ATLAS: Status and thoughts on ARM

GDB, 12 June 2024

The ATLAS experiment software on ARM, J. Elmsheuser et al., EPJ Web Conf. 295, 05019 (2024)
Accelerating science: the usage of commercial clouds in ATLAS Distributed Computing, F. Barreiro Megino et al., EPJ Web Conf. 295, 07002 (2024)
Compute Testing and ARM Provision at Glasgow's Tier 2, Dwayne Spiteri et al., HEPiX Spring Workshop, April 2024

David South (DESY) and Zach Marshall (LBNL)
On behalf of the ATLAS S&C community

First Steps

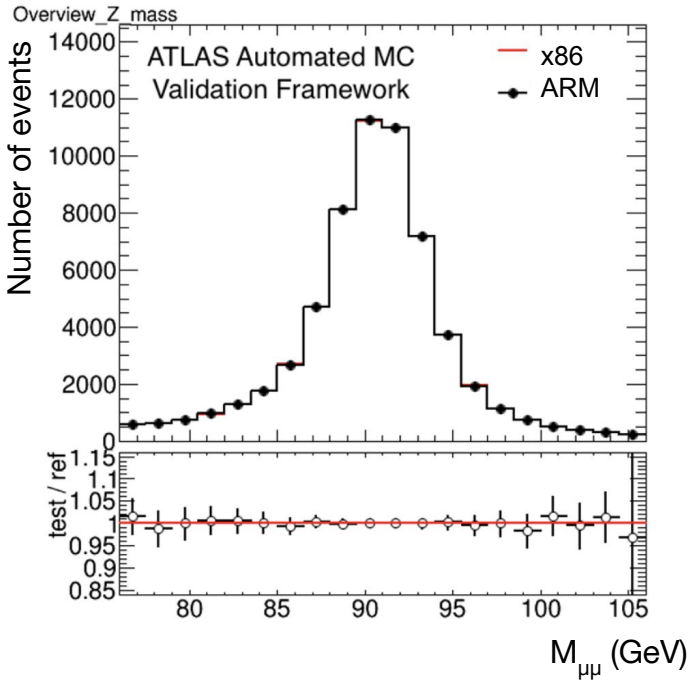
- ATLAS started to consider ARM a few years ago now

- During this period, only a limited number of ARM resources were available to ATLAS, including:
  - Very few Ampere Altra and Cavium ThunderX2 machines physically at **CERN**
  - A small public cluster with Ampere Altra CPUs provisioned through the **Oracle Cloud**
  - Some Ampere Altra/Neoverse-N1 nodes in the **Google Cloud**
  - Access to Graviton2 and Graviton3 processors in the **Amazon Web Services Cloud**

- Integration of cloud resources into two main components of ATLAS distributed computing, so that the cloud site resembles a usual grid site
  - **Compute**: PanDA queue with resources set up as a k8s cluster, interfaced via Harvester. CVMFS for ATLAS software available on the cluster
  - **Data Management**: Rucio Storage Element as an Object Store bucket, with authentication via signed URLs

