U.S. ATLAS Software and Computing R&D for the High Luminosity LHC

V3, Aug 17, 2020

Executive Summary

U.S. ATLAS has for many years been a leader within the ATLAS Collaboration in researching, designing and developing innovative software and computing approaches and systems that can both deliver ATLAS production needs in the near term and scale to meet the challenges of HL-LHC computing. U.S. ATLAS formalized the effort towards HL-LHC in June 2018 with the creation of a dedicated HL-LHC Computing activity within the U.S. ATLAS Operations Program. The program is charged with managing research and development towards affordably meeting ATLAS computing requirements in the LHC era, particularly for the US program but as part of a coherent ATLAS wide effort and wider collaborative LHC community effort. In the two years since, U.S. ATLAS has been a driving and influential force as HL-LHC computing planning has ramped up in ATLAS, WLCG, HSF and the HL-LHC directed US software projects IRIS-HEP and HEP-CCE. The influence of U.S. ATLAS in shaping the HL-LHC computing plan and R&D program is reflected in the coherence of our program with that of ATLAS overall and the collaborating projects.

This document briefly reviews the challenges of ATLAS computing at the HL-LHC, describes the present state of ATLAS plans towards HL-LHC readiness, the research and development activities underway or foreseen that are the near to mid term expression of that plan, and the specific U.S. ATLAS R&D program over the next 2-3 years within the overarching ATLAS context. Finally the overall HL-LHC computing planning process at the LHC and ATLAS levels is described, within which the US program is tightly integrated. The ATLAS and U.S. ATLAS plans we present reflect the important roles played by wider collaborations, with (U.S.) CMS, WLCG, IRIS-HEP, HEP-CCE, HSF and others.

The U.S. ATLAS HL-LHC software and computing R&D program we describe includes about 20 specific high-level deliverables/milestones over the next 3 years, across three primary activity areas making up a WBS with almost 40 sub-activities described to level 5. Of the principal R&D areas selected, since the U.S. ATLAS HL-LHC Computing activity was created two years ago, many are on track for adoption and production in Run-3, constituting a practical validation of our ongoing development activity towards Run-4. All these R&D areas have deliverables included in our milestones.

With this document we aim to demonstrate that U.S. ATLAS has a carefully planned and managed, successful and visionary program in HL-LHC S&C research and development. Our broad, mature program combines innovative research and prototyping with a practically directed development program that anchors long term R&D in near term production applications, which continually validates that we are on the right track. The R&D program is tightly embedded within and coherent with both the ATLAS wide R&D effort and the operations efforts addressing the current S&C needs of U.S. ATLAS and ATLAS. It both contributes to and strongly leverages the HEP-CCE, IRIS-HEP, and WLCG programs. It is scoped to focus on our longstanding areas of expertise, and scaled to produce deliverables that are attainable with realistic effort levels and that represent an appropriate fraction of the overall ATLAS effort, while delivering many critical components of the overall ATLAS program. Our program is built on our past successes: the U.S. ATLAS core and distributed

software efforts have had an evolutionary history free of revolutions, the legacy of which is a stable, powerful software foundation upon which our developers can very efficiently build to meet the needs of the HL-LHC.

The ATLAS computing challenge at the HL-LHC

The HL-LHC, commencing with Run-4 expected in 2027, will deliver unprecedentedly complex events, with up to 200 interactions per proton-proton bunch crossing. These events will be collected at a prodigious rate: ATLAS expects to record data at 10kHz, approximately ten times more than during previous runs. By the end of Run-5 in 2034, the HL-LHC is expected to have delivered an integrated luminosity of up to 2500 fb^{-1}, five times more than all previous runs combined. As well as the challenges involved in collecting, storing, reconstructing and analysing such a volume of data, Monte Carlo (MC) simulation events will need to be produced in similar numbers. Taken together, the data and MC requirements of the HL-LHC physics program are formidable, and if the computing costs are to be kept within feasible levels, substantial improvements must be made in both compute and storage.

While we expect technology evolution will remain a positive contributor to delivering more compute and storage per unit cost, it brings large challenges of its own, particularly the emergence of co-processor accelerators as an increasingly important component of compute power, a component that present ATLAS workloads are unable to leverage. HPCs relying on accelerators for their power are becoming prevalent at DOE HPC centers and are emerging elsewhere as well. ATLAS must be capable of using such machines to draw on all possible sources of processing power. U.S ATLAS in particular aims to include such resources in their WLCG pledge on an HL-LHC timescale. On the storage side, the situation is equally dire; disk storage is and probably will remain the largest cost component of our computing, but ‘opportunistic’ storage basically does not exist.

An extrapolation of the present computing model to the HL-LHC conditions, accounting for anticipated technology evolution, indicates a substantial shortfall in disk space and computational capacity. This is true even if the investment from funding agencies continues at the rate established in previous years. Given that substantial increases in the HL-LHC computing funding level are unrealistic, we must look to innovative new developments in software and computing as the basis for a new computing model directed at closing the gap without compromising the physics.

Resource estimates

In order to quantify (within the large uncertainties) and track the CPU and disk storage challenges of the HL-LHC, and provide the basis for mid to long range resource planning, ATLAS has developed an increasingly sophisticated model for projecting the resources needed and available during the HL-LHC period. A broad range of approximately 60 parameters supports flexible assumptions on the data-taking and computing activities that will be carried out by ATLAS, the actions taken to minimise computing costs, and the provision from the funding agencies. The model takes as input the latest available LHC schedule and estimated luminosity profile, and supports concurrently evaluating multiple scenarios based on different assumptions. U.S. ATLAS is a key contributor to this resource planning through its roles in S&C leadership and its lead role in developing and supporting the modeling software.

In the HL-LHC CDR, ATLAS has chosen to evaluate two R&D scenarios based on how aggressively we anticipate the HL-LHC S&C R&D program to deliver improvements. They are useful straw-men to discuss model uncertainties and development costs. These scenarios are measured against a baseline scenario in which we optimize the Run 2 computing model without performing additional R&D. While requiring significant effort, this baseline scenario is not a tenable strategy in itself but a benchmark against which to measure the expected return on R&D investments. These three scenarios, defined in terms of the R&D program activities described in the next section, are as follows:

- **Baseline:** ATLAS implements the new data formats foreseen by the Run-3 analysis model, the multi-threaded software framework AthenaMT, and updates to the tracking code, but otherwise continues in largely the same way as in Run-2. In particular the CPU time per event for event generation, detector simulation and reconstruction is assumed to remain at the level currently achieved by applying the current software to the Phase-II detector simulation, and the mixture of generators and simulation remains the same.
- **Conservative R&D:** The research and development activities currently underway for Run-3 are assumed to be successful, including the data carousel, fast track reconstruction, lossy data compression, and most of the detector simulation being performed with fast simulation.
- **Aggressive R&D:** ATLAS implements new developments that very significantly improve the speed and storage volumes of workflows that currently are heavy consumers of resources. For example, porting of high-precision generators to GPUs, sharing events with CMS, or speeding up the full simulation either by software efficiencies or porting parts of the code to GPUs. The development of very high quality fast simulation that could replace full simulation in almost all cases, would also fall into this category. Some R&D activities in the aggressive category can not yet be quantified in their impact and so are not yet included in the model. An example is Machine Learning (ML) models for fast simulation and reconstruction.

