ATLAS Computing
implication on storage evolution

WLCG-HSF workshop

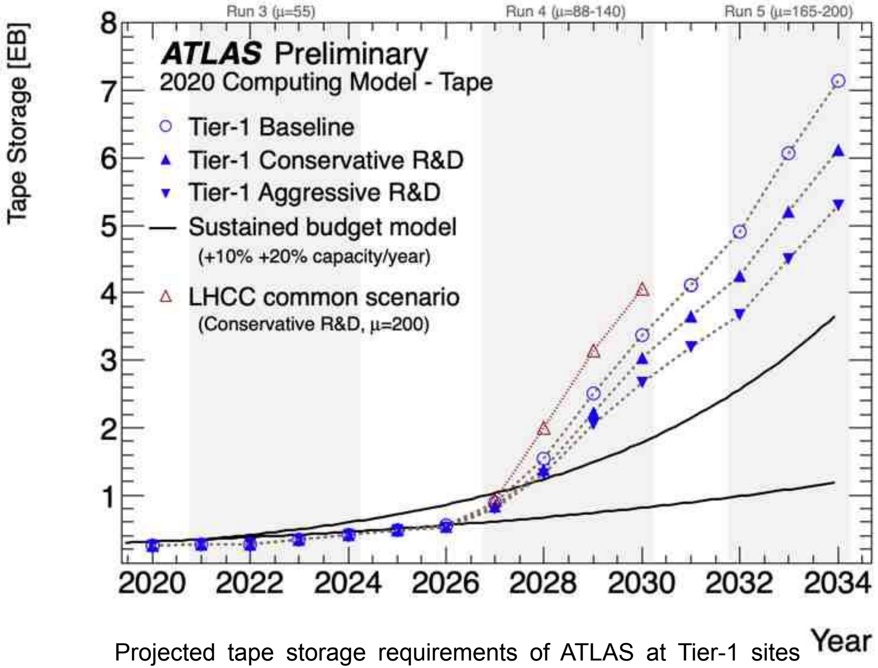
19 November 2020

This presentation has been prepared by the ADC enlarged coordination team, with suggestions and improvements of 10+ persons, and reflects the real work of the whole ATLAS Distributed Computing community

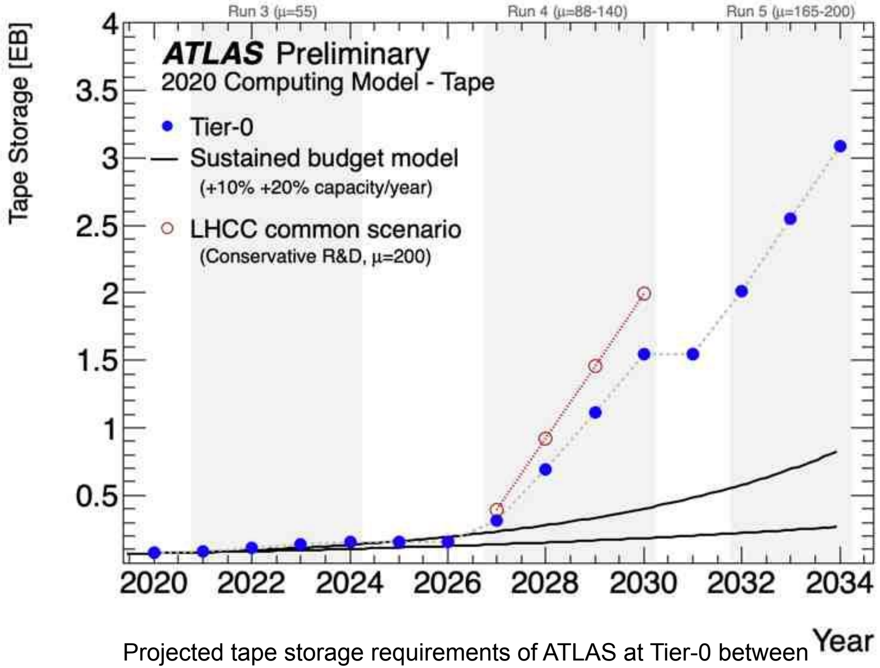
David South (DESY), David Cameron (Oslo),
Shawn McKee (AGLT2), Mario Lassnig (CERN),
Alexei Klimentov (BNL), Alessandra Forti (Manchester)
Ale Di Girolamo (CERN) & Zach Marshall (LBNL)

The HL-LHC storage challenge

The ATLAS Computing Conceptual Design Report is now a public LHCC doc.



