Science and Technology Facilities Council

# Resources Review Meeting: 2020Q3 ATLAS RAL

James Walder

XX YY 2020

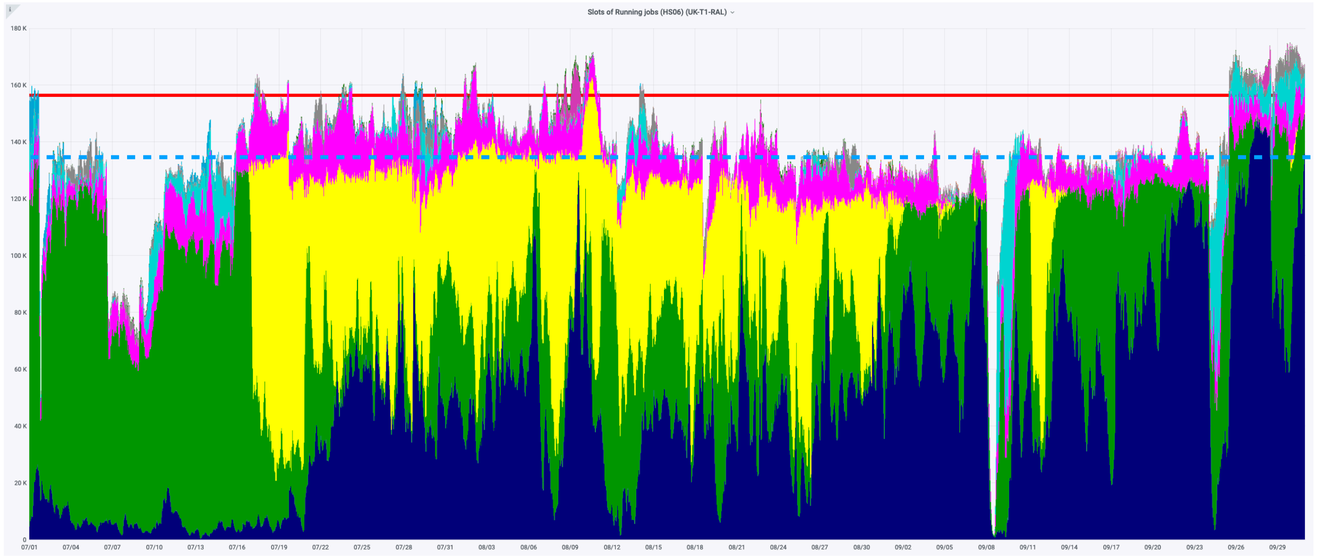
# Highlights

- Reporting period:
  - Q3 2020:  1 July – 30 September
- Highlights:
  - Deletion campaign on MCTAPE  (DATATAPE and additional MCTAPE round in Q4)
  - Data carousel performance at RAL noted again: “Staging throughput at RAL has been high and stable. What are your secrets to reach such a good performance ? :)”
  - Data reprocessing (in Data Carousel model) for full Run-2 data, for RPVLL analysis
- Major issues / events:
  - Occasional ATLAS/CERN issues, time taken for ATLAS to recover to fairshare
  - Upgrade to ARC-CE6 (smooth from VO perspective).
  - RAL (Scheduled) downtime for network switch firmware upgrades (Sept. 8th); all went well.

- New/Ongoing Issues
  - (Low-level of) failures related to Direct-io access (similar to LHCb issues).
  - RAL to be able to run sIngularity in unpriviliged mode; Experiments will then use own version of Singularity (from CVMFS)
    - Rolled into updating to Centos7 (and related batch farm services).