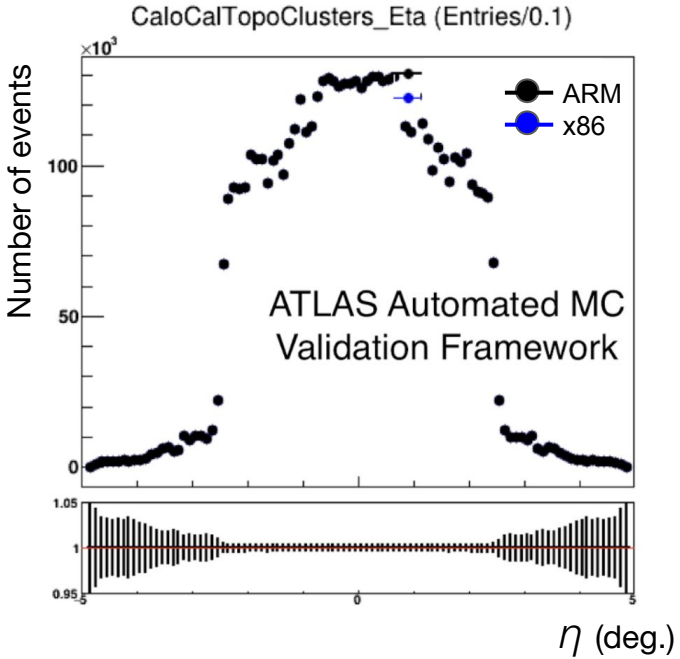
Physics Validation - Simulation

- Standard ATLAS MC workflow includes several steps
  - Event Generation, Full/Fast Simulation, Reconstruction, Derivations (analysis level input)
- Validation of ATLAS workflows by comparing dedicated production on Graviton2 ARM resources at AWS to standard production on WLCG x86 resources
  - Single workflow step on ARM; others on x86
- Simulation:
  - ARM test sample: 100M inclusive ttbar events
  - Comparison shows good agreement of all physics objects distributions between the ARM and x86 produced events
  - Shown: Invariant mass of simulated Z → μμ decays



Physics Validation - Reconstruction

- Standard ATLAS MC workflow includes several steps
  - Event Generation, Full/Fast Simulation, Reconstruction, Derivations (analysis level input)
- Validation of ATLAS workflows by comparing dedicated production on Graviton2 ARM resources at AWS to standard production on WLCG x86 resources
  - Single workflow step on ARM; others on x86
- Reconstruction:
  - ARM test samples: 13 different MC physics processes, with 100k events each
  - Good agreement of all physics objects distributions between the ARM and x86 produced events with some fluctuations in regions with low MC statistics
  - Shown: η coordinate of the reconstructed calorimeter clusters in multi-jet events

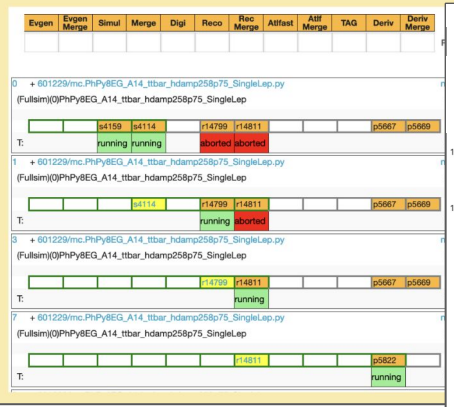


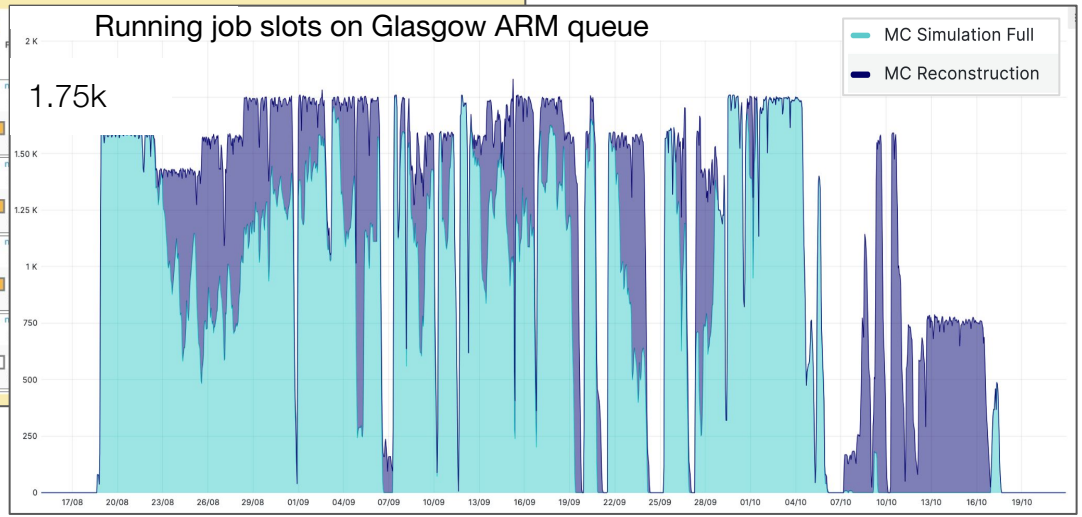
ADC, ARM and Glasgow: Jobs

- First conversations at the WLCG workshop in Lancaster in 2022
  - During 2023 ATLAS worked together with Dwayne and Dave to set up testing of integration of the ARM resources in Glasgow into our distributed computing system
  - In parallel to the physics validation efforts described on the earlier slides

ATLAS Testing

- The ARM nodes are running a job starting with 50M events and with simulation, reconstruction and derivation tasks.
- Running since Late August, and still running!





