

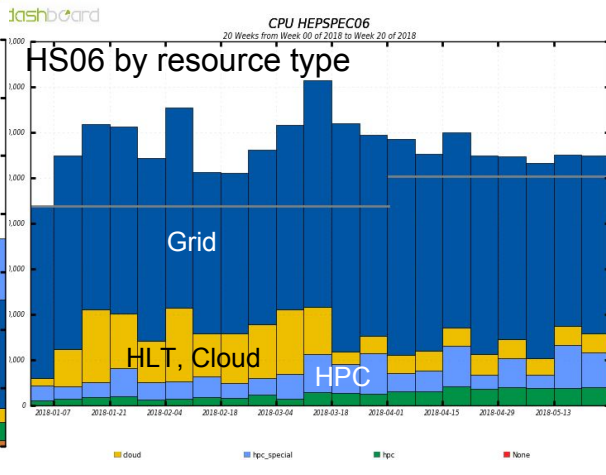
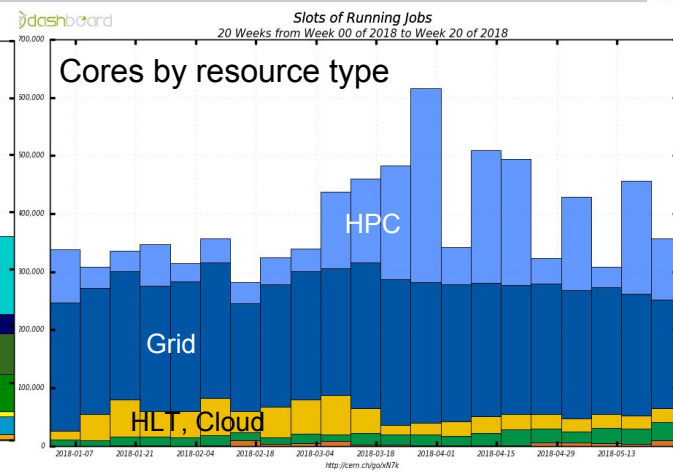
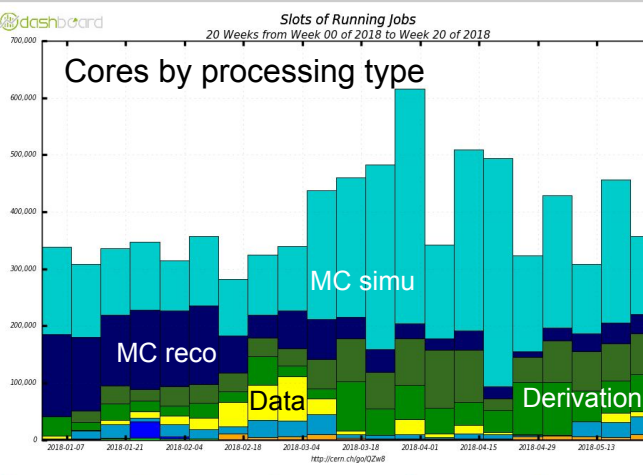


ATLAS Software & Computing Status

Torre Wenaus (BNL), Davide Costanzo (Sheffield)

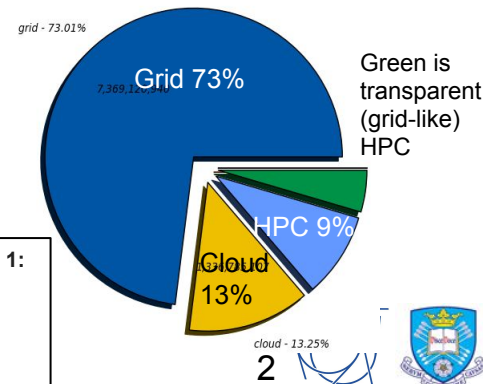
LHCC Meeting
May 29 2018

Processing since Jan 1



- Smooth Tier-0 running on 23k cores
 - Thank you CERN for a 20% bump over pledge from contingency
 - Commissioning Bphysics stream spillover to grid
- Sustained production with smooth operations, ~300-350k cores
- HPC peaks to ~900k cores (but cores are 5-10x weaker than grid)
- Derivation production of the 2018 collision data underway
- Moving >1 PB, >20 GB/s, 1.5-2M files per day