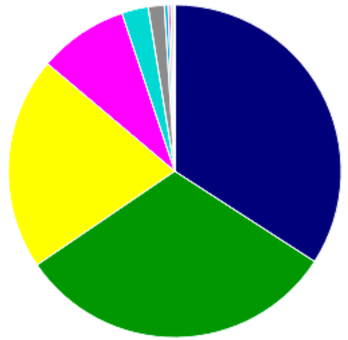
# Batch farm CPU commitment

- 2020–21 pledge: 156.4 kHS06 (13.9% of ATLAS T1 pledge)
- Average over period: 161.2 = 137.8*(11.7/10) kHS06:
  - Scaling accounts for difference between site-reported Corepower (HS06/core) and farm average.




| | min | max | avg | total |
|---|---|---|---|---|
| MC Reconstruction | 307 | 148.7 K | 47.7 K | 105.2578 Mil |
| Group Production | 200 | 128.8 K | 43.5 K | 96.0925 Mil |
| Data Processing | 0 | 116.0 K | 28.9 K | 63.8282 Mil |
| Pledges | 156.4 K | 156.4 K | 156.4 K | 57.5684 Mil |
| User Analysis | 20 | 32.7 K | 11.1 K | 24.4572 Mil |
| MC Simulation Full | 0 | 81.9 K | 3.5 K | 7.6658 Mil |
| Group Analysis | 0 | 15.4 K | 2.1 K | 4.7419 Mil |
| MC Merge | 0 | 20.8 K | 374 | 825.1 K |
| MC Simulation Fast | 0 | 21.3 K | 371 | 819.9 K |
| MC Event Generation | 0 | 3.9 K | 251 | 554.1 K |
| COVID | 0 | 773 | 49 | 107.1 K |

Wall clock time. All jobs (HS06 seconds)

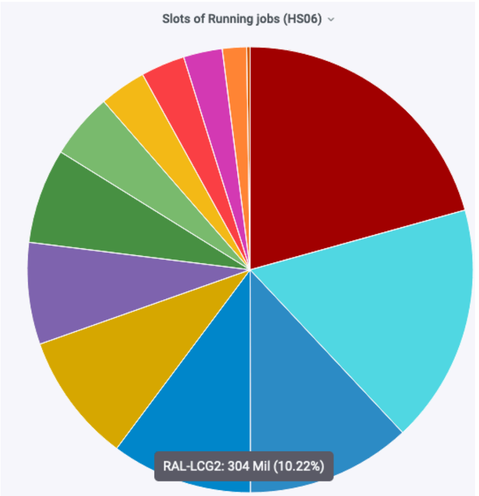


| | total | percentage |
|---|---|---|
| MC Reconstruction | 378 Bil | 34% |
| Group Production | 348 Bil | 31% |
| Data Processing | 229 Bil | 21% |
| User Analysis | 96.4 Bil | 9% |
| MC Simulation Full | 28.4 Bil | 3% |
| Group Analysis | 17.5 Bil | 2% |
| MC Simulation Fast | 3.92 Bil | 0% |

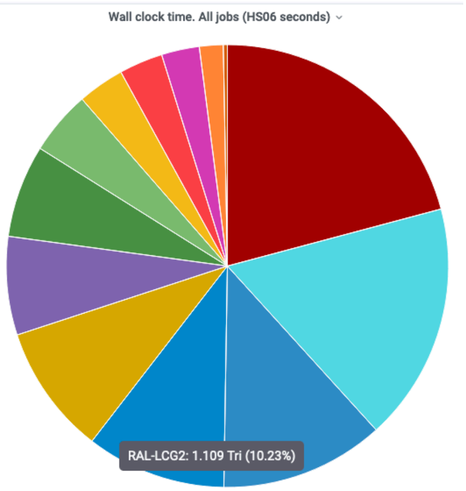
- MC reconstruction, Derivation production, RAW re-processing (for RPVLL) dominate HS06 time.

