ABSTRACT

OPTIMIZATION OF DERIVATION JOBS AND MODERNIZATION OF I/O INTEGRATION TESTS FOR THE ATLAS EXPERIMENT

Arthur C. Kraus, M.S.
Department of Physics
Northern Illinois University, 2025
Dr. Jahred Adelman, Director

The High-Luminosity LHC (HL-LHC) is a phase of the LHC that is expected to start toward the end of the decade. With this comes an increase in data taken per year that current software and computing infrastructure, including I/O, is being prepared to handle. The ATLAS experiment's Software Performance Optimization Team has areas in development to improve the Athena software framework that is scalable in performance and ready for widespread HL-LHC era data taking. One area of interest is optimization of derivation production jobs by improving derived object data stored to disk by about 4-5% by eliminating the upper-limit on TTree basket buffers, at the expense of an increase in memory usage by about 11%.

Athena and the software it depends on are updated frequently, and to synthesize changes cohesively there are scripts, unit tests, that run which test core I/O functionality. This thesis upgrades existing I/O unit tests to now exercise features exclusive to the xAOD Event Data Model (EDM) such as writing and reading object data from the previous EDM using transient and persistent data. These new unit tests also include and omit select dynamic attributes to object data during the component accumulator step.

NORTHERN ILLINOIS UNIVERSITY
DE KALB, ILLINOIS

MAY 2025

OPTIMIZATION OF DERIVATION JOBS AND MODERNIZATION OF I/O INTEGRATION TESTS FOR THE ATLAS EXPERIMENT

BY

ARTHUR C. KRAUS

A THESIS SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE
MASTER OF SCIENCE

DEPARTMENT OF PHYSICS

Thesis Director:
Dr. Jahred Adelman

ACKNOWLEDGEMENTS

Here's where you acknowledge folks who helped. Here's where you acknowledge folks who helped. Here's where you acknowledge folks who helped. Here's where you acknowledge folks who helped.

DEDICATION

26

To all of the fluffy kitties. To all of the fluffy kitties. To all of the fluffy kitties. To all of the fluffy kitties.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

Particle physics is the branch of physics that studies the fundamental constituents of matter and the forces governing their interactions. The field started as studies in electromagnetism, radiation, and further developed with the discovery of the electron. What followed was more experiments to search for new particles, new models to describe the results, and new search techniques which demanded more data. The balance in resources for an experiment bottlenecks how much data can be taken, so steps need to be taken to identify interesting interactions and optimize the storage and processing of this data. This thesis investigates software performance optimization of the ATLAS experiment at CERN. Specifically, ways to modernize and optimize areas of the software framework, Athena, to improve input/output (I/O) performance during derivation production and create new tests that catch when specific core I/O functionality is broken.

1.1 LHC and The ATLAS Detector

The Large Hadron Collider (LHC), shown in Figure 1.1, is a particle accelerator spanning a 26.7-kilometer ring that crosses between the France-Switzerland border at a depth between 50 and 175 meters underground.[2] The ATLAS experiment, shown in Figure 1.3, is the largest LHC general purpose detector, and the largest detector ever made for particle collision experiments. The detector lies in a cavern 92.5 m underground at a length of 46 m, height and width of 25 m.[3] A quadrant of the detector is shown in Figure 1.2, where η is a measure

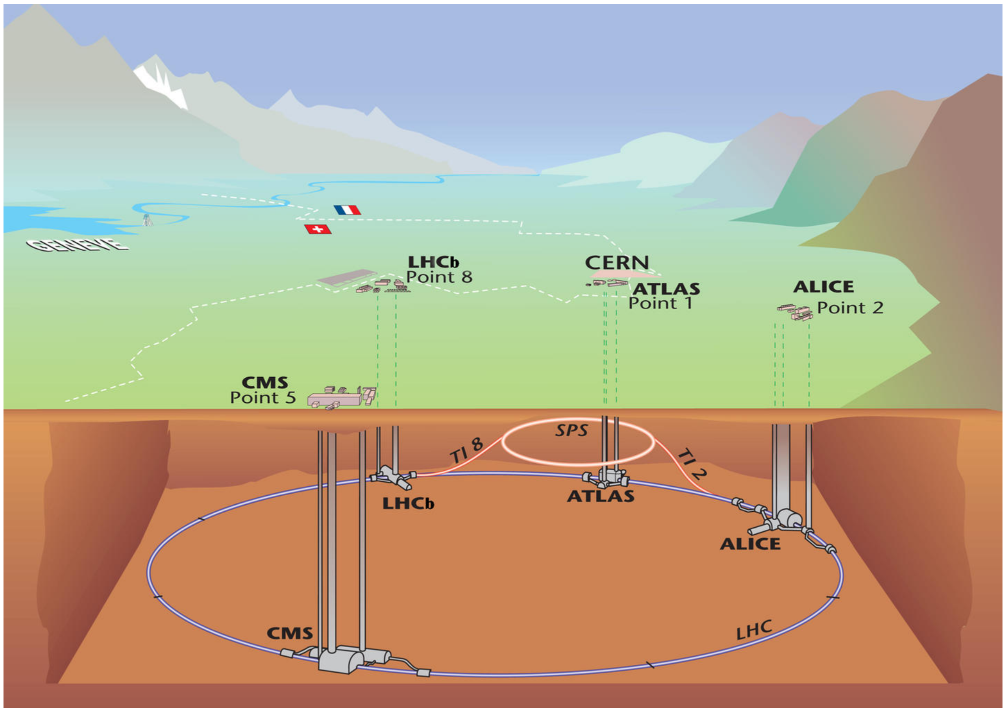

Figure 1.1: Illustration of the LHC experiment sites on the France-Switzerland border.[1]

of the pseudo-rapidity. Pseudo-rapidity is a parameter representing the the angle relative to the beamline and is defined as

$$\eta \equiv -\ln\left[\tan\left(\frac{\theta}{2}\right)\right], \tag{1.1}$$

where if $\theta = 0$ then $\eta = \infty$ and if $\theta = \frac{\pi}{2}$ then $\eta = 0$. Pseudo-rapidity is used, as opposed to traditional Cartesian angles, because it's Lorentz invariant under boosts along the beam axis, making it easier to identify tracks due to symmetry of the collision.

