ATLAS S&C toward the HL-LHC challenges

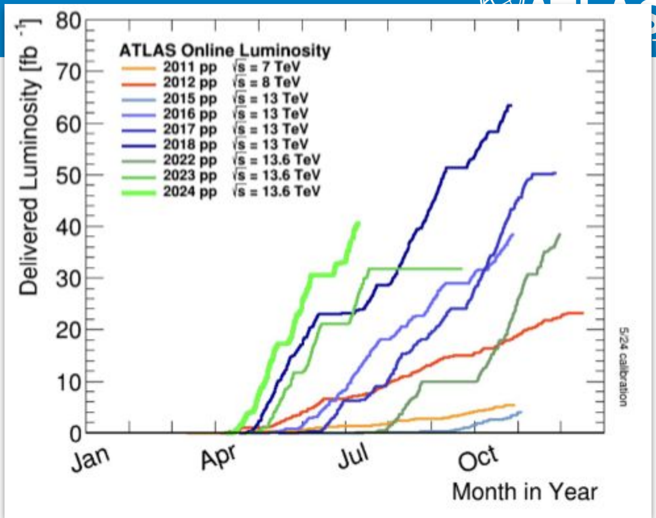
HEP-CCE All Hands Meeting

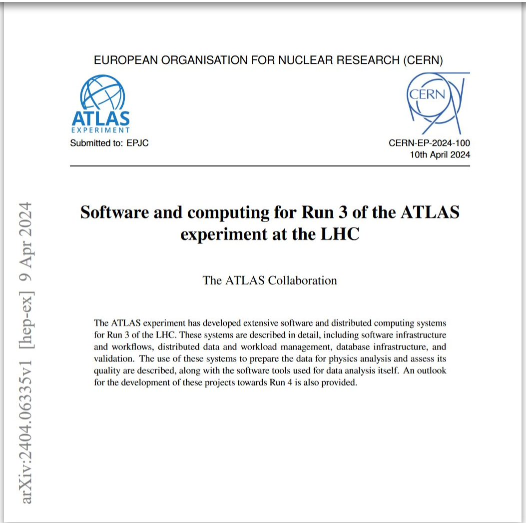
22th June 2024

Edward Moyse

Introduction

- ATLAS is currently taking data for the LHC Run 3
  - We just published a paper on our Run3 software
- But in parallel, working hard to prepare for HL-LHC (2029, though discussion of a possible ~6 month delay)
  - Have previously mentioned our roadmap document
    - Work starting on TDR, planning to publish next year
  - Many technology demonstrators
    - See later…
  - Next gen trigger project is ramping up
    - https://nextgentriggers.web.cern.ch/wp2/





EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

ATLAS EXPERIMENT

Submitted to: EPJC

CERN-EP-2024-100
10th April 2024

arXiv:2404.06335v1 [hep-ex] 9 Apr 2024

Software and computing for Run 3 of the ATLAS experiment at the LHC

The ATLAS Collaboration

The ATLAS experiment has developed extensive software and distributed computing systems for Run 3 of the LHC. These systems are described in detail, including software infrastructure and workflows, distributed data and workload management, database infrastructure, and validation. The use of these systems to prepare the data for physics analysis and assess its quality are described, along with the software tools used for data analysis itself. An outlook for the development of these projects towards Run 4 is also provided.

Milestones

- ATLAS tracks progress with extensive list of milestones
- Have "Road to Run 4" meetings every ~6 months to check on progress

