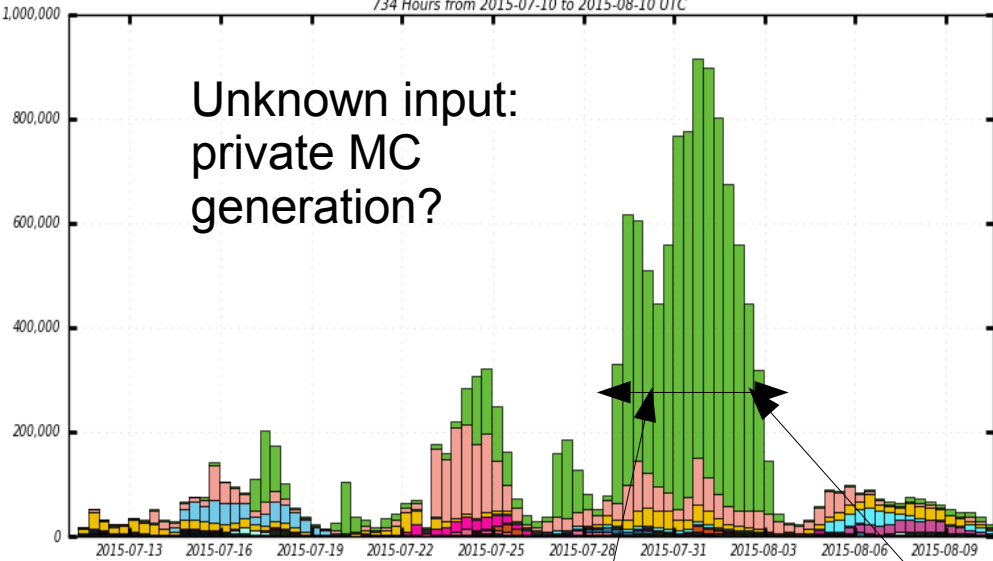
Investigating ?



Pending jobs

734 Hours from 2015-07-10 to 2015-08-10 UTC

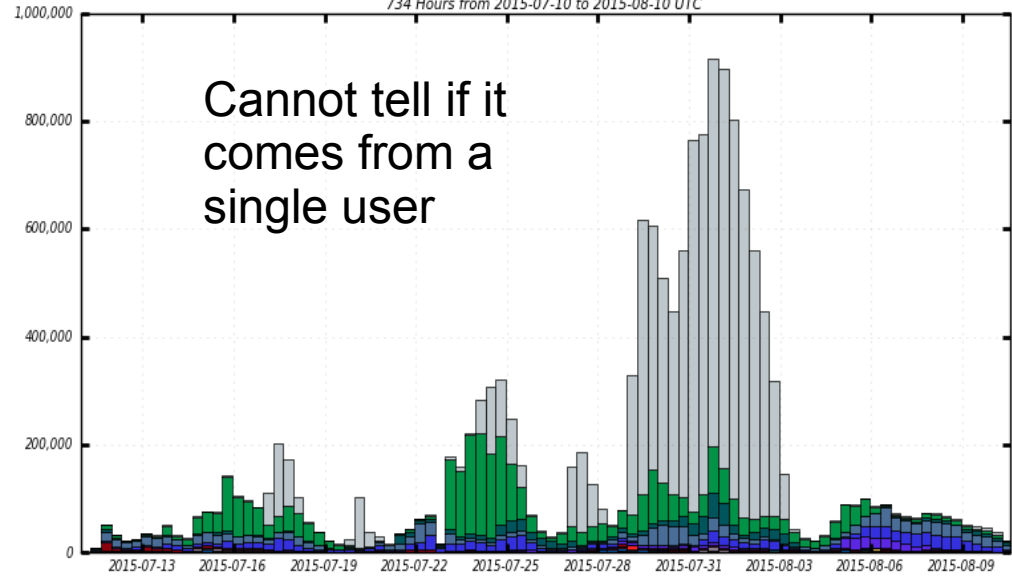
Unknown input:
private MC
generation?