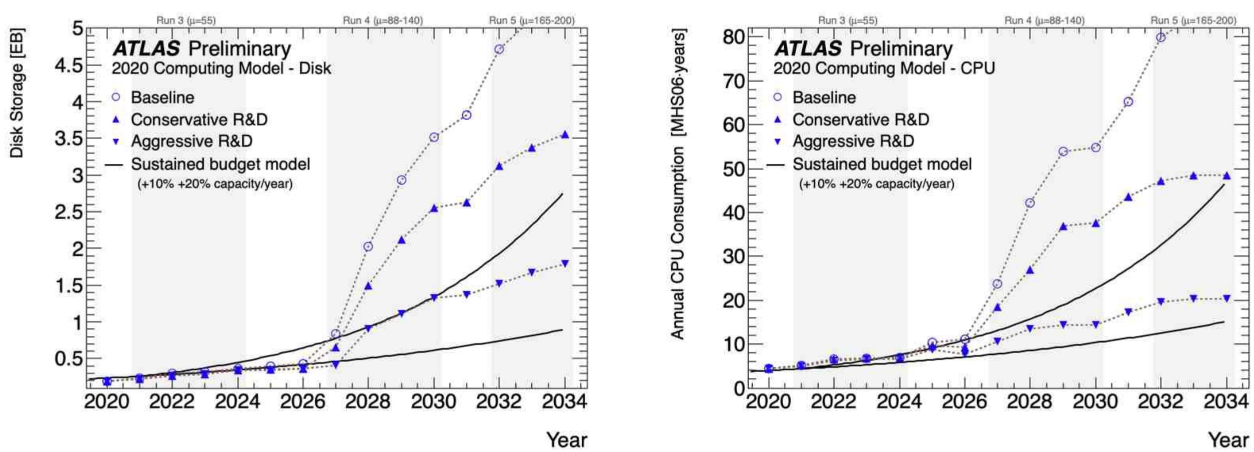


Figure 1. Estimated CPU and disk resources needed for the years 2020 to 2034 under the different scenarios. The solid lines indicate annual improvements of 10% and 20% in the capacity of new hardware for a given cost.

Figure 1 shows the estimated CPU and storage needs under these different scenarios, to be compared with provisioned resource curves based on flat funding for CPU and storage, and capacity growth per unit cost of 10% and 20% per year. U.S. ATLAS in its current best-estimate budget planning assumes 10% growth (the lower curve) for the years out to HL-LHC, based on an assessment of the technology landscape. Other estimates within the HEP community range from 5% to 20%. These studies demonstrate that ATLAS (and U.S. ATLAS) must pursue vigorously aggressive R&D options to be able to meet the HL-LHC challenges in a sustainable budget model.

Preparing ATLAS S&C for the HL-LHC

While ATLAS, and particularly U.S. ATLAS, have had the challenges of HL-LHC computing in view for many years, targeted preparation and planning in earnest have begun relatively recently. U.S. ATLAS has led the way since June 2018 with the establishment of the HL-LHC Computing activity within the U.S. ATLAS Operations Program, now a mature R&D program scoped to the U.S. ATLAS focus areas of core software, distributed software and HPC exploitation. These focus areas were selected for their impact on U.S. ATLAS analysis, and for their alignment with NSF and DOE priorities (in particular for exploiting new HPC platforms), and with U.S. expertise and long-term commitments to ATLAS software and computing. The U.S. program leverages the fact that our work has been ultimately HL-LHC directed for many years. Our core software effort has built increasingly sophisticated high concurrency capability into ATLAS offline software since the genesis of a multiprocessor framework, AthenaMP, in 2010 [AthenaMP]. Our distributed software effort has focused on extreme scalability and low operations load to support a broad range of platforms, with innovative new workflows and services targeting the specific characteristics of HPCs. On the foundation of these efforts U.S. ATLAS has for years been a leader in exploiting HPCs, including the most challenging LCFs, anticipating the central role such resources will play in the HL-LHC era.

Over the last year ATLAS has begun HL-LHC S&C planning in earnest with the writing of a Conceptual Design Report released in Spring 2020 and reviewed favorably by the LHCC. U.S. ATLAS was heavily involved in shaping and writing it, including one of the two lead co-editors. U.S. ATLAS was also a major contributor to the HEP computing roadmap prepared by the HSF that was an important input to the CDR. The CDR has given a coherent form to the present ATLAS view of the HL-LHC computing challenge and how ATLAS plans to address it. It fits hand in glove with the U.S. ATLAS program, as expected given our close involvement.

The following subsections will briefly describe ATLAS S&C at the HL-LHC, following the structure of the CDR, and highlighting US focus areas and involvement. The next section will then describe the work program, timeline and milestones towards achieving the S&C developments needed for success at HL-LHC, at the ATLAS level and in more detail at the U.S. ATLAS level.

Core software

ATLAS Core Software provides all necessary components and tools to develop and assemble online and offline data processing applications that run efficiently on a variety of platforms. U.S. ATLAS leads the ATLAS core software program including the development of the Athena application framework and the event I/O services. The principal HL-LHC challenges are to support parallel software: very high processor concurrency, and the utilization of accelerators in heterogeneous architectures. The US has led this parallel software effort since the creation of AthenaMP in 2010.

By the time of HL-LHC, most HPC facilities will rely on heterogeneous architectures, and it is likely that hardware installed at dedicated ATLAS grid sites will feature elements of these technologies as well. U.S. ATLAS is applying and extending its parallel software expertise in our HL-LHC Computing R&D program, working to support algorithms that leverage accelerators (GPUs) or machine learning (which is inherently GPU capable). The HEP-CCE project is an important contributor to this, as the first example of a cross-experiment R&D effort in portable parallelization strategies that should show the way forward [PPS].

We are also scaling up support for fine grained processing on HPCs, following the successful Yoda implementation of the event service with a second generation system based on the Ray common software package [Ray] originally from UC Berkeley. Fine grained processing ensures efficient HPC utilization, with every core fully utilized for the duration of job execution, and (almost) all productive processing time recovered no matter how abruptly the job slot is terminated; ideal characteristics for exploiting HPCs. This is an R&D path we can expect will continuously evolve towards HL-LHC as HPCs ramp up in scale and inevitably complexity.

An additional U.S. focus in the core domain is Event Data Model (EDM) and persistency. A simplified EDM is a key component of an ATLAS portable parallelization strategy, and an essential target for storage optimization, such as the development and tuning of lossy compression algorithms, and of a distributed parallel Event I/O framework for the next generation of HPCs. ATLAS has made a large and successful effort to save O(30%) in analysis data storage through more compact persistent data formats and a more efficient analysis model. Compressing yet further while scaling up concurrent I/O is an important ATLAS and U.S. R&D task towards HL-LHC.