Projected tape storage requirements of ATLAS at Tier-1 sites between 2020 and 2034 based on 2020 assessment. Three scenarios are shown for Tier-1 sites, corresponding to an ambitious ("aggressive"), modest ("conservative") and minimal ("baseline") development programme. The red triangles indicate the Conservative R&D scenario under an assumption of the LHC reaching =200 in Run4 (2028-2030).



Projected tape storage requirements of ATLAS at Tier-0 between 2020 and 2034 based on 2020 assessment. The Tier-0 line is based solely on the expected event size and LHC performance assumptions, since Tier-0 disk is used only for raw data and no other data products. The red triangles indicate the Conservative R&D scenario under an assumption of the LHC reaching

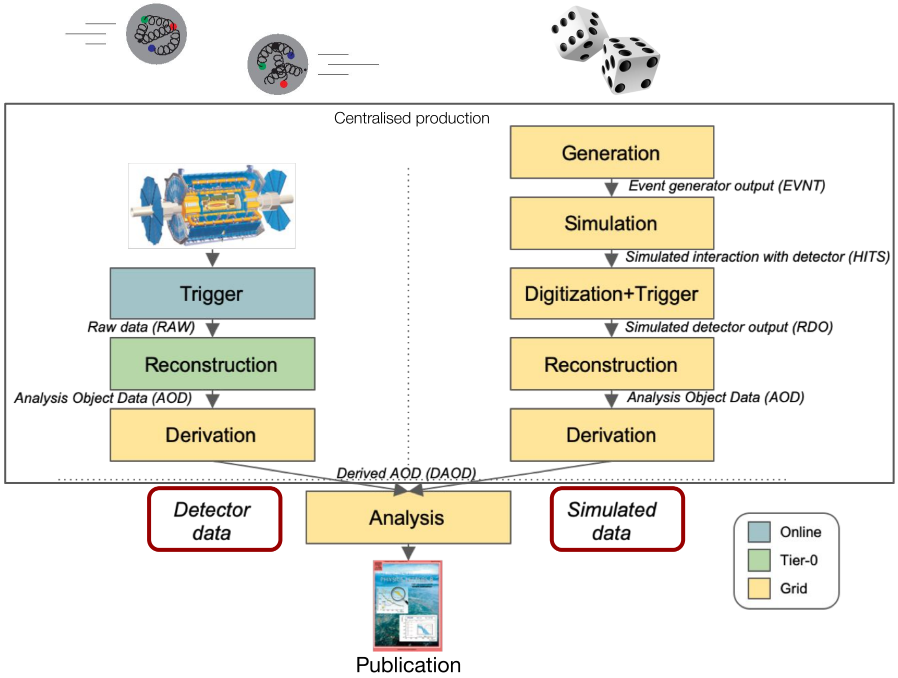
ATLAS workflows

- Raw data:
  - Sensor hits, energy deposits, timing information
- Reconstructed data (AODs):
  - Momentum of tracks (4-vectors), energy in clusters (jets), particle identification, calibration
- Derived data sets (DAODs):
  - Selected analysis level information, some of which is calibrated. Starting point for analysis
- Monte Carlo: Simulated data, comparison to theory
  - **Event generation** (EVNT): Calculated particle interactions
  - **Simulation** (HITS): Particle interactions with detector material
  - **Digitization** (RDO): Transforms simulated energy into a detector response that looks like the raw collision data
  - **Reconstruction:** Performed the same way as for real data: produce AODs, with some additional information
  - **Derivations**: as for data, produce DAOD_PHYS/PHYSLITE

No major workflows change foreseen for Run-4

- But we will have new ones, e.g. ML (in several flavours)

The Data Processing Chain



Centralised production: In essence, several steps of data processing followed by data reduction

