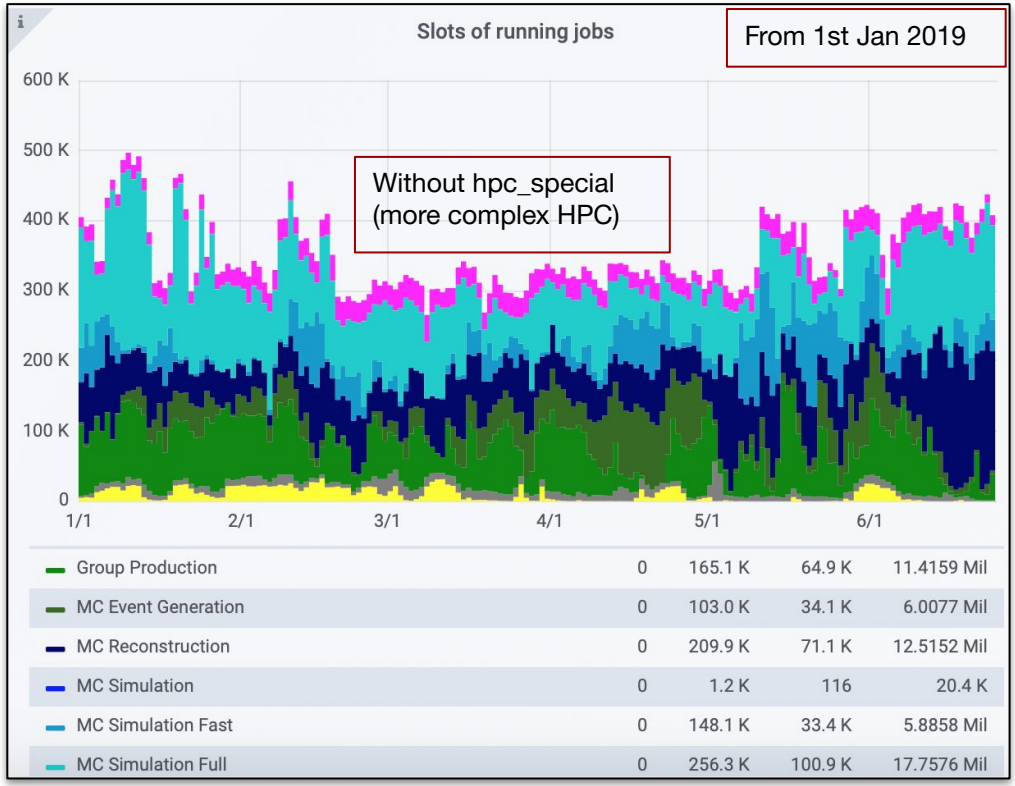
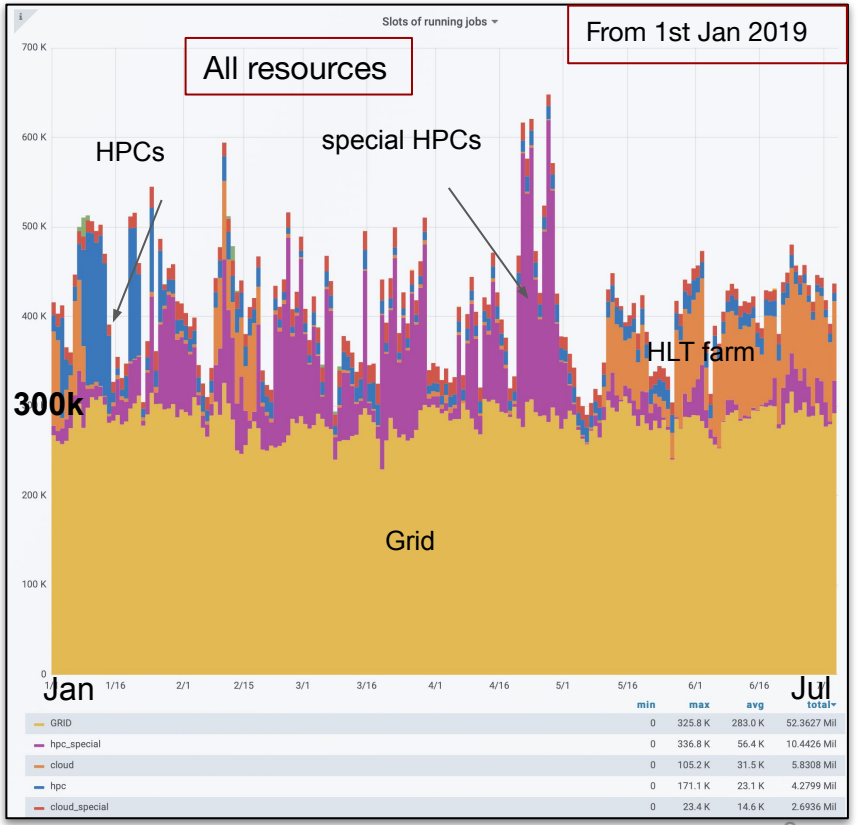
ATLAS Computing Outlook

GDB 10 July 2019

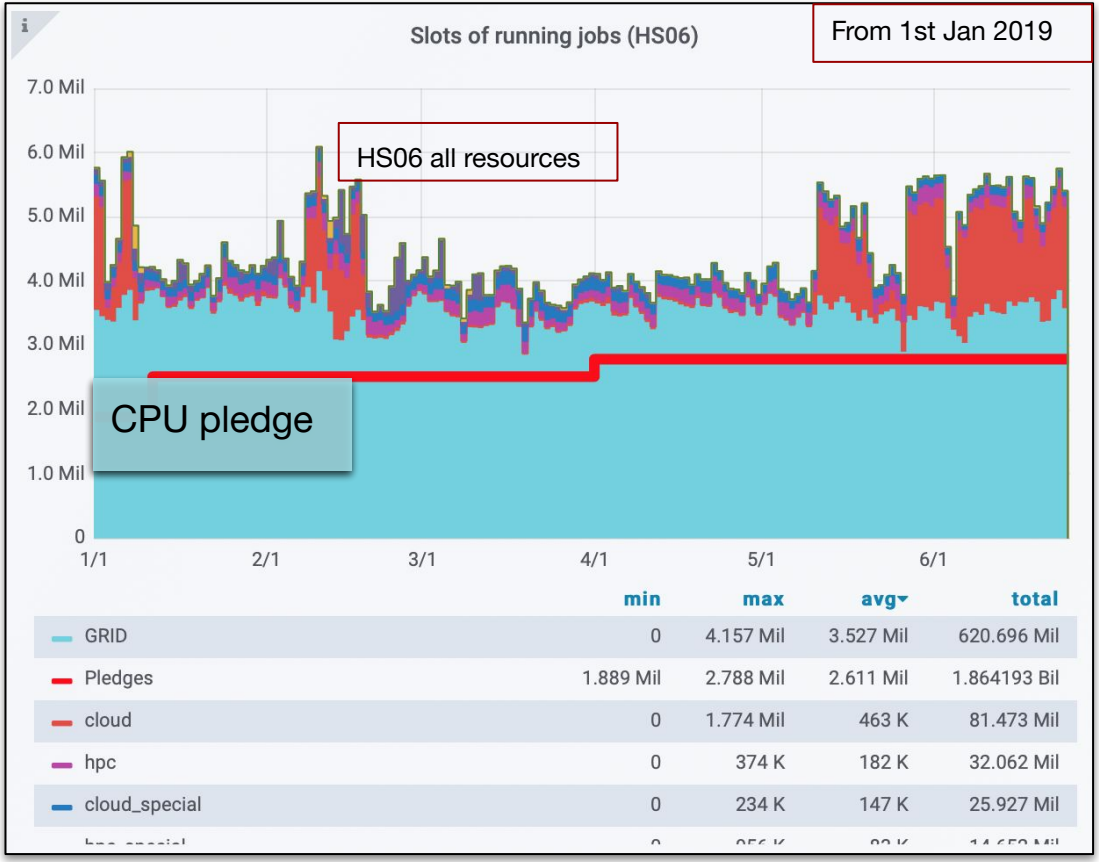
Ale Di Girolamo (CERN IT)
on behalf of the ATLAS SW& Computing community