HS06 shares



Full sim events since Jan 1:
 Grid 3.81 B
 Cloud 1.02 B
 Transparent HPC 0.553 B
 Complex HPC 0.483B



Spring 2018 resource scrutiny



- Our 2019 resource requests were approved
- It was noted that we use a lot of disk (but growth is within flat budget)
- Update on the mitigation measures mentioned in our report -- work continues and is ramping up:
 - Greater use of tape-based workflows
 - Developing a detailed plan to deploy and evaluate select workflows with tape-based inputs, will get underway during the summer
 - Follows an earlier study of Tier-1 tape service performance
 - New LZMA compression recently deployed gaining us ~10%
 - ATLAS planning a further study to optimise analysis formats/counts and their storage footprint

CRSG recommendations for ATLAS

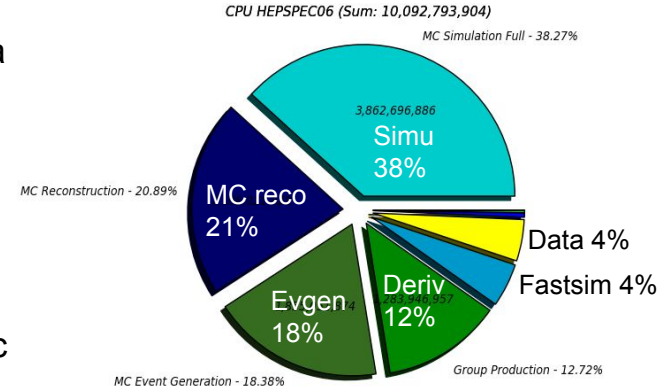
- C-RSG is concerned about the growth in the ATLAS disk usage and reliance on uncommitted CPU resources.
- C-RSG notes that ATLAS has significantly more disk resources than CMS on the Tier-1 and Tier-2 sites.
- The difference continues to grow in the 2019 requests: +29% at Tier-1 and + 38% at Tier-2.
- ATLAS explains this by differences in the computing model such as the data format and event sizes.
- C-RSG encourages ATLAS to consider investigating smaller data formats, higher compression rates and/or virtual data for fast simulation streams, similar to what CMS has achieved.

Distributed computing

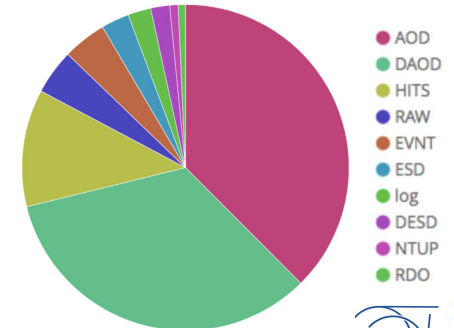


- Managing tight disk space through an active resource management board and close attention to data lifetime (e.g. older AODs)
- New ‘Harvester’ resource provisioning software in early deployment on a growing number of platforms
 - Particularly HPCs: more homogeneity and automation
 - Experiment-agnostic; discussing possible use beyond ATLAS
- Progressing on conditions DB improvements, review outcomes
 - Good progress on strategy for next-gen Crest system recommended by the review: first implement a COOL-backed distributed REST service allowing early scaling tests and adiabatic evolution/migration
 - COOL-backed Crest prototype integrated with Athena and under test
- [Rucio workshop](#) went well, keen interest among experiments
 - Discussing with WLCG its role in the ‘data lakes ecosystem’
- WLCG-HSF workshop in Naples was effective in seeding new R&D, particularly in distributed data and workload management; “data lakes” R&D program addressing both is getting under way
- *FAs are taking positive note, and they want to know where our planning is at -- would like to show them the WLCG strategy document*