TPC - Third Party Copy

- Mandate agreed 1st of August 2018
  - https://indico.cern.ch/event/747687/attachments/1696138/2730174/WLCGStorageProposal-2.pdf
  - Really a lot of work "behind the scenes"
- ATLAS is pushing forward with HTTP-TPC and has O(30) sites running in production
- Plan to completely transition by May 2021
  - Except tape, xrootd-only and GridFTP-only sites (T3s)
  - HTTP will be preferred protocol where available
  - GridFTP still available on most sites as backup
  - “Multi-hop” in Rucio in case of incompatible protocols
    - E.g. integrated also GlobusOnline to (mainly) enable US HPC
- Tape: plan to keep SRM for Run-3 for the majority of T1s
  - SRM bring online + HTTP as transfer protocol
  - Requires some development in dCache/Storm, expected to be ready by Run-3
  - CTA@CERN will use xrootd hop through EOS to outside
    - Implemented via FTS

AAI - Authentication and Authorisation Infrastructure

- ATLAS recognises the need to move to industry standards, and participates actively in AAI forums
- Effort coming from several communities
- Work has been done in many ATLAS systems to use JWT tokens at a proof-of-concept level
  - Evaluation of the token-based architecture (IAM based) and preliminary tests with Rucio and Panda
  - Planning for other services to join the test (e.g. Ami, etc)
  - Dedicated test instance of IAM being prepared
  - No need for a big-bang migration, both tokens and X509 certificates will be working during the transition phase (needs to be tested)
- Deployment and commissioning on the grid will be a much longer and harder challenge
  - Impossible before Run-3
  - Challenging for Run-4

QoS - (Storage) Quality of Service

- Currently we have two different QoS:
  - Disk: reliable, expensive, low latency
  - Tape: More reliable, cheaper, high latency
- With either disk or tape, we assume minimal data loss
- Several R&D activities looking into different (lower, cheaper) QoS
  - "Google R&D" (ATLAS working with Google on - between other topics - evaluation of hot/cold storage)
  - MAS@BNL: transparently tape-backed disk
  - Out of warranty storage@Edinburgh: Set up JBOD with expected data loss
- Most (the "expensive") ATLAS data is backed up on tape, thus we could use less reliable disk and trade-off either more disk or more tape
  - Careful: Operations will be affected
- Any QoS changes affect all levels
  - What is a site expected to provide in terms of reliability and latency?
  - QoS awareness in the WM and DDM systems, ability to transparently handle data loss
    - Up to automatic recreation of lost data
  - QoS in the high-level funding requests
    - How to account different QoS in pledge, eg 1TB of fast reliable disk = 2TB of slow unreliable disk?
    - Note: FAs are making the application now for resources up to 2027: still disk/tape, i.e. HL-LHC storage decisions start being made now
- "Expensive", "cheap":
  - do we need to add "cost"?

QoS - cost?

- "Expensive", "cheap": cost
  - Difficult even just starting the discussion about this
  - Should this be "ad-hoc" for each site? Should "someone" define grossomodo the possible ratio?

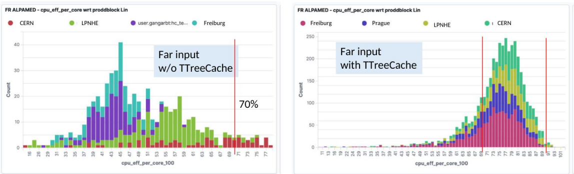
- One single uniform QoS tech is a dream. "Capability" (storage technologies), vs "activities" (experiments activities)
  - Need to define high level "simple" QoS. KIS
  - Need to plan for different implementations, need an abstraction layer
    - already architected for this possibility in Rucio and CRIC
  - Targeting the perfection is going to be a killer.

- Need to involve sites in the discussion (as we are doing!):
  - What can be really useful to have
  - What can be done in reality
  - Personpower, space, technology, etc….

Data access and caching

- Data lake definitionS
  - A compute-only resource which transfers data to and from remote storage
  - Federated storage in different physical locations exposed through a single endpoint
  - Distinct storage locations federated through a catalog or redirector
  - Multiple QoS within the same logical storage
  - → Not a single Data Lake definition
    - The reality is that we need to give space to the creativeness of the community
- Using remote storage implies using caching to hide latencies
  - Caches for latency hiding → low ops needed
  - Caches to boost data access/improve users' experience → highly challenging storage (later)
  - In the application code (root read-ahead TTreeCache)
  - Streaming data through a caching system (xcache, squid)
  - Data transfer asynchronous to the job (ARC CE)
-