Inner Detector

The ATLAS detector is comprised of three main sections, the inner detector, calorimeters and the muon detector system. The inner detector measures the direction, momentum and charge of electrically charged particles. Its main function is to measure the track of the charged particles without destroying the particle itself. The first point of contact for particles emerging from pp-collisions from the center of the ATLAS detector is the pixel detector.[6]

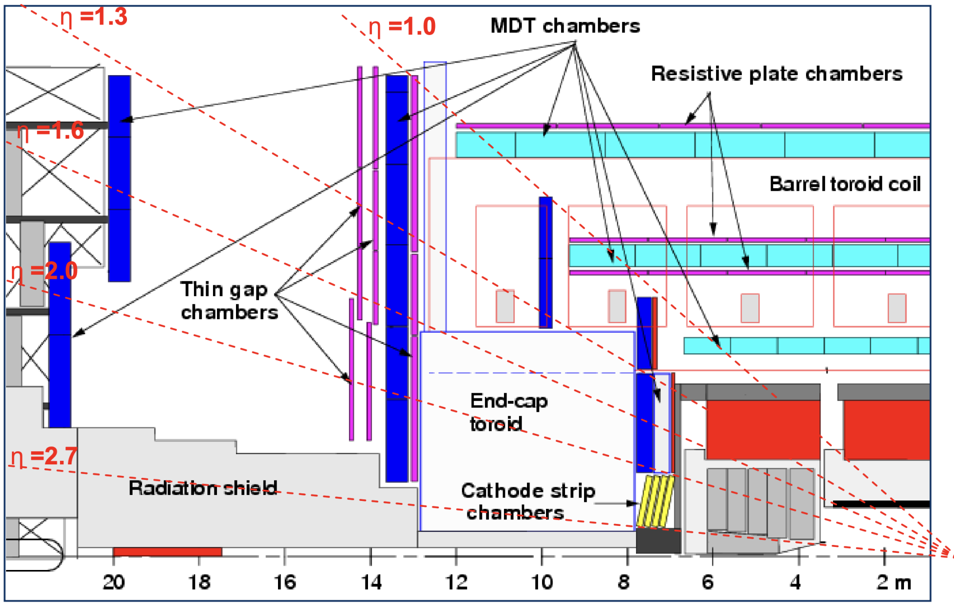

Figure 1.2: One quadrant of the ATLAS detector. The components of the Muon Spectrometer are labelled [4]

It has over 92 million pixels to aid in particle track and vertex reconstruction. Since the pixels are the first point of contact to the incident particles they have to be radiation hard so the electronics may function without fault. When a charged particle passes through a pixel sensor it ionizes the one-sided doped-silicon wafer to produce an excited electron will then occupy the conduction band of the semiconductor producing an electron-hole pair, leaving the valence band empty.[7] This hole in the valence band together with the excited electron in the conduction band is called an electron-hole pair. The electron-hole pair is in the presence of an electric field, which will induce drifting of the electron-hole pair, drifting that will generate the electric current to be measured.

Surrounding the pixel detector is the SemiConductor Tracker (SCT), which uses 4,088 modules of 6 million implanted silicon readout strips.[8] Both the pixel detector and SCT measure the path particles take, called tracks. While the pixel detector has measurement

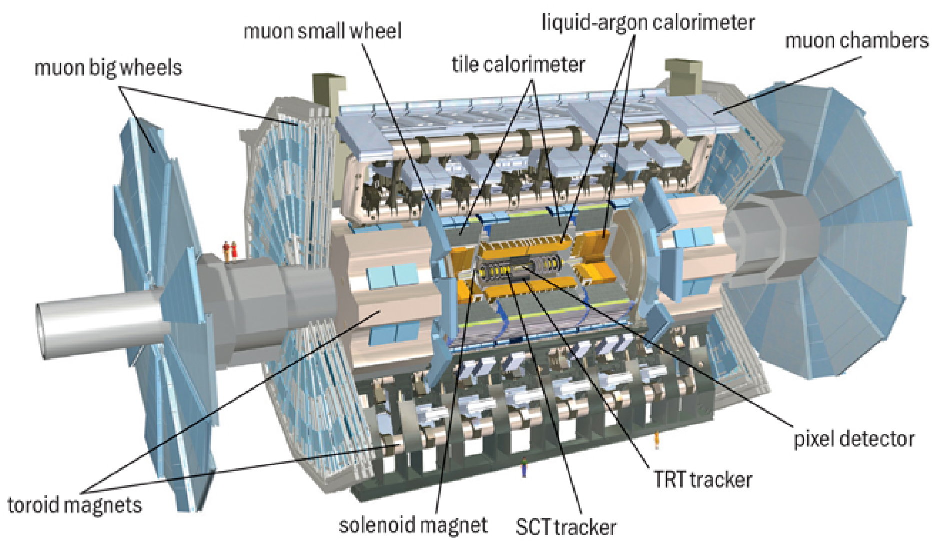

Figure 1.3: Overview of the ATLAS detectors main components, with two people in figure to scale.[5]

precision up to $10\mu m$ in the $r\phi$-direction and $70\mu m$ in the z-coordinate direction,[9] the SCT has resolution $17\mu m$ in the $r\phi$-direction and $580\mu m$ in the z-direction.

The final layer of the inner detector is the transition radiation tracker (TRT). The TRT is made of a collection of tubes made with many layers of different materials with varying indices of refraction. The TRT's straw walls are made of two $35\mu m$ layers comprised of $6\mu m$ carbon-polymide, $0.20\mu m$ aluminum, and a $25\mu m$ Kapton film reflected back.[10] The straws are filled with a gas mixture of $70\%\mathrm{Xe} + 27\%\mathrm{CO}_2 + 3\%\mathrm{O}_2$. Its measurement precision is around $170\mu m$. Particles with relativistic velocities have higher Lorentz γ-factors,

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{1.2}$$

The TRT uses varying materials to discriminate between heavier particles, which have low γ and radiate less, and lighter particles, which have higher γ and radiate more.[11]

Calorimeters

There are two main calorimeters for ATLAS, the Liquid Argon (LAr) calorimeter and the Tile Hadronic calorimeter. The LAr calorimeter surrounds the inner detector and measures the energy deposits of objects that interact via the electromagnetic force. It layers various metals to intercept the incoming particles to produce a shower of lower energy particles. The lower energy particles then ionize the liquid argon that fill the barrier in between the metal layers to produce a current that can be read out. The Tile calorimeter surrounds the LAr calorimeter and is the largest part of the ATLAS detector weighing in around 2900 tons. Particles then traverse through the layers of steel and plastic scintillating tiles. The Tile calorimeter is a hadronic calorimeter, so it interacts with particles via the strong nuclear force. When a particle hits the steel, a cascade of secondary protons, neutrons and other hadrons (quark bound states, with baryons qqq and mesons $q\bar{q}$) is produced with lower energy. Through this mechanism, these decay products will continue until the energy has entirely dissipated.

Muon Spectrometer (MS)

The MS sits at the end of the ATLAS detector and is designed to identify muon tracks and momentum to high-resolution, its components are shown in Figure 1.2. Monitored Drift Tube (MDT) chambers are used for precision measurement of muon tracks in the principle bending direction of the magnetic fields over a large η. The MDT lie in the endcaps and barrel regions covering the pseudorapidity regions $0 < |\eta| < 2.7$, where the the tubes run perpendicular to the beam and in-line with the magnetic field lines. Single cell resolution for these drift tubes can reach $60\mu m$.[3] The area of highest particle flux is the region of pseudo-rapidity $2 < |\eta| < 2.7$, here is where the cathode strip chambers lie.[12] Cathode strip chambers (CSCs) are layered to determine track vectors and use multi-wire chambers to achieve a resolution up to $50\mu m$.

The RPCs are gaseous parallel-plate detectors suited for fast spacetime particle tracking that combines the the spatial resolution (around 1 cm) of the wire chambers and the time resolution (around 1 ns) of a scintillation counter. Resistive plate chambers (RPCs) and the Thin gap chambers (TGCs) provide the trigger information for the MDTs and CSCs to then make a precision measurement, so speed takes priority over spatial resolution for the muon trigger system. Though RPCs don't have wires, their design consists of two strips separated by an insulating spacer to create a gap for the gas ($C_2H_2F_4$ plus some smaller of argon/butane) to occupy. Thin gap chambers (TGCs) exist in the forward region and are thin wire chambers that aide in muon triggering and measurement of the azimuthal coordinate to be used in compliment with MDTs. The time resolution in TCGs help identify bunch-crossings and granularity in momentum of the muon that comes within the equipotential of the wires. Since each wire can be given a position in the trigger system, any muon that passes through the TGC can be compared with greater spatial precision with the MDTs and illustrate a track later. The accuracy of identifying the correct bunch crossing with TGCs is 99% and the delivery of bunch crossing identification can be delivered within 25 ns, only a small fraction of bunch crossings arrive later than that window.