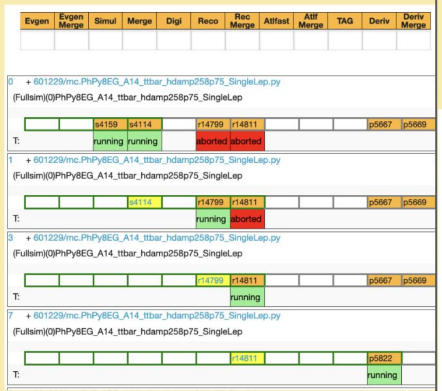
Compute Testing and ARM Provision at Glasgow's Tier2, Dwayne Spiteri et al., HEPiX Spring Workshop, April 2024

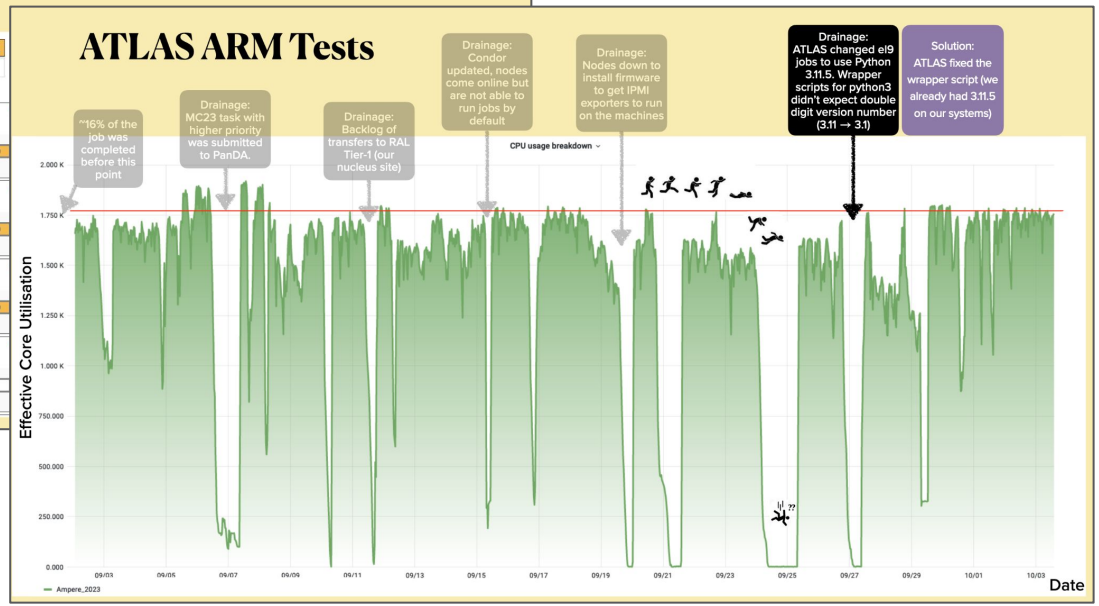
ADC, ARM and Glasgow: Jobs

- First conversations at the WLCG workshop in Lancaster in 2022
  - During 2023 ATLAS worked together with Dwayne and Dave to set up testing of integration of the ARM resources in Glasgow into our distributed computing system
  - In parallel to the physics validation efforts described on the earlier slides

ATLAS Testing

- The ARM nodes are running a job starting with 50M events and with simulation, reconstruction and derivation tasks.
- Running since Late August, and still running!