Compute resource usage

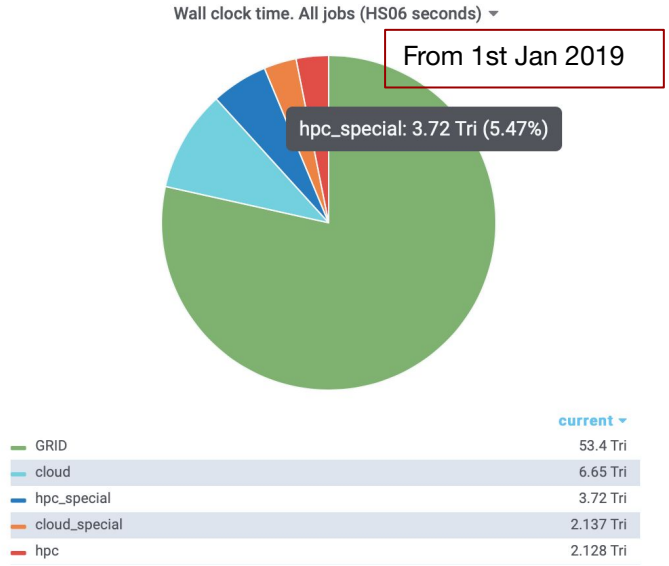




Compute usage: by resource type



	min	max	avg	total
GRID	0	4.157 Mil	3.527 Mil	620.696 Mil
Pledges	1.889 Mil	2.788 Mil	2.611 Mil	1.864193 Bil
cloud	0	1.774 Mil	463 K	81.473 Mil
hpc	0	374 K	182 K	32.062 Mil
cloud_special	0	234 K	147 K	25.927 Mil



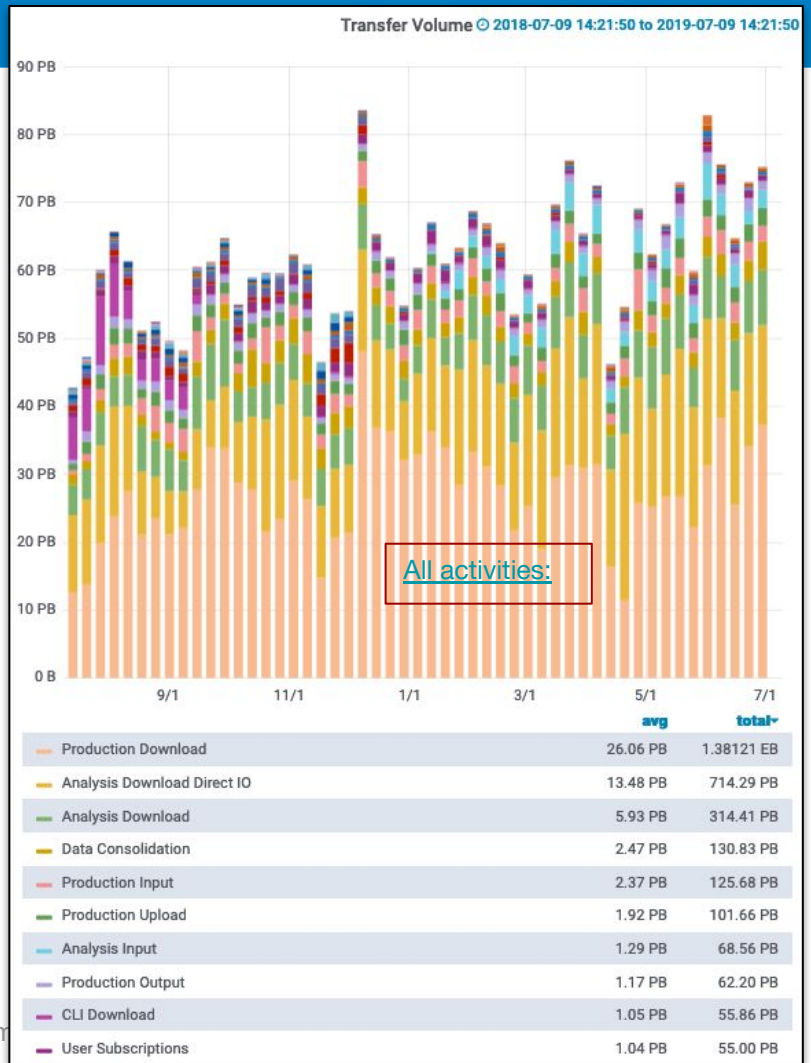
HPCs + HPC_special ~ 9%

- HPC_special: custom defined tasks, tuned ad-hoc for each HPC_special

DDM: last year

Average: 8+ PB/day, 10M+ files/day

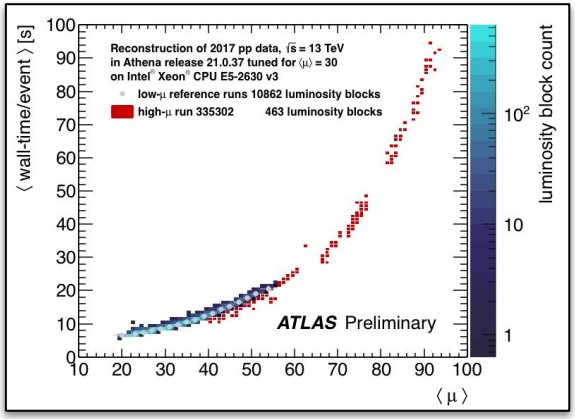
- Synchronous (storage to/from WN): ~80% of total traffic
  - Mostly LAN
  - Rucio mover in pilot to handle (almost all of) them (instead of custom scripts):
    - Benefit of traces, future possible flexibility in choosing other replicas, in general clean decoupling of DataManagement from WorfklowManagement
- Asynch, all managed by FTS through Rucio:
  - FTS is a fundamental service, very good collaboration with FTS devels
  - Even more needed in the future, e.g. deeper analytics to understand tails



Towards Run-3

- Run-3 "natural evolution" of Run-2:
  - But not at all identical! e.g. mu ~56 -> more CPU at T0 needed:
    - rough calculation: 50% more CPU in T0
  - Uncertainty on the 2021 LHC operations. LHCC guidelines:

> For contingency planning, the machine efficiency can be assumed to reach the normal value of 50%. This results in the following luminosity envelope:
>
> | | Baseline | Upper limit |
> |---|---|---|
> | ATLAS/CMS | 17/fb | 42/fb |
> | LHCb | 3/fb | 7/fb |
> | ALICE | 36/pb | 90/pb |
>
> We would like to emphasize that the upper limit is for **contingency planning only (i.e. raw data tape storage), not physics planning**.