1.2 ATLAS Trigger/Data Acquisition (TDAQ)

The LHC produces *pp*-collisions at a rate of 40 MHz, each collision is an "event". More specifically, around 10^{11} protons are accelerated in one "bunch" with around 2800 bunches per proton beam, spaced around 25 ns apart from each other. Each beam is then concentrated to the width of $64\mu m$ at the interaction point where about 20 collisions happen at one bunch crossing. "Pile-up" is the result of multiple collisions occurring from one bunch crossing.

The ATLAS Trigger system is responsible for quickly deciding what events are interesting for physics analysis. The Trigger system is divided into the first- and second-level triggers and when a particle activates a trigger, the trigger makes a decision to tell the DAQ to save the data produced by the detector. The first-level trigger is a hardware trigger that decides, within $2.5\mu s$ after the event, if the event is good to put into a storage buffer for the second-level trigger. The second-level trigger is a software trigger that decides within $200\mu s$ and uses around 40,000 CPU-cores and analyses the event to decide if it is worth keeping. The second-level trigger selects about 1000 events per second to keep and store long-term.[13] The data taken by the TDAQ system is raw and not yet in a state that is ready for analysis, but it is ready for further processing.

The amount of data taken at ATLAS is substantial, seeing more than 3 PB of raw data each year and each individual event being around 2 MB.[14] All of the data produced by LHC experiments, especially ATLAS, has to be sent to the Worldwide LHC Computing Grid (WLCG).[15] The WLCG composes of a three-tiered system, CERN serves as the Tier-0 site, there are $\mathcal{O}(20)$ Tier-1 sites, and $\mathcal{O}(200)$ Tier-2 sites.[16] Though, the numbers of each site do change over time. The raw data coming from the TDAQ systems are recorded at the CERN Tier-0 sites where a first-pass at reconstruction will take place and a copy of the raw data is sent to the Tier-1 sites. Multiple 10 Gbps capacity links streamline dataflow from the ATLAS TDAQ to the Tier-0 site. Tier-1 sites offer manage permanent storage of raw and reconstructed data and provide extensive processing capability for analysis that might demand it. Tier-2 sites provide additional computation and storage services that compliment end-user analysis.

Athena manages ATLAS production workflows which are involved with simulation of data and event generation, track reconstruction from hits, and derivation production.[17] Figure 1.4 illustrates the broadstrokes of the entire ATLAS data processing chain for both real detector data and Monte Carlo (MC) simulations. MC simulation starts with the event generation (EVNT), following simulation of events hitting the detector (HITS) and further simulation of what would be read out of the detector (RDO). The reconstructed Analysis Object Data (AOD) are then processed through derivation production jobs that reduces AODs through several steps of skimming, thinning and slimming data and from $\mathcal{O}(1)$ MB per event to $\mathcal{O}(10)$ kB per event, creating Derived AOD (DAOD). An AOD contains converted detector signals into physics objects such as particle tracks, electron and muon candidates, primary vertices, and more.[18] Further discussion on the production of DAOD can be found in Section 2.3.

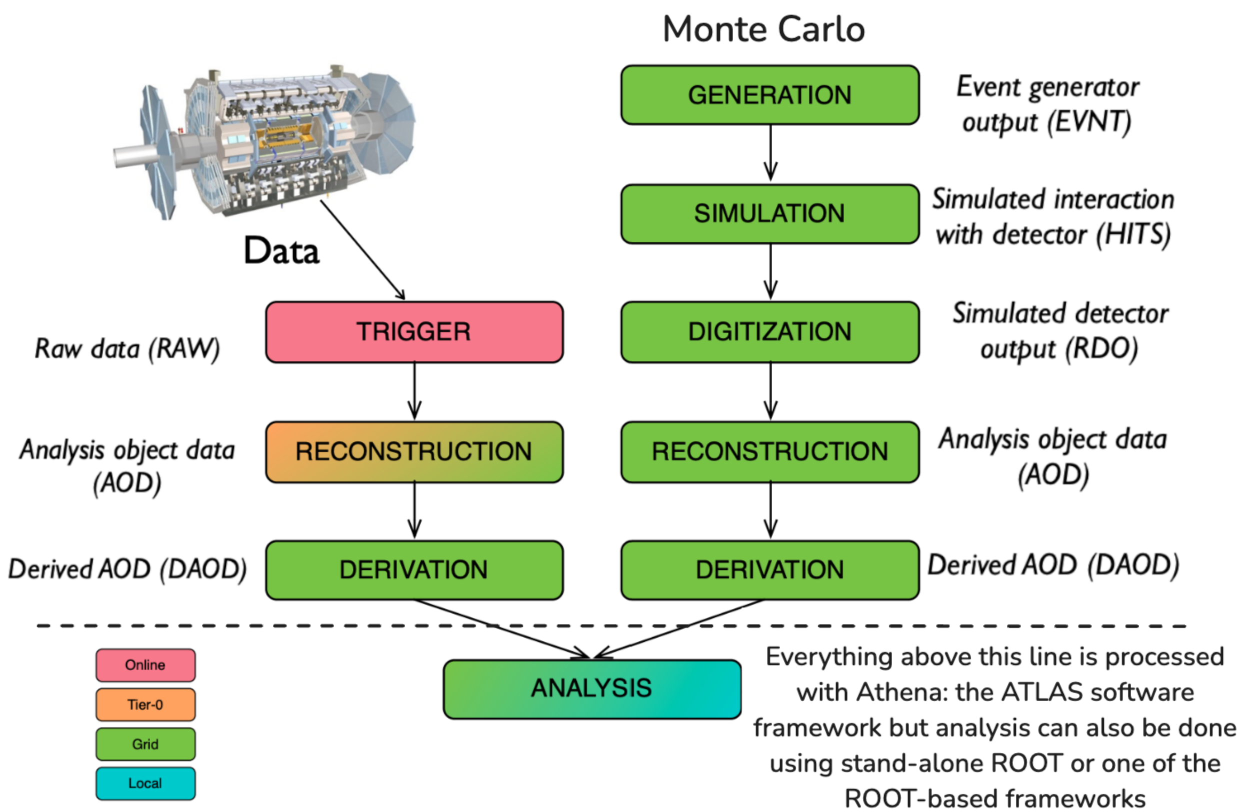

Figure 1.4: ATLAS data chain-processing for data and Monte Carlo simulation. Figure is modified from [19].

1.3 ATLAS Software and Computing Needs

The High-Luminosity LHC (HL-LHC) is the upgrade to LHC that anticipates more events and more data taken than ever before. The goal is to reach a luminosity of $350 fb^{-1}$, which is forecasted to be reached gradually by around 2040.[20] The HL-LHC era will start sooner than that, and has been projected to demand anywhere from 6-10 times data stored per year, so any attempt to save on disk storage should be investigated.[21] Increasing data means more resources from the Grid will be used, so optimization across files and software is an essential part of ensuring scalability of the data taken in by the detector. Figure 1.5 illustrates the projections of the HL-LHC era long-term storage usage for both disk and tape.