Pending jobs

734 Hours from 2015-07-10 to 2015-08-10 UTC

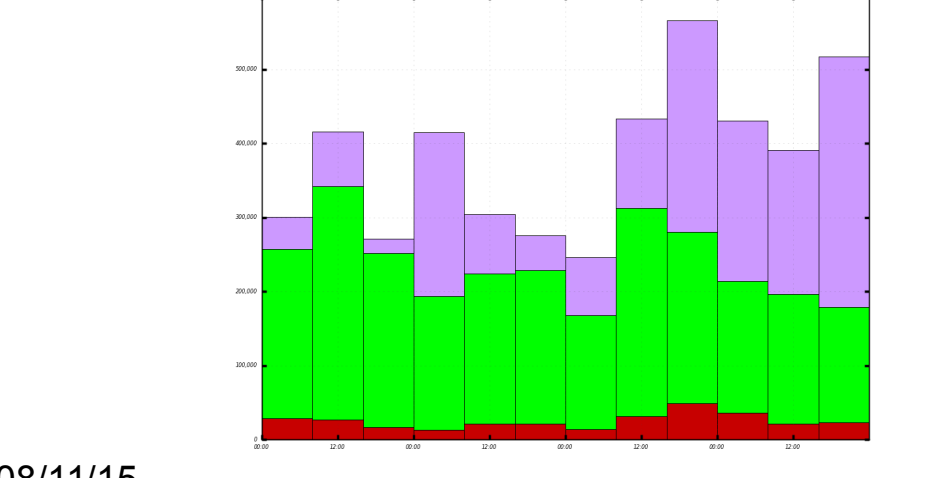
Cannot tell if it
comes from a
single user



- unknown
- NTUP_SMWZ
- NTUP_BTAGEFF
- DAOD_SUSY1
- DAOD_HIGG6D1
- DAOD_HIGG4D3
- DAOD_TAUP3
- NTUP*
- RAW
- NTUP_TOP
- DAOD_ONIAMUMU
- NTUP_COMMON
- ESD
- user
- NTUP_HSG2
- DAOD_TOPQ1
- DAOD_EXOT13
- DAOD_HIGG4D3
- DAOD_HIGG8D1
- DAOD_EXOT11
- DAOD_SUSY9
- NTUP_SUSYSKIM
- EVNT
- DAOD_JPSIMUMU
- NTUP_SUSY
- DAOD_HIGG2D1
- DAOD_HIGG4D2
- DAOD_JETM1
- DAOD_TAUP1
- DAOD_SUSY2
- DAOD_TRUTH1
- AOD
- NTUP_TAU
- RDO
- DAOD_HIGG6D2
- *
- DAOD_SUSY10
- DAOD_HIGG5D1
- DAOD_JETM1
- DAOD_TAUP1
- DAOD_SUSY2
- DAOD_SUSY3
- DAOD_SUSY4
- HITS
- NTUP_HI
- NTUP_TRUTH
- NTUP_2LHSG2
- NTUP_2LHSG5
- DAOD_HIGG5D2
- DAOD_HIGG2D4
- DAOD_EXOT4
- DAOD_JETM9
- plus 66 more

- unknown
- mc15_14TeV
- mc12_2TeV
- data15_cos
- data11_7TeV
- mc10_7TeV
- mc15
- data12_8TeV
- data15_13TeV
- mc12_14TeV
- data13_hip
- group10
- valid3
- mc15_valid
- user
- mc14_13TeV
- group
- mc11_7TeV
- valid1
- data13_2p76TeV
- mc11_14TeV
- mc15_13TeV
- data10_hi
- mc12_5TeV
- mc12_14TeV
- mc11_2TeV
- valid2
- mc15_5TeV
- data15_comm
- mc12_8TeV
- data11_hi
- mc14_8TeV
- panda
- mc12_7TeV
- mc12_valid
- mc09_7TeV