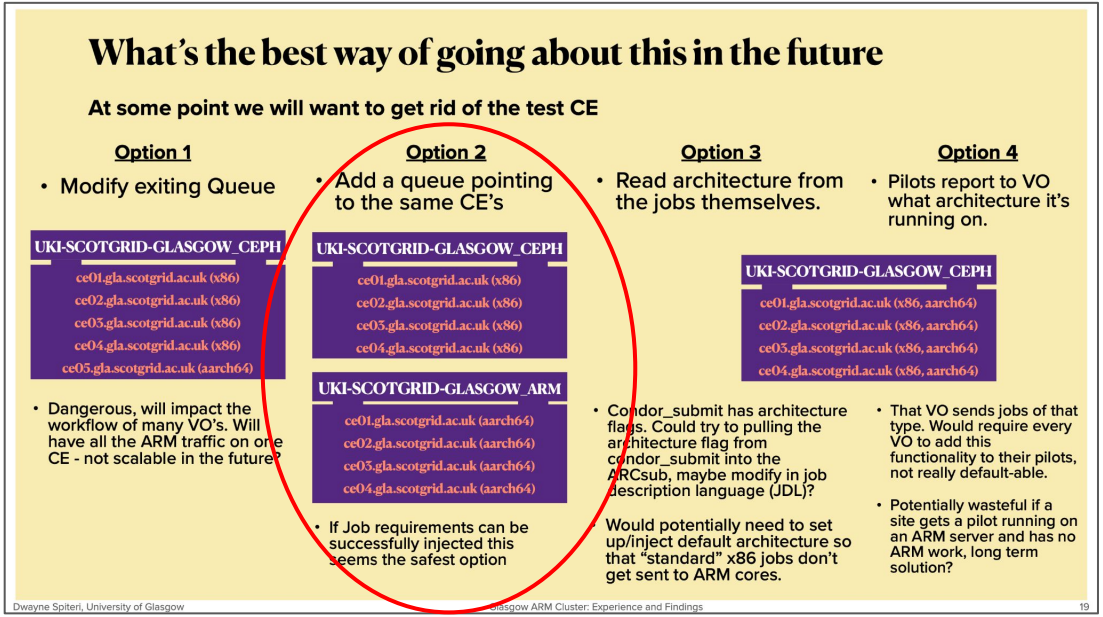


[Compute Testing and ARM Provision at Glasgow's Tier2](), Dwayne Spiteri et al., HEPiX Spring Workshop, April 2024

ADC, ARM and Glasgow: Queues

- Alternative setups considered for setting up the ARM Compute Element

- For ATLAS, the simplest option was to add a separate PanDA queue for ARM jobs, pointing to the same CE
  - No real limit or scaling problems with additional PanDA queues