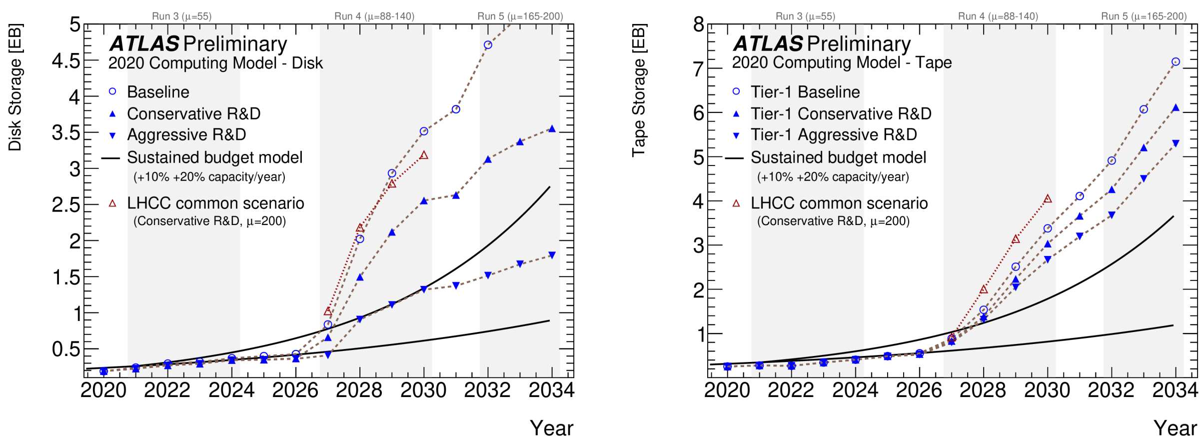

Figure 1.5: HL-LHC computing model projections on the future disk and tape usage compared to the expected budget increases.[22]

One avenue optimization is being investigated is in the method of storing data to file. The traditional method of storing event information for AOD/DAOD is with ROOT TTrees. ROOT TTrees (referred to as “TTrees” or “trees”) have been the standard data storage format for over two decades, and they provide a clear means of organizing and accessing physics objects for processing and analysis. The development of the ROOT N-Tuple (RN-

Tuple) I/O subsystem updates areas to support multi-thread processing, asynchronous I/O, object stores, and more. It's been shown to outperform the TTree I/O subsystem and other storage formats in file size (by about 15%), throughput, and compression, but still has more development before full implementation into the analysis pipeline.[23][24] While RNTuple is in development, there are still insights regarding resource usage optimization that are found by using TTree in its current state.

CHAPTER 2

I/O TOOLS

The Trigger/DAQ system sends and saves data from the detector to a persistent data storage solution. The data at this stage needs to be reconstructed and consolidated into physics objects, or Analysis Object Data (AOD) files. Creating AODs from data requires significant computation power and is undertaken by a software framework maintained by ATLAS, Athena. This chapter will cover important tools and concepts used by ATLAS to run derivation jobs, as well as introduce data structures that represent event information.

2.1 Event Data Models

An Event Data Model (EDM) is a collection of classes and their relationships to each other that provide a representation of an event detected with the goal of making it easier to use and manipulate by developers. An EDM is how particles and jets are represented in memory, stored to disk, and manipulated in analysis. It's useful to have an EDM because it brings a commonality to the code, aiding developers who reside in different groups often with various background experience. An EDM allows those developers to more easily debug and communicate issues when they arise.

2.1.1 Transient/Persistent (T/P) EDM

ATLAS used an EDM schema for Run-1 which had a separate transient and persistent status of the AOD. AODs would often be converted to an "ntuple" based format that allowed for fast readability and partial read for efficient analysis in ROOT, though it left the files disconnected from the reconstruction tools found in Athena.[25] When transient data was present in memory, it could have information attached to the object and gain in complexity the more it was used. Transient data needed to be simplified before it could become persistent into long-term storage (sent to disk). ROOT had trouble handling the complex inheritance models that would come up the more developers used this EDM. Before the successor to the T/P EDM was created, ATLAS physicists would convert data samples using the full EDM to a simpler one that would be directly readable by ROOT. This would lead to duplication of data and made it challenging to develop and maintain the analysis tools to be used on both the full EDM and the reduced ones. Additionally, converting from transient to persistent data was an excessive step which was eventually removed by the adoption of using an EDM that blends the two stages of data together, this was dubbed the xAOD EDM.

2.1.2 xAOD EDM

While the T/P EDM still remains functional in Athena, the xAOD EDM has been adopted as of Run 2. The xAOD EDM is an iteration to the T/P EDM and brings a variety of improvements.[26] This EDM, unlike T/P, is not strictly separated into transient or persistent data. Rather, xAOD primarily separates data into interface objects and its corresponding auxiliary data stores. The xAOD EDM has built-in functionality to add and remove dynamic attributes configured during job steering. These dynamic attributes to xAOD objects are called decorations.

The xAOD EDM use two types of objects that handle data, interface objects and payload objects. Interfaces act as an interface for the user to access the object but without its stored data. This differs from T/P where the user wants to load an object into memory to access the object. If the user wanted to read the data, they could use the interface object to do so, protecting the user from changes to the payload in the process. The payload object contains the data for the interface object and allocates contiguous blocks of memory. Payload classes are often referred to as auxiliary storage.

The specific data structure used by ATLAS is the ROOT TTree, but the EDM is agnostic to the type of data structure used. The ROOT TTree is discussed further in the next section. ATLAS specific libraries are not required to handle files written in the xAOD format since the payload can be read directly from the contiguous allocation of memory, a central tenent of the xAOD EDM. This allows for the separation of ATLAS specific analysis frameworks and the preferred analysis tool of the user. More information on how the xAOD EDM is deployed into unit tests in Section 5.1.

2.2 Athena and ROOT

Athena is the open-source software framework for the ATLAS experiment.[27] It is based off the Gaudi project and uses ROOT and other software as part of the LHC Computing Grid (LCG) software stack.[15] The LCG software stack is a set of software frameworks that provide general solutions for the LHC experiment's computing needs. It contains on the order of 500 packages, which include binary builders and compilers, language libraries and dependencies, simulation and analysis software, and more. Athena also provides some in-house based analysis tools as well as tools for specifically ROOT based analysis.

An Athena application relies on *components*: Algorithms, Tools, Services and Properties. [18] Each component plays a role in executing an Athena application or job, where PYTHON is used. ATLAS uses PYTHON for job configuration and steering.[1] Specifically, an Algorithm

[1]Job transforms are PYTHON scripts that steer Athena production jobs by configuring arguments that would alter low-level behavior of the entire job.

accesses data objects in the event store, as shown with the solid lines in Figure 2.1, but does not own or provide any data itself. Algorithms can "own" Tools, which serve as helpers exclusive to Algorithms or other components that call them.[2] Services are not as exclusive with its access, as they can be used by other components to provide a service such as Athena-ROOT conversion, random number generators, and others. Properties are able to be called at initialization of the job configuration and include flag definitions, input and output file names, and other algorithm specific options. `ComponentAccumulator` (CA) is a python class that put into Athena production as a way to prevent extra calls of setting flags during configuration.