Event generation

Event generator workflows are the first step in the production chain for MC events. They are a large consumer of computing resources (approximately 10-15% of CPU cycles), likely to increase in Run-4 as many measurements and searches will require more accurate physics modelling, implying the use of NLO with many legs or NNLO event generators. ATLAS is exploring two approaches to controlling event generation processing needs in Run-4: physics optimizations, and re-engineering/parallelizing event generators, particularly for HPCs.

Parallelizing event generators important for HL-LHC, like Sherpa and MADGRAPH, is a clear interest for the U.S. particularly, where LCFs are a major resource. The LHC experiments cannot undertake this work on their own, it needs to be a collaboration with the phenomenologists who develop the generators. U.S. ATLAS has leading R&D involvement in Sherpa and MADGRAPH parallelization efforts through HEP-CCE and the ALCF A21 Early Science Program, in collaboration with the authors of the generators.

Detector simulation

ATLAS currently expends around 40% of its CPU resources on detector simulation. Approximately half of the events are produced with ``full'' simulation, using the Geant4 simulation toolkit, with the rest coming from ``fast'' simulation, which uses a parameterised model of the calorimeter response and Geant4 elsewhere. Full simulation uses around five times more computing resources than fast simulation, and ATLAS aims to use mostly fast simulation for Run-4 and beyond. Studies are ongoing to improve the quality of the fast calorimeter simulation, extend the fast simulation concept to all parts of the detector, including the inner tracker as well as the calorimeter. Novel simulation techniques made possible by deep learning may also prove highly effective at improving the precision of the fast simulation. Use of full simulation will remain unavoidable for certain applications, and to train the fast simulation. For this reason ATLAS will continue to collaborate closely with the Geant4 team to ensure the best possible physics fidelity for the lowest possible expenditure of resources. Finally, the transition to multithreaded processing may make it possible to significantly increase the memory and I/O efficiency of pileup digitization workflows, by sharing and reusing minimum bias accidental pileup events across all threads.

When the U.S. ATLAS HL-LHC Computing program was first established in 2018, R&D to adapt fast simulation (particularly the complete fast chain) to accelerators and HPCs was judged to be the most promising relatively early path to establishing an important production workflow on GPU HPCs and LCFs. In contrast to Geant4, which many projects have tried to adapt to GPUs, fast simulation on GPUs -- specifically fast calorimeter simulation which is by far the slowest part of the simulation in ATLAS -- is an untouched and inherently simpler problem. It is also amenable to machine learning, another route to GPU and LCF exploitation. This choice has shown its value in the two years since. FastCaloSim and FastChain have become ever more important in the ATLAS HL-LHC strategy, the ML approach is showing promise, and FastCaloSim's GPU port has been adopted as a standard parallelization testbed by the HEP-CCE project.

Event reconstruction

Dealing with the increased event complexity from unprecedented levels of pile-up poses a large challenge to the offline event reconstruction and physics object identification. While the calorimeter dominates the simulation time, tracking dominates the reconstruction time. A naive scaling of reconstruction time with the increased combinatorics would lead to untenable CPU consumption. Controlling CPU utilization without compromising physics performance has been and will continue to be an important R&D activity at least until Run-4. Several optimizations have been identified, beginning with the design of the Phase-II inner tracker itself, which was optimized not only for physics performance but also for reconstruction processing time under HL-LHC conditions. The fast track seeding and finding approaches enabled by the detector design have borne fruit in reduced reconstruction time.

Nonetheless, much improvement remains to be achieved, and much scope exists to achieve improvements through novel tracking approaches. The physics impact of a parallel algorithm that scales quadratically or better with the average pileup would be dramatic, particularly for precision physics studies, as they would enable lowering online and offline pT thresholds and consequently gaining physics acceptance. Believing a fully new tracking software is called for, a group of ATLAS tracking experts initiated the ACTS project [acts], a new cross-experiment multithreaded tracking software, to exploit the optimizations made possible by a modern C++ design with a light-weight event data model, including active efforts to port time-critical parts of the code to GPUs. By making it cross-experiment, ATLAS is both sharing the benefit of a new tracking package and reaping the benefit from a wider community of developers (which includes FCC, sPHENIX, and EIC among others).

U.S. ATLAS is a strong supporter of this R&D path and is leveraging a number of community projects (beyond ACTS itself) to support our involvement in ACTS and next-generation tracking in general. There are active R&D investigations in parallel tracking algorithms both in the ACTS and HLT communities. Machine Learning track finding algorithms like similarity hashing, metric learning and graph neural networks are particularly promising because they run very efficiently on accelerators, and show promising physics performance on simplified detector simulations. U.S. ATLAS members are leading or contributing to all of these approaches, ensuring that investigations are relevant for ATLAS tracking. We're involved in Exa.TrkX [exatrkx] exploring ML approaches, IRIS-HEP has with our encouragement and involvement become an ACTS collaborator, and ACTS has, like FastCaloSim, been adopted as a HEP-CCE parallelization testbed. ACTS is a prominent example of an open software approach to a new project catalyzing growth by attracting other common software projects.

Analysis

The largest disk consumer in ATLAS is and will remain the Derived Analysis Object Data (DAODs). The expected DAOD volumes at HL-LHC make it imperative for ATLAS to implement a new analysis model with a smaller DAOD disk footprint, and to reduce the "time to physics" by streamlining the data reduction and analysis workflows. Modern analysis tools that can "scale out", running an analysis in parallel on distributed resources, and data formats that are amenable to be used efficiently by these modern tools will be key in this R&D program.

As a first step, ATLAS in a U.S.-led effort defined two new analysis formats, one with a maximum size of 50KB/event (DAOD_PHYS) and the other 10KB/event (DAOD_PHYSLITE). The second smaller format contains precalibrated reconstructed quantities, and in consequence the variables needed to apply the calibrations do not need to be stored. Both these new formats will be used in Run-3 in preparation for them taking over as the main format at HL-LHC, no doubt after some evolution informed by the Run-3 experience. DAOD_PHYSLITE will replace most of the dedicated analysis formats currently in use during Run-2, and will reduce the disk footprint of the analysis formats overall by approximately half. A yet more aggressive option is to forego persistent storage of most AODs and instead rely on the ability to regenerate AODs on demand in a fast simulation full chain. U.S. ATLAS R&D work on FullChain for LCFs will help answer whether this potentially very powerful approach will be practical. It could turn LCFs into a prolific simulation AOD factory. The U.S. ATLAS distributed computing R&D program described below is also delivering a key part of the strategy to substantially reduce the DAOD disk footprint, the data carousel.

The DAOD_PHYSLITE ntuple-like columnar data may be served to the analyst over the network to reduce replica counts and thereby storage footprint, in which case it is much more efficient to send only the data needed by the analyst, in an optimized form. U.S. ATLAS in collaboration with IRIS-HEP is developing tools that perform basic transformations near the data source, the intelligent data delivery service [iDDS] and its data transformation component ServiceX [ServiceX]. These tools, together with high bandwidth access to the data, could filter and apply basic transformations to efficiently serve a personalized data stream.