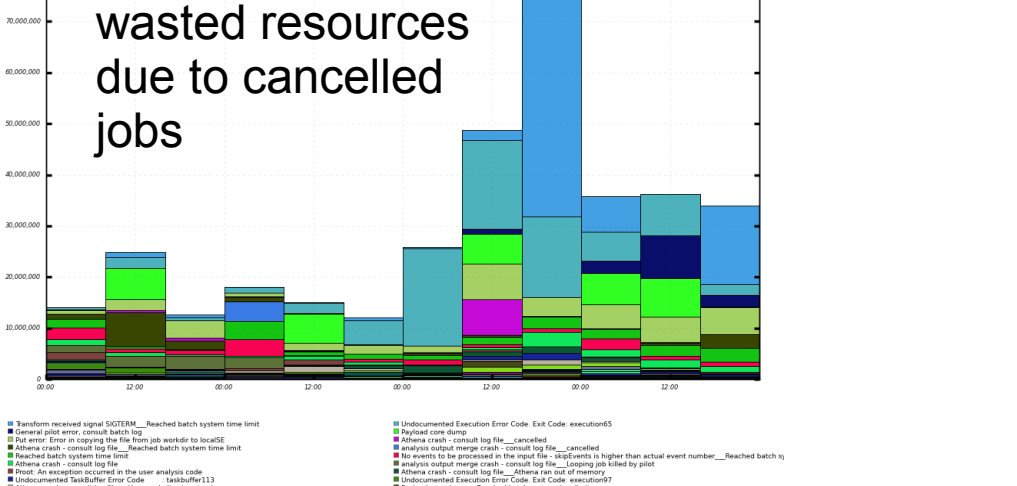
95 Hours from 2015-07-29 00:00 to 2015-08-01 23:59 UTC