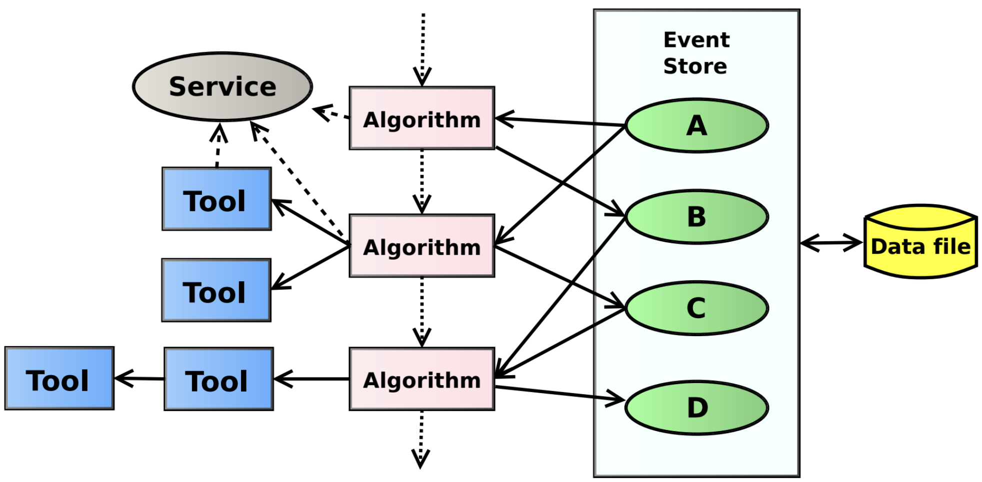

Figure 2.1: An Athena application's general structure.[18]

An important step throughout the development of Athena is to ensure any new changes to the codebase will not overrule the functionality of core features to the present workflows. One of the areas needed to be tested before and upon merging of any new changes to Athena is the I/O functionality, or the performance of reading and writing of stored objects within a broader context of various jobs, i.e. reconstruction or derivation. While CA is a more general mechanism to run many kinds of job with Athena, the scope of this thesis is using

[2] "Ownership" here refers to the components' exclusive access or control of a Tool or Service.

CA to test core I/O functionality of the new event data model. An example Athena job configuration is found in Appendix B.

ROOT is an open-source software framework used for high-energy physics analysis at CERN.[28] It uses C++ objects to save, access, and process data brought in by the various experiments based at the LHC, the ATLAS experiment uses it in conjunction with Athena. ROOT largely revolves around organization and manipulation of TFiles and TTrees into ROOT files. A TTree represents a columnar dataset, and the list of columns are called branches. A TTree is a ROOT object that organizes physically distinct types of event data into TBranches, or just branches. Event data could range from information about a specific type of interaction, this includes tracks, position of particles at one point in the detector.

```
==========================================================================================================
 |  Mem Size       Disk Size        Size/Evt      MissZip/Mem  items  (X) Container Name (X=Tree|Branch)
==========================================================================================================
  108286.649 kb    75465.794 kb       0.539 kb        0.000   140000  (B) EventInfoAuxDyn.mcEventWeights
  703839.521 kb    75806.374 kb       0.541 kb        0.000   140000  (B) AntiKt4TruthDressedWZJetsAux.
  937529.397 kb    84669.190 kb       0.605 kb        0.000    12816  (T) DataHeaderForm
  156560.056 kb   136608.917 kb       0.976 kb        0.000   140000  (B) InDetTrackParticlesAuxDyn.definingParametersCovMatrixOffDiag
 1907707.847 kb   447106.466 kb       3.194 kb        0.000   140000  (B) HLTNav_Summary_DAODSlimmedAuxDyn.decisions
```

Figure 2.2: A snapshot of the TBranches composing a TTree, from a PHYSLITE DAOD

One function relevant to TTree is `Fill()`. `Fill()` will loop over all of the branches in the TTree and compresses the baskets that make up the branch.[29] This initiates the data in memory to start filling a branch's basket buffer (or just "baskets"). While this first buffer is always unoptimized, it allows opportunity to calculate an optimal basket buffer size. At regular intervals, dictated either by number of bytes written or by number of entries written, `AutoFlush` will start moving basket buffers from memory and saving them to disk. It's this "flushing" mechanism that allows for easy access to the branch data as each of the baskets will be stored contiguously in memory. The Athena default basket maximum size at present is 128 kB, and the default minimum number of entries is 10. The minimum number of entries helps reduce processing on every entry which might be empty, and the maximum basket size is in place to prevent baskets from using too much memory throughout a Grid job. Prior to this thesis, the original implementation of both the basket size and minimum number of entries had not yet been fully investigated for avenues of optimization, this is explored in Section 4.1.

CMake and Make are open-source software that is used to build Athena, ROOT, and other software. A sparse build is a way to make changes to an individual package of code without having to recompile the entire framework at once, which saves time and resources. A user can create a text file identifying the path to the package modified, and the sparse build for Athena will proceed upon issuing the following commands:

```
cmake -DATLAS_PACKAGE_FILTER_FILE=../package_filters.txt ../athena/
Projects/WorkDir/
make -j
```

The POOL framework is part of a larger framework known as the Persistency Framework (PF). [30] The PF was developed with the intent to be independent of any individual experiment, and the goal was to address data access requirements of LHC experiments in different ways. POOL was in charge of C++ object storage, collection of metadata, and file catalogs by using streaming and relational technologies. POOL provided highly scalable object serialization to framework evolving PF files. It was eventually discontinued by other experiments in favor of a newer persistency mechanism that uses ROOT in a more streamlined way. ATLAS then became the sole supporter of POOL and integrated it within Athena to support persistent navigation of the ROOT storage layer. Now, Athena has both the original PF POOL functionality and a separate modern AthenaPool functionality. AthenaPool resides in the ATLAS I/O framework and controls ROOT TTree and TBranch properties such as compression and basket buffer sizing.

2.2.1 Continuous Integration (CI) and Development

CI is a software development practice where new code is tested and validated upon each merge to the main branch of a repository. Every commit to the main branch is automatically built and tested for specific core features that are required to work with the codebase. This helps to ensure that the codebase is working as intended and that any new code is compatible with the existing codebase.

Athena is hosted on GitLab and developed using CI with an instance of Jenkins, called ATLAS Robot, which builds and tests the new changes within a merge request interface.[31][32] ATLAS Robot will then provide a report of the build and test results. If the build or test fail, ATLAS Robot will provide a report of which steps failed and why. This allows for early detection of issues before the nightly build is compiled and tested.

2.3 Derivation Production Jobs

A derivation production job takes AODs, which comes from the reconstruction step at $\mathcal{O}(1\ \mathrm{MB})$ per event, and creates a derived AOD (DAOD) which sits at $\mathcal{O}(10\ \mathrm{kB})$ per event. Derivation production is a necessary step to make all data accessible for physicists doing analysis as well as reducing the amount of data that needs to be processed. While derivations are reduced AODs, they often contain additional information useful for analysis, such as jet collections and high-level discriminants.[33] The two mainstream output file formats Athena is capable of handling are PHYS and PHYSLITE. Figure 2.3 shows the object composition of a PHYS and PHYSLITE $t\bar{t}$ sample.

