ATLAS Software & Computing Status

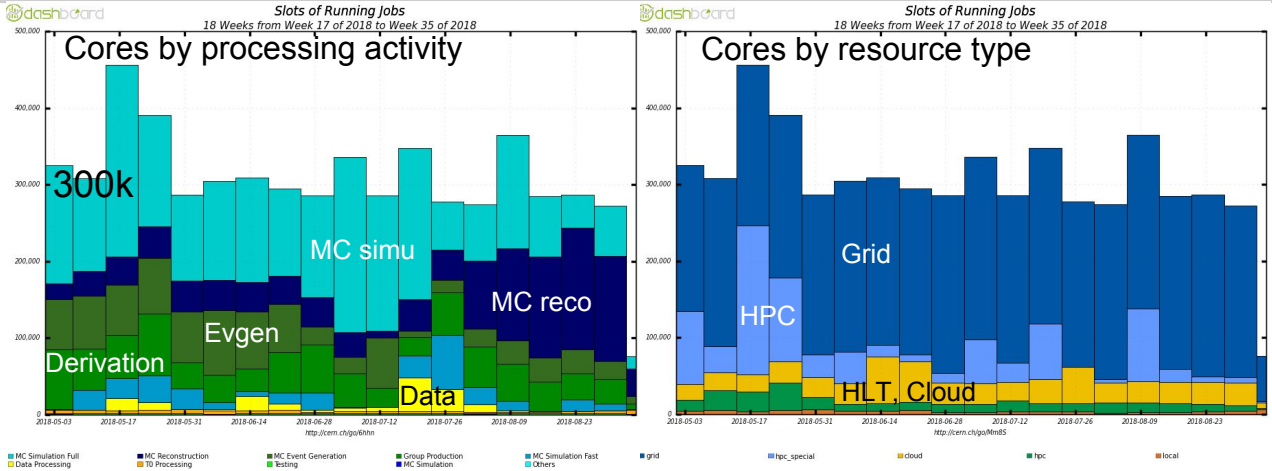
Torre Wenaus (BNL), Davide Costanzo (Sheffield)

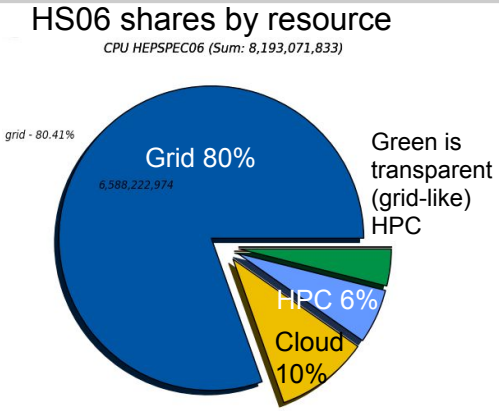
LHCC Meeting
September 11 2018

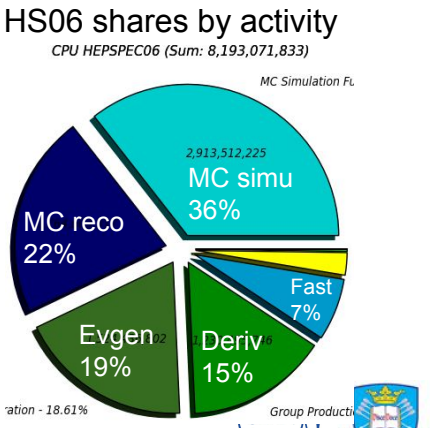
BROOKHAVEN

The University Of Sheffield.

Processing since May 1 (ev#s need update)

