ATLAS S&C Week #70

October 4 - 8, 2021

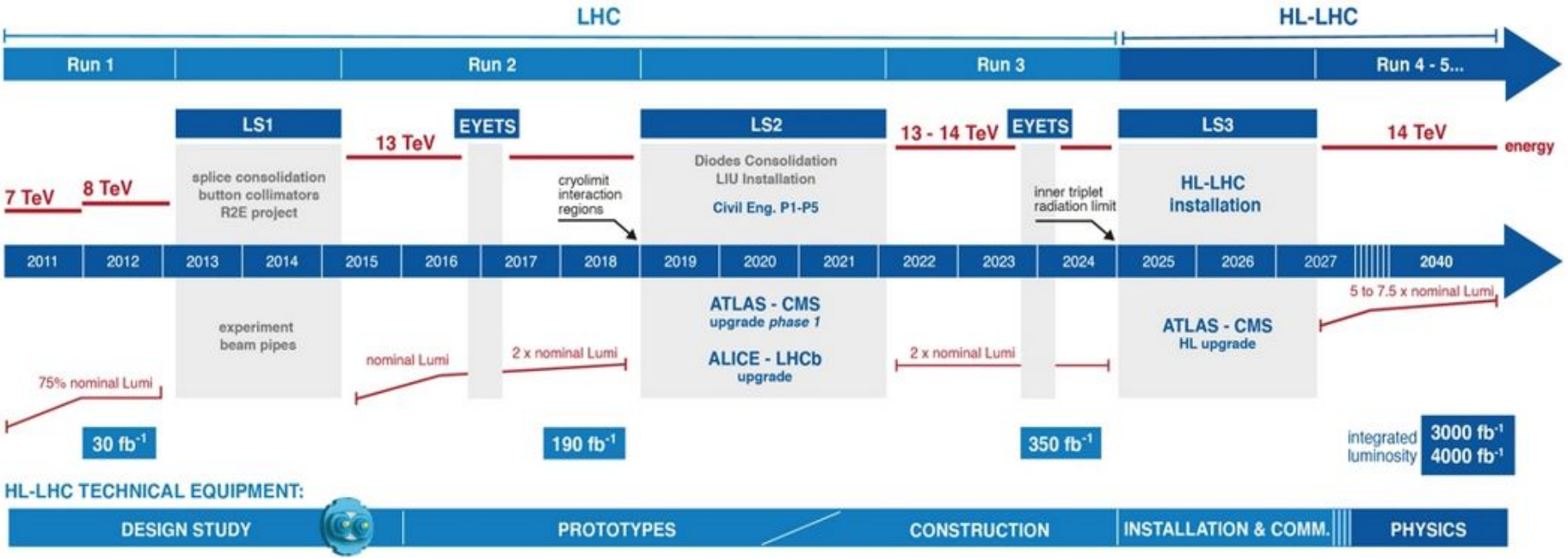
Highlights

Alex Undrus

NPPS meeting #62

October 22, 2021

Focus: readiness for LHC Run 3 (2021-2025)



At 15:46 October 19 ATLAS observed the first splash event of the Pilot Run (first time since the end of Run 2)

Major Presentations

- Readiness reports from 3 major domains
  - Software
  - ADAM (ATLAS Database and Metadata)
  - ADC (ATLAS Distributed Computing)
- Developments towards HL-HLC
  - ATLAS-Google R&D Project (GCP4HEP) led by Alexei (previously presented at this forum)

SOFTWARE: development workflow

Developers: 514 git contributors in 2021

63 new authors since June 2021

7420 MRs opened in 2021

MR shifters (2 levels)

GitLab

CONTINUOUS INTEGRATION
Jenkins Server

BUILD FARM
24 CC7 nodes

20,000 builds annually, tested with 2000 int., unit tests

Oracle-based in BigPanDA display

Multi-platform, multi-project
(Athena, AthSimulation...) builds

LXR Code Browser (NPPS/SDCC)