PHYS output files, at 40.0 kB per event, is predominantly made of jet collections, while PHYSLITE, at 16.1 kB per event, has more trigger and track information.

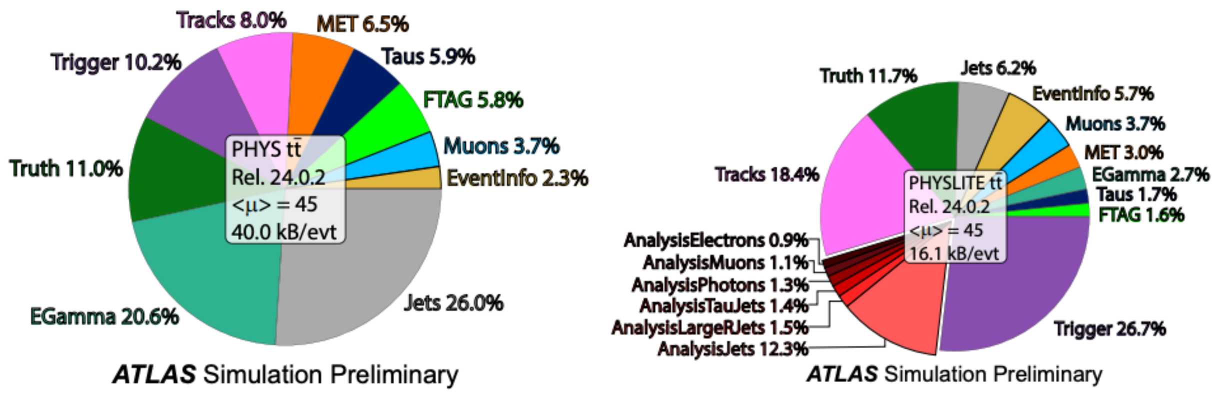

Figure 2.3: Object composition of a PHYS and PHYSLITE $t\bar{t}$ MC simulated sample from Run 3.

These jobs can demand heavy resource usage on the GRID, so optimization of the AOD/-DAODs for derivation jobs can be vital.

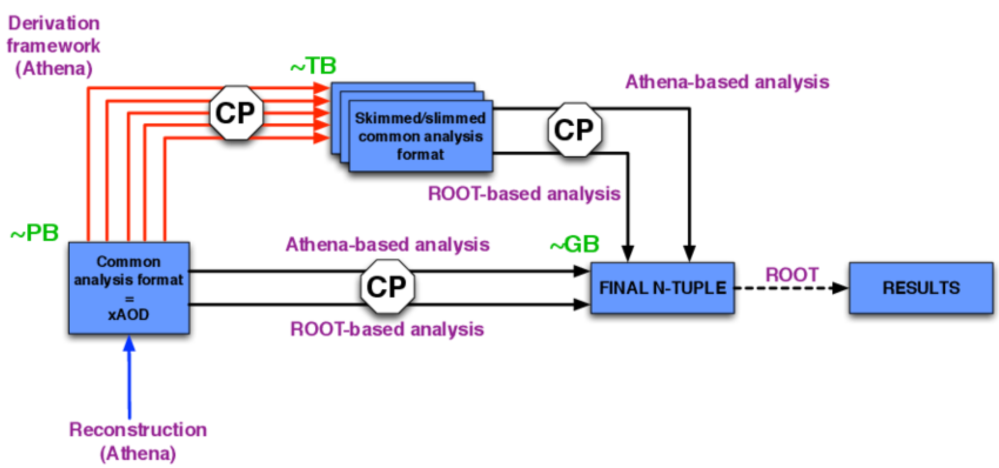

Figure 2.4: Derivation production from Reconstruction to Final N-Tuple[34]

The derivation framework is sequence of steps that are performed on the AODs to create the DAODs. Skimming is the first step in the derivation framework, and is responsible for removing whole events based on pre-defined (or augmented) criteria. Thinning is the second step, and it removes whole objects based on similarly pre-defined or augmentable criteria. Lastly slimming removes variables from objects uniformly across events.

CHAPTER 3

TOY MODEL BRANCH STUDY

Building a toy model for derivation production jobs offers a simplified framework to effectively simulate and analyze the behavior of real and Monte Carlo (MC) data. A toy model collection of data can mimic commonalities in both data and MC by filling in branches with a mixture of position coordinate information, momenta, or other details about the detector.

Integers that are repeated can be easier to compress than floating-point numbers of increased precision, so adjusting the ratio of integers to floats creates a mixture which can yield compression ratios closer to real and MC data. Replicating this mixture in a branch give us an effective model that resembles how current derivation jobs act on real and MC simulated data. These toy model mixtures provide an avenue to test opportunities for optimizing the memory and storage demands of the GRID by first looking at limiting basket sizes and their effects on compression of branches.

3.1 Toy Model Compression

3.1.1 Random Float Branches

There were a number of iterations to the toy model, but the first was constructed by filling a TTree with branches that each have vectors with varying number of random floats to write and read. Vectors are used in this toy model, as opposed to arrays, because vectors are dynamically allocated and deallocated, which allows for more flexibility when synthesizing

AODs. This original model had four distinct branches, each with a set number of events (`N=1000`), and each event having a number of entries, vectors with 1, 10, 100, and 1000 floats each. The script can be compiled with `gcc` or `g++` and it requires all of the dependencies that come with ROOT. Alternatively, the script can be run directly within ROOT.

The following function `VectorTree()` is the main function in this code. What is needed first is an output file, which will be called `VectorTreeFile.root`, and the name of the tree can simply be `myTree`. The toy model starts variable initialization with the total number of events in the branch, i.e. the number of times a branch is filled with the specified numbers per vectors, `N`. Additionally the branches have a number of floats per vector, this size will need to be defined as `size_vec_0`, `size_vec_1`, etc. The actual vectors that are being stored into each branch need to be defined as well as the temporary placeholder variable for our randomized floats, `vec_tenX` and `float_X`, respectively.

```
void VectorTree() {
  ...
  const int N = 1e4; // N = 10000, number of events
  // Set size of vectors with 10^# of random floats
  int size_vec_0 = 1;
  int size_vec_1 = 10;
  int size_vec_2 = 100;
  int size_vec_3 = 1000;

  // vectors
  std::vector<float> vec_ten0; // 10^0 = 1 entry
  std::vector<float> vec_ten1; // 10^1 = 10 entries
  std::vector<float> vec_ten2; // 10^2 = 100 entries
  std::vector<float> vec_ten3; // 10^3 = 1000 entries

  // variables
```

```
    float float_0;
    float float_1;
    float float_2;
    float float_3;
    ...
  }
```

From here, branches are initialized so each one knows where its vector pair resides in memory.

```
  void VectorTree() {
    ...
    // Initializing branches
    std::cout << "creating branches" << std::endl;
    tree->Branch("branch_of_vectors_size_one", &vec_ten0);
    tree->Branch("branch_of_vectors_size_ten", &vec_ten1);
    tree->Branch("branch_of_vectors_size_hundred", &vec_ten2);
    tree->Branch("branch_of_vectors_size_thousand", &vec_ten3);
    ...
  }
```