Physicists are relying more and more on the scientific python ecosystem to run their analysis workflows. Jupyter notebooks, python data analysis and visualization libraries, not to mention several machine learning frameworks are used more and more frequently. U.S. ATLAS has started a development program aiming at integrating these tools with ATLAS workflow and data management systems, as well as to provide direct access to ML training and inference tools from ATLAS software. U.S. ATLAS also started an aggressive R&D program, described below, with the goal to process very large datasets and to train and optimize complex ML models on DOE and NSF HPCs.

ATLAS has made significant progress in the adoption of industry tools to ease analysis development, deployment and preservation. U.S. ATLAS has been a leader in this program, and is collaborating with NSF DIANA and IRIS-HEP to bring the RECAST framework [RECAST] into production usage. A major use case for this is the preservation of data analysis pipelines for later reuse in combined summary analyses, such as an ATLAS-wide assessment of SUSY. As the LHC experiments enter an era of high-precision measurements such combinations of multiple disparate analyses will be crucial to maximally exploit the physics potential of the HL-LHC.

Distributed computing

The ATLAS Distributed Computing (ADC) system and organization is a sophisticated ensemble of software systems, computing facilities and people that make it possible for ATLAS processing to run 24x7x365 on about 450k cores across over 100 worldwide sites and a wide range of resource types including grid clusters, HPCs, LCFs and clouds, processing in total about 1.5 exabytes of data per year. The automated tools developed by ADC include the ProdSys production system [ProdSys], PanDA workload manager [PanDA], Rucio data manager [rucio] and others. ProdSys and PanDA were developed by U.S. ATLAS and designed to support scaling up to the use cases and operating scale of HL-LHC, underway in such projects as the data carousel and iDDS described below. Rucio also is designed to scale to HL-LHC, and has been adopted as the LHC and

WLCG standard; it is primarily a CERN product, but ProdSys and PanDA are its biggest clients and work in close concert with it.

Many of the upgrade plans require tighter future integration among tools to improve efficiency in the more complex environment of the HL-LHC. Intelligent scheduling of jobs based on improved coupling between job characteristics and available resources, for instance CPU and accelerators, is important to allow full exploitation of the heterogeneous resources that are likely to be provided in the coming years. Intelligent automation coupled with emerging technologies (such as container orchestration) will increase the efficiency with which we will exploit Grid and opportunistic resources. Use of new storage paradigms, such as cloud storage, data lakes and caches, might also yield further benefits, especially for smaller sites. On-demand production of custom data formats (particularly DAOD) will be a natural complement to dynamic data storage. The U.S. ATLAS driven iDDS and ServiceX projects are examples of cross-experiment R&D activities in this area, run in collaboration with IRIS-HEP and overseen by the HSF and the DOMA group within WLCG.

Economizing storage is an important goal for HL-LHC computing R&D. Unlike CPU, the requirements for which will become approximately constant once the LHC reaches its design luminosity, storage needs will continue to increase during the lifetime of the HL-LHC. And as mentioned, while opportunistic computational resources exist, opportunistic storage does not. Innovative ideas for optimising storage by breaking out of the disk/tape paradigm to a finer grained spectrum of storage cost-reliability-latency are being pursued. In particular, the U.S.-driven ATLAS data carousel project is developing a mechanism to stage data from tape when they are required for processing, reducing by 50% or more the input sample volume resident on disk. The data carousel needs a tightly integrated orchestration of workflow and workload management that processes data with file-level granularity as soon as it appears from tape, keeping its disk residency to a minimum.

The success of the LHC and ATLAS computing to date has its foundation in powerful networking, and this will not change in the future. Strategies for HL-LHC computing are based on extensive use of powerful networks to reduce data replication by streaming over the net, and consolidating distributed resources into cohesive virtual federations such as data lakes. HL-LHC data and processing scales are large enough that optimizing for efficient bandwidth use will be essential. As part of this optimization, the ability to mark network packets to identify sources of traffic, and to route traffic to control speed and cost may become vital. To this end, ATLAS initiated the HEPiX Network Function Virtualization working group in collaboration with other experiments and national entities including ESnet. U.S. ATLAS will play central roles in the R&D in this area, drawing on our close relationship with ESnet, our own long standing networking expertise within the facilities, and our distributed computing leadership in PanDA and its Rucio integration; shaping network traffic through Rucio and PanDA will be crucial.

Analysis data will dominate storage needs at HL-LHC, and group and user analysis workflows are expected to play a larger role than in the past. Currently, Tier 3 sites, funded by U.S. ATLAS and managed outside the WLCG framework, provide the necessary resources for end-user analysis. For the HL-LHC, it is desirable that users have access to resources that support running both asynchronous, batch-like workloads as well as synchronous interactive, scalable workloads. Significant R&D work is needed to determine the size, scale and cost of analysis facilities, and crucially the set of analysis tool services that they will offer. Metrics of reduced cost, increased usability, efficiency, and scalability need to be demonstrated. U.S. ATLAS, with its expertise in established analysis facilities, leading edge tools such as Jupyter, and distributed computing support for analysis, will play a large role in the R&D.

HL-LHC S&C R&D in U.S. ATLAS

In this section, having set the overall ATLAS context, we describe how the R&D is organized and managed in U.S. ATLAS, with the plans and milestones for the program over the next 2-3 years developed in the context of overall ATLAS planning. To ensure tight integration with ATLAS in executing the R&D work, U.S. technical activities take place under the ATLAS technical reporting lines of each topical area. The following section concludes with the overall HL-LHC computing planning at the WLCG and ATLAS levels.

Organization and planning of U.S. ATLAS R&D

U.S. ATLAS plays leading roles in many ATLAS S&C areas including Core Software, Distributed Computing and Operations; runs the largest Tier 1 and Tier 2 centers in the ATLAS computing grid; and has pioneered the usage of HPC systems for HEP production. U.S. ATLAS has for many years been a leader within the ATLAS Collaboration in researching, designing and developing innovative software and computing approaches and systems that can both deliver ATLAS production needs in the near term and ultimately scale to meet the challenges of HL-LHC computing. Scaling to HL-LHC demands extensive R&D, and U.S. ATLAS kicked off its R&D program by contributing strongly to the HSF community roadmap in 2017.

U.S. ATLAS formalized its R&D effort towards HL-LHC in June 2018 with the creation of the HL-LHC Computing activity at Level 2 within the U.S. ATLAS Operations Program. The activity is charged with managing the U.S. research and development program towards affordably delivering ATLAS computing requirements in the LHC era. While addressing particularly the U.S. program, it is structured as part of a coherent ATLAS wide and in many respects LHC community wide effort. In the ensuing two years U.S. ATLAS has been a driving force as HL-LHC computing planning has ramped up in ATLAS, WLCG, HSF and the HL-LHC directed U.S. software projects IRIS-HEP and HEP-CCE. The influence of U.S. ATLAS in shaping the HL-LHC computing plan and R&D program is reflected in the coherence of our program with those of ATLAS overall and the collaborating projects. U.S. ATLAS holds leadership positions in IRIS-HEP and HEP-CCE, and several US ATLAS L3 managers are actively involved in both projects.

