ADC Priorities During LS3

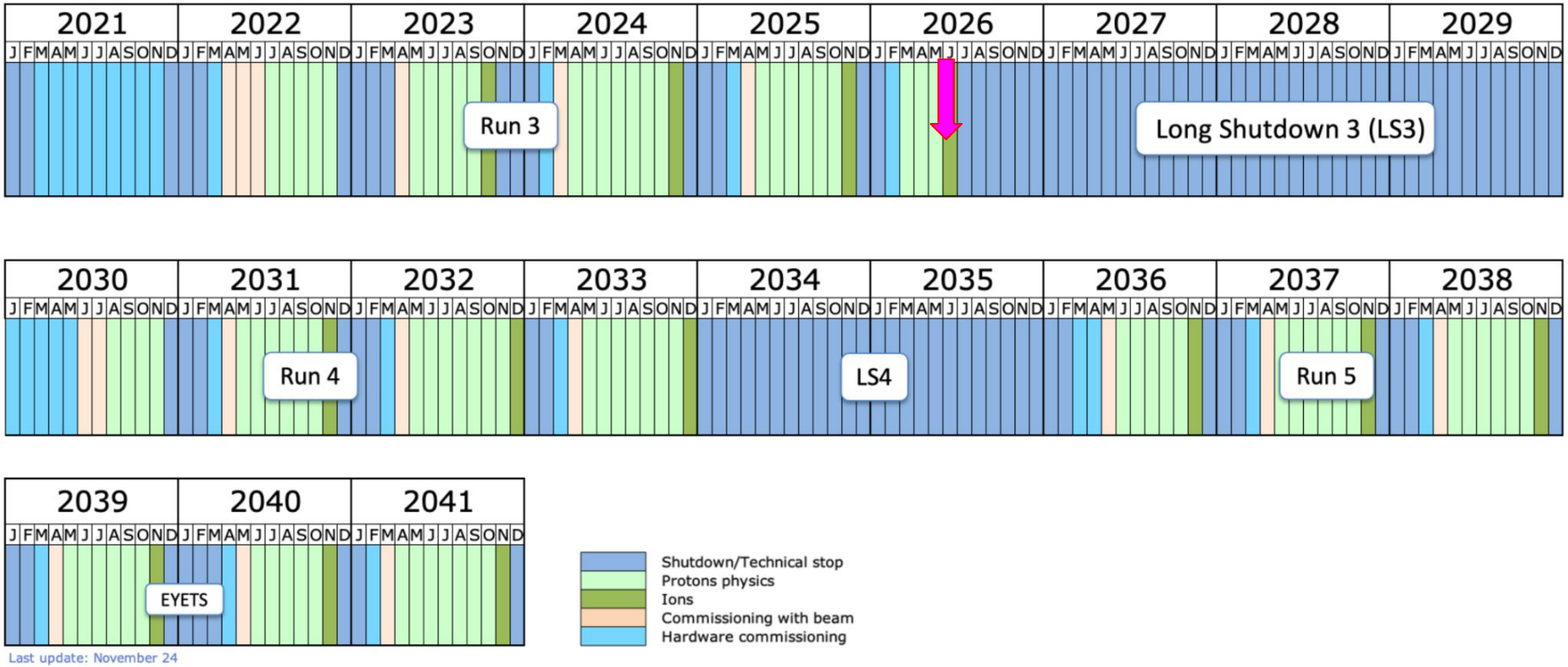
and site procurement implications

Throughput Computing Week 2026

9th June 2026

Ivan Glushkov (BNL)
For ATLAS Distributed Computing

LHC Schedule



ADC Priorities for LS3

- HLT@P8 fully operational ASAP
  - By the end of 2026
- Token-based authentication / retirement of X.509
  - FTS-side tape
  - Rucio by Q1/28
- FTS4 qualification
- DC27: 2.8 Tb/s at 50% of HL-LHC rates
  - February 2027
- Panda
  - Client modernization
  - Kubernetes migration
- Rucio
  - Client modernization
  - Accounting pipeline modernization
- Tape metadata
  - Deployed and in use an all ATLAS tape sites

