



# ATLAS Software & Computing Status

Torre Wenaus (BNL), Davide Costanzo (Sheffield)

LHCC Meeting  
November 29 2017

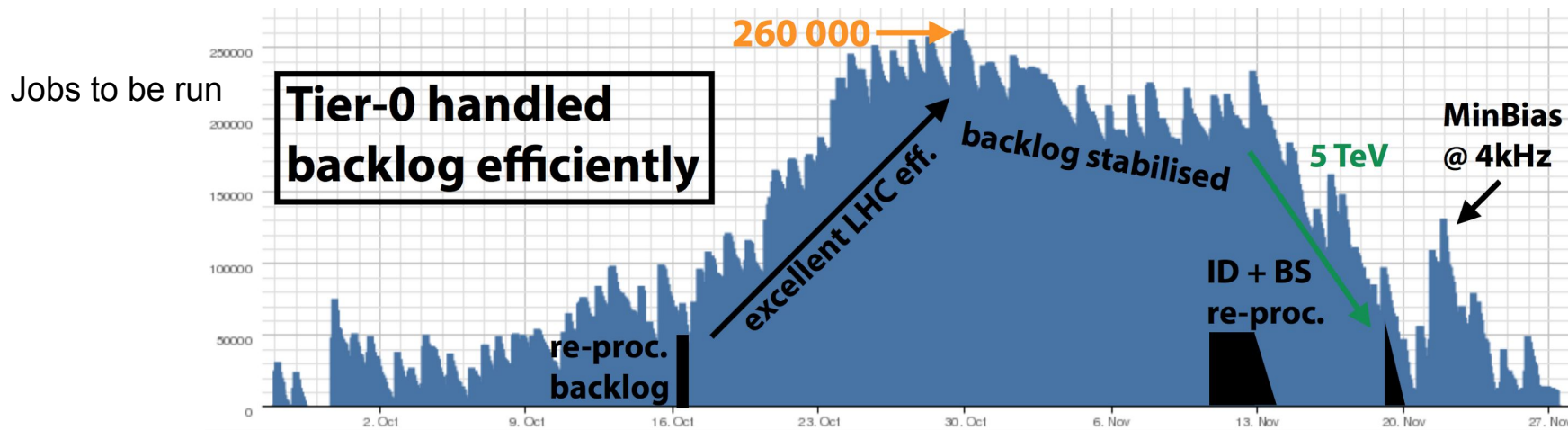
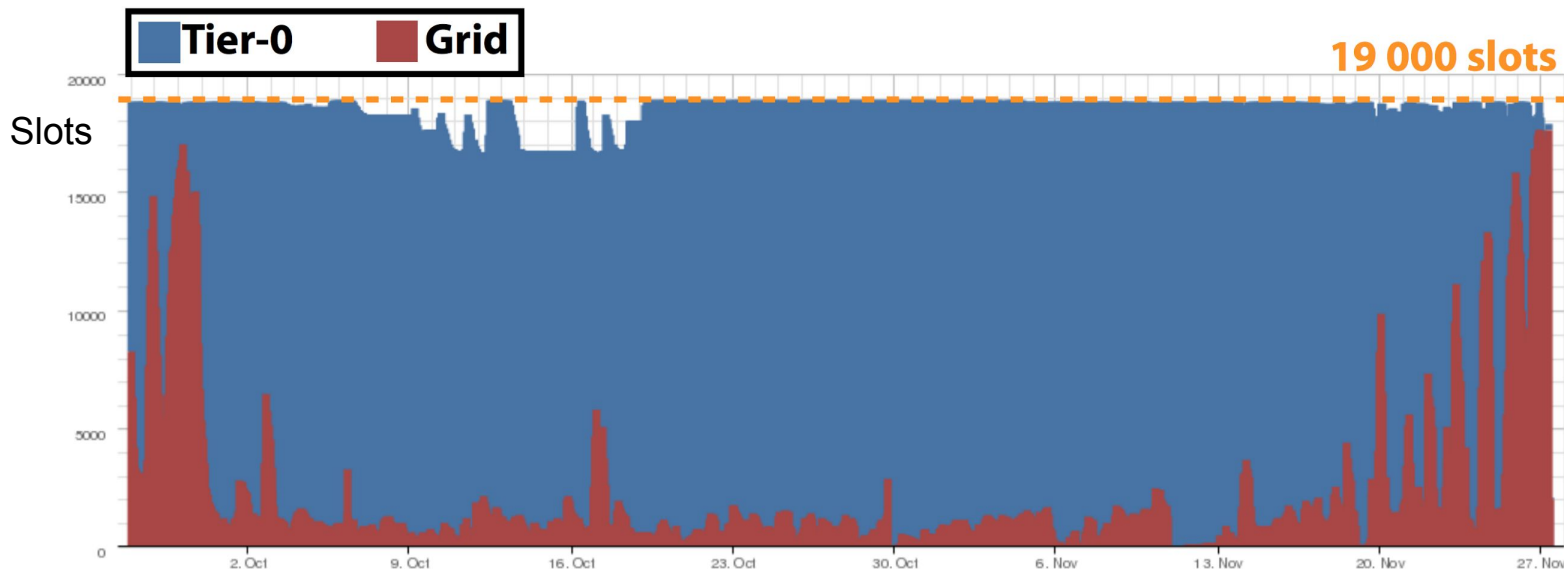
# Computing operations & resources