The US R&D effort is organized in three activity areas selected for their impact on U.S. ATLAS analysis, their alignment with NSF and DOE priorities (in particular for new platforms, most prominently HPCs), and their alignment with US expertise and long-term commitments to ATLAS Computing and Software. The activity areas are (1) ***software re-engineering and algorithm development***, (2) ***workflow porting to new platforms*** and (3) ***distributed computing***. The full WBS to level 5 includes almost 40 sub-activities and currently engages the (fractional) effort of about 50 people. The organization chart to level 4 is described in Appendix 2. The carefully planned program over the next 3 years includes about 20 specific deliverables and milestones. The 7 principal R&D areas selected since the activity was created two years ago are all on track for adoption and production in Run-3 (Appendix 4), constituting a practical validation of our activity: we are on track. All these R&D areas have deliverables included in our milestones.

Here follows an overview of activities, plans and milestones across our R&D activity areas. The effort invested specifically in the R&D program (WBS 2.4) -- which strongly leverages the personnel and expertise of other operations program areas -- is summarized in Appendix 3.

Software re-engineering and algorithm development

Objective: Prepare for efficient use of next generation heterogeneous architectures across the full scope of the ATLAS offline software.

Event generation

Objective: Help catalyze the generator developer community to deliver next-gen LCF capable generators that are key to ATLAS, before Run-4.

This is planned as a small activity. Workplan and milestones will be established shortly now that a dedicated person has been hired. We will also draw on HEP-CCE and SciDAC efforts. Support from DOE, NSF, HEP-CCE.

Simulation

Objective: Commission fast simulation techniques to help solve the HL-LHC CPU problem, leveraging accelerators and HPCs.

The ATLAS fast simulation is an order of magnitude faster than the full Geant4 simulation, and the fast chain (including reconstruction) is two orders of magnitude faster than the full chain. The challenges here are in commissioning these approaches, particularly on HPCs and leveraging accelerators. It is a major activity, ~2 FTEs. Support from DOE, NSF, HEP-CCE.

Milestones

- Oct 2019: FastCaloSim first version working on LCF. *Completed.*
- Jul 2021: FastChain first version on LCF. *Delayed from Jan 2020 due to LCF complexity.*

Reconstruction

Objective: Contribute to making the ACTS tracking package the ATLAS Run-4 fast parallelized tracking software that meets performance requirements and leverages accelerators.

This is a new activity in FY21 catalyzed by IRIS-HEP involvement in ACTS as a cross-experiment common tracking software. Will reach ~1.5 FTEs in FY21. Support from NSF, IRIS-HEP, HEP-CCE.

Milestones

- Sep 2021: Demonstrate the physics and technical performance of the ACTS reconstruction algorithms for ATLAS.

Analysis

Objective: Leverage primarily external activity, common projects, common software, and workflow management strengths to contribute to a state of the art HL-LHC analysis software suite and toolkit.

No dedicated FTEs or milestones in current planning. Benefits from other activity such as parallel fitter development by IRIS-HEP and the workflow management effort (milestone below).

Framework and services

Objective: Make leading ATLAS contributions to core software support for concurrency, accelerators and ML, directed particularly at utilizing accelerators and LCFs in Run-4.

This is a major activity to support concurrency, accelerators, and ML in the framework. It includes leveraging the Ray common software package to develop a second-generation event service system (Raythena) and intra-node scheduling system for HPCs, establishing how to develop portable parallelized software, integrating the capability to use GPUs and ML within the Athena framework, and high concurrency upgrades to the common software Gaudi framework underlying Athena. About 3.5 FTEs. Support from DOE, HEP-CCE.

Milestones

- Jan 2021: Deploy Raythena for running Event Service simulation production.
- Aug 2021: Evaluation of GPU management techniques and infrastructure, with integration into Athena.
- Sep 2021: Develop a prototype of the HPX-based task scheduler in Gaudi.
- Oct 2021: Develop a prototype Ray-based mechanism for intra-node scheduling of Athena reconstruction components on heterogeneous platforms.
- Mar 2022: Deliver a portable parallelization package recommendation to ATLAS software for migration to heterogeneous architectures.

Event I/O and persistency

Objective: Make leading ATLAS contributions to event persistency for Run-4 supporting all use cases and particularly HPCs, accelerators and fine grained workflows.

This is a major activity to scale up I/O and persistency services to highly parallel HPCs, support storage-efficient fine grained workflows, and investigate modern storage and data compression technologies. About 2 FTEs. Support from DOE, HEP-CCE.

Milestones

- Oct 2021: Complete I/O framework capable of handling EventService Simulation in AthenaMT mode, with correct handling of in-file metadata.
- Mar 2022: Prototype implementation of non-ROOT parallel event storage backend for fine-grained processing. First R&D target is HDF5.
- Oct 2022: Implement I/O roadmap recommendations to improve metadata support for fine-grained workflows.

Workflow porting to new platforms

Objective: Porting workflows to new HPCs, and porting new workflows to HPCs. Particularly directed at LCFs, which are the machines that take substantial case by case porting effort.

At the present time the effort level on this activity is low, because we have no workloads that the GPU based LFCs can run. The effort will ramp up as HPC/LCF utilization ramps up once more when we have compatible workloads. The first to arrive will be ML services for hyperparameter optimization (HPO) and training, which will bring a small effort increase in FY21. ML services porting to LCFs has suffered delays because of LCF complexity and lack of expertise. *The effort would be greatly aided by LCF expert assistance.*

Milestones will be established in this area once our HPC/LCF utilization plan is clearer.

Distributed computing

Objective: Develop the distributed software required to deliver HL-LHC storage and processing needs within available budgets.

Data management

Objective: Leading contributions to data management innovations to solve the HL-LHC storage challenge, particularly through dataflow/workflow orchestration and fine grained dataflows.

This is a major activity in cold/hot storage optimization, distributed fine grained dataflows, dataflow/workflow orchestration, intelligent data distribution, and caching. This activity works in close concert with workflow management, below. All the workflow management milestones involving iDDS are joint milestones with data management. About 3 FTEs. Support from DOE, NSF, IRIS-HEP.

Milestones

- Jul 2021: Complete integration of the new Sci-token identity management system into XRootD.
- Dec 2021: iDDS - Data Carousel integration complete for Run-3 production, with iDDS based fine grained orchestration of JEDI, PanDA, and Rucio to minimize task execution time tails.

Workflow management

Objective: Leading contributions to workflow and workload management to solve the HL-LHC CPU challenge, leveraging all possible resources and architectures, and supporting all ATLAS workflows on appropriate platforms.

This is a major activity in workflow and workload management, fine grained workflows, intelligent brokering and scheduling, new workflows such as ML services and active learning, and HPC workflows. About 2.5 FTEs. Support from DOE, NSF, IRIS-HEP.