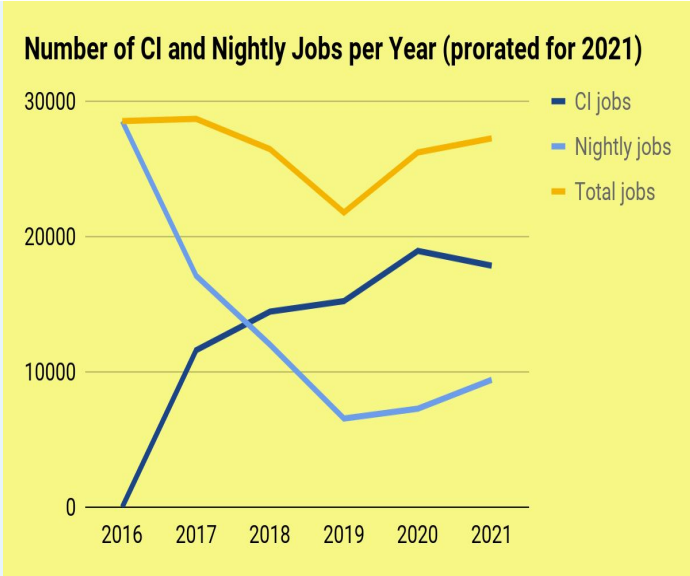
NIGHTLY SYSTEM
Jenkins Server

BUILD FARM
22 CC7 nodes
NEW: ARM thunderx

Short CI and Nightlies status summary

- **Stability:** <1% of jobs with technical problems
- **Regular software updates**, security patches applied
- **Best possible Linux VMs** (no CPU steal, local SSD)
- **ARM/aarch64** machine recently added
- **Sufficient reserve:** can run 20% more jobs
- **Excellent BigPanDA-based monitoring**
- **Good progress** of OpInt CI project