CPU usage (HS06) in 2018



Disk usage by format





- AthenaMT milestones are late; on the present trajectory, AthenaMT will not be ready for Run-3
 - Central problem is insufficient sw effort in several detector and combined performance groups
 - Mandated by ATLAS Management, S&C is addressing the problem with each subsystem
 - Not easy! Since the mandate was established the situation has worsened considerably (a key developer is leaving)
 - Probing the collaboration for new and renewed software effort
- Planning the (incremental) integration of the new ACTS tracking software -- initiated by ATLAS, now a multi-experiment common project
 - A study of needed sw effort with/without ACTS integration is informing the planning
 - Integrated effort is less with ACTS

Software milestone highlights

- **2017 Q4: Finish migration to MT compliant event data access (DataHandles) - late**
- **2018 Q1: Start ACTS integration in Athena - in progress**
- **2018 Q2: MT compliant conditions data retrieval - behind schedule**
- 2018 Q4: Public algorithm tools thread-safe
- 2018 Q4: Make Services thread-safe
- **2018 Q4: TDAQ milestone: First integration with online, concurrent data access demonstrated**
- 2019 Q1: First phase [ACTS](#) integration done
- 2019 Q2: Start physics validation of MT vs. ST vs. R21
- 2019 Q3: MT compliant data quality monitoring
- **2019 Q4 TDAQ milestone: Algorithms migrated & tested, multiple threads working and in use**
- 2020 Q1-Q4: Bug fixes, optimization & full validation
- 2021 Q1: Release 22 in production for Run 3



- ATLAS management is recognizing and working to address the wider problem in software: supporting, recognizing, sustaining software effort as a crucial part of the ATLAS experiment (and HEP in general)
 - Applies to computing/operations also but the biggest problem is in software
- Steady progress on fast simulation and fast chain, albeit still tight on effort
 - Goal remains to start physics validation this year
- Good progress on pileup pre-mixing / overlay, addressing (possibly) the last few issues in validation, should be in production this year

ATLAS management has set up a working group...

Working Groups on software contributions in HEP

Recognition of software contributions in HEP

The group analyses ways to achieve a mind change at the university/lab and funding agency level towards the recognition of software work by physicists targeting a university or lab career as research physicist.

The group also analyses how the number of specialist positions for software experts (physicists) can be increased at the university and lab level.

...with brainstorming session at June ATLAS Week

ATLAS is benefitting from long-term release planning and strategy emphasising stability

- Started 2018 on same release as 2017 (first time!)
- All Run-2 data and corresponding MC are reconstructed fully consistently
- Resulting in rapid analysis turnaround on 2017 and 2018 data and inclusion into full Run-2 dataset
- The same coherence is applied to the trigger selection.



- Open sourcing the Athena software
 - Working on settling licenses via dependency analysis, cleaning the repo
 - End June target date for opening it should be about right
 - Already useful: will engage HPC experts in code study in July
- Infrastructure upgrade: ATLAS Release Tester (ART) took over from RTT for larger scale testing of software
 - More maintainable, efficient resource usage (no dedicated resources)
 - Leverages standard PanDA processing infrastructure and monitoring
- Documentation improvement: easy JIRA-based reporting embedded in every page, to gather needed improvements; then a 'docathon' to address them
- Developing I/O & persistency roadmap to guide evolution to Run-3 and towards Run-4
 - Draft in late June, review the complete roadmap in late September
- Software re-engineering for future architectures is surging in attention, including from funding agencies; taking this as an opportunity to build effort
 - Encouraging HSF to grow its role as a common clearing house for such work

Summary



- Smooth operations across Tier-0, grid and opportunistic resources with full utilization
- Managing tight disk space and working in several directions to reduce disk demands in the future
 - In the near term and all the way out to HL-LHC
- Insufficient AthenaMP migration-directed software effort levels are a major and growing problem
 - Now clear that we cannot deliver AthenaMT based Release 22 for Run-3 on schedule with present effort levels
 - Being addressed as a top level ATLAS priority, recognizing also it's a symptom of a wider community issue
- A new round of R&D is taking shape, most visibly so far in storage and data management (Data lakes / DOMA)
 - DOMA = data organization, management and access
- There is interest e.g. from FAs in our plans and strategy -- we'd like to show them the WLCG strategy document