LS3 Plans: Software and Computing

ADC Priorities During LS3 and Site Procurement implications

Jun 9, 2026, 9:05 AM
25m
Monona Room - 8th Floor (Fluno Center)

Speaker

Ivan Glushkov (BNL)

Description

What should be we planning on doing for infrastructure updates?
- ARM compute servers?
- GPU servers?
- Storage servers, specifically what are the needed charactistics?
- Netwoking: how do we get to 400 Gpbs links by HL-LHC startup?
- Infrastructure: power, management, resiliency, capacity; what do we need?
- Services: what middleware and are changes needed in services we deploy?
- Applications: provisioning, monitoring, management, planning ?

ATLAS activities in 2027: Summary

Data sample	Activity	2025		2026				2027			
		Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Run 2 Data	DAOD Production and User Analysis										
	DAOD_PHYS Production										
Run 2 MC	New Production for Ongoing Analyses										
	DAOD_PHYS Production										
Run 3 Data	Tier-0 Reconstruction										
	Partial Reprocessing										
	Full Reprocessing										
	Delayed stream Reconstruction										
	DAOD Production and User Analysis										
Run 3 MC	Generation										
	Simulation										
	Reconstruction										
	Re-Reconstruction										
	DAOD Production and User Analysis										
Upgrade MC	Production and Analysis										
Heavy Ions	First pass Reconstruction										

Full steam Partial Minimal

ATLAS activities in 2027: Highlights

Partial reprocessing of the 2026 collision data

- "Fast" reprocessing of 2026 (+2025?) data with finalised data quality, alignment and calibrations etc likely to extend into 2027 (including delayed streams)
- Unlike as was done in Run 2, currently not planning full reprocessing of all Run 3 data with release 25 (Run 4 release)

Continuous event generation and simulation and reconstruction of MC samples for other Run 3 years

- Estimating 40 billion events simulated in 2027
- Still aiming for ⅔ fast simulation by the end LS3, but this will require further adoption of AF3

Upgrade sample production for physics studies for the HL-LHC expected towards the end of the year

- Focus on tuning and validation of Run 4 reconstruction, with large scale event generation expected in 2027

Continuous production of derived formats for data and MC

- DAOD_PHYS, DAOD_PHYSLITE, and DAODs for analysis and for combined performance studies