➢ READY FOR RUNS

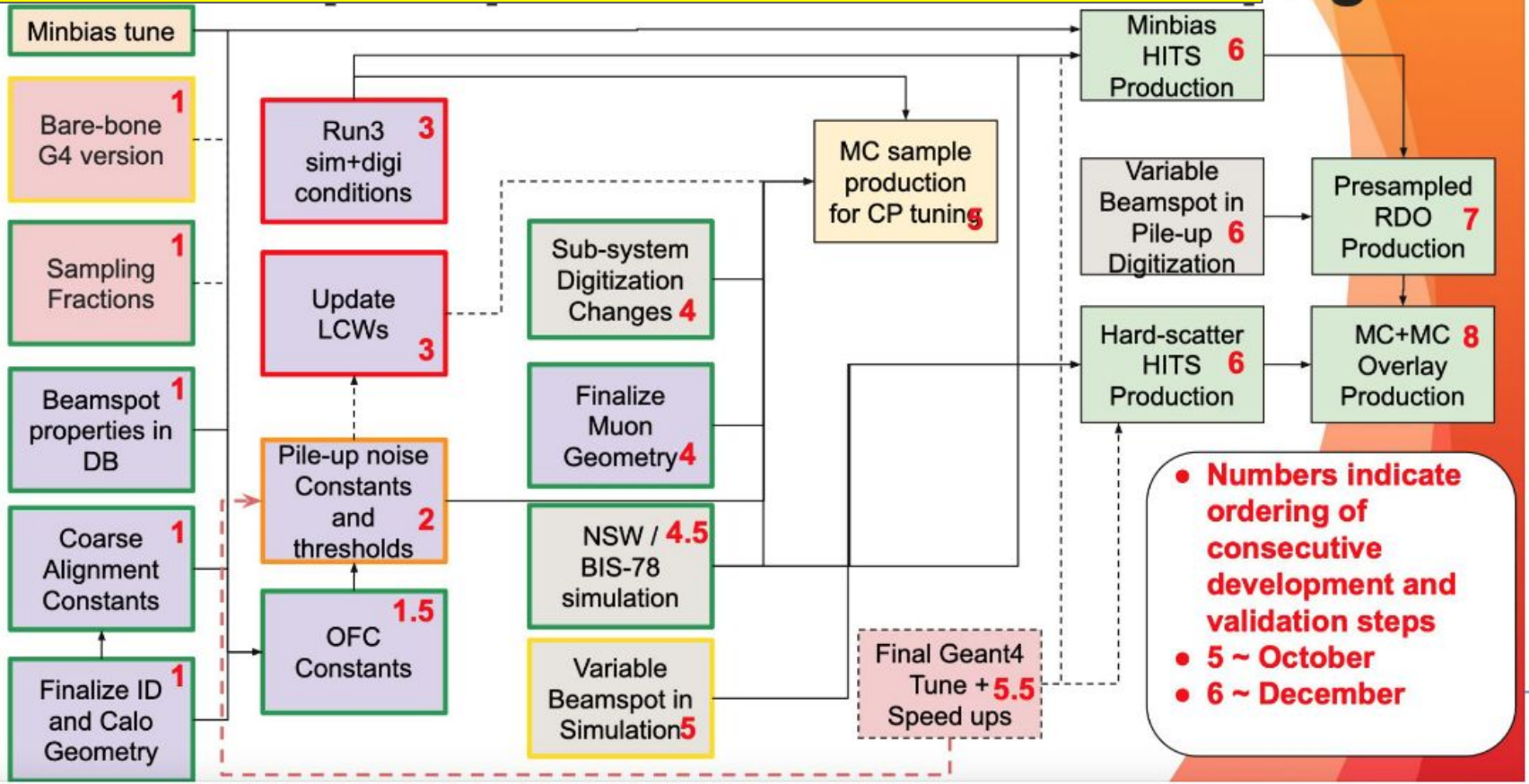


Nightly Group	Branch	Recent Release	Build time	Compilation errors (w/warnings)	CTest (or CI) test errors (w/warnings)	ART LOCAL	ART GRID	CVMFS (on client)
MASTER	master_AthSimulation_x86_64-centos7-gcc8-opt	2021-09-29T2101	29-SEP 22:16	0 (0)	3 (3)	6,0,0	0,3,7,7	29-SEP 22:51
MASTER	master--HepMC3_AthGeneration_x86_64-centos7-gcc8-opt	2021-09-29T2001	29-SEP 20:55	0 (0)	0 (0)	9,0,27	0,8,3,29	29-SEP 21:21
MASTER	master_AthGeneration_x86_64-centos7-gcc8-opt	2021-09-29T2101	29-SEP 22:09	0 (0)	0 (0)	13,0,23	1,13,7,19	29-SEP
MASTER	master--dev4LCG_Athena_x86_64-centos7-gcc11-opt	2021-09-30T0635	30-SEP 14:12					
MASTER	master--dev4LCG_Athena_x86_64-centos7-gcc8-opt	2021-09-30T0600	30-SEP 12:15					

aggregate results from several frameworks and subsystems

Simulation and Digitization

Multi-step developments and validations for Run3 are well planned, on schedule



Software: Run3 Preparation Highlights

- Run2 reprocessing with Athena MT - major production campaign - launched
- Simulation and Digitization validation for Run 3 on schedule
- Geant 4 optimization task force:
  - Investigating configuration options and simplified geometries and magnetic-field descriptions
  - Improving the ATLAS interface code to G4
  - 30% speedup achieved w.r.t. Comparable Run 2 simulation
- New NSW muon spectrometer
  - Detector description is ready
  - Digitization is extensively validated (needs tuning as time resolution measured from cosmic data is worse than expected from design)
  - Trigger efficiency studies are underway
- WIP to use A Common Tracking Software (ACTS) by Run 4 (means GPU use)
- Data formats for Run3 are being finalized (see next slide)

Analysis Model Group - Derivations

- Derivation production of DAOD_PHYS with AthenaMP ready for production
  - Repeat validation when first samples from reprocessing become available
- Keep an eye on the memory usage (see maxPSS/core plot for last 60 days for R22 derivations - see right plot)
- DAOD_PHYS size in range of 30-40 kB/event, below initial target of 50 kB/event
- Still up to 80 residual formats expected (largely CP groups with all data and selected MC) which take up large DAOD disk share in comparison to DAOD_PHYS (see right table)
- Discussions about:
  - assumed/planned 4 reproductions of the DAODs per yea
  - who enforces PHYS-unless-you-have-a-good-reason ?
  - continual upward pressure on the contents of PHYS ?
  - Early 2022 analysis data format requirements

	Number of formats
B-physics	22
Exotics	9
Higgs	2
Standard Model	5
SUSY	2
Jets	13
E-gamma	10
Muons	7
Flavour tagging	5
Tau	5
SUM	80

"Derivation Framework" takes the petabyte-scale AOD output from ATLAS reconstruction and produces samples, typically terabytes in size, targeted at specific analyses.