CPU Consumption of Panda Failed jobs by ExitCode

95 Hours from 2015-07-29 00:00 to 2015-08-01 23:59 UTC

wasted resources
due to cancelled
jobs

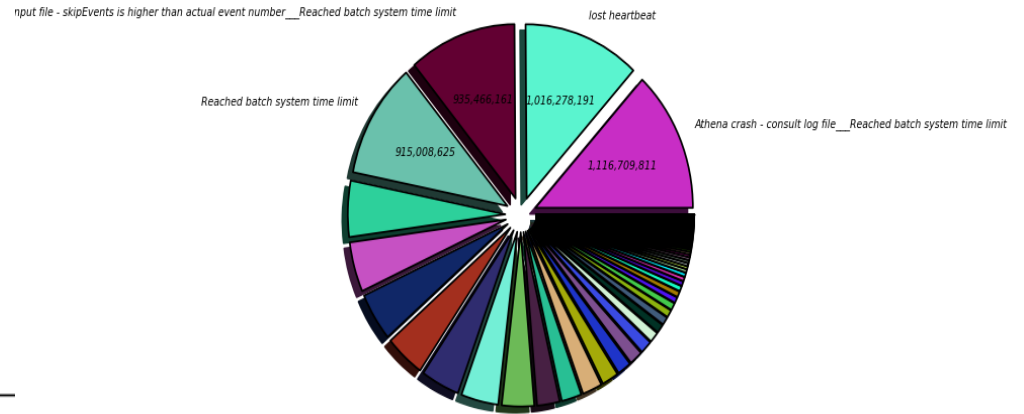


Analysis jobs – last month – all T1s and T2s

For once, highest point of failures does not seem to be storage



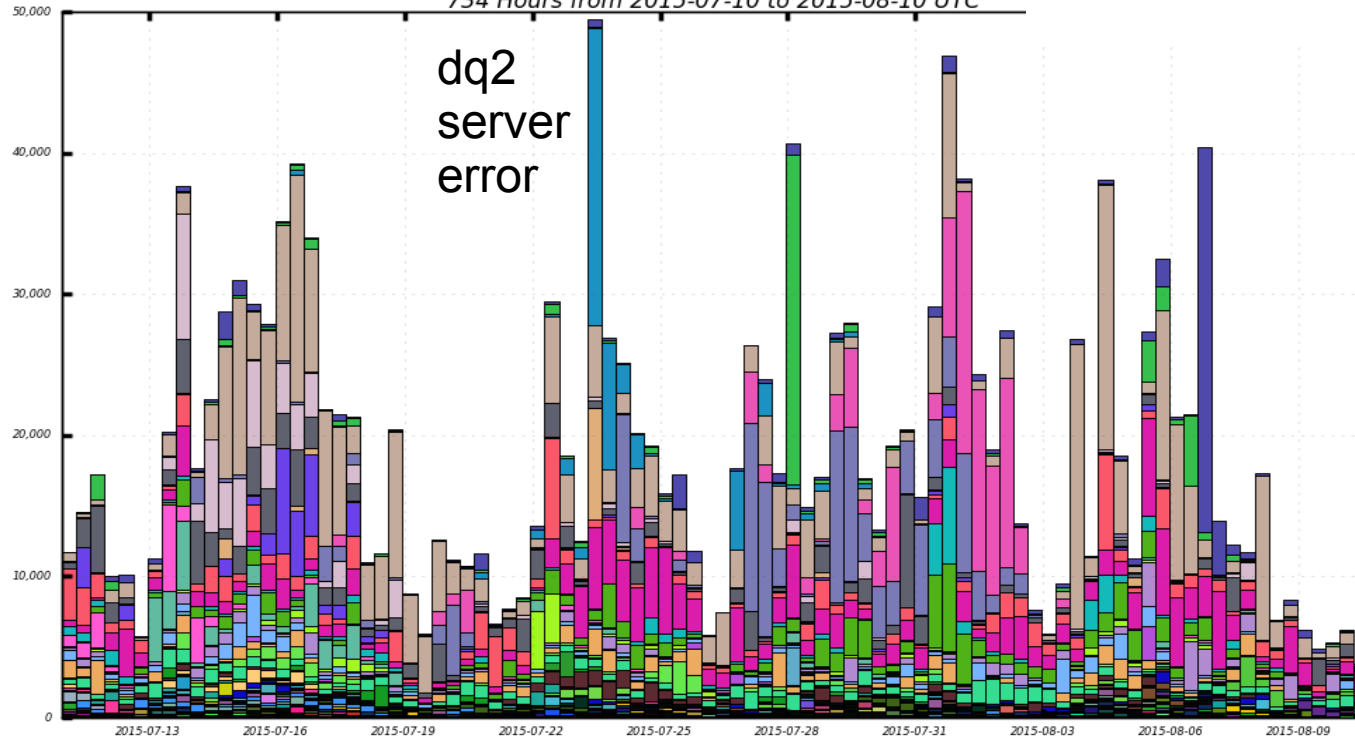
WallClock Consumption of Panda Failed jobs by ExitCode (Sum: 8,494,642,973)



Most wallclock consumption of failed jobs comes from cancelled jobs and lost heartbeat

Panda Failures by ExitCode

734 Hours from 2015-07-10 to 2015-08-10 UTC



dq2
server
error

- Undocumented Execution Error Code. Exit Code: execution65
- DQ2 server error
- Payload core dump
- TRF_SEGVIO - Segmentation violation__Athena ran out of memory
- Athena crash - consult log file__cancelled
- Athena crash - consult log file
- Undocumented TaskBuffer Error Code : taskbuffer113
- Athena ran out of memory
- Prout: An exception occurred in the user analysis code__Athena ran out of memory
- General pilot error, consult batch log
- TRF_SEGVIO - Segmentation violation
- Unspecified error, consult log file
- Error when handling transform output file
- Put error: Local output file missing
- New trf: Transform received signal SIGABRT; Old trf: Athena core dump or timeout, or conddb DB conne
- Get error: Staging input file failed
- Put error: Error in copying the file from job workdir to localSE
- New trf: Transform received signal SIGABRT; Old trf: Athena core dump or timeout, or conddb DB conne__Ather
- Prout: Input files cannot be opened
- ... plus 213 more