# T1 Comparison

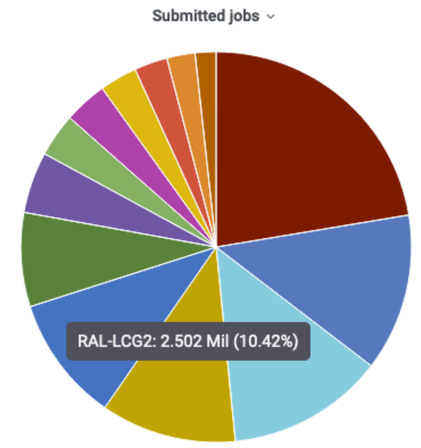
- RAL providing ~10% of ATLAS slots, jobs and walltime
- CPU efficiency rises slightly from Q2:
  - ~70% -> 75%:
    - Efficiency varies substantially across job type; would also need to be extracted



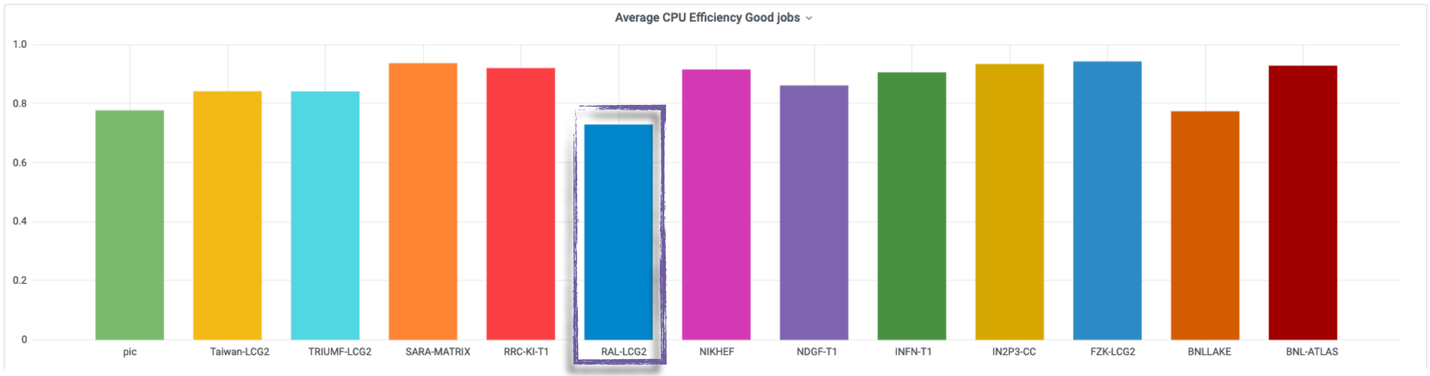

| | total | percentage |
|---|---|---|
| BNL-ATLAS | 616 Mil | 21% |
| TRIUMF-LCG2 | 516 Mil | 17% |
| FZK-LCG2 | 356 Mil | 12% |
| RAL-LCG2 | 304 Mil | 10% |
| IN2P3-CC | 279 Mil | 9% |
| NDGF-T1 | 220 Mil | 7% |
| INFN-T1 | 205 Mil | 7% |
| pic | 141.5 Mil | 5% |
| Taiwan-LCG2 | 100.6 Mil | 3% |
| RRC-KI-T1 | 95.1 Mil | 3% |
| NIKHEF | 83.7 Mil | 3% |
| SARA-MATRIX | 52.1 Mil | 2% |
| BNLLAKE | 7.51 Mil | 0% |