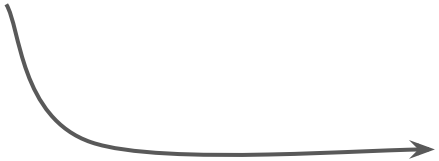


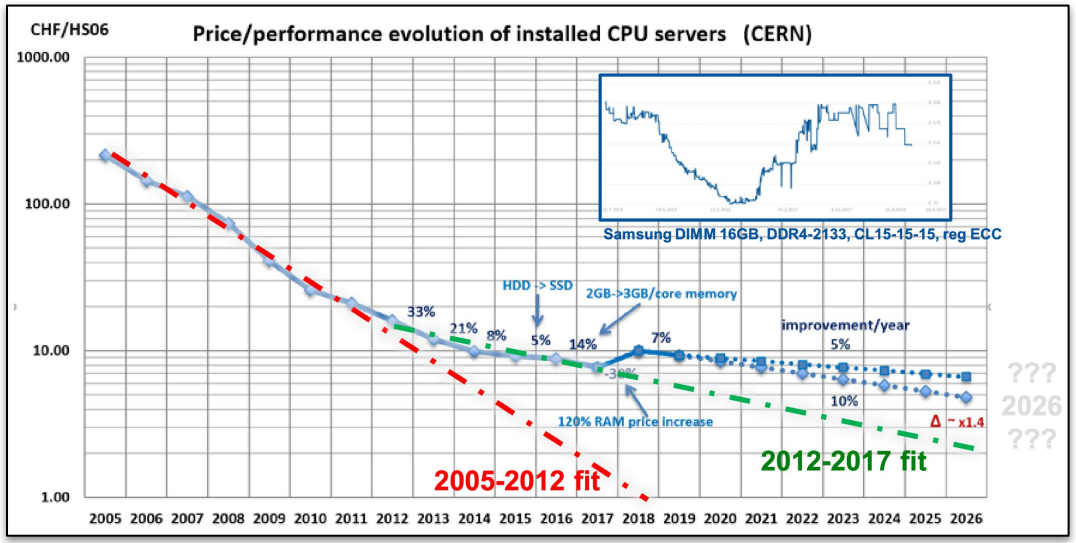
- Preparation of Release 22 ongoing
  - Use the new multithreaded framework to optimise memory usage (driving hardware costs)
  - Upgraded detector, improved reconstruction, better physics performance
  - Run-2 reprocessing with release 22 to provide a consistent dataset for Run-2 and Run-3
- Evolution of the infrastructure, e.g. CentOS7, IPv6, Pilot2, Containers, rucio mover migration…
  - N.b. this above is just one line…. But a lot of work from many people
- Use a REST interface for access the conditions via COOL
  - Optimise DCS data volume

Computing roadmap for the HL-LHC

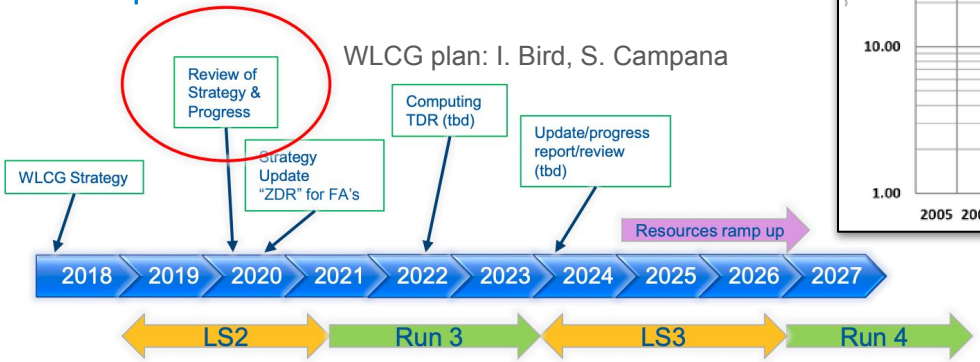
HL-LHC computing is a challenge

- More resources needed. Higher lumi, trigger rates (10KHz)
- No silver bullet! Every cm counts….
- Hardware getting cheaper, but not enough to compensate our growing needs!
- General loss of predictability of hardware costs, not clear how much the "flat budget model" can be for real

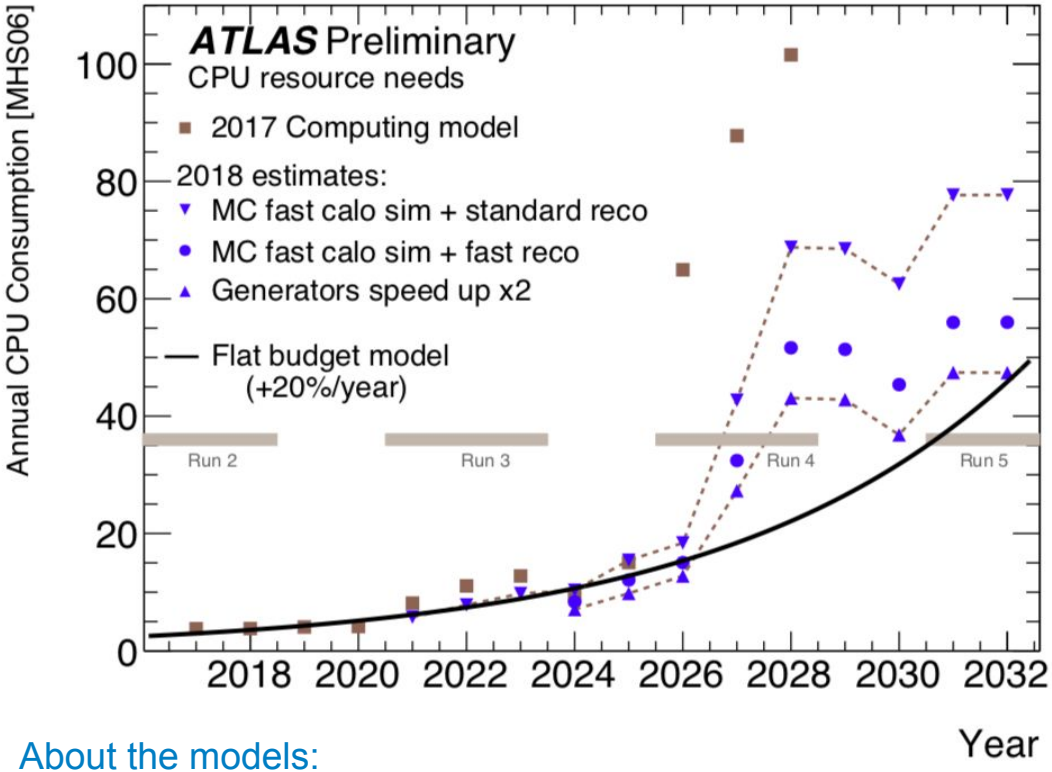


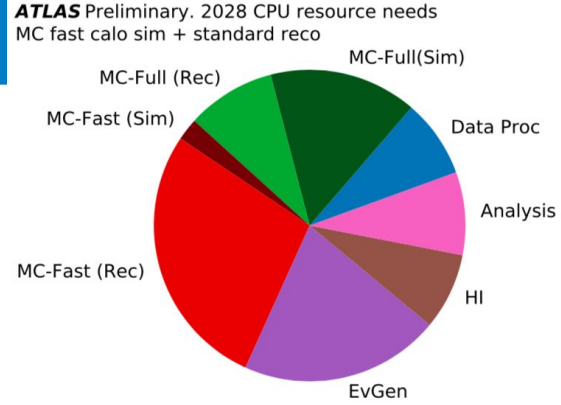


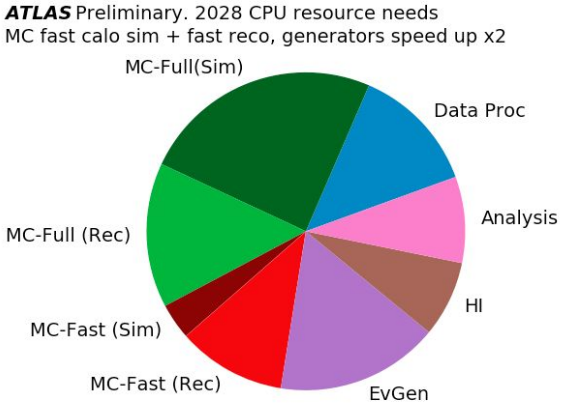
A strategy review, followed by a CDR planned for 2020