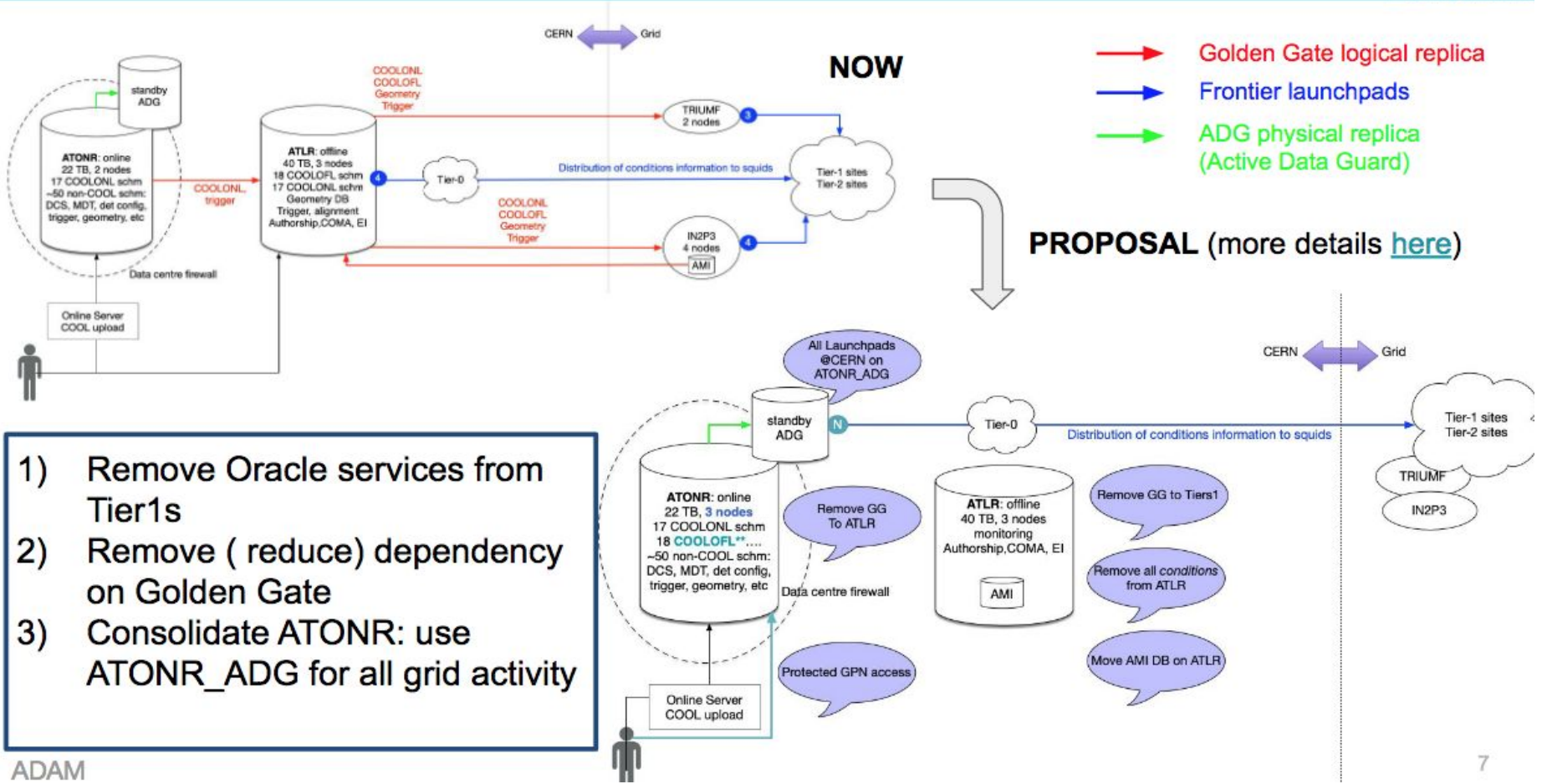
Database&Metadata: Run3 Preparation Highlights

ATLAS uses Oracle

- Three production clusters - 115 TB
- Golden Gate replication
- Current 19c version for entire Run3 (supported until 2027)
- New licensing: cost proportional to number of cores deployed in DB servers
  - Goal is to minimize the required Oracle processor licenses
  - Oracle nodes will not be available for ATLAS (outside CERN) anymore *probably*

ATONR consolidation

ATLAS Online DB server (alignments, calibrations)



1) Remove Oracle services from Tier1s
2) Remove (reduce) dependency on Golden Gate
3) Consolidate ATONR: use ATONR_ADG for all grid activity