- **Tier-0 reached capacity limits** due to excellent LHC performance
  - Longer processing times due to extended high mu values from levelling
  - ~2 days/week backlog during high duty cycle periods, recovered during quiet periods
  - Tier-0 spillover to grid commissioned and on standby but in the end not used
- **Smooth (mostly!) grid production operations** with ~300k concurrent cores
- **CNAF flooding mitigation** underway
  - Represents ~7% of ATLAS Tier-1 resources
  - Re-replicating CNAF's RAW share to other Tier-1s; 2017 data complete, 2016 earlier will be next, a few PB in total
  - Unique data at CNAF didn't present any crises, reprocessing where needed
  - Grid-Ka has given us extra resources, CERN has offered as well, to help mitigate -- thanks very much!
- **HLT farm transitioning to offline** production ~now for much of YETS
- At Supercomputing 2017, Oak Ridge folks were heralding the fact that ATLAS/PanDA have enabled Titan **utilization** to reach 96% (very high for an HPC) thanks to our **backfill**

# Tier-0 performance

Last two months



Jaroslav Guenther (CERN)

# Efficiency: An example

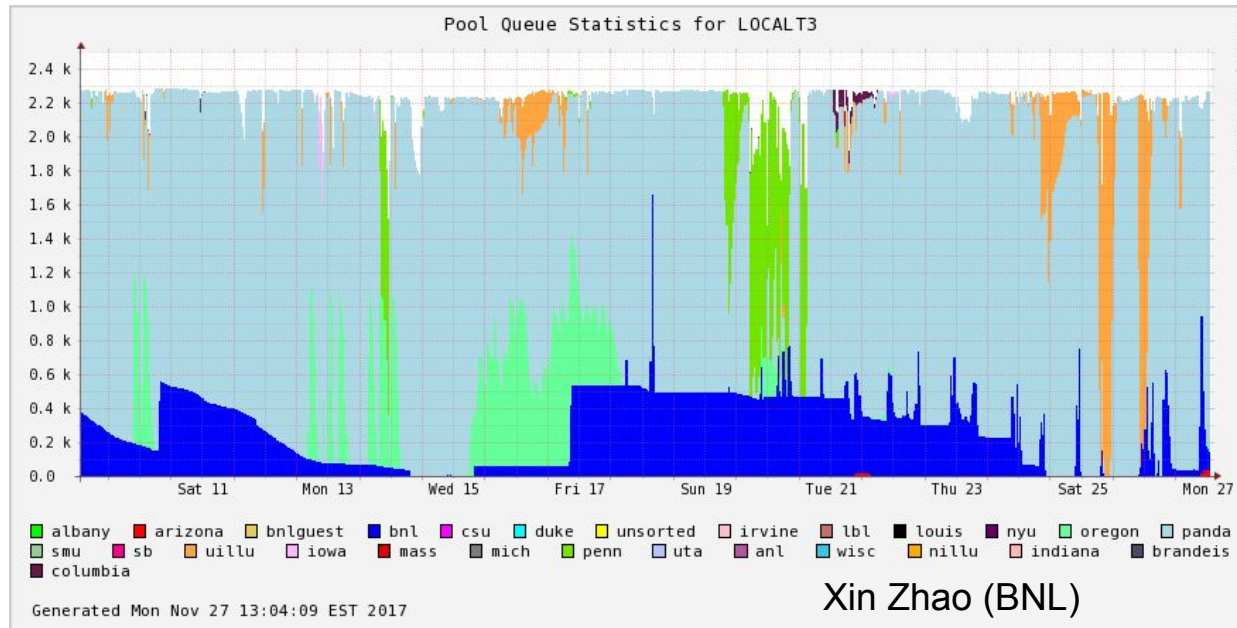


In past meetings we've shown our **CPU efficiency** is good, with inefficiencies understood and improvement measures underway (event service, init from checkpoint image, less serial merging time, ...)