CPU projections for HL-LHC



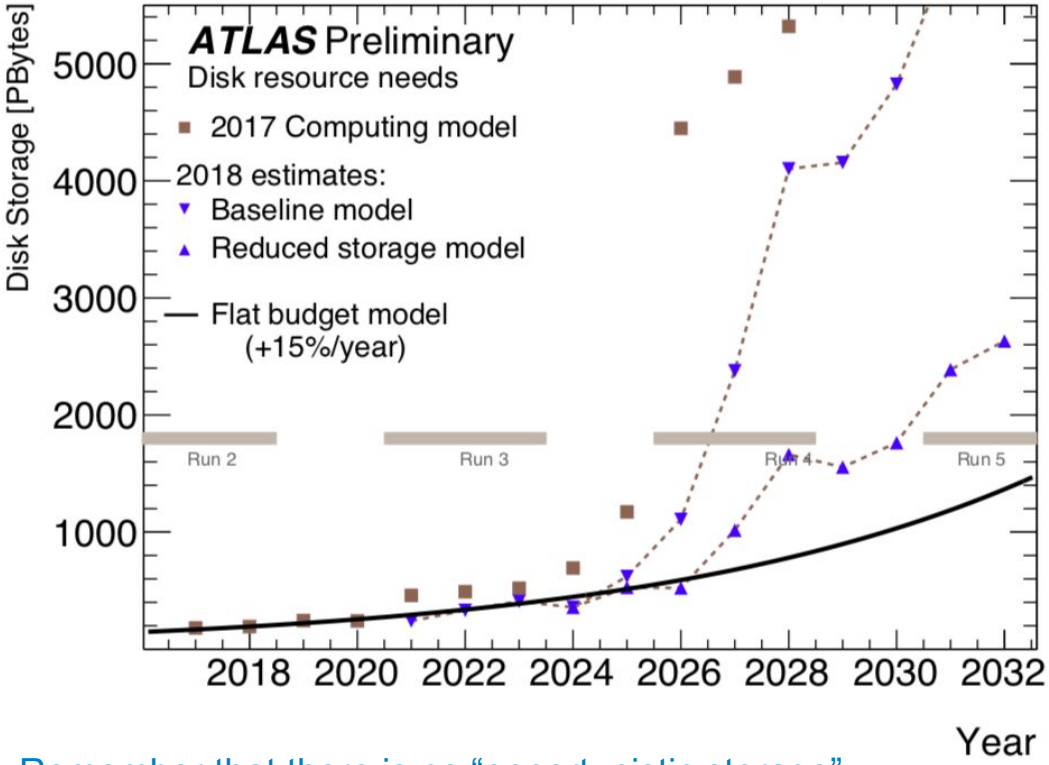


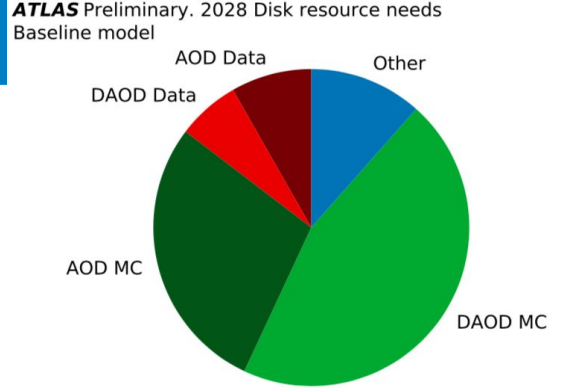


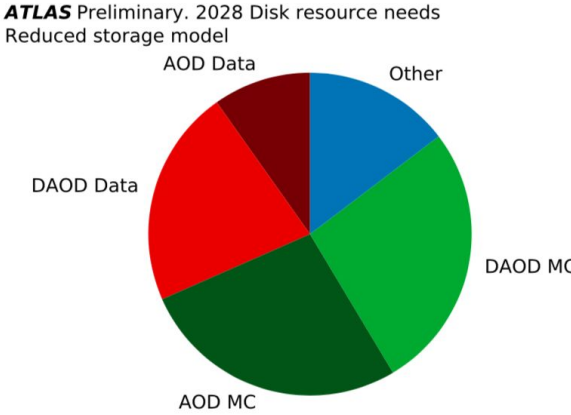
About the models:

- Parameters stored in an Excel file
- Read in parameters and process in python
- Parameters are “Internal”

Disk storage projections for HL-LHC





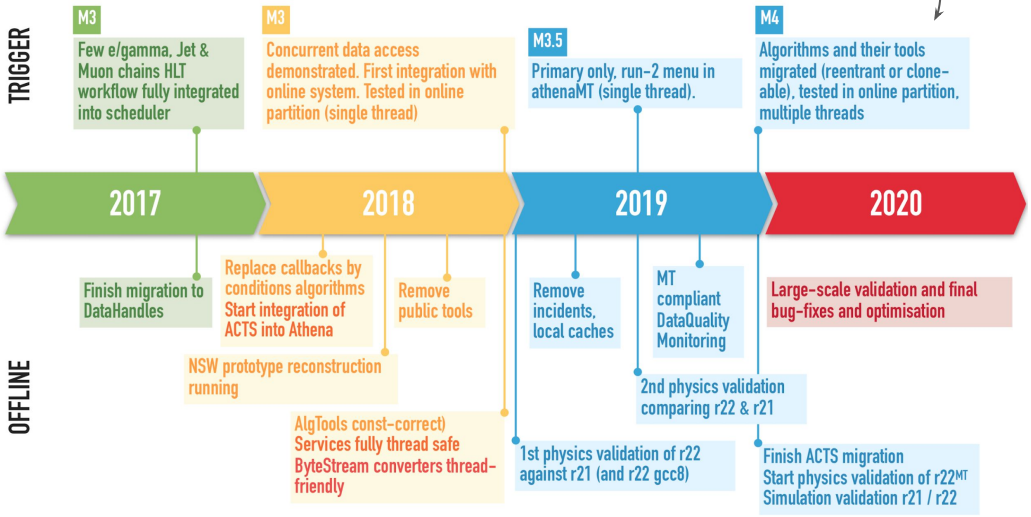


Remember that there is no “opportunistic storage”

- Could we replace some storage with cheaper and slower options? Tape?
- Draconian solution is to have (much) less Monte Carlo
- Parameters are “Internal”

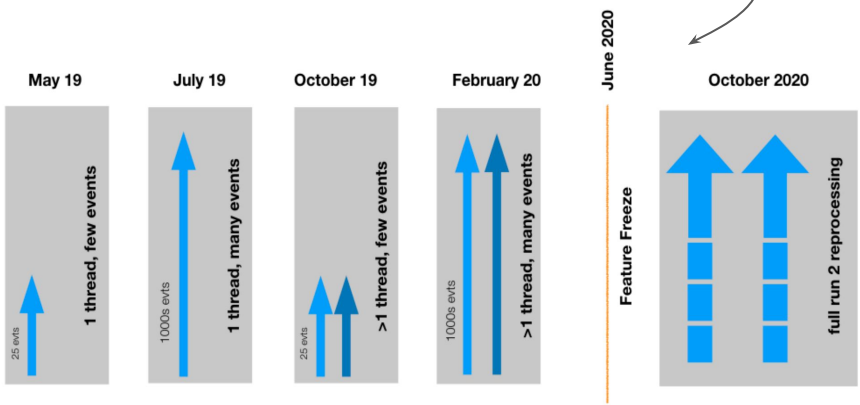
SW readiness for Run-3

- First round of validation of release 22 completed
  - Differences, as expected, being worked on
  - HLT test planned for end-June
  - Further validation round planned for July



Migration strategy:

- Complementing the Component focused plan
  - (made good enough progress now, able to run "something")
- with the Functionality focused plan
  - (let's run and fix the failures!)