CPU efficiency with direct access : Far sites



TTreeCAche mechanism mandatory at European Datalake level

Caches - more thoughts

- Useful to get industrial technologies and see how can fit our needs
  - But we need to keep in mind that our needs are different
  - Our data is colder, the access pattern is completely different than e.g. Facebook/Youtube
  - R&D ongoing: VP (VirtualPlacement) inspired at the Facebook caching model.
- If we can get funds with these project very good
  - Important to use the personpower in and impactful way
  - We have challenges and we are looking for solutions
    - We should be careful in not ending up in the "we have a solution, let's find a problem we can solve with this"
- We like intelligence!
  - important to know workflow-dependent data access patterns
  - Intelligence must come from the experiment side knowing what and when to cache
  - What's about the outputs? They also suffer of the same issues of the inputs
    - And in addition when we have issues in writing out the output we lose *a lot* of CPU
    - Solutions like ARC and Harvester are solving this issue: stageout to remote site is asynchronous and robust(w retries), not keeping the CPU idle, and orchestrated over the site
- Need to have our frameworks flexible to adapt to these solutions
  - non-negligible work

Can't forget Network 1/2

Networking needs to be considered in any storage discussion

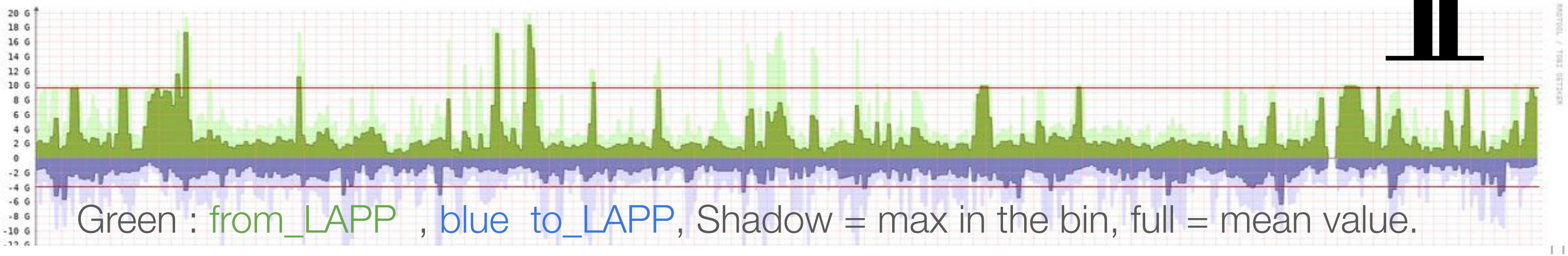
- Evaluating the concrete impact of swapping more networking for less storage (via remote access)
- Discussions ongoing about creating a set of network challenges
  - Ensure we have the networking capabilities we need for HL-LHC
  - Will involve both production and prototyping networks (FABRIC/FAB, trans-oceanic research and education links, ESnet, GEANT, Internet2, etc)
  - Involving big sites (why not also HPCs?) to understand operational modes and network access
  - Critical to define realistic production milestones of each challenge.

Computing never stops: network/data challenges at scale are paramount, but they require significant personpower!

Overall goal: scale to ~Tbps networks for HL-LHC

Can't forget Network 2/2

- Remote access within lake is foreseen, but without (properly sized) network we would be doomed
- LAPP WAN over last month (LHCONE):
  - 2 weeks with 20 Gb/s : No saturation ; 3 weeks with 10 Gb/s (network issue): saturation
- ESCAPE DataLake remote upload exercise degraded by saturation
  - Just an example!!
  - Need global monitoring of activities, including queue of "work to be done"!
  - Global orchestration? → FTS



Green : from_LAPP , blue to_LAPP, Shadow = max in the bin, full = mean value.