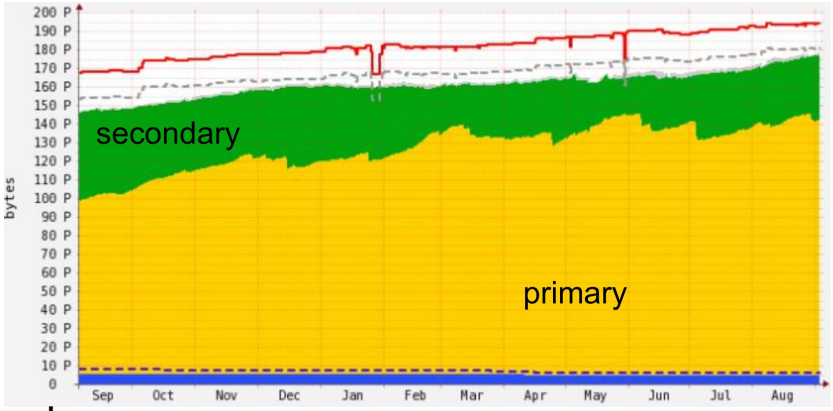






- Smooth Tier-0 running on 23k cores
  - Bphysics stream spillover to grid commissioned early but physics validation took time, ultimately concluding differences are primarily due to chip differences between Tier-0 and grid resources (Intel variants, AMD)
  - Tier-0 is keeping up OK so we can live without spillover (until Run-3)
- Sustained production with smooth operations, ~300k cores
- HPCs relatively quiet; exhausted allocations and wait for more
- Moving >1 PB, >20 GB/s, 1.5-2M files per day

Storage status



- Jagged primary (pinned) data plot tells the story of continuous management of our essentially full disk space
- Secondary (unpinned, deletable) should be ~25% of total for smooth ops, it is squeezed
- Several deletion and optimization campaigns in recent months
- Driven by the management mechanisms we have in place
  - Proactive resource management board monitoring usage, implementing deletion policies and scrutinizing space and exception requests
  - Lifetime model ascribing an enforced (with scrutinized exceptions) lifetime for each file
  - Obsolescence rules for data product versions
- With this close management, keeping up with the storage needs of data taking and ongoing MC16 simulation production

Heavy Ion datataking in 2018

- Plans in progress for the 2018 data taking
  - Plan 10^6 s of running and 10 different data streams
  - 2.1B events expected, 2.5 PB for RAW and 1.5 for AOD
  - Use partial event build (only ID and forward detectors) for calibration streams to save disk space
  - Expect 2 kHz rate with 2.5 GB/s bandwidth
  - Plans still evolving

- Participation in the end-to-end test with CMS and ALICE
  - Expected bandwidth of 2.5 GB/s

October 2018 CRSG report

- Submitted August 30
- Resources needed in 2019 were agreed at April RRB
  - T0 flat; no datataking
  - T1+T2 below flat budget (11% CPU, 19% disk, 4% tape)
  - C-RSG noted ATLAS uses more disk than CMS and with faster growth rate
    - We've been working to improve, and the comments were impetus for further action
      - A new computing model study group mandated to achieve at least 30% storage reduction relative to the Run-2 model
- Initial request for 2020 assumes 60 fb^{-1} with possible increase to 80fb^{-1}
- Summary of usage in 2018 and final request to be submitted in February

2020 request to C-RSG

	2018 Agreed @ Oct2017 RRB	2018 pledges	2019 Agreed @ April2018 RRB	2020 Request @ Oct 2018 RRB	Balance 2020 wrt 2019 request
T0 CPU (kHS06)	411	411	411	411	0%
T1 CPU (kHS06)	949	969	1057	1079	2%
T2 CPU (kHS06)	1160	1136	1292	1320	2%
SUM CPU	2520	2516	2760	2810	2%
T0 DISK (PB)	26	27	27	27	0%
T1 DISK (PB)	72	80	88	96	9%
T2 DISK (PB)	88	86	98	106	9%
SUM DISK (PB)	186	193	213	229	8%
T0 TAPE (PB)	94	105	94	94	0%
T1 TAPE (PB)	195	196	221	221	0%
SUM TAPE (PB)	289	251	315	315	0%

- Assumptions made:
  - Tier0 to be used for offline processing
  - HLT farm partially available for MC
  - Availabilty of HPC and extra CPU resources

No increase at the Tier0

~No increase in CPU needs

Some increase in disk resources

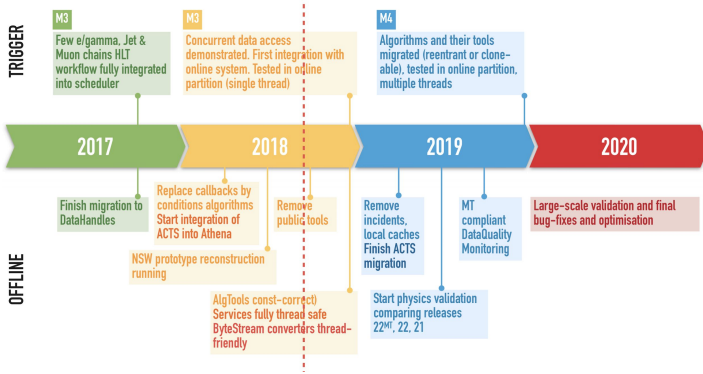
No increase in tape (the model was adjusted as it is currently under used)