	Category	Milestone	Deliverable	Description	Completion Date Baseline	Responsible	September 2023 Current/Estimated Effort	Completion Dat estimate
Core Software, Heterogeneous Computing and Accelerators								
B	CS	1		Pileup-digitization in AthenaMT production ready	12/2022	Beojan, John C., Tadej		12/2023
B	CS		1.1	Ensure reproducibility of MT production of presampled MB RDO files	6/2022	Beojan, John C., Tadej		6/2022
A	CS	2		Complete investigation of lossy compression techniques	12/2023	Serhan		12/2023
A	CS		2.1	Lossy compression of the ID track covariance matrix in the primary AODs	12/2021	Thomas S.		12/2021
A	CS		2.2	Lossy compression of DAOD	12/2021	James C.		12/2021
A	CS		2.3	Lossy compression of primary AODs	12/2023	Peter vG		12/2023
B	CS	3		Implement I/O roadmap metadata recommendations	12/2022	Peter vG, Vakho		12/2023
B	CS		3.1	Multi-threaded in-file metadata handling	6/2022	Maciej S.		6/2022
C	CS		3.2	Redesign of the metadata handling infrastructure (better support for fine-grained workflows)	12/2022	Maciej S.		12/2023
C	CS		3.3	Migrate current metadata framework to use single propagation tools per type	12/2024			
C	CS		3.4	Allow metadata tools to be executed by the SharedWriter via plug-in mechanism	12/2025			
C	CS		3.5	Evaluation of RNTuple for metadata storage and migration	12/2024			
A	CS	4		Evaluation of data formats well-suited for massively parallel I/O (HPCs)	3/2022	Peter vG		3/2022
A	CS		4.1	Storing intermediate EventService Simulation data in HDF5	3/2022	Peter vG		3/2022
C	CS	5		Migration to ROOT 7	12/2026	Marcin, Peter vG		6/2025