Milestones

- Dec 2019: PanDA based distributed training and hyperparameter optimization service operating on grid. *Completed.*
- Sep 2020: PanDA based distributed training and hyperparameter optimization service operating on an LCF HPC. *Delayed beyond Sep 2020 due to LCF complexity. Would be aided by direct OLCF/ALCF expert assistance.*
- Feb 2021: PanDA/iDDS based ML hyperparameter optimization and training services in production, including user client tools and centralized visualization of training results.
- Jun 2021: PanDA/iDDS support for directed acyclic graph (DAG) based workflows to support complex diverse experiment-agnostic workflows including active learning.
- Sep 2021: PanDA/iDDS Ray integration supporting checkpoint-capable ML hyperparameter optimization and training on HPCs and other large preemptible resources.
- Oct 2021: PanDA/iDDS based function-as-a-service platform allowing users to run fitting and active learning workflows on remote GPU, TPU, and HPC resources. Effectively a joint milestone with the analysis activity area.

Common infrastructure

Objective: Deliver the information services, performance metrics, monitoring tools and containerization infrastructure needed to support the other activities.

This is a small activity to maximize the commonality of infrastructure used across all platforms. About 0.75 FTEs. Support from DOE.

Milestones

- Jan 2021: Develop an automated procedure and software tools for ATLAS software release containerization and validation.

The technical milestones are summarized below, together with planning milestones discussed in the next section.

HL-LHC computing plan

Within the framework of a WLCG driven planning process, ATLAS recently completed a conceptual design report (CDR) for HL-LHC computing, a 41 page document describing (as has been summarized here) the computing challenges, resource estimations and R&D program to meet the challenge. Befitting a CDR, the technical content, planning and milestones are at a relatively high level. The CDR is the first ATLAS input to a newly established HL-LHC computing review panel and process established by the LHCC (and was well received). The next step will be an assessment of U.S. ATLAS R&D activities in 2021, with all R&D activities delivering a report on their achievements and the steps leading to the 2022 demonstrator, to be assessed and ranked by priority and risk. The major step of a technical design report (TDR) will come in 2023; it will go into more detail informed by the R&D program. Like the CDR, the TDR will be developed in close concert with WLCG and CMS, with the oversight of the LHCC.

As described here, the (U.S.) ATLAS R&D activities are well underway, with an established program of milestones and deliverables leading up to the TDR. In many cases the milestones include production level deployments of new software, i.e. not only research, but development tested in the real world. A number of milestones target readiness for Run-3 (2022). These include developments in multithreaded Athena, the new storage-efficient DAOD formats, an improved fast calorimeter simulation, the first deliverables of ACTS tracking, the data carousel, and ML services supporting GPU and HPC utilization. R&D and milestone performance in advance of the TDR will allow for well informed decisions in the TDR on what development paths should be carried forward. Our goal is to complete most conservative R&D projects before the TDR.

Aggressive R&D projects have a significant risk of failure, and their schedule, performance, and budget impact is hard to quantify until they are close to completion. Many aggressive R&D activities are collaborations between ATLAS and other projects like HEP-CCE and IRIS-HEP. U.S. ATLAS has leadership roles in these projects, but their schedule and priorities are not under direct U.S. ATLAS control. Also, successful aggressive R&D projects adapt their scope and schedule to their findings. For these reasons our R&D managers, who closely monitor technical progress on R&D projects through detailed biweekly reports from U.S. ATLAS contributors, have asked U.S. ATLAS management to allow for some flexibility with the schedule and scope of their technical milestones prior to the TDR.

Following the TDR, the goal is to develop in 2023-2024 fully functional versions of all R&D projects endorsed by the TDR as high priority. As noted above, the Run-4 directed R&D program contains a substantial development component throughout, reflected in production deliverables and milestones. From 2025 the focus will shift fully to development, guided by a specific 2025-2027 development program towards HL-LHC readiness. A lower priority, longer term R&D program will continue towards Run-5 directed objectives.

High level milestones for the ATLAS HL-LHC software and computing program are as follows.

- 2020: **CDR released.** Identifies HL-LHC R&D needs, and the projects attempting to address them. *Completed.*
- 2020: **LHCC review.** First appraisal of HL-LHC computing plans by the LHCC review panel. *Completed.*
- 2021: **Run-3 deployment.** All Run-3 directed milestones delivered, including some conservative R&D items.
- 2021: **U.S. ATLAS R&D assessment.** All R&D activities deliver a report on their achievements and the steps leading to the 2022 demonstrator. Plans are assessed and ranked by priority and risk.
- 2022: **Run-3 starts.** R&D focus shifts fully to HL-LHC R&D.

- 2022: **Run-4 Proof of concept demonstrators delivered** by R&D projects targeting Run-4.
- 2023: **TDR released.** Prioritizes successful R&D projects for development aimed at Run-4. Endorses a limited amount of Aggressive R&D projects for Phase 3
- 2024: **Run-4 projects approval.** Go/no-go decision for R&D projects targeting Run-4. Every project must provide functionally complete prototypes.
- 2025: **Run-4 development planning.** Each Run-4 project provides a plan to complete development in time for commissioning prior to Run-4, with a WBS inclusive of effort and risk estimates. It also provides training and documentation tailored for the target developers community.
- 2025: **First Run-4 deployment.** Includes production quality implementations of all new developments to be included in Run-4. Not all physics code migrated yet.
- 2026: **Second Run-4 deployment.** Includes production versions of all Run-4 code and integration. Validation starts.
- 2026: **Run-4 dress rehearsal.** Test software and distributed computing readiness for datataking.
- 2027: **Run-4 starts.** R&D focus shifts to Run-5.

Several of these milestones are included in US milestones below, to explicitly couple US planning to the overall plan.

U.S. ATLAS HL-LHC S&C R&D milestone summary

Here follows a time ordered summary of U.S. ATLAS HL-LHC S&C R&D milestones Technical milestones constitute the objective metrics by which R&D project progress is assessed. As an R&D project approaches the decision point on whether it will transition to an approved development project, we will add a technical milestone expressing the metrics that will inform the decision.

