WBS 2.3.1

Eric Lançon / BNL Tier-1, WBS 2.3.1

US ATLAS Software and Computing Operations - pre-Scrubbing

June 2022

2.3.1 FTE Summary

	FY22	FY23	Comment
Benjamin	0.30	0.40	
Burstein	0.30	0.30	
Chou	0.30	0.30	
Destefano	0.50	0.15	Increased B2 activity
Fontana	0.10	0.10	
Frith	0.50	0.25	Split with new comer
Gamboa	0.50	0.85	dCache operation
Garonne		0.50	New comer
Hancock	0.85	0.90	
Hollowell	0.40	0.40	
Huang	0.20	0.60	dCache operation
Ito	0.50	0.25	Increased B2 activity
J. Smith	0.50	0.30	
Kandasamy	0.20	0.20	
Lancon	0.45	0.45	
Latif	0.25	0.15	
Lepore	0.50	0.50	
Liu	0.75	0.30	Effort reassigned

	FY22	FY23	Comment
Lukasczyk	0.20	0.20	
McCarthy	0.25	0.25	
Metz		0.30	New comer
Misawa	0.60	0.50	
Novakov	0.30	0.30	
Pelosi	0.30	0.40	
Poat		0.30	New comer, replaces Will
Rao	0.50	0.45	
Snyder	0.35	0.50	
Spradley		0.35	New comer
StreckerKellogg	0.30		Left SDCC
T. Smith		0.20	New comer
Throwe	0.10		Left SDCC
Wong	0.25	0.15	
Wu	0.50	0.30	Effort reassigned
Zaran	0.30		Left SDCC
Zaytsev	0.35	0.35	
Zhao	0.05		Left SDCC

Left SDCC: 4
New comer: 4

all SDCC staff - work on either HEP, NP, BES, Lab computing projects

2.3.1 FTE Summary

	FY22	FY23	Comment
WBS 2.3.1.1 (Administration)	0.45	0.45	No change
WBS 2.3.1.2 (Tier-1 Infrastructure)	3.00	3.10	Slightly more work at data center(s)
WBS 2.3.1.3 (Linux Farm)	1.15	1.20	Slight increase to account for less experienced staff
WBS 2.3.1.4 (Storage)	4.75	4.75	Same amount while new team
WBS 2.3.1.5 (Services for (US)ATLAS)	2.10	1.95	Slightly less work on CVMFS and Tokens
Total	**11.45**	**11.45**	**No change**

Achievements since last scrubbing

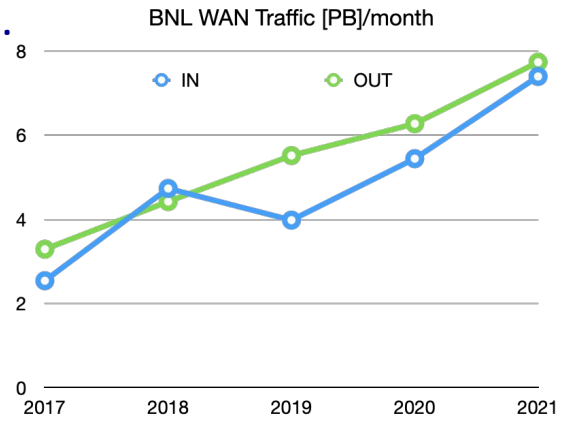
- Tier-1 did not shut down during pandemic, staff have been working mostly remotely for more than 2 years