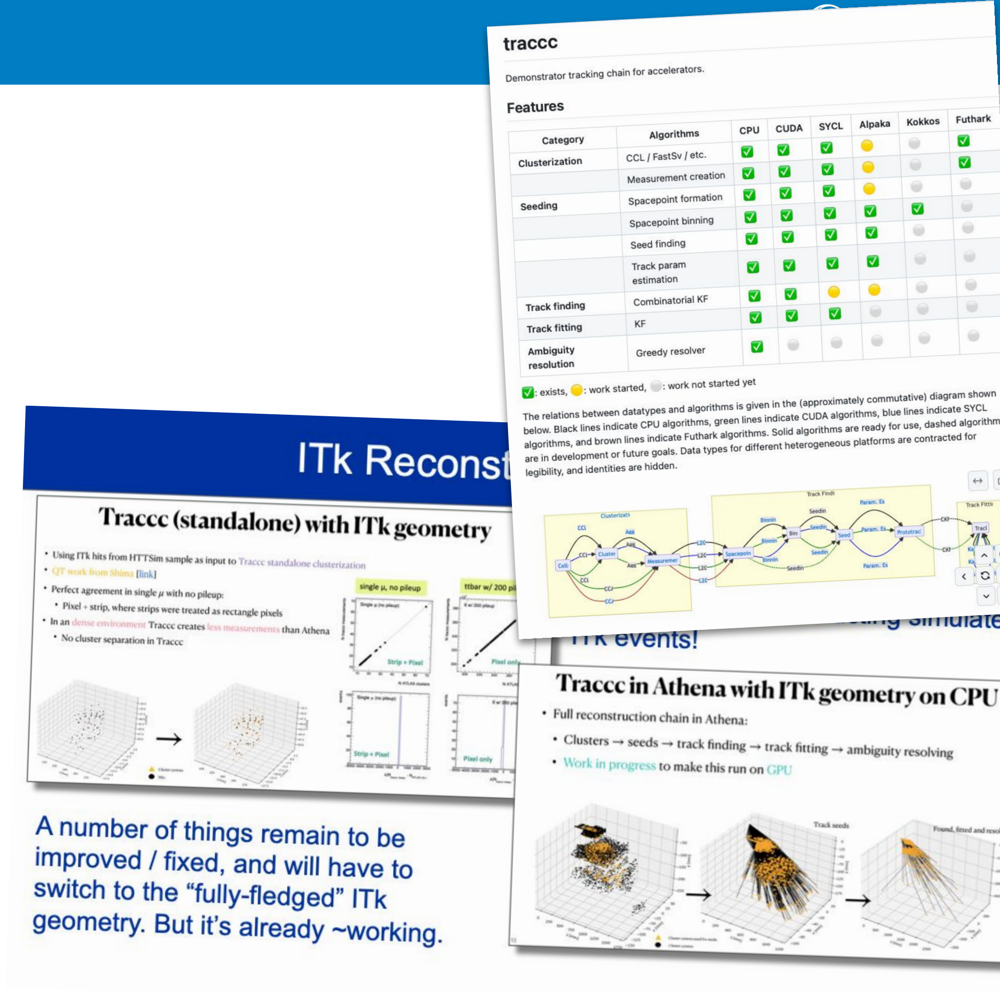
Demonstrators

- ATLAS also has ~30 demonstrators
  - Idea is that these will help inform the TDR
- Most of these have been submitted to CHEP
- In the next slides, I will highlight a few…

Name / Brief Topic	Responsible	ATLAS Milestone(s)	Timeline	Google Doc Link
ACTS tracking: demonstrate that performance (physics, CPU) will be sufficient for HL-LHC needs	Noemi Calace, Andreas Salzburger	RE-4		http://cds.cern.ch/record/2845248
Performance advantage of running components of tracking algorithms on GPU and/vs of running end-to-end tracking on GPU	Attila Krasznahorkay, Andreas Salzburger	RE-5		https://docs.google.com/document/d/1TelP96bd2c4krZhOSeJNiNVejo-GG7pYLSwfGZWrGZ8/edit?usp=sharing
GNN Pipeline for Particle Tracking	Paolo Calafiura, Jan Stark	RE-5	Next MS: Oct 22	Doc
Remote read of selected data vs copy of whole file	Ilija Vukotic, Paul Nilsson	DC-11 (relevant?)		
RNtuple fully capable of storing DAOD-PHYSLITE with advantages on storage footprint and I/O performance	Marcin Nowak	CS-5.1, AN-2		https://docs.google.com/document/d/1MeaB3nw8ZUD7fmtmrWQM8vhQOuLnkX-7DP6XffXwH0c/edit#
Demonstrate that lossy compression preserves physics performance	Serhan Mete	CS-2		Lossy compression preserves physics performance Demonstrator - Google Docs
Broaden DAOD-PHYS usability while limiting size increase for analysis by adding custom event augmentation.	Peter van Gemmeren	AN-5.4b		Broaden DAOD-PHYS usability by adding custom event augmentation performance Demonstrator - Google Docs
Stress-test xcache by running many (or all) the US analysis jobs at a single site and have all inputs done through xcache.	Ilija Vukotic	DC-11 (kind of)		
HPX for scheduling tasks in distributed heterogeneous environment	Beojan Stanislaus, Julien Esseiva, Vakho Tsulaia	CS-13.1		Vertically Integrated Next Generation Task Scheduler for the Athena Framework - Google Docs
Process data from multiple events on accelerators (event batching)	Beojan Stanislaus, Charles Leggett	CS-10.1		Event Batching Demonstrator
FastChain demostrator	FastChain group	SI-6		FastChain demonstrator doc: this is written as three demonstrators
Adaptive data placement/job scheduling using Intel's Loihi2 neuromorphic computing platform.	Ilija Vukotic	DC - no milestone		
Use ATLAS-Google site as "bursty" resources (demonstrate ATLAS bulk campaign [reprocessing or MC production] on this site)	Fernando Barreiro Johannes Elmsheuser Alexei Klimentov Tadashi Maeno	DC-10.2		GoogleDoc
ATLAS-Google site for analysis workflows and new technologies for physics analysis	Fernando Barreiro Alexei Klimentov Tadashi Maeno	DC-10		**ON HOLD**
ARM based PanDA queue setup and operation for Athena porting and Physics Validation	Fernando Barreiro Johannes Elmsheuser	DC-6.1		GoogleDoc
Recreate DAOD datasets on demand by PanDA using Data Carousel (Delete DAOD datasets from the lifetime model exception list and reproduce them on demand if needed)	Alexei Klimentov Tadashi Maeno Xin Zhao	New for DC-7 (was 11)		Recreate DAOD on demand demonstartor
Using Xrootd to seamlessly integrate S3 storage with ATLAS DDM system, including Google and Amazon cloud S3 storage.	Andy Hanushevsky Wei Yang	DC-10?	23Q4	
~~Large scale AI/ML services across diverse resources, complex analysis workflows~~	Tadashi Maeno	DC - no milestone		AI/ML Services Demonstrator