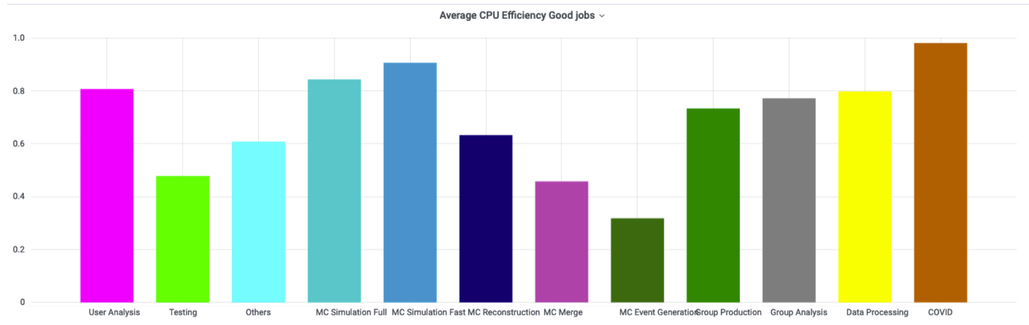



| | total | percentage |
|---|---|---|
| BNL-ATLAS | 2.260 Tri | 21% |
| TRIUMF-LCG2 | 1.888 Tri | 17% |
| FZK-LCG2 | 1.300 Tri | 12% |
| RAL-LCG2 | 1.109 Tri | 10% |
| IN2P3-CC | 1.028 Tri | 9% |
| NDGF-T1 | 779 Bil | 7% |
| INFN-T1 | 734 Bil | 7% |
| pic | 506 Bil | 5% |
| Taiwan-LCG2 | 370 Bil | 3% |
| RRC-KI-T1 | 346 Bil | 3% |
| NIKHEF | 300 Bil | 3% |
| SARA-MATRIX | 188 Bil | 2% |
| BNLLAKE | 31.8 Bil | 0% |



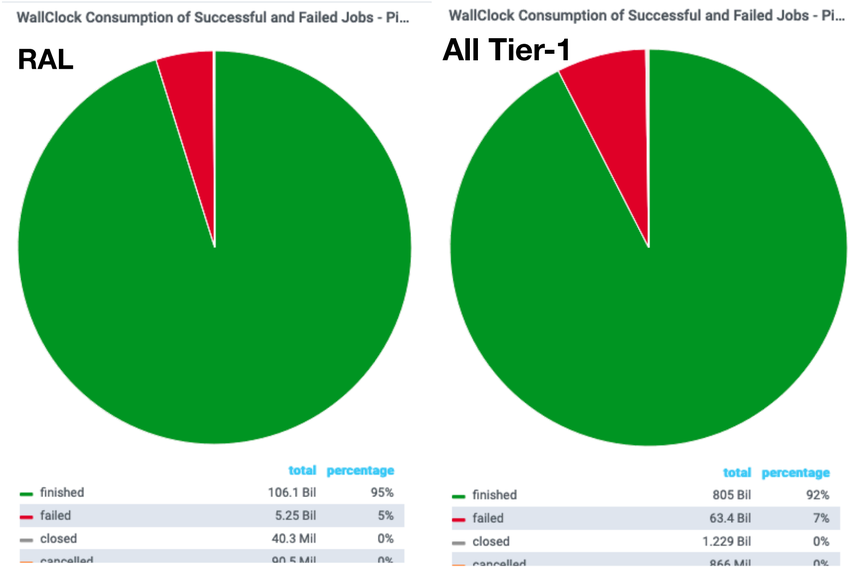

| | total | percentage |
|---|---|---|
| BNL-ATLAS | 5.37 Mil | 22% |
| FZK-LCG2 | 3.13 Mil | 13% |
| TRIUMF-LCG2 | 3.13 Mil | 13% |
| IN2P3-CC | 2.694 Mil | 11% |
| RAL-LCG2 | 2.502 Mil | 10% |
| INFN-T1 | 1.870 Mil | 8% |
| NDGF-T1 | 1.219 Mil | 5% |
| pic | 862 K | 4% |
| NIKHEF | 857 K | 4% |
| Taiwan-LCG2 | 746 K | 3% |
| RRC-KI-T1 | 660 K | 3% |
| SARA-MATRIX | 561 K | 2% |