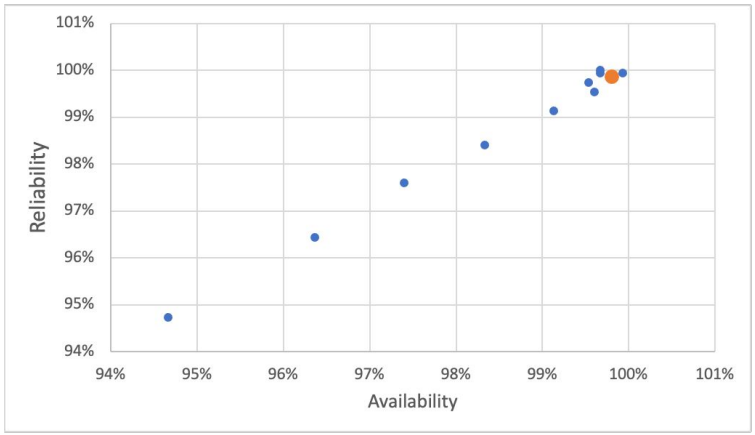
- Better operational efficiency: improved WLCG Tier-1 key parameters values compared to last year
  - Ready for Run 3: Above parameter values defined for LHC data taking period

- 2022 Highlights (short list)
  - Deployment of new equipment in new data center, transparent operation between 2 data centers
  - Commissioning of new tape library and migration to HPSS 8.3.10 allowing for increased capabilities
  - Best performances among ATLAS Tier-1s during WLCG data & tape challenges
  - First Tier-1 to migrate to tokens
  - New US ATLAS website
  - Comprehensible cost model developed

2.3.1: Overview of Activities

- The Tier-1 contributed to ATLAS distributed computing at the level of performance and availability that is commensurate with the agreed upon share: 23 %
- The Tier-1 is being used by ATLAS
  - **For I/O intensive and critical workloads** (no data loss)
  - **As a central repository** (about 200PB were imported/exported by the Tier-1 last year)
  - Networking capabilities and storage reliability are key
- Leadership role on storage activities acknowledged by WLCG-DOMA
- Numerous presentations in various venues ATLAS S&C weeks, Grid Deployment Board, HEPiX, Snowmass white paper on future analysis facilities, ...





Major FY22 deliveries

- Successful operation within new data center (together with the old one)
- Deployed pledges to meet ATLAS needs (M113)
- Successful WLCG challenges (M59) & (M60)
- Stabilise & improve dCache system:Gamboa/Garonne/Huang
- HPSS migration (M54): Chou/Novakov/Liu
- Evaluation of Ceph as a storage solution (M115): Zaytsev/Rao
- Cost model for data center: Hollowell/Misawa/Zaytsev
- Increase of tape usage efficiency: Misawa
- Development of Rucio QoS: Ito/Snyder
- Deployment of Kubernetes Platform: Hollowell/Rao

Milestone status

WBS No.	WBS Title	Milestone #	Milestone	Baseline Date	Actual/Estimated Completion Date
2.3.1	OK	54	Tier 1 Migration to HPSS v9	Sep 2021	Aug 2021	Completed.
2.3	OK	59	Facility completes the first WLCG Data Challenge	Oct 2021	Oct 2021	Completed
2.3	OK	60	Facility completes the first WLCG Tape Challenge	Oct 2021	Oct 2021	Completed
2.3	OK	113	Complete deployment of T1/T2 2022 pledged resources	Apr 2022	Apr 2022	Completed
2.3.1		115	Decision about Ceph as primary T1 storage for FY23	Jul 2022	Jul 2022	On Schedule

FY23 and beyond

WBS No.	WBS Title	Milestone #	Milestone	Baseline Date	Actual/Estimated Completion Date
2.3		117	US facility provides network site monitoring (real-time data IN/OUT of the site) for our Tier-2s and Tier-1	Feb 2023	Feb 2023	On Schedule
2.3		118	Complete deployment of T1/T2 2023 pledged resources	Apr 2023	Apr 2023	On Schedule
2.3		119	Facility completes second WLCG Network Data Challenge	Nov 2023	Nov 2023	On Schedule
2.3.1		120	Old BNL data center decommissioned for compute	Jan 2024	Jan 2024	On Schedule

New milestones

- 2023Q1: Medium term central disk storage model
- 2023Q2: Alternative to SL7 as Operating System at T1 batch farm