Demonstrators(2)

- [Traccc](#) is the ACTS demonstrator for Tracking on accelerators
- Can run on the “OpenDataDetector” and gives reasonable results (even without optimisation)
- … and also now on ATLAS’s ITk (HLK-LHC tracker)



[Update from Attika Krasznahorkay in ATLAS S&C week 78](#)

Demonstrators(3)

RNTuple demonstrator

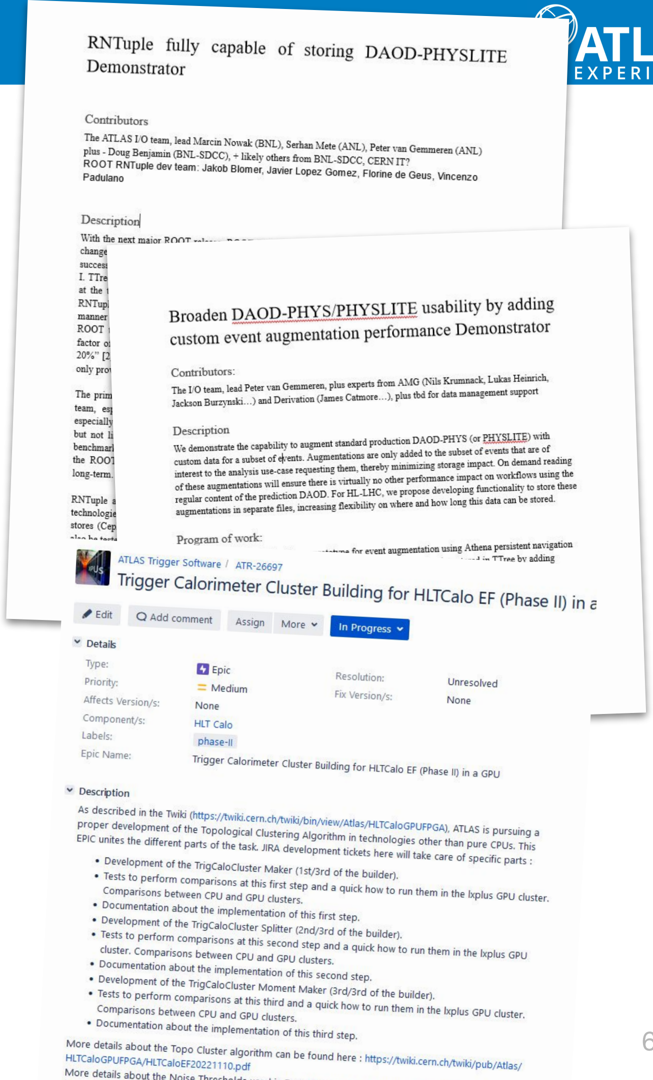
- ATLAS has RNTuple support for all official formats, e.g. HITS, RDO, ESD, AOD, DAOD etc
  - ACAT presentation

Event Augmentation

- Adds ability to add custom data for subset of events, for particular analysis
- On demand reading, in order not to impact other workflow
  - ATL-SOFT-PROC-2023-003