Note: These will likely change this winter based on actual LHC delivered luminosity

Software

- Emphasis on stability paying off: continuing with Release 21 in 2018 contributes to rapid analysis turnaround and a consistent Run-2 dataset
- Still behind on AthenaMT migration milestones in the subsystems, but the concerted efforts to enlist more development effort are showing effects
  - Four subsystems moved to ‘OK’ column, effort levels up by ~3 FTEs
- Open sourcing the Athena software: we were on schedule to do so in June but decided it would be too disruptive to datataking, so postponed to the end of year stop
  - Cleaning up the repo entails creating a new repo, so everyone must re-fork
- Comprehensive I/O roadmap document towards Run-3 and (sketchily) Run-4 written and will be reviewed later this month, very useful in planning event I/O development
- Good progress on new workflows: new fast simu, fast chain, and MC pileup overlay appear to be on track for completing physics validation in 2019 (possibly 2018 for pileup overlay)

Release 22 milestone status



- **Q4 2017**: Finish migration to DataHandles. MOSTLY DONE (MAINLY MONITORING LEFT)
- **Q4 2017**, Trigger M2: Few e/gamma, Jet & Muon chains HLT workflow fully integrated into scheduler. Single thread
- **Q1 2018**: Start integration tests of ACTS into Athena & review DONE
- **Q2 2018**: Replace all callbacks by conditions algorithms. That will also kill a number of public AlgTools NOT DONE
- **Q2 2018**: NSW prototype reconstruction running ALMOST DONE
- **Q3 2018**: First validation tests of e.g. ACTS integration IN PROGRESS
- **Q3 2018**: Get rid of remaining public tools (make them private or convert to services) IN PROGRESS
- **Q4 2018**: Making AlgTools const-(correct), implies killing any local cache IN PROGRESS
- **Q4 2018**: Make Services fully thread safe STARTING SOON
- **Q4 2018**: Make ByteStream converters thread-friendly
- **Q4 2018**, Trigger M3: Concurrent data access demonstrated. First integration with online system. Tested in online partition (single thread).
- **Q1 2019**: Cleaning out remaining thread-hostile constructs: Usage of incidents, local caches in Algorithms
- **Q2 2019**: Finish migration of tracking code (+ clients) to ACTS
- **Q2 2019**: Start physics validation comparing release 22 MT vs release 22 ST vs release 21
- **Q3 2019**: MT compliant DataQuality Monitoring
- **Q4 2019**, Trigger M4: Algorithms and their tools migrated (reentrant or clonable), tested in online partition, multiple threads
- **Through 2020**: Large-scale validation and final bug-fixes and optimisation
- **Q1 2021**: Start of Run 3

Actively looking for developers to supplement the effort, with ATLAS support available under the right circumstances

Common software, common projects

- An eventful period for common software and common projects
- ATLAS adopted the ACTS tracking software, a multi-experiment common project initiated by ATLAS
  - Thorough review process recommended incremental adoption
  - Contingent on ACTS readiness milestone this fall
- CMS joined the community of experiments adopting Rucio for data management -- welcome!
  - Bodes well for a coherent, efficient WLCG DDM infrastructure
- ATLAS's software infrastructure guru Emil Obreshkov joined the EP-SFT group, placing our expert where he can leverage and contribute to common work

Computing usage in LS2 and Run-3

- No plan to change our model for Run-2 analysis during LS2, but changes are coming for Run-3
- During LS2, computing will be dominated by
  - Completing MC16 simulation (statistical extensions, new generators)
  - Some reprocessing of specific samples (no full reprocessing anticipated)
  - Run-3 preparation (validation, samples preparation)
  - HL-LHC studies
- Preliminary expectations for Run-3
  - Datataking: leveling at mu ~65, trigger rate similar to Run-2, 6.5M seconds in 2021
  - Tier 0: prompt processing 1.5-2x Run-2 in CPU; spillover to grid may be essential
  - Big change for CPU: Greater use of fast simulation (FastCaloSim, FastChain)
  - Big change for Disk: New analysis model (study group underway)
  - AthenaMT in Release 22 and software optimisation (eg require more hits per track)
  - New, more efficient workflows: pileup overlay, less merging, more tape usage, more remote data access
  - Detector upgrades: New Small Wheels, Level-1 Calorimeter trigger, FTK
  - Resource needs in 2021: ~1.5x 2018, within flat budget