FY22 Tier-1 Operation & Equipment

	FY22 Proj. 2021 scrubbing	FY22 Actual June 2022 proj.	*Guidance*
Operation	**$2,025,000**	**$2,116,000**	*$2,115,000*
Equipment	**$1,970,000**	**$1,981,000**	*$2,068,000*

FY22-25 Projection

Details & explanations are available here and here

	FY22	FY23	FY24	FY25	Guidance	FY22-25 <av>
[1.1] Operations	$2,115,000	$2,020,000	$2,000,000	$2,031,000	$2,115,000	$2,041,500
[2.1] Equipment	$1,981,000	$2,188,000	$2,306,000	$2,321,000	$2,068,000	$2,199,000
					Sum	
Delta from Guidance	*$87,000*	*-$25,000*	*-$123,000*	*-$169,000*	***-$230,000***	
Delta from Guidance [%]	*2.1%*	*-0.6%*	*-2.9%*	*-4.0%*	*-1.4%*	

2.3.1: FY23 activities

Deliver services at level of performance & availability according to MoU

By order of priority

- Migration to tokens (storage): Hancock/Wong (0.2 FTE)
- Mid-term storage evaluation: Garonne/Hancock/Huang/Rao (0.4 FTE)
- Increased tape usage efficiency: Misawa (0.2 FTE)

- Improved Data Analytics: Poat/Snyder (0.3 FTE)
- Commissioning of Rucio QoS: Ito/Snyder (0.2 FTE)
- ADC Dynamic Data Handling working group: Ito (0.1 FTE)

- Fed. Id. access to Kubernetes platform: Smith/Rao (0.1 FTE)
- Improved cost model: Hollowell/Misawa/Zaytsev (0.2 FTE)
- HEPScore: Hollowell/Wong (0.2 FTE)

2.3.1 priorities aligned with ADC roadmap

DC-1 Transition to tokens

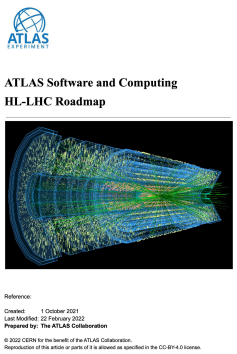
DC-2 Storage Evolution

DC-3 Next operating system version

DC-4 Network infrastructure ready for Run-4

DC-8 Storage optimization

DC-9 Disk management: secondary (cached storage)



ATLAS Software and Computing
HL-LHC Roadmap

Reference:

Created: 1 October 2021
Last Modified: 22 February 2022
Prepared by: The ATLAS Collaboration

Distributed Computing

	MID	DID	Description	Due
	DC-1		Transition to tokens	Q4 2025
		1.1	Submission from Harvester to all HTCondor CEs with tokens	Q1 2022
		1.2	All users move from VOMS to IAM for X509	Q4 2022
		1.3	All job submission and data transfers use tokens	Q4 2025
	DC-2		Storage evolution	Q4 2025
		2.1	No GridFTP transfers at any site	Q1 2022
		2.2	SRM-less access to tape	Q4 2025
		2.3	Recommended transition plan from DPM completed	Q4 2021
		2.4	Transition plan from all DPM sites	Q4 2022
		2.5	All sites moved away from DPM	Q2 2024
	DC-3		Next operating system version	Q2 2024
		3.1	Ability to run on "future OS" on grid sites	Q4 2022
		3.2	Central services moved to "future OS"	Q4 2023
		3.3	(CentOS 7/8 EOL)	Q2 2024
	DC-4		Network infrastructure ready for Run 4	Q4 2027
		4.1	Network challenge at 10% expected rate	Q4 2021
		4.2	Network challenge at 30% expected rate	Q4 2023
		4.3	Network challenge at 60% expected rate	Q4 2025
		4.4	Network challenge at 100% expected rate	Q4 2027
	DC-5		Integrating next generation of HPCs	Q2 2023
		5.1	Integration of at least 2 EuroHPC sites	Q4 2022
		5.2	Integration of next generation US HPCs for production	Q2 2023
	DC-6		Exploratory R&D on GPU-based workflows for next generation HPC	Q4 2023
	DC-7		HL-LHC datasets replicas and versions management	Q2 2024
		7.1	Replicas and versions detailed accounting	Q4 2022
		7.2	DAOD replicas reduction	Q4 2023
		7.3	DAOD versions reduction	Q2 2024
	DC-8		Data Carousel for storage optimization	Q4 2023
		8.1	Investigate with sites the cost of Tape infrastructure and the estimated cost in case of sensible increase of read/write throughput	Q4 2022
		8.2	Reduce the AOD on disk to 50% of the total AOD volume, using Data Carousel to orchestrate the stage from tape for DAOD production.	Q4 2023
	DC-9		Disk management: secondary(cached) dataset	Q2 2023
		9.1	Evaluate the impact on job brokering and task duration if disk space for secondary data is reduced	Q2 2023