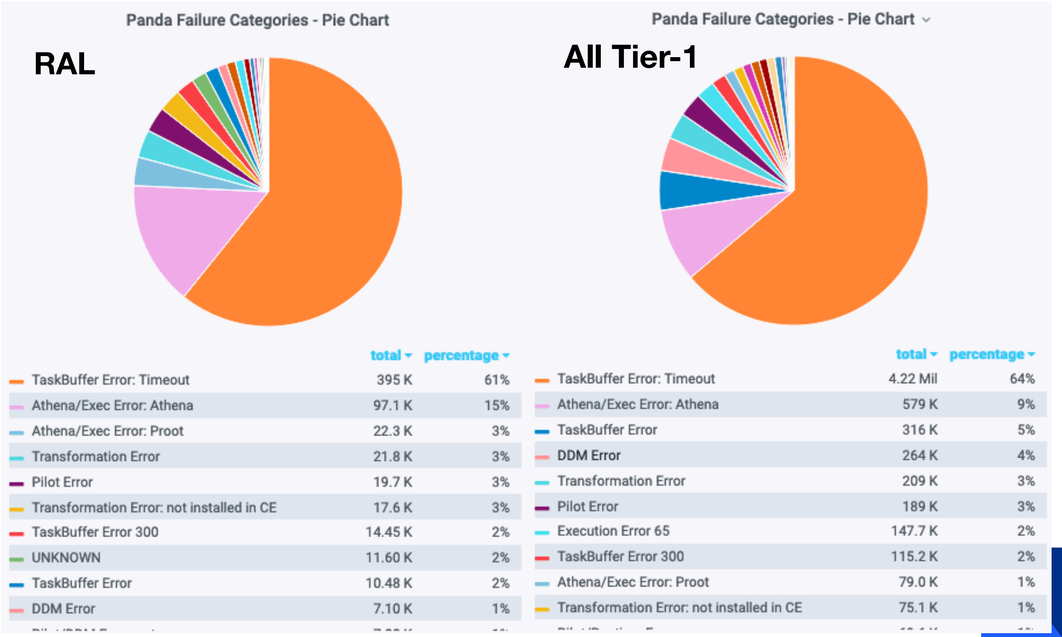


# Job Failures

- Overall improvement (in all Tier-1s) of lost wall time:
  - RAL's success rate (in walltime) 95% cf 92% average.
- All Tier-1s: 29% walltime lost in User Analysis jobs; cf. RAL: 25%  (all failure types)
- Work on ATLAS side to improve resilience to Stage-in, stage-out failures
- Direct-io failures still observed at low-level (relevant mainly to user jobs)
  - Implementation of vector reads should help
  -