Analysis Facilities

- Typical analysis:
  - Bulk batch processing on the grid. Then final plots form ntuples on institute resources or lxplus
  - Operationally easier and improved user experience if consolidate grid analysis at few large sites
  - However there is a common desire of the sites to run analysis type jobs
  - We therefore see bulk organized analysis continuing on several well supported sites
- Other types of analysis:
  - GPU-based
    - A handful of sites expose GPUs as grid resources. Not huge demand as yet
  - Columnar data
    - Data formats more friendly to common data science tools (R, Pandas, spark etc)
  - Interactive analysis
    - Notebook-style with in built support for above tools and data access (eg EOS access in SWAN)

- Clearly there is a demand and potential paradigm shift towards interactive data science-like analysis
  - Analysis workflows and use cases are evolving:
    - too early to think about committing real money on specific dedicated solutions → AF: natural performance evolution of the existing (good) sites
    - But Person Power needed to make serious evaluations of tools and technologies
  - We want to be fair
    - careful to avoid walled-off resources only available to a certain region → Federated Identity
  - Fairshare and Priority

Areas of concern

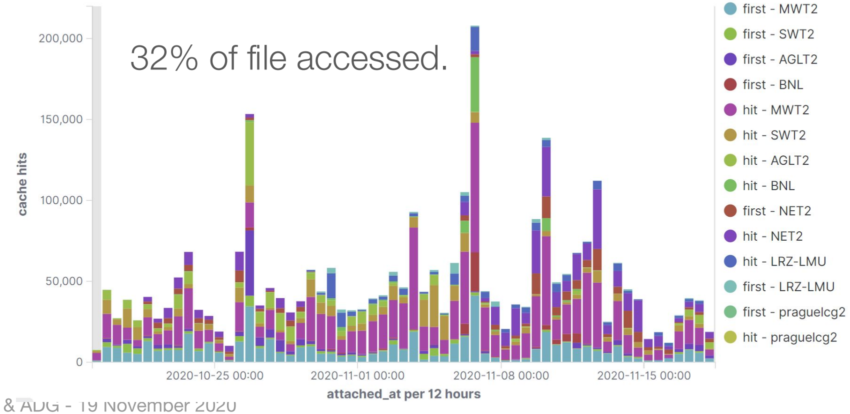
- Run a 10-times bigger infrastructure w/o comparable increase in personpower
  - Automation is paramount
- Long-term development and support of FTS
  - Engagement in strategic evolution.
- Future of DPM and alternatives for “small” sites
  - Discussed in several places, it is a concern.
- Growing diversity of storage technologies (CTA, Echo, CEPH), and increasing use of heterogeneous resources
  - Very good to engage in new technologies, but need development and integration efforts on experiments side! Would be useful to drop the "old" ones, tails are lethal.
- Engagement of experts (sites and storage)
  - Very few experts (needed for the real work!), too many meetings, impossible to have all the correct relevant people in the right meeting → need consolidation.
- Sustainability
  - Some of our tools have become/are becoming "common" beyond HEP. And we are using tools "common" beyond HEP. Still, despite not being sexy, packaging / documentation / education / dissemination beyond our infrastructure is vital for sustainable computing
-

BACKUP


Honey ... I, for one, suggest your company needs a better file backup storage plan.
Blazek

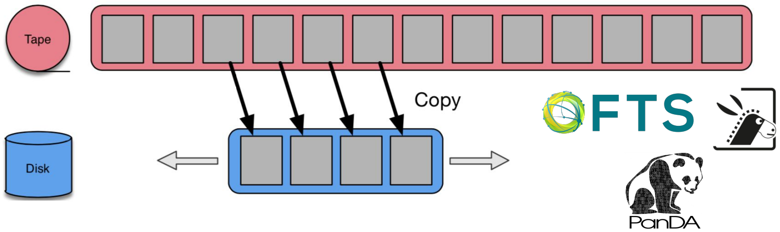
ATLAS, caches, diskless sites, and lakes

- ARC cache, experience, numbers. (multi node shared FS, single node)
  - O(15) sites using it: NDGF, HPCs, BOINC, …
  - From few TB (sites running simulation-only) to 500TB (all workflows)
  - Controlled data transfer asynchronous to the job, but only possible through “push” mode
- Xcache
  - MWT2, AGLT2, BNL, SWT2, NET2, LRZ-LMU, Prague2
  - Used through VP mechanism by ANALY_VP panda queues.
  - All single node SLATE deployments. Multinode deployments will be supported by VP.
- Diskless:
  - Birmingham (with cache)
  - Romania
- Lakes:
  - Israeli lake, French lake