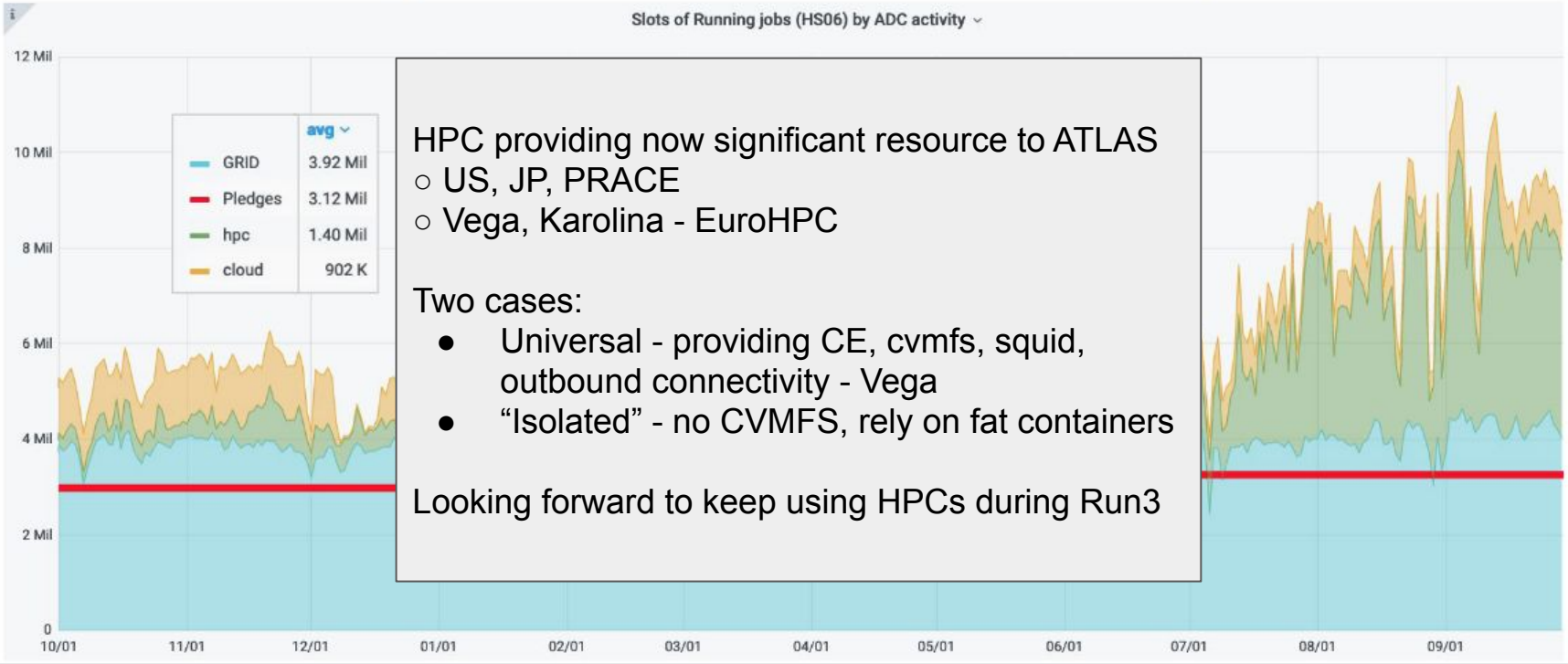
ADC: Run3 Preparation Highlights

- In particular, as it has been a while since ADC has been in "data taking mode": *Do we remember how to do it?*
  - Or perhaps phrased a little better: *Do we (still) have everything in place to efficiently provide the run 3 data as well as the means to perform run 3 analysis to the collaboration?*

- This means:
  - Recording the collisions and getting the data out of the Tier-0
  - Having the rescources to process and distribute the data
  - Understanding and validating the relevant workflows
  - Being able to sufficiently monitor the infrastructure
  - Preparing what needs to be done during data taking, both for the day-to-day activities and for the longer scale activities such as migrations

Resource Usage - Compute - last 12 months

- Excellent performance of our distributed computing infrastructure
  - Opportunistic resources (HPCs and HLT farm) are doubling our pledge
    - **CPU only** – no additional disk from these sites. Opportunistic – may disappear at any time.