Date	**Milestone**	**Activity**	**Coll.**
Oct 2019	FastCaloSim first version working on LCF. **Completed.**	Simulation	
Dec 2019	PanDA based distributed training and hyperparameter optimization service operating on grid. **Completed.**	Data & workflow management	
Sep 2020	PanDA based distributed training and hyperparameter optimization service operating on an LCF HPC. **Delayed** beyond Sep 2020 due to LCF complexity. Would be aided by direct OLCF/ALCF expert assistance.	Data & workflow management	
Jan 2021	Deploy Raythena for running Event Service simulation production.	Core software	
Jan 2021	Develop an automated procedure and software tools for ATLAS software release containerization and validation.	Software infrastructure	
Feb 2021	PanDA/iDDS based ML hyperparameter optimization and training services in production, including user client tools and centralized visualization of training results.	Data & workflow management	
Jun 2021	PanDA/iDDS support for directed acyclic graph (DAG) based workflows to support complex diverse experiment-agnostic workflows including active learning.	Data & workflow management	
Jul 2021	FastChain first version on LCF. Delayed from Jan 2020 due to LCF complexity.	Simulation	
Jul 2021	Complete integration of the new Sci-token identity management system into XRootD.	Data management	IRIS-HEP
Aug 2021	Evaluation of GPU resource management techniques and infrastructure, with integration into Athena.	Core software	

Sep 2021	Demonstrate the physics and technical performance of the ACTS algorithms for ATLAS.	Reconstruction	IRIS-HEP
Sep 2021	Develop a prototype of the HPX-based task scheduler in Gaudi.	Core software	
Sep 2021	PanDA/iDDS Ray integration supporting checkpoint-capable ML hyperparameter optimization and training on HPCs and other large preemptible resources.	Data & workflow management	
Oct 2021	Develop a prototype Ray-based mechanism for intra-node scheduling of Athena reconstruction components on heterogeneous platforms.	Core software	
Oct 2021	PanDA/iDDS based function-as-a-service platform allowing users to run fitting and active learning workflows on remote GPU, TPU, and HPC resources. Effectively a joint milestone with the analysis activity area.	Data & workflow management, analysis	IRIS-HEP
Oct 2021	Complete I/O framework capable of handling EventService Simulation in AthenaMT mode, with correct handling of in-file metadata.	Event I/O	
Dec 2021	iDDS - Data Carousel integration complete for Run-3 production, with iDDS based fine grained orchestration of JEDI, PanDA, and Rucio to minimize task execution time tails.	Data management	
Dec 2021	U.S. ATLAS R&D assessment. All R&D activities deliver a report on their achievements and the steps leading to the 2022 demonstrator. Plans are assessed and ranked by priority and risk.	Planning	
Mar 2022	Deliver a portable parallelization package recommendation to ATLAS software for migration to heterogeneous architectures.	Core software	HEP-CCE
Mar 2022	Prototype implementation of non-ROOT parallel event storage backend for fine-grained processing. First R&D target is HDF5.	Event I/O	HEP-CCE
Oct 2022	Implement I/O roadmap recommendations to improve metadata support for fine-grained workflows.	Event I/O	
Dec 2022	Deliver Run-4 R&D proof of concept demonstrators for US R&D projects.	Planning	
Dec 2023	Technical design report completed.	Planning	
Jun 2024	Deliver functionally complete prototypes for Run-4 directed US development projects.	Planning	

Conclusion

U.S. ATLAS has a carefully planned and managed, successful and visionary program in HL-LHC S&C research and development. Our broad, mature program combines innovative research and prototyping with a practically directed development program that anchors long term R&D in near term production applications for validation. The R&D program is tightly embedded within and coherent with both the ATLAS-wide R&D effort and the operations work addressing the current S&C needs of U.S. ATLAS and ATLAS. It both contributes to and strongly leverages the IRIS-HEP, HEP-CCE and WLCG programs. It is scoped to focus on our longstanding areas of expertise, and is scaled to produce deliverables that are attainable with realistic effort levels and that represent an appropriate fraction of the overall ATLAS effort, while delivering many critical components of the overall ATLAS program. Our program is built on our past successes: the U.S. ATLAS core and distributed software efforts have had an evolutionary history free of revolutions, the legacy of which is a stable, powerful software foundation upon which our developers are efficiently building the software to meet the needs of the HL-LHC within realistic budgets.