SW grants

- New initiative to increase effort and expertise in SW

- 6 people to be based at CERN for ~6 months
  - Working on general MT migration, global track-fitting, job configuration and muon reconstruction
- Commitment from their groups to continue involvement in SW
  - Long term support
- Training
  - Athena tutorial specifically aimed at the grant awardees in September
    - (but open to everybody willing to participate)
  - Expanding ATLAS SW Docs in the meanwhile

DBs

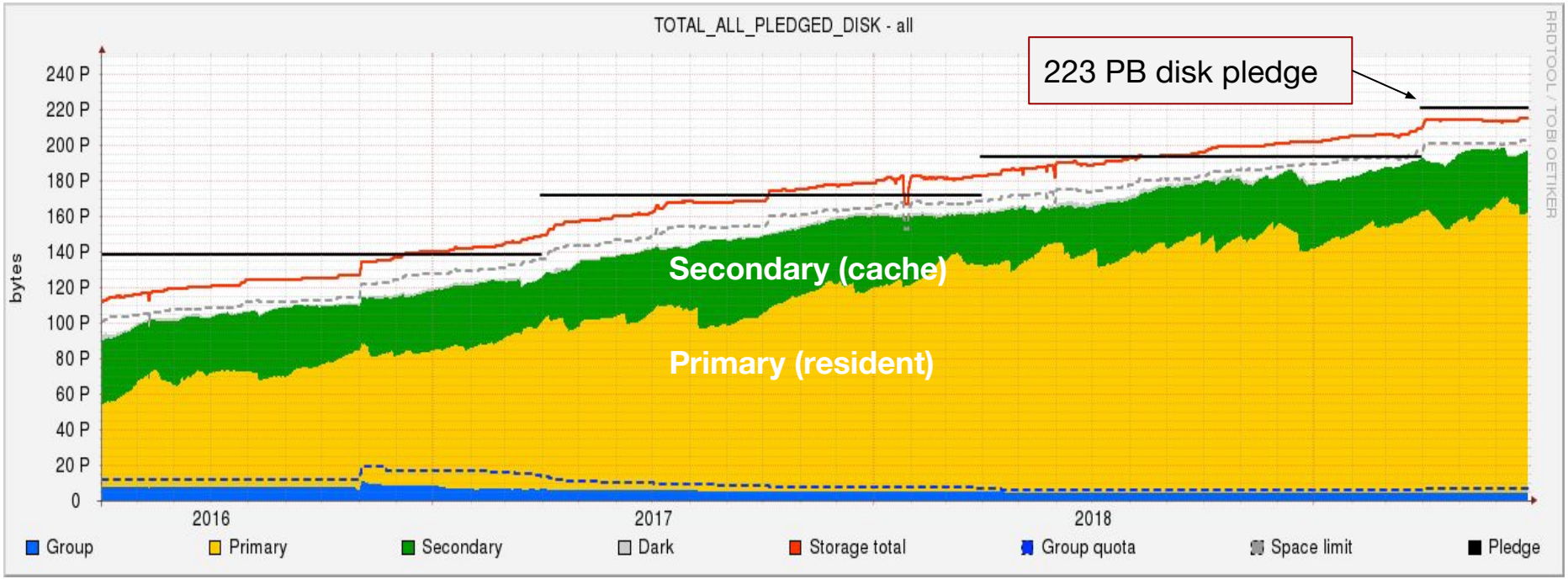
- Oracle 19c upgrade
  - Long term support release - should be ok for the entire Run-3
  - Target end of the year (validation with users and upgrade)
- REST access to COOL data
  - help to make more cacheable queries
  - Prototype ongoing, stress testing to understand HW needs
  - Trying to evaluate possible advantage of this approach wrt existing clients to understand if worth to deploy some service during Run-3
- DCS Hackathon in September
  - Main goal: reduce the amount of DCS data
  - Very much subsystem dependent
  - Share code and expertise
- Analytics for DB Ops
  - Goal: monitor conditions data usage to spot the problematic data
    - Huge progress done to find and report Frontier intensive workflows
  - Running at CERN Kubernetes in test mode, to be deployed in prod

Monte Carlo & Reprocessing plans during LS2

- Monte Carlo simulation in 2019 (and 2020) discussed by Physics Coordination
  - "Business as usual"
  - … driven by physics needs and publication deadlines

- Full Run-2 reprocessing planned to start in Autumn 2020
  - Using release 22 as for full Run-3 processing.
  - Consistent SW for Run-2 and Run-3. Concentrate expertise on a single release.
  - Reconstruction in 2018 (Data and MC) of good physics quality
- Various "smaller" scale reprocessings for Heavy Ions, long lived particles
  - smaller in output but in Sept 2019 already foreseen RPVLL repro which requires all the Run-2 RAW
  - "Good" for Data Carousel!

Disk pledge and utilization

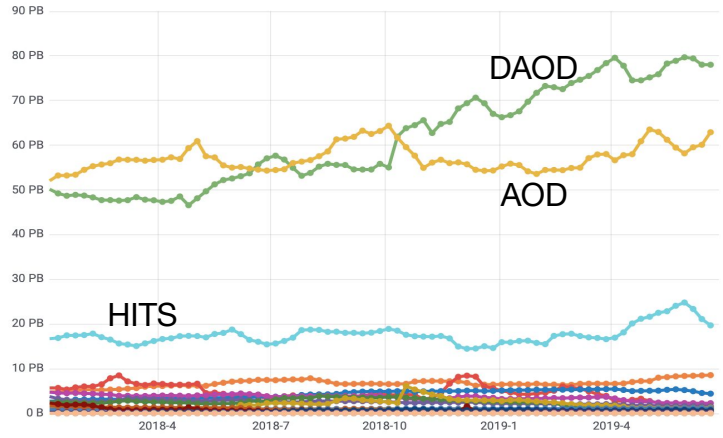


- No pledge increase until 2021
- Disk space will be very tight later this year and esp. in 2020
  - Too tight at this rate!
  - Will need to be even more proactive than now in managing what we have on disk

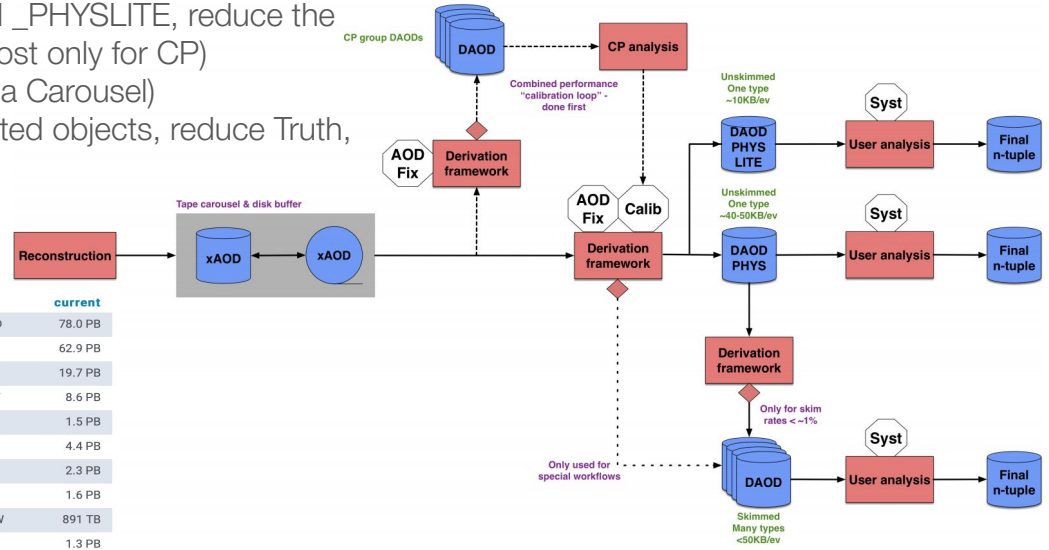
Analysis Model for Run-3 (AMSG-R3)

Most of the ATLAS disk usage for analysis formats

- Study Group established with the goal to reduce disk footprint by >~30%
- Very detailed study, circulated in the collaboration. Shorter (and public) presentation planned for CHEP
- Changes to the analysis model advised. Important to implement them during LS2. In a nutshell:
  - Introduce 2 new formats DAOD_PHYS and _PHYSLITE, reduce the number of all the others DAOD (DAOD almost only for CP)
  - increase the (organized) usage of tape (Data Carousel)
  - AOD/DAOD size reduction (e.g. use calibrated objects, reduce Truth, track and trigger info)



	current
DAOD	78.0 PB
AOD	62.9 PB
HITS	19.7 PB
EVNT	8.6 PB
RAW	1.5 PB
log	4.4 PB
ESD	2.3 PB
DESD	1.6 PB
DRAW	891 TB
RDO	1.3 PB
HIST	1.1 PB
NTUP	756 TB
no_name	316 TB
TXT	720 TB
user	297 TB
group	217 TB
TAG	14 TB



Data Carousel

‘data carousel’ - an orchestration between workflow management (WFMS), data management (DDM/Rucio) and tape services whereby a bulk production campaign with its inputs resident on tape, is executed by staging and promptly processing a sliding window of X% (5%?, 10 %?) of inputs onto buffer disk, such that only a small fraction of inputs are pinned on disk at any one time.