The Data Carousel

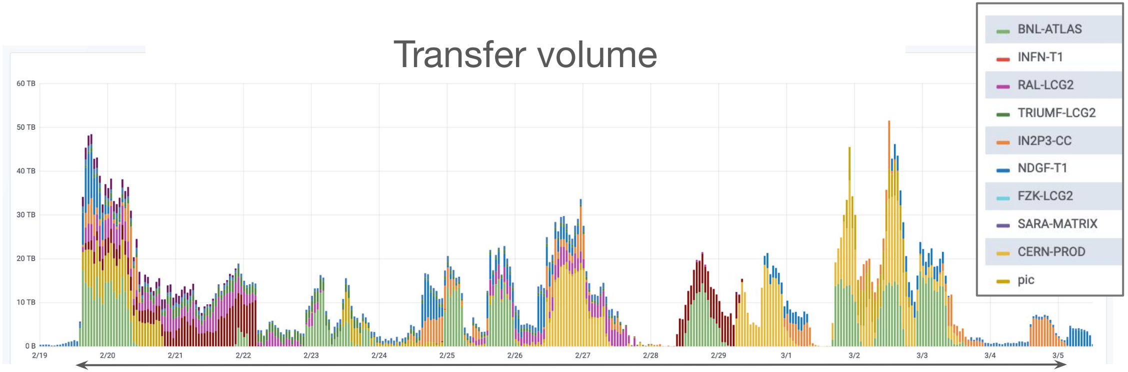
- New workflow: treats expensive fast storage (disk) as a cache
  - Data is staged from cheaper media (tape) on demand
- On demand reading from tape without pre-staging
  - Rolling disk buffer with adjustable size, tuned to suit available resources and production requirements



- Used by ATLAS production in a recent Run 2 DRAW_RPVLL reprocessing (skimmed RAW format)

- Full Run 2 input RAW data (18 PB) staged over several weeks rather than all at once: less disk space needed

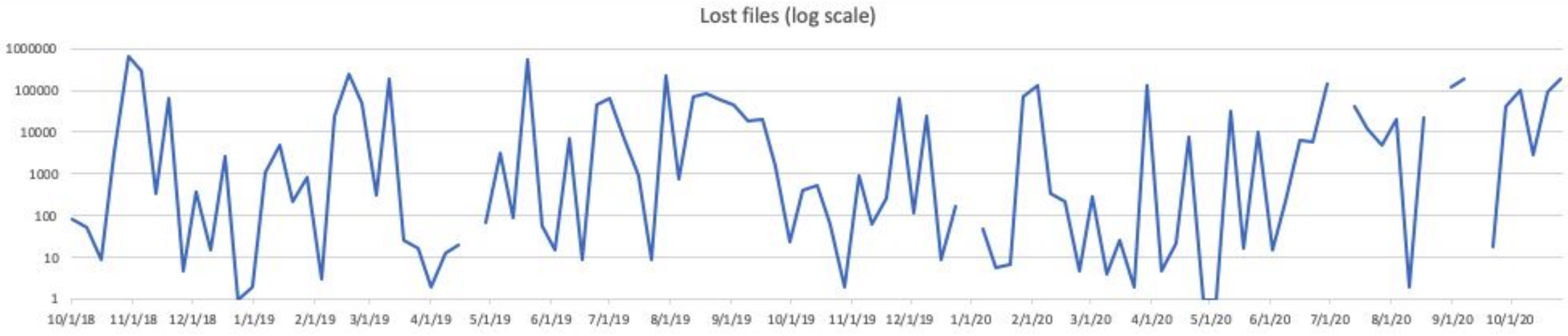
- As an example: 2017 (total 5.7 PB)
  - < 1 PB on disk at one time
  - Several 100TB processed and removed from disk after ~ days