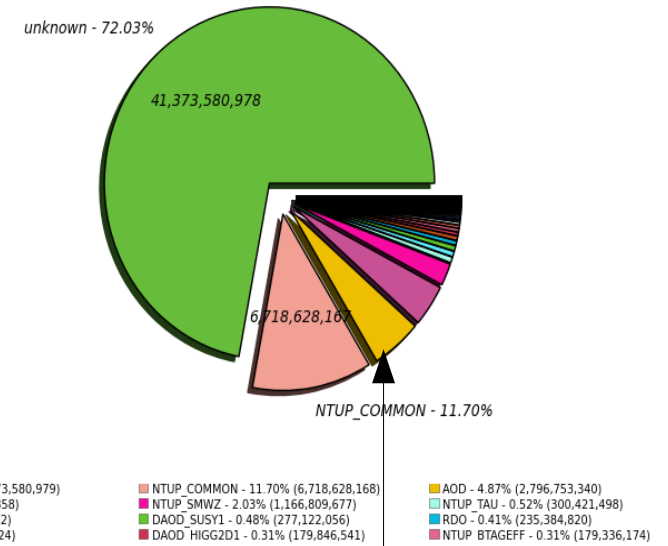
Maximum: 49,478 , Minimum: 0.00 , Average: 18,665 , Current: 3,831

Analysis jobs – last month – all T1s and T2s



Wall Clock consumption Good Jobs in seconds (Sum: 57,436,379,639)

Most wallclock consumption from successful jobs are likely private event generation (unknown)