User analysis on Run 2 and Run 3 data and MC samples

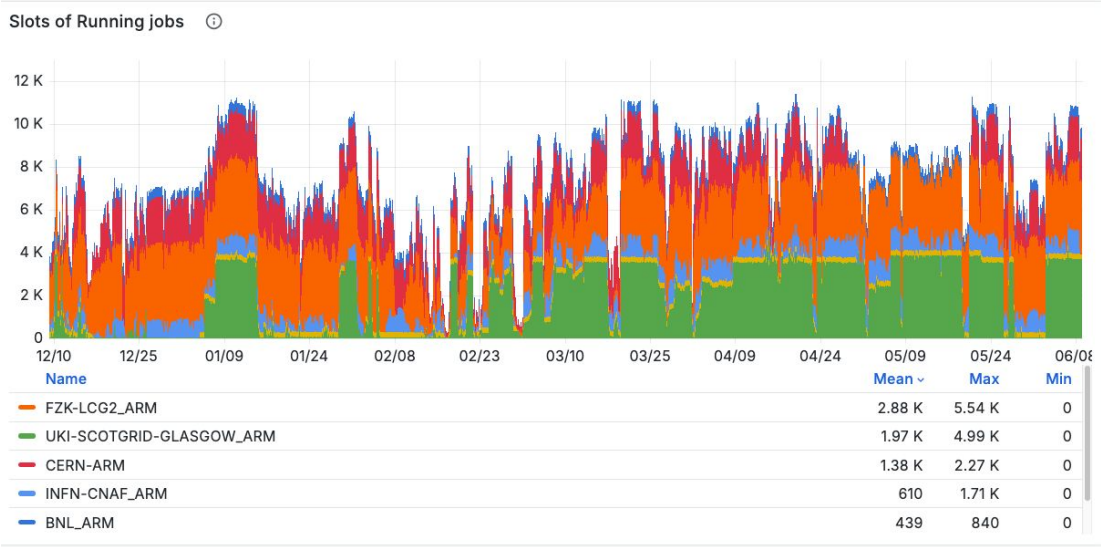
Resource Projections

- 2027
  - The ATLAS Resource Request was accepted by the C-RSG
  - CPU: No increase
  - Disk 8% increase:
    - T2s: 15.7 PB (US: 23%: 3.61 PB)
    - T1s: 22.4 PB (US: 23%: 5.15 PB)
  - Tape:
    - 55.8 PB (US: 20%: 11.1 PB)
- LS3 (2028 - 2029)
  - We do expect increase of the CPU resource request for 2028 and 2029 (despite the use of the HLT farm at P8)
- Run4 (2030)
  - Much more significant request increase to be expected
- Long term projections
  - Last done in 2022
  - Updated one is expected in autumn for the TDR

ATLAS		2025			2026		2027		
		RRB approved	Pledged	Used	RRB approved	Pledged	Request	2027 req. / 2026 RRB	C-RSG recomm.
CPU	Tier0	1100	1100	1019	1265	1265	1265	100%	1265
	Tier1	1635	1639	1760	1802	1827	1802	100%	1802
	Tier2	1998	2297	4078	2202	2381	2202	100%	2202
	Total	**4733**	**5036**	**6857**	**5269**	**5473**	**5269**	**100%**	**5269**
	HLT			*257*					
	Others			*1399*					
Disk	Tier0	56.0	56.0	51.2	65.0	65.0	65.0	100%	65.0
	Tier1	186.0	186.7	194.0	199.0	198.3	214.0	108%	214.0
	Tier2	227.0	218.9	191.0	243.0	239.6	262.0	108%	262.0
	Total	**469.0**	**461.6**	**436.2**	**507.0**	**502.9**	**541.0**	**107%**	**541.0**
Tape	Tier0	258.0	258.0	268.0	302.0	302.0	312.0	103%	312.0
	Tier1	561.0	567.6	560.0	692.0	695.2	741.0	107%	741.0
	Total	**819.0**	**825.6**	**828.0**	**994.0**	**997.2**	**1053.0**	**106%**	**1053.0**

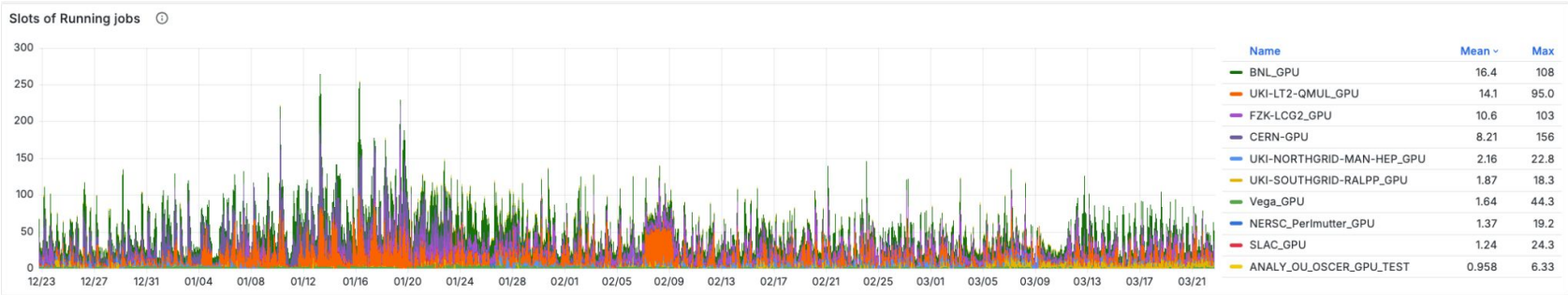
ARM

- ATLAS accepts ARM as pledge for up to 50% of the resources
- All software releases are build on ARM
- ARM is in production on the GRID: 11k slots at 5 sites
- All major ATLAS workflows run on ARM