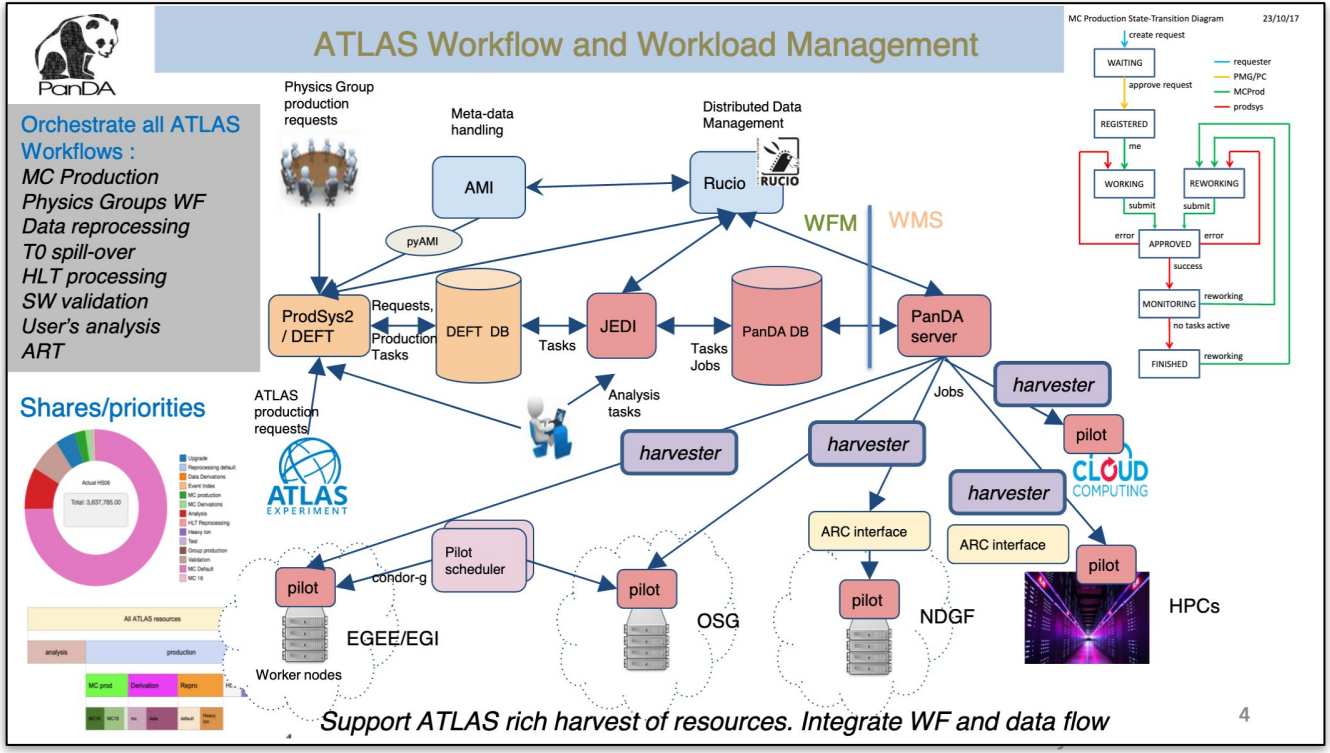
Phase 1 - Tape systems stress test at Tier1s. Concluded

Phase 2 - End-to-end: analyse 2 PB of AOD data from 2016 "directly" from tape

- ProdSys2 requested prestaging without destination
- Throttling to be done by Rucio
- Very encouraging results, Tape read rate increased of a factor ~2-3 wrt of first tests done 1y ago
- Data grouping on tape during writing is not doable yet
    - Would increase efficiency, discussion (with dcache and others) ongoing, difficult to find balance between optimization and overdoing.

WFMS and DDM for Run-3 and Run-4

- ATLAS relies on Panda and Rucio ecosystems
  - Clear separation of responsibilities but tightly interfaced/communicating
  - Never stopping evolutions



SCIENTIFIC DATA MANAGEMENT
RUCIO

No need to go here into details with Rucio:

- will be at GDB soon
- 2nd Rucio community forum few months ago

Panda & Rucio evolution: just some bits

- Panda ecosystem: several tools/frameworks evolving
  - Harvester, complete re-architecture of payload scheduling:
    - Uniform approach for Grid, Cloud, HPCs
  - Pilot2
    - Flexible design with modular architecture
    - Support for containers for users and production
    - Improved error identification
  - Data Carousel
  - Jumbo jobs on HPCs (whole node EventService jobs + grid spill-over)
  - PandaQueues GrandUnification
    - Merge Prod SCORE/MCORE/HighMem first and later also Analysis -> one per site
- Rucio (just a short selection)
  - Deletion agent (Reaper) rewrite,
  - Archive support,
  - kubernetes deployment of all services,
  - Global quotas for users
  - Optimize usage of Tape
    - For reading AND for writing
    - SRM-less tapes (e.g. CERN Tape Archival - CTA)

GPUs and ML

- Opportunistic Resources available to us on acceleration hardware
  - Technical challenge to access them efficiently
  - Much debated topic (e.g. HPC X-Experiments meeting on 10 May, more later)
  - Accelerating something is better than accelerating nothing
    - But not trivial

- ML training workloads submitted via panda on a few sites via containers
  - Expanding the list of sites with GPU resources
  - Increase the user base (currently flavour tagging training)
  - Should increase our ML ideas
  - Still small compute time overall for now, might evolve, need to be ready

- GPU-enabled framework on a longer time scale
  - Athena interfaces defined, tests being made. Target Run-4

Running on non-x86 HW

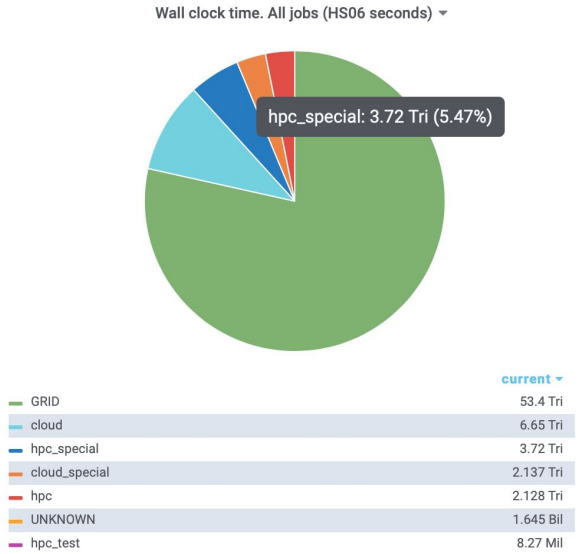
- For instance some of the new big HPC coming out
  - Summit has Power9

- Doable with some effort
  - ATLAS SW stack is mainly C++
    - Based on externals components such as root and Geant4, maintained as LCG-externals
  - Working examples ARM-build in 2016 and just few months ago with Power9
    - Only subset of code, AthSimulation
    - Most of the challenges related to building externals

- Long-term sustainability
  - Need to integrate in build system, both ATLAS and LCG ones
  - Most probably need build cluster dedicated
    - Isn't it needed also by other experiments?

HPCs

- Useful first HPC X-Experiments meeting on 10 May
- "Grid like" HPC exploited with all workloads
  - E.g. SuperMuc, UIO, CSCS, HPC2N, Marenostrum ...
- HPC_special (mostly US big machines such as Titan, Cori, Theta)
  - Very successfully exploited with EventService and Harvester
  - So-called Jumbo (an co-Jumbo) jobs
    - Whole node ES jobs, many node running in parallel, Grid resources to spill-over
  - No free lunch, each huge HPC requires special dedicated attention
- How do we use future HPCs with GPUs and eventually non-x86 ?
  - ML training is not enough for now
  - Try hybrid approach today ? ML on GPU + regular simulation on CPU



HS06 walltime in 2019

- HPC_special 5.5%
- HPC 3.1%

? How many HPC resources available in the future?