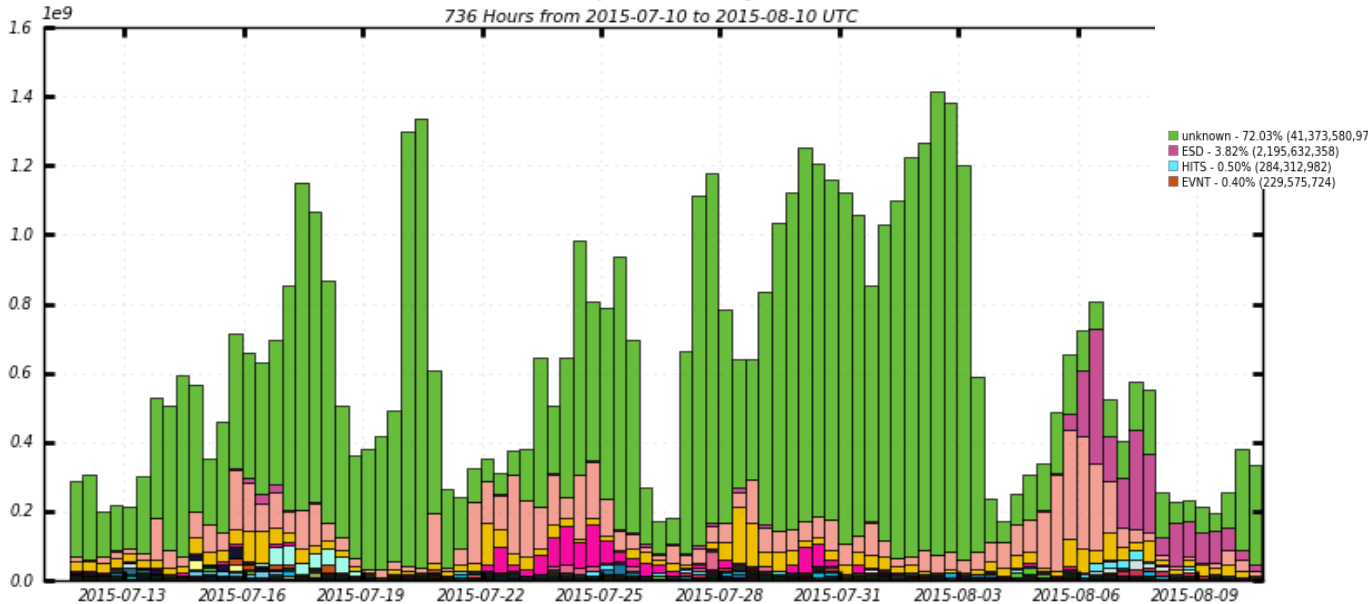


AOD analysis still plays a big role



Wall Clock consumption Good Jobs in seconds

736 Hours from 2015-07-10 to 2015-08-10 UTC



- unknown
- NTUP_TAU
- DAOD_JPSIMUMU
- DAOD_SUSY1
- NTUP_TRUTH
- DAOD_HIGG5D2
- DAOD_HIGG8D1
- DAOD_HIGG5D1
- DAOD_SUSY2
- DAOD_STDM3
- ESD
- NTUP_BTAGEFF
- EVNT
- NTUP_SUSYSKIM
- NTUP_FTKIP
- NTUP_SUSY
- DAOD_HIGG3D1
- DAOD_SUSY9
- DAOD_HIGG4D3
- DAOD_HIGG6D2
- NTUP_COMMON
- user
- NTUP_2LHSG2
- RDO
- NTUP_TOP
- DAOD_EXOT4
- DAOD_HIGG2D4
- DAOD_SUSY10
- DAOD_JETM4
- DAOD_EGAM1
- AOD
- HITS
- DAOD_HIGG2D1
- NTUP_HSG2
- DAOD_TOPQ1
- DAOD_SUSY5
- *
- DAOD_MUON0
- DRAW_ZMUMU
- DAOD_EXOT12
- NTUP_SMWZ
- DESD_SGLMU
- RAW
- NTUP_HI
- DAOD_HIGG4D2
- DAOD_STDM4
- DAOD_UPSIMUMU
- DAOD_EXOT13
- DAOD_HIGG6D1
- ... plus 66 more

User jobs consuming high memory

[from Andrej]

- Example task: <http://bigpanda.cern.ch/task/5918422/>
 - Unfortunately logs not available anymore
 - Successful jobs with reasonable memory consumption (< 2GB)
 - Successful jobs with very high memory consumption (> 4 GB + lot of swap), not killed because the site has not set **max memory**
 - Failed jobs killed because of memory limits (various)
- Bottom line:
 - User jobs may have very different memory requirements (but larger than 2GB hint to memory leaks...)
 - No common policy for memory limits at sites
 - Complicated to debug/monitor this kind of issues
 - Hard to educate users
 - Should a “reasonable” limit be set? Use the minimum current max memory value? Or closer to the average?

DAST report

[Farida and Alden]

- Even during holiday period stable load on DAST, in general smooth operations
 - 3 shifters per timezone + Alden and Farida to cover August
- DaTRI → Rucio rules transition (July 27th)
 - Users notified by announcement via DAST mailing list on [July 22nd](#)
 - However when DaTRI was switched off on July 27th, user tasks failed due to one (forgotten) call to DaTRI from Panda clients (checkDestSE)
 - [Should have been tested before](#)
 - Fixed by Tadashi in jedi clients 0.5.52
 - In the meanwhile, DaTRI resurrected for a few more days
 - Now all requests going through rucio

HammerCloud: new AFTs

- Candidate AFTs
 - two new templates based on QuickAna (analysis package supported by ASG, root-based analysis with EventLoop)
 - Athena-based: [AthAnalysis 2.3.11](#)
 - **Still failing at LRZ and Pavia**
 - **LRZ**: issue due to running SLC6 on SuSe: experts working on it, but no solution found up to now
 - **Pavia**: no reply from site admin
 - Root-based: [RootAnalysis 2.3.14](#)
 - **Still failing at OU and Pavia**
 - **OU**: issue due to old openssl library. No plan to upgrade in the near future (site admin waiting for major downtime)
 - **Pavia**: no reply from site admin
 - Input dataset pattern for both templates:
 - [mc15_13TeV:mc15_13TeV.*.merge.AOD.*_r6630_r6264*](#),
[mc15_13TeV:mc15_13TeV.*.merge.AOD.*_r6633_r6264*](#)
 - One dataset manually replicated by Rucio team to a **few** sites
 - Proposal by David to automatize the procedure by replicate one (small) dataset to **all** sites that run the tests