Here's an example of the good **utilization efficiency** we're obtaining as well: the event service based backfill on a BNL Tier 3 queue.

Light blue = event service MC backfill that would otherwise go unutilized

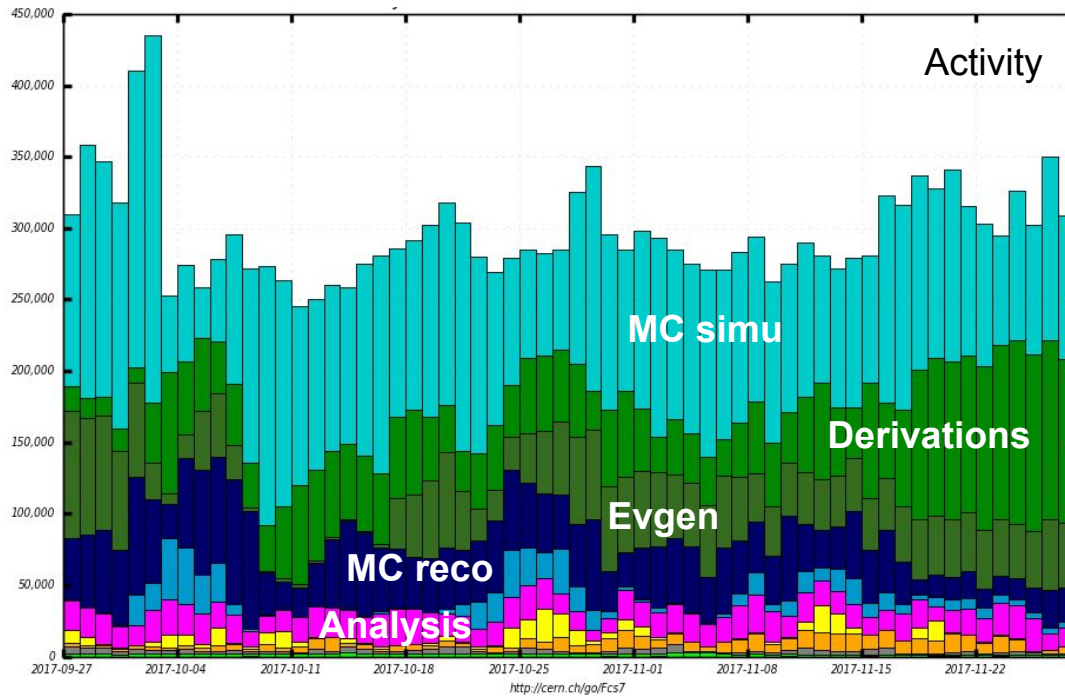
All other colors = Tier 3 jobs



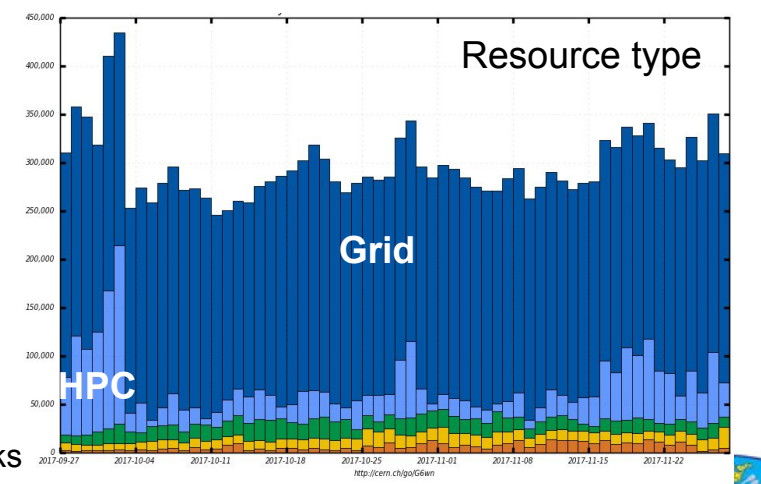
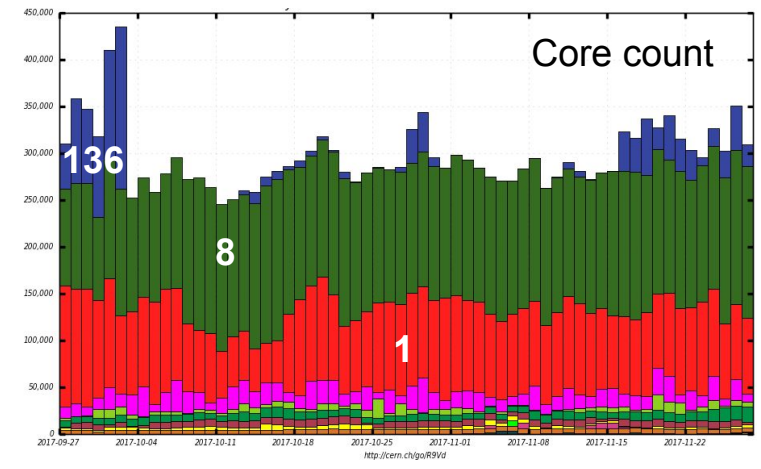
# Distributed processing