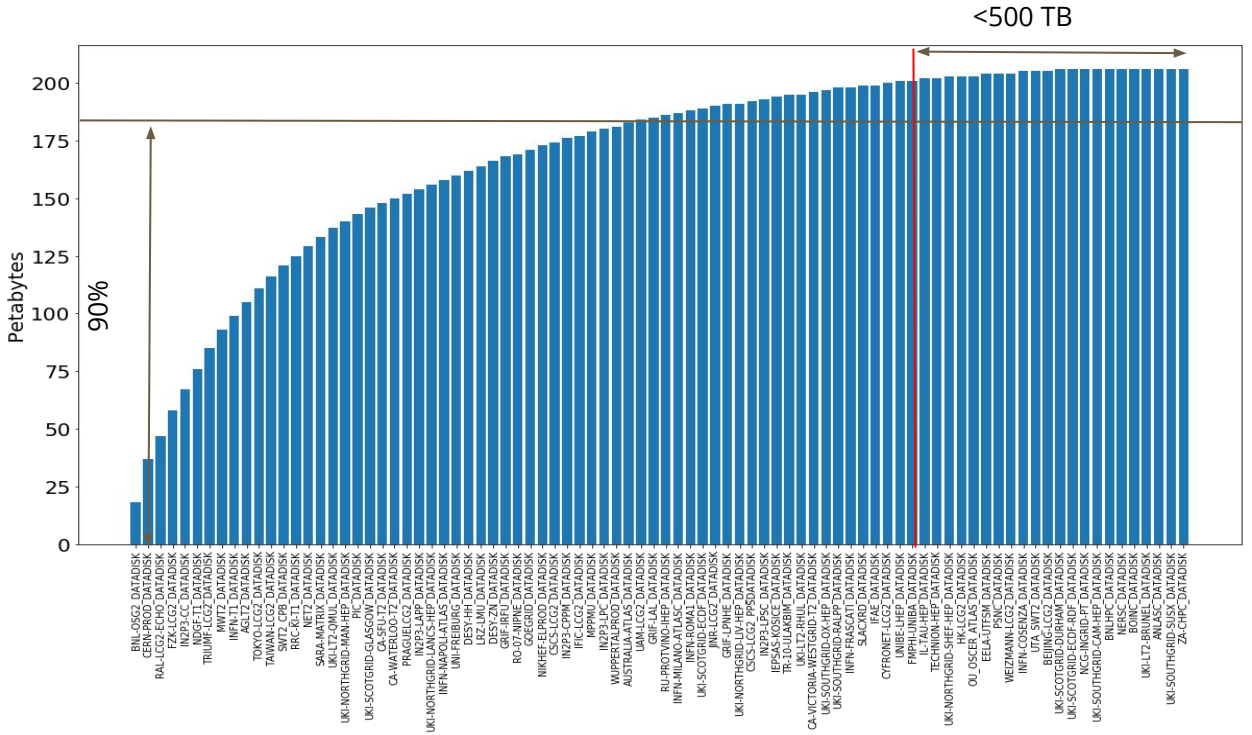
Facilities evolution: Site Consolidation

Still a significant amount of problems coming from the smallest sites

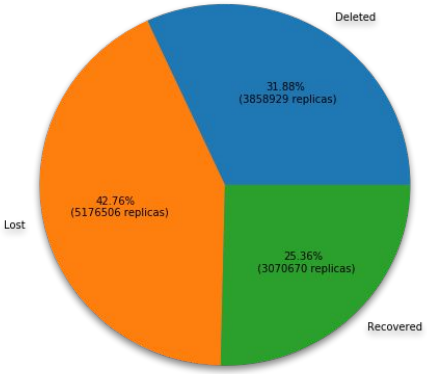
- E.g. one site was offline for more than 2 months. Decision taken to decommission

Strategy not changing: decommission storages <500TB (<0.5% T2 pledge)

- Some sites already decided to go storageless or to use Xcache (e.g. Birmingham)
- Some other UK following
- Israel 3 sites decided to try too.
- Very slow process

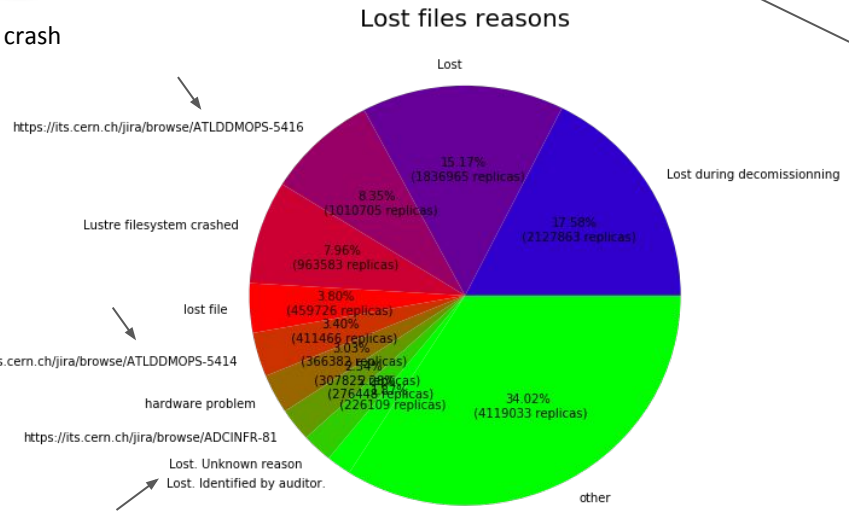


DDM: Lost file statistics



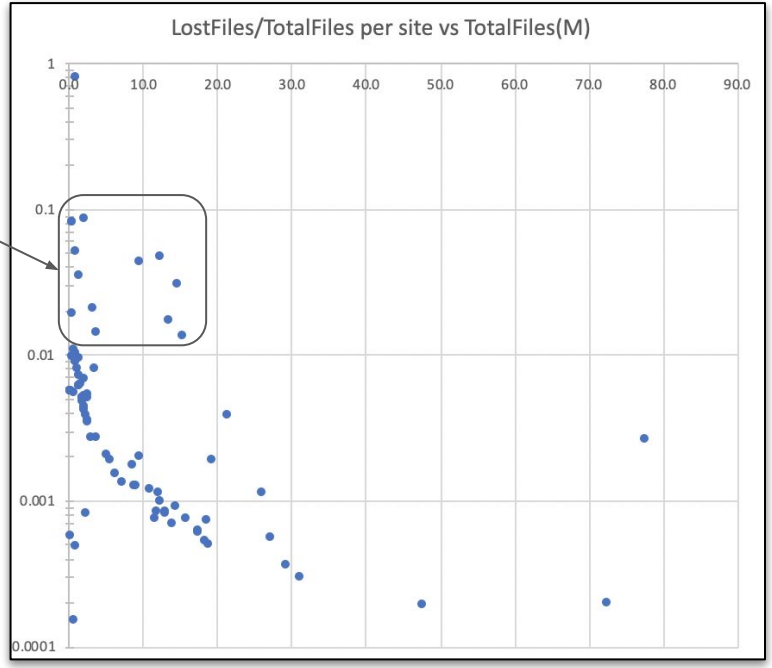
Statistics since the beginning of Rucio (2015): 12M files total