Supplemental

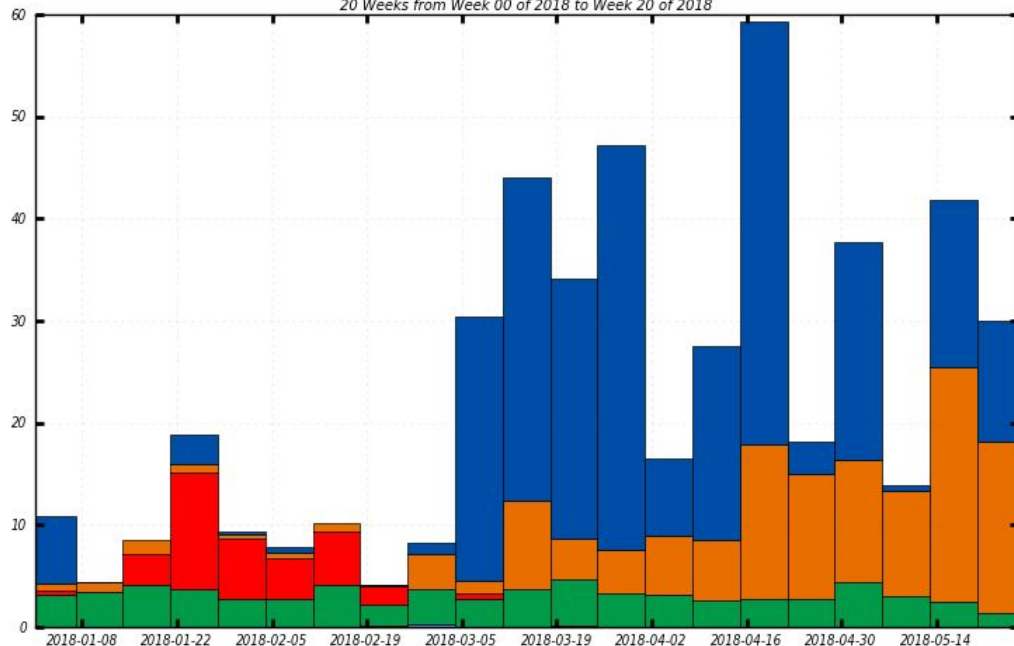


Event production on non-transparent HPCs



NEvents Processed in MEvents (Million Events)

20 Weeks from Week 00 of 2018 to Week 20 of 2018



NERSC_Cori_p2_mcore Titan_long_MCORE ALCF_Theta ORNL_Titan_MCORE Titan_Harvester_MCORE
NERSC_Cori_p2_ES

Maximum: 59.33, Minimum: 0.00, Average: 21.96, Current: 29.98

Event production in the 'hpc_special' (complex HPC) category, requiring custom infrastructure (not grid-like)

- NERSC Cori at LBNL
- Titan at Oak Ridge
- Theta at Argonne

Missing effort needed for a timely AthenaMT migration



- The missing effort *rough* estimates:
 - HLT: ~ ok
 - Pixel: 1 - 2 FTE needed
 - SCT: ~ ok
 - TRT: ~ 0.5 FTE needed
 - Tracking: needed effort estimated in the context of ACTS planning, dependent on integration plan to be decided in late June
 - LAr: ~ ~ok
 - Tile: ~ok
 - Calo: ~ ok
 - Muon: 2 - 3 FTE needed
 - Top level expert leaving -- a major problem
 - egamma: 0.5 - 1 FTE needed
 - Jet/MET: 0.5 -1 FTE needed
 - Tau: 0.5 - 1 FTE needed
 - Flavor tagging: 0.5 - 1 FTE needed