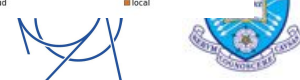
Concurrent cores, daily average, last 2 months



■ MC Simulation Full    ■ Group Production    ■ MC Event Generation    ■ MC Reconstruction    ■ MC Simulation Fast  
■ User Analysis    ■ Data Processing    ■ TO Processing    ■ Group Analysis    ■ Testing



hpc: automatically assigned tasks  
 hpc\_special: manually assigned tasks





# Processing campaigns



- **Derivation reprocessing** campaign recently started and running smoothly
  - Re-derivation of DAODs from AODs with updated software
- **2017 data reprocessing** campaign with updated conditions now in validation, expected to start shortly, first half before/during holiday, second half in Jan/Feb
- Commissioning **new memory reduction measures** to incorporate early in 2018
  - Expanding the pool of resources available for memory intensive processing, e.g. high-mu events
- **MC16d campaign** reconstructing MC with final mu profile (very different from spring estimation) to begin shortly
- In addition, steady flow of **new/extended evgen/MC** sample requests
  - High priority top samples will fill the rest of the grid
- **Plenty to keep resources busy over the holidays!**
- Production system's support for **automated shares** allows priorities to be managed when running multiple large campaigns like this

# Software developments



- Phase-I and Phase-II **upgrade software efforts integrated** into overall software effort and working towards code convergence (on schedule compliant with software stability through the TDR period)
- Detailed **Release 22 (Run 3) schedule** developed (next slide)
- Examined **software effort levels** across subsystems
  - Ensure timely development of reconstruction, simulation and digitisation software towards Release 22 and the integration of the Phase-I upgrade systems
  - Common conclusion: effort is very tight but Release 22 milestones should be sustainable
  - More effort would put schedule on more solid ground
  - Looking at mechanisms to boost the effort by boosting the appeal and recognition of software work
  - *Software work is too often seen as a career minus*
- A new **Software Performance Optimization Team (SPOT)** created with a broad, resourced mandate to measure and drive improvements in core, I/O, subsystem software and workflow performance

# Release 22 milestones



- 2017 Q2: First AthenaMT developers workshop inviting subsystem developers
- 2017 Q4: Finish migration to MT compliant event data access (DataHandles) - **on target**
- 2018 Q2: MT compliant conditions data retrieval
- 2018 Q4: Public algorithm tools thread-safe
- 2018 Q4: Make Services thread-safe
- 2018 Q4: Tracking code migrated to next-gen MT infrastructure [ACTS](#)
- **2018 Q4: TDAQ milestone: First integration with online, concurrent data access demonstrated**
- 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**



# Conditions DB for Run-3 and beyond



- **Review next week** of ATLAS conditions database plans for Run-3 and beyond
  - Experts from other experiments, CERN IT in addition to ATLAS
  - Assessing a development and migration plan for a successor to COOL/Frontier
- Applying the **lessons of 15 years** experience to produce a successor **tailored to our use cases**
- And to the **high scalability demands** of distributed operation in Run 3+
- CERN IT **support for COOL ends** during Run 3 if, as expected, ATLAS is by then the sole user
- With available effort in the subsystems extremely tight, **migration has to be gradual and as close to transparent as possible**