Maintenance & Operations | Conservative R&D | Aggressive R&D

Covid impacts: Personals

- Loss of key personals as IT companies increase remote & telework
- Last year two very talented staff members left SDCC for private companies
  - **One time** opportunity to hire 2 young local staffs from another directorate following a reorganization

- Competitive job market environment
  - Hiring process takes months
  - Staff receive job offers

Covid impacts: cost & delays

Delays in delivery chain and increased cost/unit,

- **CPU**: 6-7 months of lead time (no reversal of the cost trend for the CPU only systems (yet?), GPU costs went up 20% in 2022Q2)
- **Disk**: 8-9 months of lead time, storage cost went up ~7% in 2021Q4 (reversal of the cost trend observed)
- **Network** equipment 6-**12** months for the new placed orders, equipment cost went up ~15% in 2022Q2

Delays for each category vary with time

Additional delays when complex systems need to be assembled

- Storage systems: JBODs are delivered but head nodes arrive months later
- CPU: partially integrated racks arrive months before the set of Top of Rack (ToR) switches needed to connect them

Risk mitigation strategy: Advanced purchase for all components that can be integrated on site (CDUs, ToR switches, line cards, transceivers, DAC cables) - implies increased cost increase by ~10%

2.3.1: Risks

- Increased usage by ATLAS of tape system and its associated components not covered/anticipated by current facility planning - Owner: US ATLAS - Low
- New dCache setup do not handle Run 3 load - Owner: US Tier-1 - Low
- Delays in hardware supply chain - Owner: US ATLAS - High
- Increase of hardware cost - Owner: US ATLAS - High
- Lack of a stable programmatic budget for Analysis Facility - Owner: US ATLAS - Medium
- Increase of electricity cost - Owner: US ATLAS - High
- Retention of key personnel - Owner: US Tier-1 - High/Medium

2.3.X: Coordination & Dependencies

- Cross-cutting WBS 2.3
  - Close interactions with 2.3.4 (Analysis Facilities Operation) and 2.3.5 (CIOPS)
- WBS 2.2, 2.4, and WBS 5
  - WBS 2.2:
  - WBS 2.4: Data Carousel, Co-coordinator Xin Zhao (0.5 FTE)
  - WBS 5: Close collaboration. Analysis Facility operation depends on effectiveness of Analysis Facility ‘user support’
- External
  - Rucio
  - DOMA (JWT)
  - dCache
  - Operating System
  - HPSS
  - OSG
  - ESnet
  - IRIS-HEP
  - HSF

What if 10% cut in FY23

- Strategy:
  - Explore ways to reduce spending
  - Reduce scope
- Priority is reduction of hardware spending, scope reduction implies loss of expertise and capability and involves BNL HR
- Hardware budget reduction:
  - Delay deployment of new CPU (expand lifetime of existing hardware by 1y) gain of ~$1M - may be a long term option given the new data center remove limitations of previous one
- Scope reductions do not bring substantial saving and will put the facility at risk
  - Reduce contribution to CIOPS : 0.4 FTE
  - Cost model : 0.2 FTE
  - HEPScore: 0.2 FTE