First 2 weeks of 2017 DRAW_RPVLL reprocessing

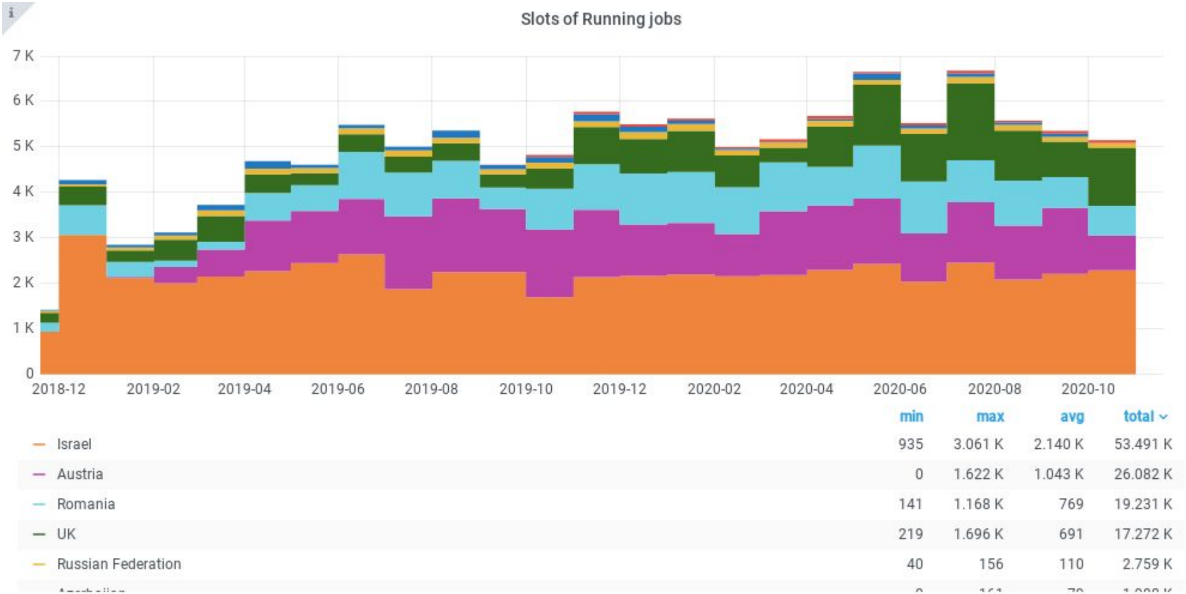
Lost files

- Complexity of auto-recreation
  - Automated a lot, but still manual
  - Devil is in the corner cases… if you "recreate" more events is *not* good.
  - Disk servers down only temporarily



CPU only Grid sites

- Access remote Grid storage
  - IO activity : From *simulation only* to *production+analysis*
  - Associated to 1 Grid storage for Read and Write
  - Long term experience now
  - Aim : keep all ATLAS activities as much as possible (network bandwidth, remote storage resilience,...)