One extra step taken during this phase of testing is the disabling of `AutoFlush`.

```
  void VectorTree() {
    ...
    tree->SetAutoFlush(0);
    ...
```

Disabling `AutoFlush` allows for more consistent compression across the various sizes of branch baskets. If `AutoFlush` were enabled, then across the various branch types, as in 3.1, ROOT would decide when to compress each branch basket preventing a consistent compression configuration for the toy model. The derivation production jobs tested in Chapter 4 were tested with **AutoFlush** enabled because those tests are focused on memory and disk usage as opposed to mimicking real or MC data, which they are already using. Following branch initialization comes the event loop where data is generated and emplaced into vectors.

```
void VectorTree() {
  ...
  // Events Loop
  std::cout << "generating events..." << std::endl;
  for (int j = 0; j < N; j++) {
      // Clearing entries from previous iteration
      vec_ten0.clear();
      vec_ten1.clear();
      vec_ten2.clear();
      vec_ten3.clear();

      // Generating vector elements, filling vectors
      // Fill vec_ten0
      // Contents of the vector:
      //    {float_0}
      //    Only one float of random value
      float_0 = gRandom->Rndm() * 10; // Create random float value
      vec_ten0.emplace_back(float_0); // Emplace float into vector

      // Fill vec_ten1
      // Contents of the vector:
      //    {float_1_0, ... , float_1_10}
      //    Ten floats, each float is random
      for (int n = 0, n < size_vec_1; n++) {
          float_1 = gRandom->Rndm() * 10;
          vec_ten1.emplace_back(float_1);
      }
```

```

        // Do the same with vec_ten2 and vec_ten3, except for
        //     vectors with size 100 and 1000 respectively.

        // After all branches are filled, fill the TTree with
        //     new branches
        tree->Fill();
    }
    // Saving tree and file
    tree->Write();
    ...
  }
```

Once the branches were filled, ROOT then will loop over each of the branches in the TTree and at regular intervals will remove the baskets from memory, compress, and write the baskets to disk (flushed).

As illustrated, the `TTree` is written to the file which allows for the last steps within this script.

```
  void VectorTree() {
    ...

    // Look in the tree
    tree->Scan();
    tree->Print();

    myFile->Save();
    myFile->Close();
  }

  int main() {
```

```
    VectorTree();
    return 0;
}
```

Upon reading back the ROOT file, the user can view the original size of the file (Total-file-size), the compressed file size (File-size), the ratio between Total-file-size and File-size (Compression Factor), the number of baskets per branch, the basket size, and other information. Filling vectors with entirely random values was believed to yield compression ratios close to real data, but the results in Figure 3.1 show changes needed to be made to bring the branches closer to a compression ratio of $\mathcal{O}(5)$.[1] It is evident that branches containing vectors with purely random floats are more difficult to compress due to the high level of randomization.

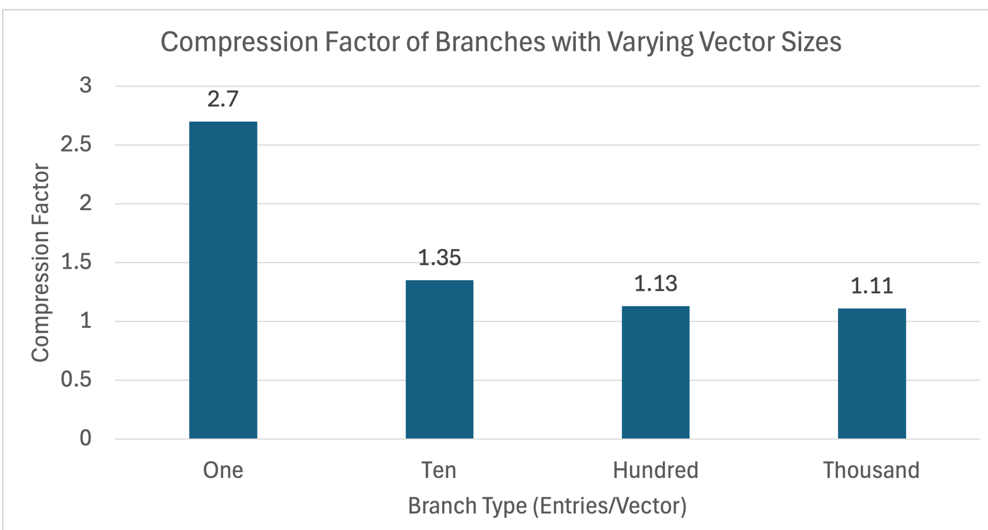

Figure 3.1: Compression factors of $N = 1000$ entries per branch with random-valued vectors of varying size.

[1] This compression factor comes as the average branch compression factor post-derivation job, which is discussed in Section 2.3

Figure 3.1 shows compression drop-off as the branches with more randomized floats per vector were present. This is the leading indication that there needs to be more compressible data within the branches.

3.1.2 Mixed-Random Float Branches

The branches needed to have some balance between compressible and incompressible data to mimic the compression ratio found in real data. How this was achieved was by filling each vector with different ratios of random floats and repeating integers, which will now be described in detail.

The first change was increasing the total number of events per branch from $N = 10^4$ to $N = 10^5$. Mixing of random floats and repeated integer values takes the same script structure as Section 3.1.1 but adjusts the event generation loop.

```
void VectorTree() {
  ...
  // Events Loop
  for (int j = 0; j < N; j++) {
      // Clearing entries from previous iteration
      vec_ten0.clear();
      vec_ten1.clear();
      vec_ten2.clear();
      vec_ten3.clear();

      // Generating vector elements, filling vectors
      // Generating vec_ten0
      // Contents of the vector:
      //    {float_0}
      //    Only one float of random value
```

```
        // And since there's only one entry, we don't mix the entries.
        float_0 = gRandom->Gaus(0, 1) * gRandom->Rndm();
        vec_ten0.emplace_back(float_0);


        // Generating vec_ten1
        // Contents of the vector:
        //    {float_1_0, float_1_1, float_1_2, float_1_3, float_1_4, 1,
    1, 1, 1, 1}
        //    5 floats of random values, 5 integers of value 1.
        for (int b = 0; b < size_vec_1; b++) {
            if (b < size_vec_1 / 2) {
              float_1 = gRandom->Rndm() * gRandom->Gaus(0, 1);
              vec_ten1.emplace_back(float_1);
            } else {
              float_1 = 1;
              vec_ten1.emplace_back(float_1);
            }
        }

        // Do the same with vec_ten2 and vec_ten3, except for
        //     vectors with size 100 and 1000 respectively.


        // After all branches are filled, fill the TTree with
        //     new branches
        tree->Fill();
    }
    // Saving tree and file
    tree->Write();
    ...
```

```
46   }
```

As shown in the `if`-statements in line `25`, if the iterator was less than half of the total number of entries in the vector then that entry had a randomized float put in that spot in the vector, otherwise it would be filled with the integer `1`. Having a mixture of half random floats and half integer `1` led to the larger branches still seeing poor compression, so a new mixture of 1/4 random data was introduced. Even though $N = 10^5$ had the larger branches closer to the desired compression ratio, testing at $N = 10^6$ events improves the accuracy of the overall file size to more closely resemble real data.