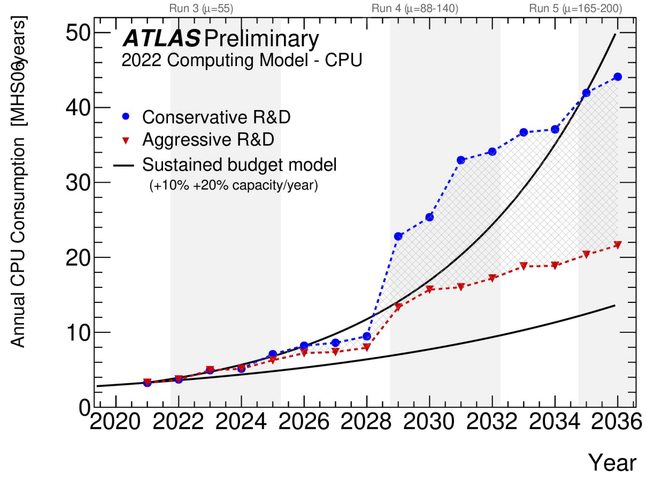
GPU-based TopoCluster reconstruction

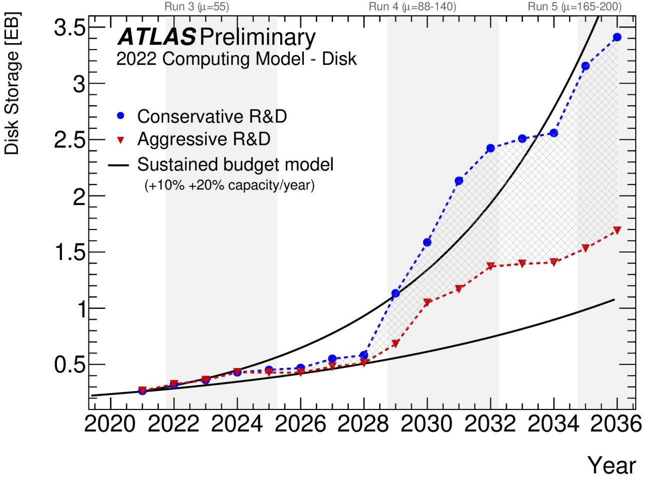
- Topo Cluster is one of the most resource demanding algorithms for HLTCalo in Run ⅔
  - doi:10.1088/1742-6596/2438/1/012044



Resources

- Hopefully these computing model plots are familiar now!
- Will be updating these soon™, adding GPUs for the first time
  - Requires estimates from each area for how much work they think can be done on a GPU
  - Requires model for incorporating GPUs into the resource estimates





ATLAS feedback

General

- ATLAS is grateful for HEP-CCE’s awareness of our timeline and needs
  - e.g. in Paolo’s talk at a recent ATLAS S&C week, he linked our R2R4 milestones to HEP-CCE goals
- Focus on usability, and real-world applications is great!

HEP-CCE and ATLAS Road 2 Run 4

Many areas of active or potential collaboration (refer to R2R4 milestones)

		R2R4 milestone	CCE activity
RE	3	New Reconstruction EDM	EDM for heterogeneous computing and columnar data formats
RE	5	Accelerator and machine learning (R&D)	Inference + ACTS/traccc as a Service, Scaling up training and inference for large models (GNNs, transformers)
EG	1	Assessment of CPU costs using Run3 MC setup	by Run 4, CCE MadGraph on GPU, + Pepper NLO can help move some of EvGen workloads to GPU systems
EG	2	Development of nominal MC setups	

11

ATLAS feedback (2)

ATLAS EXPERIMENT

PAW

- Good example of activity that is very nicely focused on real world applications
- Tasks and plan for next year(s) looks both interesting and feasible/realistic
- “Our” Traccc is getting closer to production quality.
  - We think that this might be good to add this as a mini-app?
- We also support the idea of using an ATLAS HPC workflow
  - We regularly run on HPC systems, so we know this works well!
  - **Update**: I gather from Charles’ slides earlier this is happening? Great!