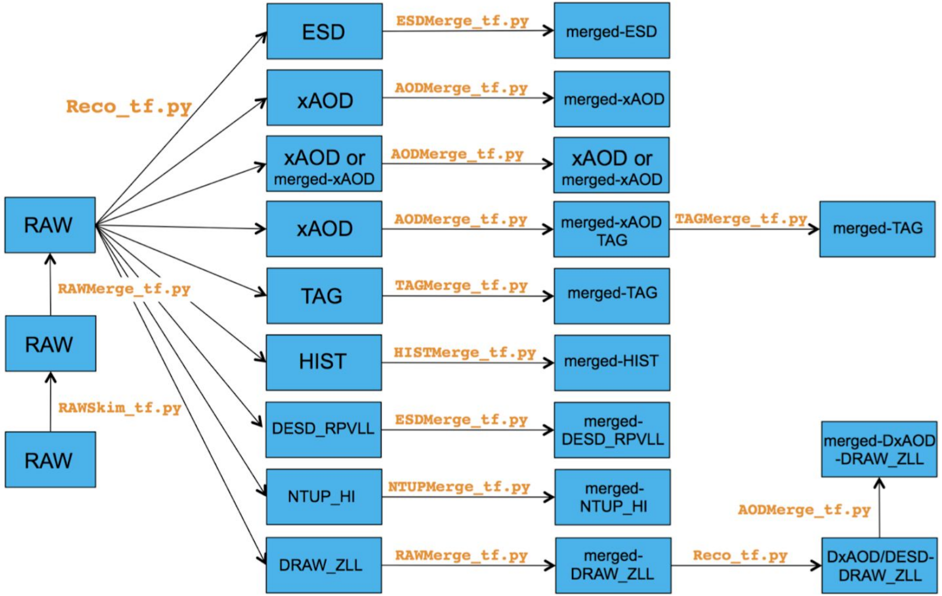


Data reprocessing workflow and formats

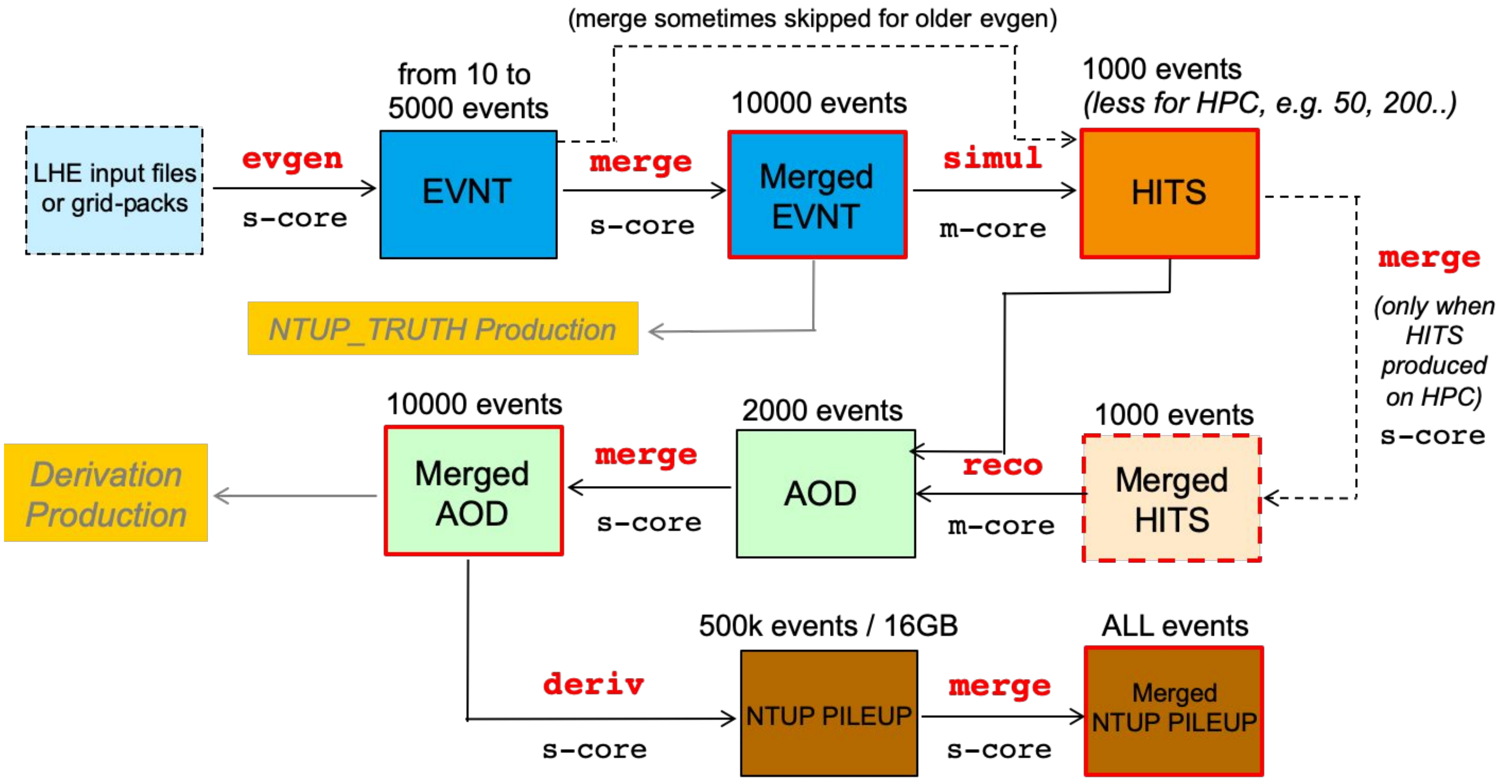
Generally, a data reprocessing comprises a reco step, followed by merges

Request is not single one-to-one steps: we have multiple outputs, varying levels of merging

- Reco produces multiple outputs using the RAW inputs: AOD, HIST, DRAW*, DESD(M)*, DAOD*, and (very rarely) ESD
- In the on-going 2015 data reprocessing campaign: 14 formats, including 4 DRAW, which subsequently run reco again to produce further DESD(M)* and DAOD* formats
- That makes total of 3 different DAOD types in ATLAS: from RAW, from DRAW and later from AOD (derivations)



MC production workflow



Keep only the (merged) dataset at each step, according to the relevant rules of the lifetime model for each data type