HPC providing now significant resource to ATLAS
- US, JP, PRACE
- Vega, Karolina - EuroHPC

Two cases:
- Universal - providing CE, cvmfs, squid, outbound connectivity - Vega
- “Isolated” - no CVMFS, rely on fat containers

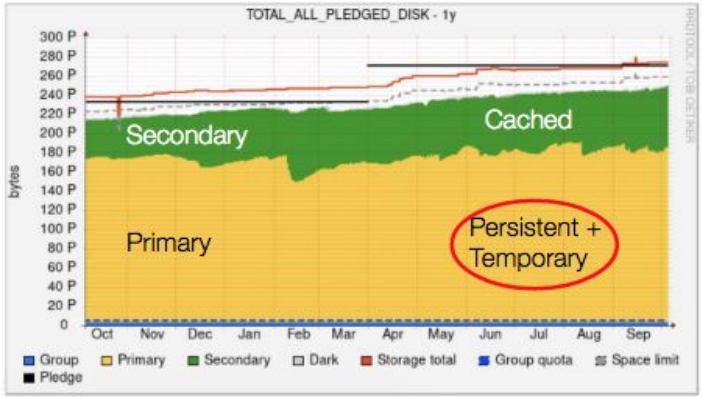
Looking forward to keep using HPCs during Run3

Rucio and DDM

- Many improvements on different major Rucio components during LS2
  - Rule daemon (aka judge) now able to handle very large rules
  - Major rewrite of the transfer daemon (aka conveyor), new features (e.g. multihop) to support new use cases like CTA
  - Major rewrite of the deletion daemon (aka reaper) to improve scalability, still some issues to fine tune
  - Rebalancing daemon is being improved to ensure proper balance between cached and permanent data

- To be done before Run 3 data taking begins
  - New scopes need to be created in AMI DB, which are read by rucio probe (ensure read access)
  - TAPE families: Sites should implement them before the start of the run if needed
  - Define and approve (CREM) the replication policies (including checking MoU weights) similarly to ReplicationPolicy2018
  - Creation of the subscriptions

- New, alternative categorisation of files on disk, finer granularity to Secondary/Primary, to aid in control of the available space
  - To help understand what is pinned on our disks, and why
    - **Custodial** data: rule with lock (precious data on tape)
    - **Persistent** data: rule without lifetime (for purpose)
    - **Temporary** data: rule with lifetime (typically from PanDA)
    - **Cached** data: no rule (can disappear at any time)

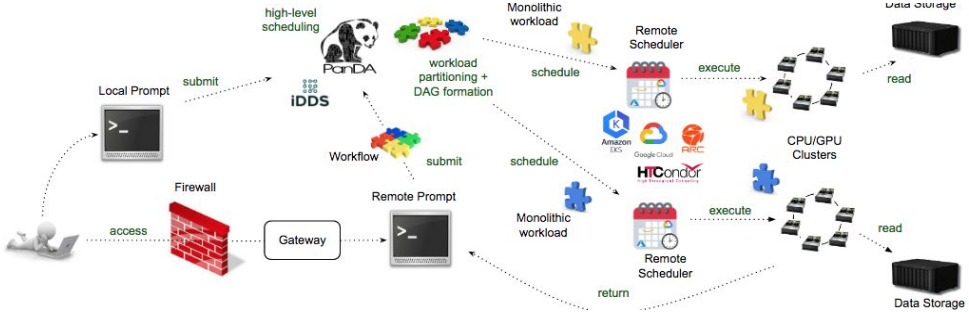


ADC: Challenges for Run3

- Tape: Run 3 data-taking rate will be ~10GB/s. Tier 1 site resources should be sufficient for writing the whole run to the system of tape disk buffer and tape drives
  - Tests ongoing
  - BNL T1 optimization effort (talk of Shigeki Misawa): grouping and accessing multiple data sets together will improve access to data types with small data set
- Network: when shipping between Tier1s the bandwidth achieved in recent test reached the target of 120 GB/s (~ 10% of the bandwidth required at HL-LHC)

ADC: PanDA Developments

- iDDS Development for User Analysis and New Workflows (Tadashi Maeno) - Pseudo-Interactive Analysis
  - Focus on reducing latencies



- Real-time Logging (Paul Nilsson, Shuwei Ye)
  - Implemented in Rubin, possible use in ATLAS case
  - User can follow the progress of the payload as it runs in near real time

ADC: Migrations for Run3

- TPC: Move away from gsiftp to https (80% of sites migrated)
- Transition to tokens issued by ATLAS IAM by mid 2022 as HTCondor is dropping all support for X.509
  - *IAM - Identity and Access Management web service providing security control access to resources*