HEP-CCE Phase 2 Plan

HEP-CCE

- In HEP-CCE Phase 2, our goal is to provide the experiments with both a validated, ready-to-use portability solution and a suite of portability tools that can be integrated into their production systems.
  - To reconcile different services and tools provided by HEP and ASCR.
  - To reduce the operation and maintenance overhead of deploying HEP workflows on HPC systems
- Building on the experience of **PPS** and **CW** groups in HEP-CCE Phase I, we will have two main tasks in Phase 2:
  - **Task 1:** apply lessons learned in PPS to help HEP experiments develop portability solutions in their applications
  - **Task 2:** develop portable, experiment-agnostic, workflow overlays to interface existing HEP workflows with HPC centers

10 U.S. DEPARTMENT OF ENERGY Office of Science Argonne NATIONAL LABORATORY Brookhaven National Laboratory Fermilab BERKELEY LAB

Year 1 Plan

HEP-CCE

Task 1 - Application Portability:

- Develop a **cookbook** for portability layers based on Phase 1 findings
- Outreach to experiments for portable solution implementation (**workshops/hackathons**, followed by regular office hours)
  - Understand the experiments’ timescales for portable accelerator uses
- Create **mini-apps** based on two of the Phase I PPS testbeds that can be executed at NERSC, OLCF and ALCF, preferably with the same software environment (FCS, p2r)
- Use mini-apps to extract **figures of merit** for ASCR facilities and LCFs to use as baselines

Task 2 - Workflow Portability:

- Complete **survey** of existing HEP experiment workflow technologies on HPC; also look into workflow technologies used by **other experiment facilities such as light sources**.
  - Find commonalities between experiment workflow systems
- Explore the needs of HEP in terms of **ML workflows/pipelines and microservices** (synergistic with the distributed ML activity)
- Investigate common layers and interfaces (batch scheduler, policies, pilots, …) to facilitate portability and interoperability across ASCR facilities in collaboration with **IRI testbeds**
- Create **2 representative HEP experiment workflows** to run two different HPC systems. Candidates include: LSST/DESC, LZ, **DUNE**, LHC Experiments (**ATLAS**/CMS).

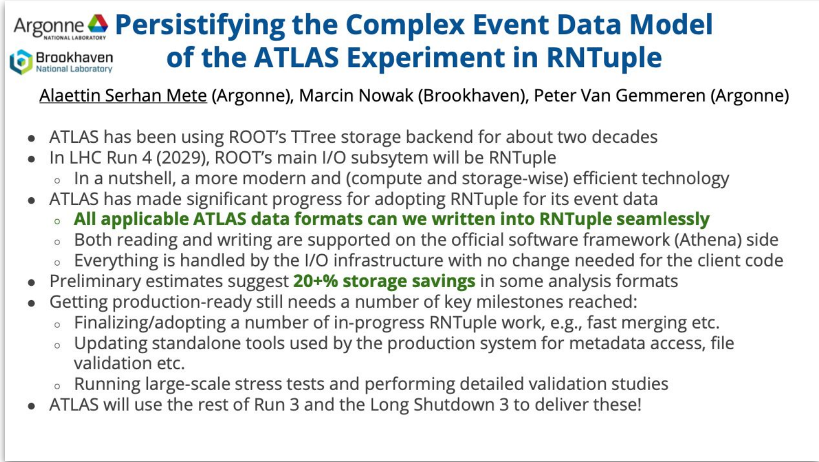
13 U.S. DEPARTMENT OF ENERGY Office of Science Argonne NATIONAL LABORATORY Brookhaven National Laboratory Fermilab BERKELEY LAB

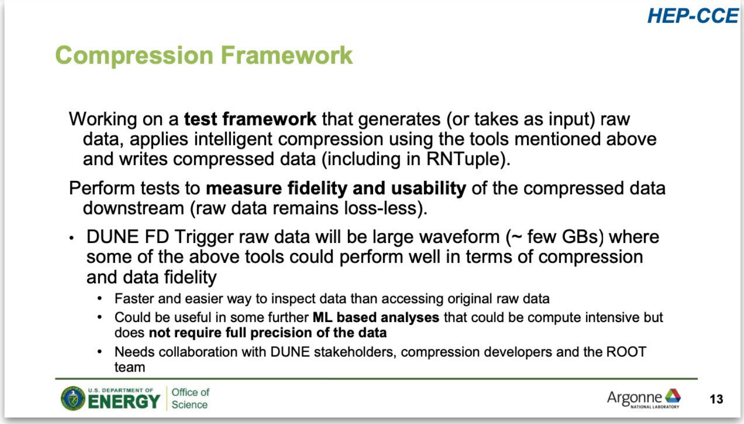
Charles's talk at BNL workshop

ATLAS feedback (3)