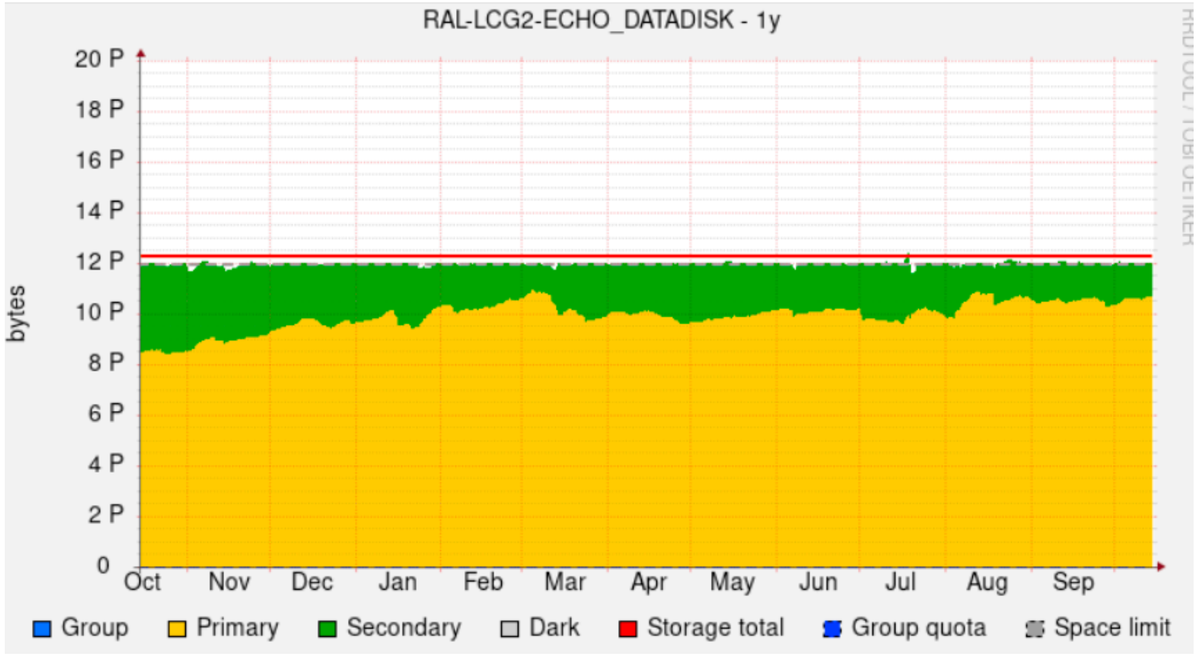


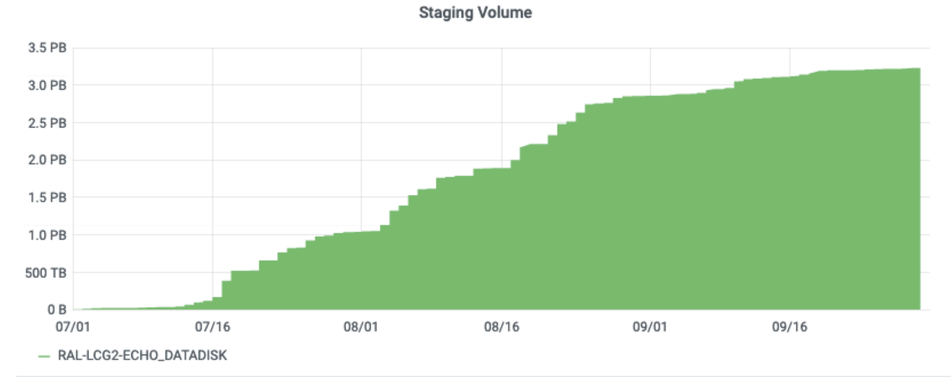
# Disk usage

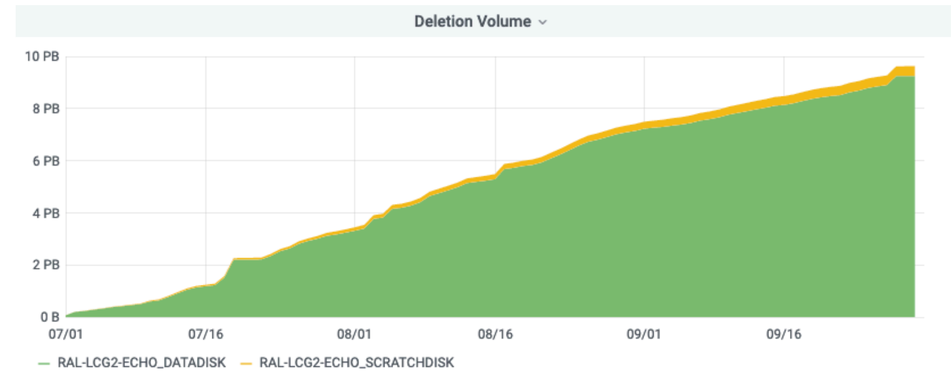
- 12PB on DATADISK
- 9.6PB deleted

- More regular use of Lifetime deletion model started







# Tape

- 2.9 PB added to TAPE
- 2PB deletion on MCTAPE (secondary data)
  - Deletion on DATATAPE just starting in end Q3
  - Further MCTAPE deletion in Q4
- 26.7 PB total tape usage (~32PB pledged).