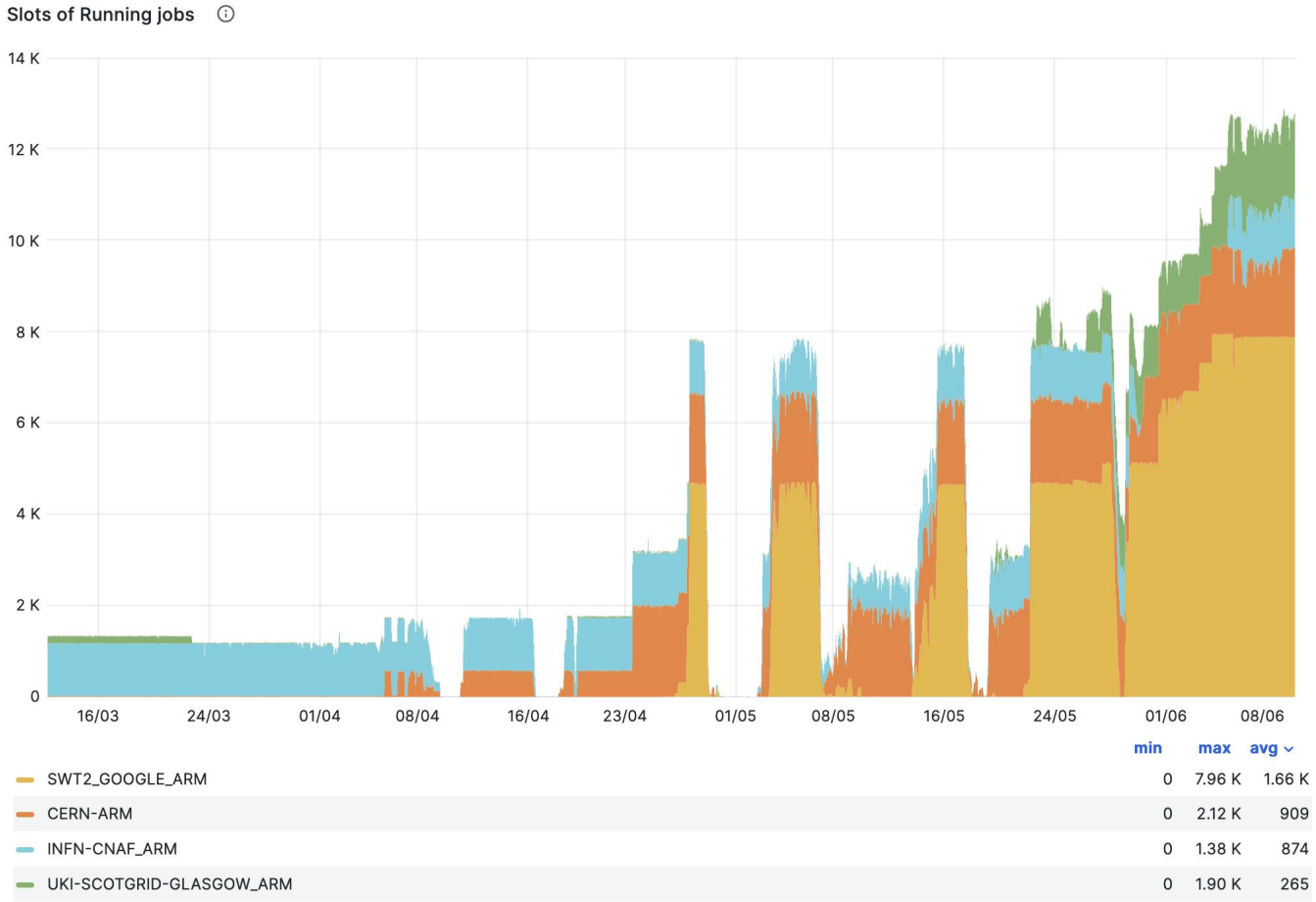
- As the number of sites with ARM increases, we may want to consider a multi-architecture queue, with the relevant information handled by the pilot
  - On a mixed PQ, pilot probes WN arch or instruction set and chooses the right setup
  - Still some work to do to enable this, some syntax *task.architecture='#x86_64 | #aarch64'* handled in brokerage

ATLAS readiness for using ARM today

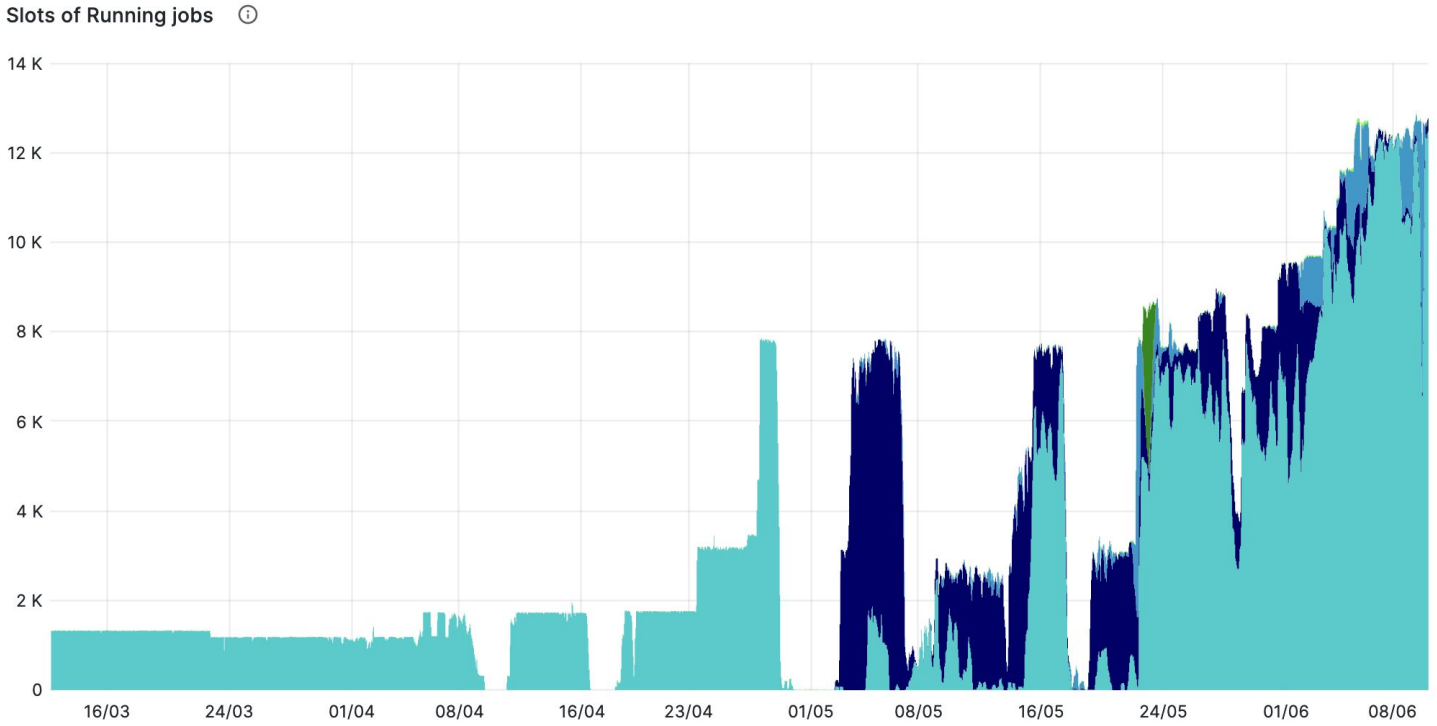
- Workflows and validation
  - Full simulation and Reconstruction are physics validated 😎
  - Event generation is not there yet and intrinsically more varied
  - Derivations (DAOD production) technically runs, but not yet fully validated
  - Data reprocessing not looked at yet (although relatively little impact in total resource usage
  - User analysis is wild-west; we need some dedicated effort there