# White paper, R&D, common solutions

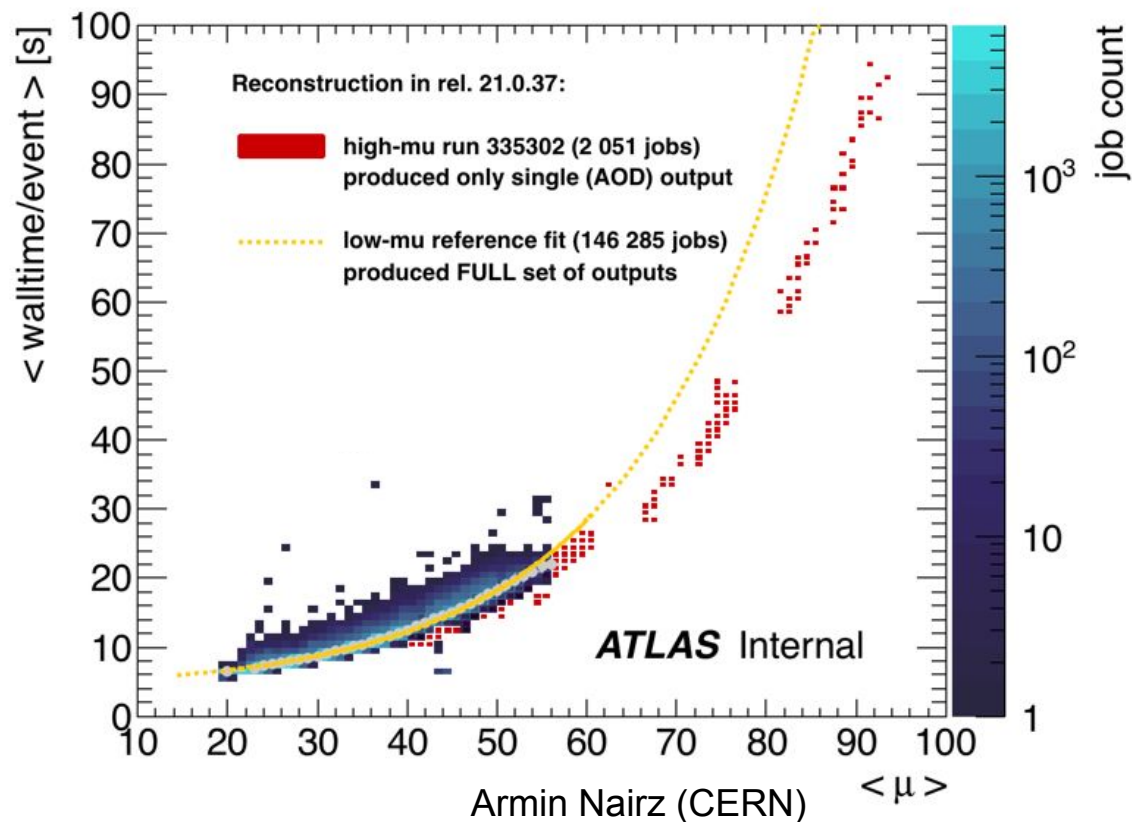


- With HSF community white paper process completing ([summary report](#)), now turning attention to **identified R&D topics and common solutions**
  - Well positioned with **hardened, scalable systems**
    - First [Rucio Community Workshop](#) coming in March
    - PanDA adoption beyond ATLAS is proceeding, e.g. Compass
  - A good start already made on **R&D efforts towards Run 3 & 4**
    - **Event service** produced 20% of all ATLAS MC last week
    - **Event streaming service R&D** underway, planning **data lake R&D** collaboratively with IT et al
  - **Discussions of R&D and common solutions** will take place at WLCG/HSF meetings in March 2018 and at CHEP 2018
- ATLAS holding a **documentation workshop** in December, looking to expand use of standard community tools
- Gathering input on a draft **software policy [document](#)** towards -- all being well -- **opening the offline software** (selectively) in 2018 Q1
- Good progress on **analysis preservation** in collaboration with the CERN Analysis Preservation (CAP) activity
  - Working to leverage the tools during analysis as well as post analysis

# 2018 expectations and preparations



- ATLAS successfully handled intense 8b4e data taking in 2017, we're confident we can handle the 2018 scenarios
- That said, ATLAS strongly favors 25ns BCMS, with lower  $\langle \mu \rangle$  for same luminosity
- Studying computing impact of the scenarios, implications for ATLAS and IT infrastructure, how to sustain headroom
- Tier-0 in 2018 ~same size as 2017
- Tier-0 load linear with rate, exponential with pile-up
- Comparing BCMS and 8b4e runs in 2017, the high pile-up and high rates of the latter result in ~90% higher CPU usage



# Summary



- 2017 was a challenging datataking year at Tier-0 and down the processing chain with high lumi + pile-up + duty cycle
  - Tier-0 saturated but performed well and handled the load
  - Valuable data and experience to plan and ready ourselves for 2018
- Smooth operations across grid resources, modulo the unfortunate CNAF incident we're mitigating successfully
- A full slate of campaigns keeping resources busy
- Software has moved into a new phase with development towards Release 22 underway detector-wide, working to a coordinated schedule and incorporating upgrades
- In community level work, moving from white papers to R&D and common software development, where the real fun is!