Disk server crash



Identified by consistency checks

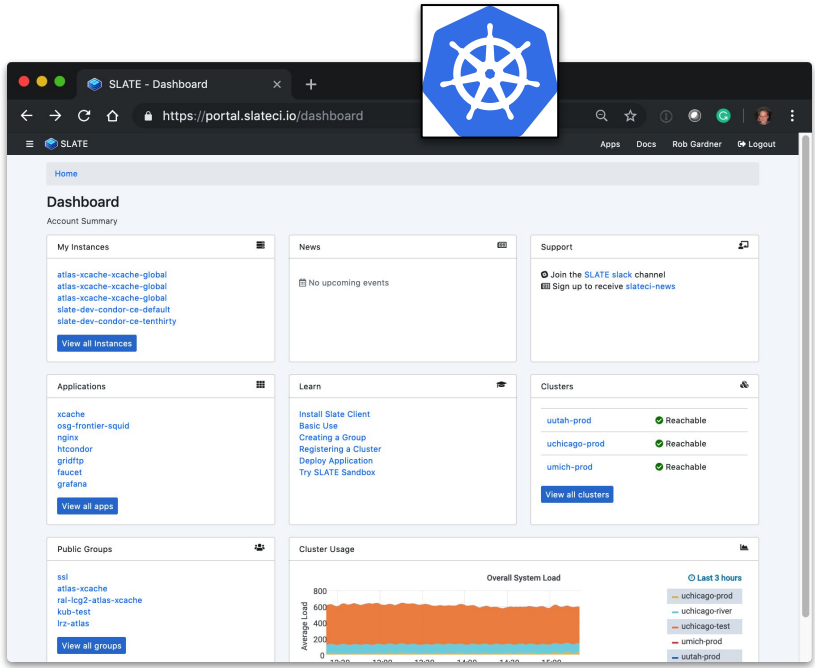
Small sites (10M files ~ 1-2PB) are not all behaving similarly - but we see a trend



Facilities Operations - automation - new ideas

Changing the Computing Ops paradigm → minimize burden: from "central + sites" support to "distributed teams" supporting specific services

- Remotely manage edge services at sites by expert teams from trusted organizations
  - Quickly deploy updates & easily introduce new services
  - Save time and effort for local site admins, still relying on the precious distributed expertise unique to our community
- SLATE - Edge federation via lightweight server/client overlay using Kubernetes
  - Support for CVMFS, ingress controller (multi-tenant, scoped privileges), Prometheus monitoring, curated application catalog w/ Jenkins CI
- Site security & policy conscious
  - SLATE works as an unprivileged user
  - Single entrypoint via institutional identity
  - Site owner controls group whitelists & service apps; retains full control
  - Now in contact also with WLCG Security team
- "Trust model" is definitely still a ?, all to be understood
  - Start "simple", e.g. PerfSonar and Squid is a possibility
  - Evolving once time are mature to more complex ideas, e.g. caches and why-not batch



Understanding our needs: benchmarking

Benchmarking: long-term key activity which can allow to optimize our needs minimizing the costs

SW is changing:

- shift of paradigms, architecture, heterogeneous resources (GPU, non-x86…)

What do we want?

- get "enough" resources to accomplish our physics mission!

is HS06, a decade-old suite, still the best tool we can use?

- For now we use it (been using it since 10+years) for pledging the WLCG computing resources …

… what's about the future???

- HS06 lack of correlations with the HEP workloads [CHEP18 proceedings], and HS17 not doing better...
- not so easy to change to "events/sec" (or related metrics), need to be careful, but we need to start!

Working on a field specific suite:

- X-experiment activity (within Hepix WG)
- containers encapsulating experiments workloads
- flexible, state of the art in terms of tech used - maintainable and extensible, and realistic …
- E.g. GPUs: containeraize workload, submit, study performance; similar for new workflows, e.g. FastChain

Several interesting aspects related:

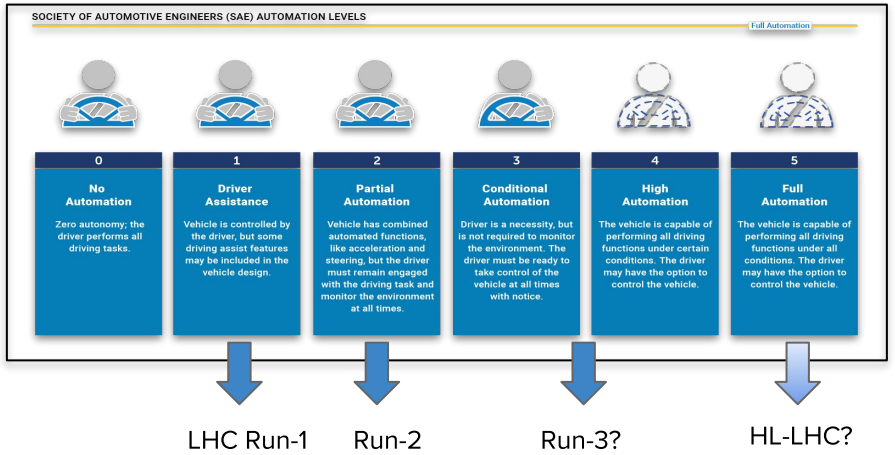
- Performance understanding: which are the key metrics? -> in collaboration with ATLAS Sw Performance Optimization Team (SPOT)
- Accounting evolution (and pledge evaluation): more flexibility?
- Understand sites performance

In general: "Understanding"

- Understanding (might) allow us to avoid both overcomplicating and oversimplifying
- Need (even more) involvement in understanding our complex environment
  - Understand performance,
  - Understand errors
  - Understand data placement and data transfers,
  - ….
- Analytics is (one of) the key
  - Need to invest brainpower in
    - building up/consolidate big data infrastructure with tools,
    - build know-how and
    - sharing expertise (not to reinvent the wheel each time)
    - … and it's taking time
  - One example focused on optimizing computing operations, the Operational Intelligence x-experiment activity (next slide)

Operational Intelligence

- Operational Intelligence - X-experiment effort aiming to:
  - Streamline and automate computing operations
  - **save manpower & improve resources utilization**
    - Increase level of automation in operation tasks
    - Cost reduction metrics: needed number of operators, reduce/optimize number (and location) of files/replicas
- By:
  - Identifying common projects
  - leveraging **common** tools/infrastructure
  - Collaborate, share expertise, tools & approaches
    - Across experiments
    - Across teams (operations, monitoring, analytics)
  - **Bottom-up** approach:
    - For now CMS, ATLAS, HammerCloud, Rucio, MONIT, DUNE/FNAL, LHCb

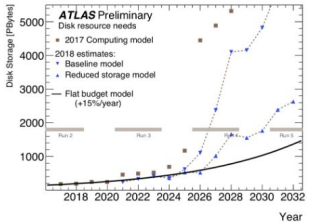


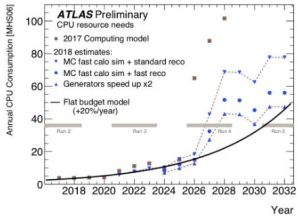
Distributed Computing R&D and Evolution activities

ADC is (pro)actively participating and sometimes lead many R&D and Evolution activities

- https://twiki.cern.ch/twiki/bin/view/AtlasComputing/ADCEvolutionActivities
- Bureaucracy to the bare minimum
- Still for ATLAS related activities need to be able to plan together goals and milestones

ADC Developments and Evolution





How do we address the projected CPU and Disk shortages ?

Examples:

- Effectively use current/new resources like e.g. HPCs and enable new workflows on GPUs
- Combine Disk/Tape efficiently with Hot/Cold storage models: Data Carousel, Data lake, Caches
- Reduce size/event and number of analysis formats: AMSG-R3 with DAOD_PHYS/PHYSLITE

Further developments to enable new technologies like e.g. Containers, New Monitoring etc.

We are doing a lot of R&Ds which should be discussed in the TCB meeting and ideally have a brief entry in the ADC evolution twiki (link) so that it's known to everybody

6

Summary

- Computing never stops: lot of activities during LS2!
  - Many commissioning ongoing
  - Fully using our resources for MC and analysis
- Main focus is the preparation for Run-3
  - Athena-MT is progressing. Challenges ahead
  - Resource request in preparation.
    - More disk needed to accommodate new data and new MC
  - Analysis Model StudyGroup recommendation provided, to be implemented
- R&Ds and Evolution activities in several areas
  - We are not alone, working in collaboration with many others
  - Need involvement of top experts, careful in focusing on most promising ones
- HL-LHC: strategy doc (internal) early 2020

"IT MAKES NO SENSE TO WORRY ABOUT THE FUTURE.
BY THE TIME YOU GET THERE, IT'S THE PAST!"