GPUs

- Status:
  - Infrastructure:
    - ~280 GPUs on 8 GRID sites
    - Utilization: **Very low**
    - TDAQ for Run4: “The Technology Choice Committee recommends a heterogeneous EF farm based on CPUs and GPUs”
  - Software (expectations):
    - Simulation with GPU (up to 50% offloading), Reconstruction: probably for online
- Accounting:
  - No CPU-equivalency cost model
  - No slot definition
  - No nVidia-AMD-Intel equivalency
  - No benchmark
- Considerations:
  - Power consumption and CO2 per event
  - Half-used GPU is a net loss



Network

- WLCG Technical Evolution

Site Type	2024	2027	2029	2030-32 (HL-LHC Ramp-up)	Full HL-LHC Rate (2033+)
Tier-2(small)	10 Gb	20 Gb	20-40 Gb	20-40 Gb	40 Gb+
Tier-2 (normal)	40 Gb	100 Gb	100-200 Gb	100-400 Gb	200-400 Gb
Tier-2(large)	100 Gb	200 Gb	400 Gb	400-800 Gb	800-1200 Gb
Tier-1	200 Gb	400 Gb	400-800 Gb	800 Gb+	800 Gb+ (multi links)

- ATLAS requirements
  - WAN: 100 Gb/s
  - LAN: 10-20 MB/s
- Network capacity estimation
- "- Networking: how do we get to 400 Gpbs links by HL-LHC startup?"

Network Capacity Estimation for Run-4

- Assume analysis on an HT-Core (job-slot) consumes 1.2 MBytes/sec
  - WAN access to remote storage at 20%
  - **Minimal Tier-2**: 2000 job slots => 2.4 GBytes/s, WAN 4.8 Gbits/sec
  - **Nominal Tier-2**: 10000 job slots => 12 GBytes/s, WAN 24 Gbits/sec
  - **Leadership Tier-2**: 20000 job slots => 24 GBytes/s, WAN 48 Gbits/sec
  - **NOTE**: Run-4 will have ~3 times the data and we want "overhead" of about a factor of 2
- **Tier-2 Summary Network Capacity Recommendation for HL-LHC:**
  - Average numbers implied for Run-4 (x6 from above):
  - **Minimal Tier-2 WAN**: 4.8 Gbps x 6 = 28.8 Gbps => **40G link**
  - **Nominal Tier-2 WAN**: 24 Gbps x 6 = 144 Gbps => **2x100G link**
  - **Leadership Tier-2 WAN**: 48 Gbps x 6 = 288 Gbps => **400G**
- **Tier-1s should plan on increasing connectivity by a factor 6-10 as well (if they have not recently upgraded significantly).**
  - **For current ~100G Tier-1s, this implies 2x400G or 800G**

Q&A

- Storage servers, specifically what are the needed characteristics?
  - No finer grade storage tiering (HDD-SSD-NVMe)
  - Contemporary storage servers are limited by network
  - ADC priority: Uptime!
  - Site priority: Protect against overload
- Infrastructure: power, management, resiliency, capacity; what do we need?
  - All boils down to (pledged) capacity
  - All of the above
- Services: what middleware and are changes needed in services we deploy?
  - Varnish, CVMFS - no changes foreseen/needed
- Applications: provisioning, monitoring, management, planning ?
  - Not clear