Towards HL-LHC

- Working on an update and revision of resource needs out to HL-LHC
  - Incorporating updates for coming model changes, e.g. more fast simulation and smaller analysis format disk footprint
  - Also making the move from spreadsheet to a model in python
    - Drawing in part on the generalization of the CMS model undertaken in the cost model WG's hackathon a few months ago
- Active on HL-LHC R&D initiated following the Naples WLCG-HSF meeting
  - DOMA R&D program coordinated by WLCG
  - DOMA caching subgroup
  - ATLAS tape carousel study, now involving most ATLAS Tier-1s
- Reengineering for new architectures and HPCs is seeing more attention: e.g. a dedicated workshop at BNL in July, a new 'HL-LHC Computing' activity area in US ATLAS
- Establishing an ATLAS computing strategy study group to write and evolve a strategy document complementing the WLCG strategy with ATLAS specifics
  - Start discussion at the next Software and Computing week (3 Oct)

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, now also with a new computing model study group targeting 30% reduction
- AthenaMP migration-directed software effort levels continue to be tight but are improving, with more efforts underway, including support for new developers (when we find them)
- Common project successes: ACTS in ATLAS, Rucio in CMS
- Requesting a modest resource increase in 2020 with respect to 2019, and anticipating needs compatible with flat funding to Run-3
- A host of changes coming for Run-3 to use our resources more efficiently: more physics per unit resource
- Very active already in HL-LHC directed R&D, and planning an ATLAS strategy doc

Supplemental

Release 22 main technical deliverables

- Multithreaded Simulation, Digitization and Reconstruction using the athenaMT/GaudiHive infrastructure
- Common to both offline and online (HLT)
  - HLT with its hard deadlines is a major driver for timely completion of Release 22 core deliverables, and so far HLT's requirements are being met
- A new job configuration system, still python based, replacing the over-complicated current system (RecExCommon) with a more modular and maintainable one
- Simulation, Digitization and Reconstruction for upgraded detectors foreseen for Run-3 and beyond

Monte Carlo simulation for Run-2 (MC16)

- Current (July 18) status of MC16. Minimal set:

Sub-campaign	Data taking year	Luminosity (fb^{-1})	Events needs (2017 estimate)	Events needs (2018 estimate)	Status (July 2018)
MC16a	2015+2016	39.5	6.5 B	7.8 B	7.7 B
MC16c/d	2017	46.9	8.0 B	6.6 B	6.5 B
MC16e/f	2018	60 (80)	8.1 B	6.5 B (8.7 B)	3.7 B (Sim only)
TOTAL		**146.4 (166.4)**	**22.5 B**	**20.9 B (23.1 B)**	**17.9 B**

- Expect to complete by the end of the year
- Fewer simulated events per fb-1 as we progress in Run-2
  - Better understanding of MC use (better slicing, fewer systematic samples, …)

- Number of simulated events in minimal set roughly similar to number of data events
  - Some analysis suffer from poor MC statistics

Further MC simulations in 2019-20

- Statistical extensions of current samples:

Process	MC16 minimal #events	Extra Statistics factor	Extra Statistics #events
W/Z/ γ (+jets)	6.6 B	x2.5	16.5 B
Diboson	1716 M	None	0 M
Inclusive ttbar	900 M	x5	4.5 B
ttbar+V	74M	None	0 M
single top	80 M	x5	0.4B
Multijet (x2 generators)	175 M	x5	1.7 B
Total (extra stat)			**23.1 B**

- 6B events for new generator studies and 1B events new signal samples
  - Total increase to MC16 of 20.1B events (final MC/data ratio of ~2)
- 5B events for Run-3 preparation in 2020
  - Validation, initial samples preparation
- Reprocessing of Run-2 data with Release 22 not included

Computing model parameters

Event size (MB/Event)			Processing Time (HS06s/Event)		
Format	**2018**	**2020**	**Processing**	**2018**	**2020**
RAW	1.0	1.0	Full Simulation	3500	3250
Sim HITS	1.0	1.0	Fast Simulation	300	300
Real AOD	0.31	0.28	Generators	980	980
Sim AOD	0.41	0.37	Data Reconstruction	230	230
			MC Digi+Reco	567	508
			Derivations	0.4	0.4
			User Analysis	0.4	0.4

G4 static linking, geometry improvements

MC Overlay

LZMA compression scheme

Heavy Ion assumed to need ~10% of pp resources

Upgrade samples (HL-LHC) small contribution to the total