SOP

- Again, effort seems nicely focussed on real-world practicality
  - i.e. actionable results
- ATLAS is expecting RNTuple to be vitally important in Run-4, so we strongly support the SOP efforts here e.g.
  - RNTuple workshop
  - RNTuple API review
- Interested to see the outcome of other investigations, e.g. metadata, compression and object stores





Peter's summary at CAF

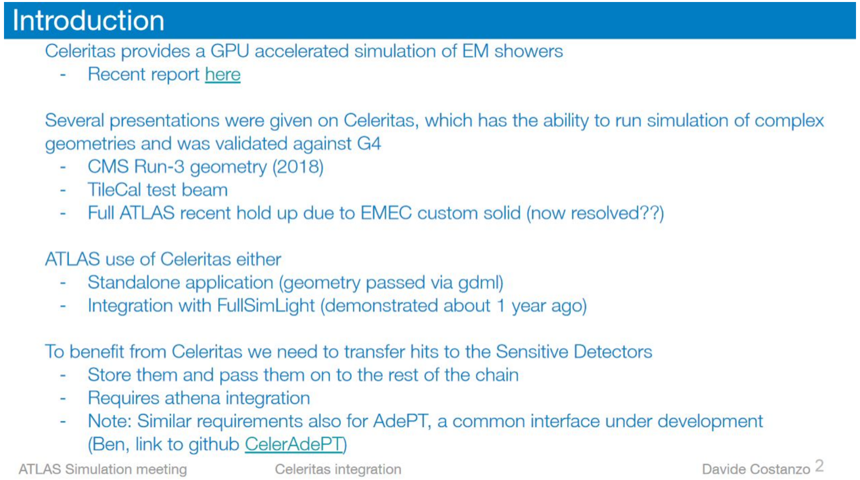
ATLAS feedback (4)

SIM

- Not so obvious what has been happening in this area
  - Perhaps the work could be better publicised?
  - (I heard from Julien Esseiva about work on optical physics/Orange)
- ATLAS has been working on Celeritas integration
  - Planning hackathon in October

SML

- Recent meeting discussing Simulation based inference in ATLAS
  (“ATLAS Data Analysis using a Parallel Workflow on Distributed Cloud-based Services with GPUs” paper)
- Given expected increasing importance of ML in Run4, of definite interest to ATLAS



Introduction

Celeritas provides a GPU accelerated simulation of EM showers
- Recent report here

Several presentations were given on Celeritas, which has the ability to run simulation of complex geometries and was validated against G4
- CMS Run-3 geometry (2018)
- TileCal test beam
- Full ATLAS recent hold up due to EMEC custom solid (now resolved??)

ATLAS use of Celeritas either
- Standalone application (geometry passed via gdml)
- Integration with FullSimLight (demonstrated about 1 year ago)

To benefit from Celeritas we need to transfer hits to the Sensitive Detectors
- Store them and pass them on to the rest of the chain
- Requires athena integration
- Note: Similar requirements also for AdePT, a common interface under development (Ben, link to github CelerAdePT)

ATLAS Simulation meeting Celeritas integration Davide Costanzo 2

Slides from Davide Costanzo (in ATLAS simulation meeting)

Conclusions

ATLAS strongly supports the HEP-CCE efforts

Some areas apparently already very active, some .. there is apparently work ongoing, but not-so well publicised (regular summary talks are very useful).

We look forward to seeing