- Automation and software build nodes
  - ATLAS currently has 4 ARM build nodes at CERN
  - Releases built from nightlies and we have a total of 8 ARM nightlies set up
    - *AthSimulation* and *Athena* for **24.0** (latest stable release series)
    - *AthGeneration*, *AnalysisBase*, *AthSimulation*, *Athena* and *DetCommon* for **main**
    - *Athena* for **main--dev4LCG4**
  - If ARM releases are required for other branches (for example 23.0) they are built on-demand

ARM queues today - sites



ARM queues today - workflows

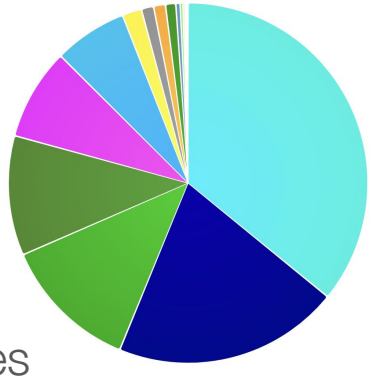


	min	max	avg
MC Simulation Full	0	12.5 K	2.74 K
MC Reconstruction	0	7.80 K	799
MC Simulation Fast	0	6.05 K	137
Group Production	0	3.55 K	20.0

Towards pledged ARM resources for ATLAS

- Are we ready?
  - There are still workflows left to validate, but full simulation and reconstruction make up **almost 60%** of the used wallclock of the last 6 months (which is also typical going back a longer period)
  - For now, restrict which workflows we run on ARM, which is not strictly in the spirit of pledges
  - The brokerage to ARM resources maybe done differently when more sites appear, but for now separate PanDA queues are working well and are running jobs

Wallclock seconds last 6 months of ATLAS jobs



MC Simulation Full	3.57 Tri	36%
MC Reconstruction	2.02 Tri	21%
Group Production	1.22 Tri	12%
MC Event Generation	1.06 Tri	11%
User Analysis	812 Bil	8%
MC Simulation Fast	650 Bil	7%
Data Processing	167 Bil	2%
Group Analysis	103 Bil	1%
t0_processing	98.1 Bil	1%
MC Resimulation	84.0 Bil	1%
MC Merge	31.3 Bil	0%

- ATLAS is ready for pledged ARM resources!
  - This will naturally be a slow influx, not a sudden replacement of all x86
  - Something like “up to 50% of pledge is ARM” is therefore completely fine for us