HammerCloud: new AFT and stress template

- Under development:
 - new template (could also be used as AFT in the future) using the new reco_tf to run **official derivation**
 - [AtlasDerivation 20.1.5.7](#): `Reco_tf.py --inputAODFile %IN --outputDAODFile test.pool.root --reductionConf HIGG3D1`
 - Using the same DS pattern as previous tests, could also add data15
 - Currently working on optimal job splitting
- New stress templates to be used for heavy I/O tests (for **analysis queues**):
 - **710**: Athena-based: [AthAnalysis 2.3.11](#)
 - **720**: Root-based: [RootAnalysis 2.3.14](#)
 - **728**: Official derivation: [AtlasDerivation 20.1.5.7](#)

HammerCloud: robot page

PanDA-Production

CANADAPROD



FZKPROD

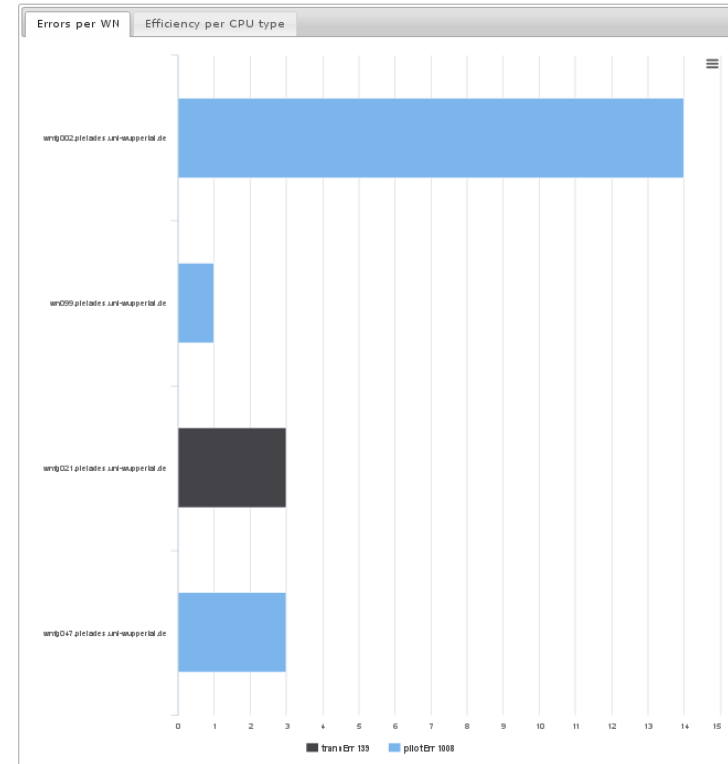


wuppertalprod

Improved monitoring for AFTs/PFTs

WorkerNode Error Overview:

The worker node error overview is still in prototype phase. Only errors with worker node id are classified. Since the PFTs are still provided without wn id, they are not considered in this plot!



621 (functional) ... 489 (functional) ... 571 (functional) ...

Template: 621 (functional) - PFT mc14 Sim_tf 17.7.3.12

Time	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Eff[%]
10/08/15	0	0	0	0	c	0	0	c	c	c	c	0	c	c	c	0	c	0	0	0	0	0	0	0	100
09/08/15	c	c	c	c	0	c	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	c	c	100
08/08/15	0	0	c	c	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	c	c	100
07/08/15	c	c	c	f	f	f	c	f	c	c	0	c	c	c	c	0	c	0	0	0	0	c	c	0	86
06/08/15	0	0	0	c	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	c	c	97
05/08/15	c	c	c	c	c	0	c	0	c	0	0	0	0	0	0	0	0	0	0	0	0	0	f	f	93

Error Report

Show All entries

Showing 1 to 10 of 10 entries

Search:

Previous Next

Links to single jobs on bigpanda monitor

Outlook

- **Analysis issues:**
 - Peak of submitted jobs with unknown input (July 30th – August 1st)
 - Mostly cancelled
 - From one user? Limit the number of daily jobs from a single user to avoid flooding the system?
 - wallclock of failed jobs dominated by long jobs hitting batch system limits and lost heartbeat
 - How to treat high memory user jobs?
- **HammerCloud:** new AFTs on the way, new stress templates
- **DAST:** usual load, no major issues to report, DaTri → rucio migration done, some disruptions to users, quickly fixed

Backup