Supplemental Material

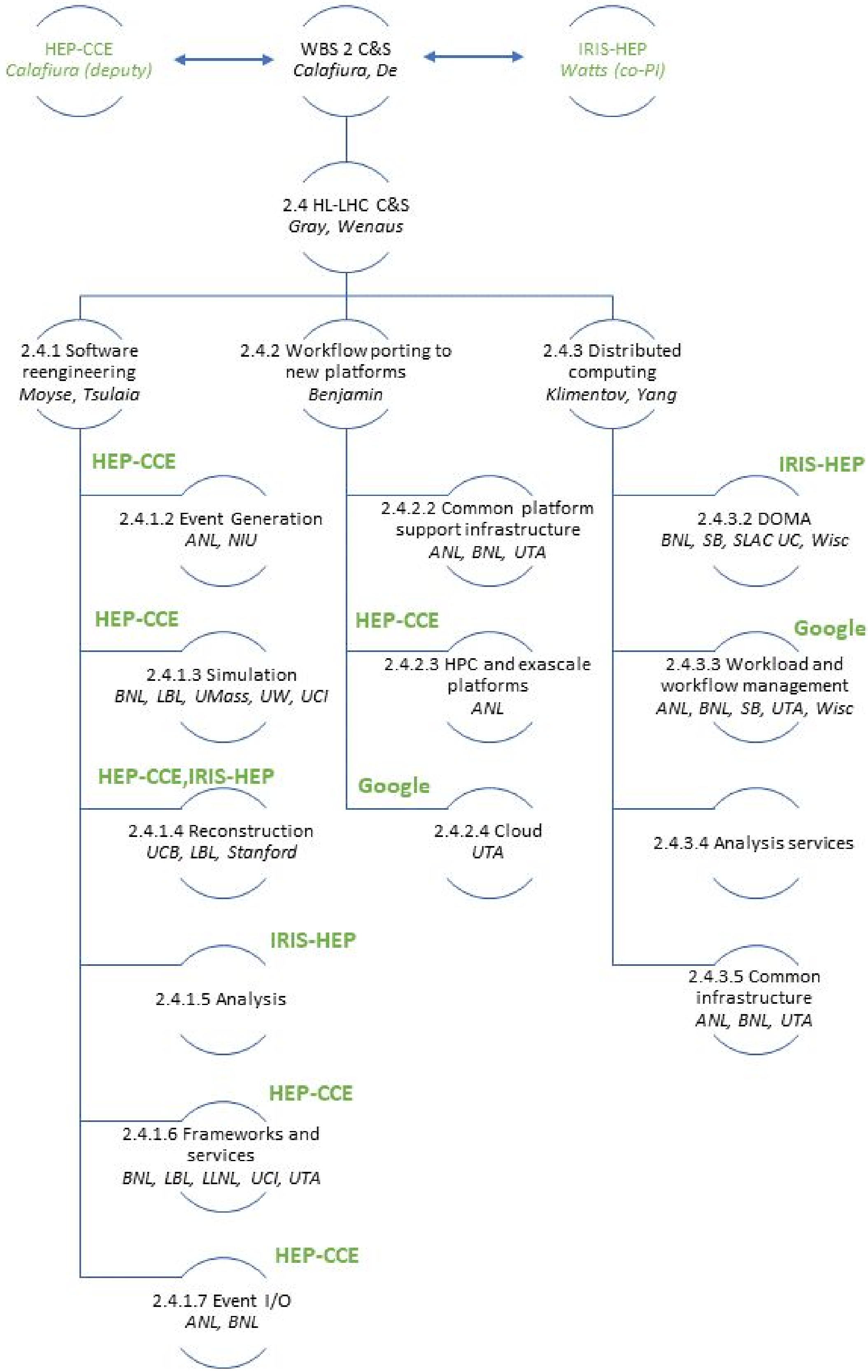
References

1. [AthenaMP] Harnessing multicores: Strategies and implementations in ATLAS, S. Binet, (Orsay, LAL), P. Calafiura, (LBL, UCB), S. Snyder, (Brookhaven), W. Wiedenmann, (Wisconsin U., Madison), F. Winklmeier, (CERN), 2010. J. Phys. Conf. Ser. 219:042002, 2010.
2. [PPS] Portable Parallelization Strategies in the Era of Heterogeneous Architectures, HSF WLCG Virtual Workshop on New Architectures, Portability, and Sustainability. May 2020. https://indico.cern.ch/event/908146/contributions/3826737
3. [Ray] https://rise.cs.berkeley.edu/projects/ray/
4. [acts] Acts Common Tracking Software https://github.com/acts-project/acts
5. [exatrkx] N Choma, D Murnane et al, Track Seeding and Labelling with Embedded-space Graph Neural Networks, 2020, arXiv:2007.00149
6. [iDDS] intelligent Data Delivery Service. https://idds.cern.ch.
7. [ServiceX] IRIS-HEP Service X https://iris-hep.org/projects/servicex.html
8. [RECAST] The RECAST framework. https://iris-hep.org/projects/recast.html
9. [ProdSys] F H Barreiro et al. "The ATLAS Production System Evolution: New Data Processing andAnalysis Paradigm for the LHC Run2 and High-Luminosity". In: Journal of Physics: ConferenceSeries898 (Oct. 2017), p. 052016. doi:10.1088/1742-6596/898/5/052016.
10. [PanDA] K De et al, The future of PanDA in ATLAS distributed computing, 2015 J. Phys.: Conf. Ser.664 062035
11. [Rucio] Martin Barisits et al. "Rucio: Scientific Data Management". In: Computing and Software forBig Science3.1 (Aug. 2019), p. 11. doi:10.1007/s41781-019-0026-3. https://doi.org/10.1007/s41781-019-0026-3.

Appendix 1 - Acronyms

- **ACTS** ATLAS tracking software for Run-3 and Run-4, and a cross-experiment common project
- **ADC** ATLAS Distributed Computing
- **AMSG3** Analysis Model Study Group for Run-3
- **AOD** Analysis Object Data
- **CDR** Conceptual Design Report
- **CP** Combined Performance
- **CUDA** Nvidia Compute Unified Device Architecture
- **DAOD** Derived AOD
- **DESD** Derived ESD
- **DCS** Detector Control System
- **DOMA** Data Organisation and Management
- **DQM** Data Quality Monitoring
- **DPC++** Intel Data Parallel C++
- **EDM** Event Data Model
- **ESD** Event Summary Data
- **HLT** High-Level Trigger
- **HPC** High-Performance Computing
- **HPO** Machine learning hyperparameter optimization
- **HS06** HEP-SPEC06
- **HSF** HEP Software Foundation
- **iDDS** Intelligent Data Delivery System
- **ITk** ATLAS Inner Tracking detector
- **LCF** DOE Leadership Computing Facility
- **NLO** Next to Leading Order
- **NNLO** Next to Next to Leading Order
- **PEB** Partial Event Building
- **RDO** Detector digitization format
- **QoS** Quality of Service
- **TBB** Intel Thread Building Blocks
- **TDAQ** Trigger and Data Acquisition
- **TDR** Technical Design Report
- **TLA** Trigger-object Level Analysis
- **WBS** Work Breakdown Structure
- **WLCG** Worldwide LHC Computing Grid

Appendix 2 - U.S. ATLAS WBS 2.4 organization chart



Appendix 3 - U.S. ATLAS HL-LHC S&C R&D effort

Other: non-ops FTEs contributing to WBS activities supported by IRIS-HEP, HEP-CCE

	FY19 Ops	FY20 Ops plan	FY20 Ops actual	FY21 Ops	FY22 Ops	FY19 Other	FY20 Other	FY21 Other	FY22 Other	Comment
2.4.1 - Software engineering and algorithm development	2.65	6.45	5.65	6.95	6.95	0.1	1.3	2.2	2.2	FY20 ops actual < plan because of missing 50/50 hires (UIC visa problem, LLNL hire still pending)
2.4.2 - Workflow porting to new platforms	0.48	1.0	0.7	0.8	0.8	0.45	0	0.5	0.5	FY20 actual ~ plan. More common infrastructure work (performance metrics, containerization), less HPC-directed dev (ML services dev moved to grid). FY20 Other is zero because of reassignment of non US ATLAS effort to other priorities. Activity restored with HEP-CCE support in FY21
2.4.3 - Distributed computing development	2.6	2.95	4.05 (+ 0.6 RBT)	5.55	5.35	0.3	1.9	2.2	2.2	FY20 actual > plan mainly because of rapid rise of new workflows (data carousel, iDDS, HPO)
Total	5.73	10.4	10.4	13.3	13.1	0.85	3.2	4.9	4.9	

FTEs engaged in the U.S. ATLAS HL-LHC S&C R&D effort (WBS 2.4). The 'Other' category includes contributions from outside the operations program such as IRIS-HEP and HEP-CCE.

Appendix 4 - U.S. ATLAS HL-LHC S&C R&D status

Activity	Status	Milestones
ML training and hyperparameter optimization (HPO) services for GPUs and LFCs	Steady rise in ATLAS. Taken up by ATLAS ML Forum. Encouraged by Computing Coordination.	Grid deployment: met and meets the need. LCF deployment: late but not needed at the present usage scale.
Fast simu/Fast chain for GPUs and LFCs	Taken up by ATLAS simu, core software, BNL CSI, HEP-CCE as prime use case for GPU porting	Fast simulation on LCF demonstrated. Fast chain work is coupled to the long ATLAS development/validation timeline.
Fast tracking	ACTS taken up by IRIS-HEP and mainstreamed in ATLAS	Demonstrate ACTS meets ATLAS requirements, Sep 2021.
Data carousel in production for Run-3	Already demonstrated in production, accepted as production workflow	Fully deployed for Run-3 production, on target.
Intelligent data delivery system (iDDS)	Already a production component of the data carousel, several additional use cases in progress	First production use case delivered (data carousel) Feb 2020.
Ray based next-gen event service on HPCs	Successful prototyping, now the mainstream replacement for Yoda	In production Jan 2021, extend to Ray based intra-node HPC reconstruction Oct 2021.
Generalized iDDS based workflows: ML HPO, active learning, DAG based workflows, parallel fitting	Enthusiastic interest from ATLAS ML and analysis communities. Potential beyond-ATLAS DAG apps (important given the experiment-agnostic IRIS-HEP collaboration)	Four 2021 milestones established.

U.S. ATLAS HL-LHC S&C R&D status summary. Every activity undertaken since the program was created in 2018 is on track for adoption and production, and is reflected in our milestones.