Figure 3.2 shows the difference between compression between the two mixtures at $N = 10^6$ events. When the number of events is increased from $N = 10^5$ to $N = 10^6$, at the 1/2 random-mixture, the branches with more than one entry per vector see their compression factor worsen. Figure 3.3 shows a compression ratio hovering around 3 for the larger branches, whereas Figure 3.2 shows the same branches hovering around 2.

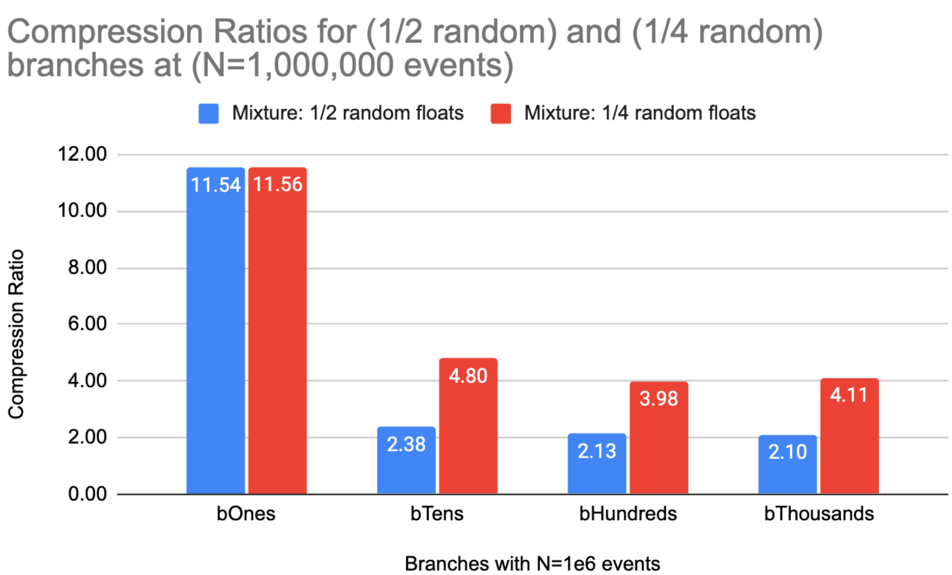

Figure 3.2: Compression Ratios for ($\frac{1}{2}$ random) and ($\frac{1}{4}$ random) branches at ($N = 10^6$ events)

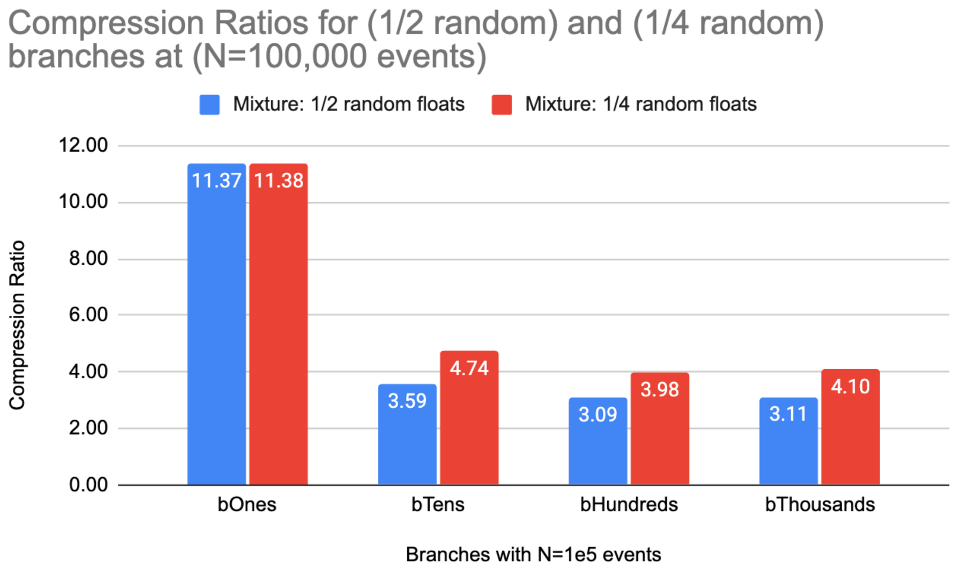

Figure 3.3: Compression Ratios for ($\frac{1}{2}$ random) and ($\frac{1}{4}$ random) branches at ($N = 10^5$ events)

Unlike the mixture of branches having 1/2 random data, the 1/4 mixture does not see the same compression effect, but with this mixture we see a compression ratio that is in-line with real data. This is inline with expectation, more repeated integers within the mixture makes the branch more compressible, and the more random floats in the mixture will make the branch more difficult to compress. With these mixtures added to the toy model, we can start looking at varying the basket sizes to see how they affect compression.

3.2 Basket-Size Investigation

Investigating how compression is affected by the basket size requires us to change the basket size, refill the branch and read it out. Changing the basket buffer size was done at the script level with a simple setting after the branch initialization and before the event loop the following code:

```
int basketSize = 8192000; // 8 MB
tree->SetBasketSize("*",basketSize);
```

This ROOT-level setting was sufficient for the case of the toy model; testing of the basket size setting both at the ROOT- and Athena-level would be done later using derivation production jobs in Section 4.1. The lower bound set for the basket size was 1 kB and the upper bound was 16 MB. The first branch looked at closely was the branch with a thousand vectors with half of them being random floats, see Figure 3.4.

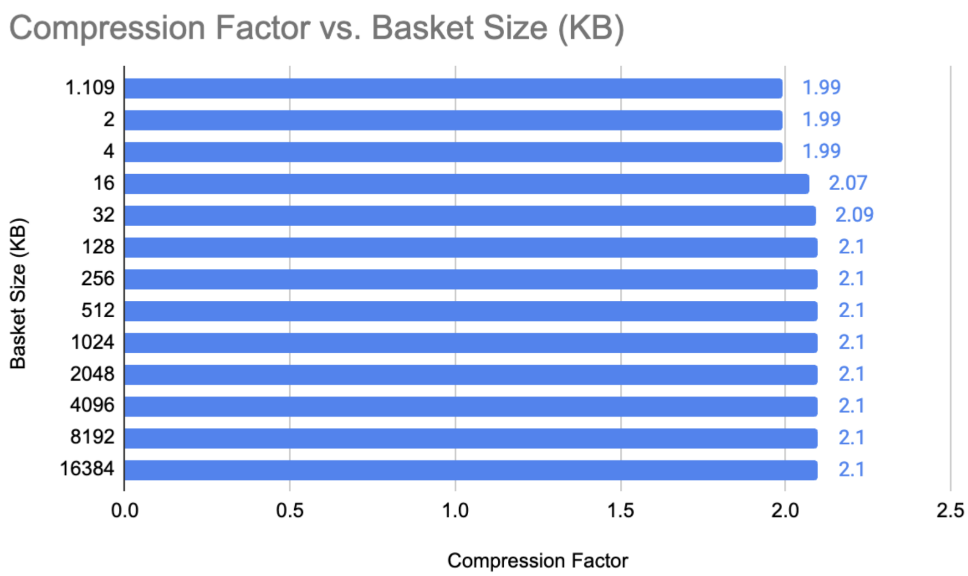

Figure 3.4: Compression Factors vs Branch Size (1000 entries per vector, 1/2 Mixture $N = 10^6$ events)

Figures 3.4 and 3.5 are the first indication that the lower basket sizes are too small to effectively compress the data. For baskets smaller than 16 kB, it is necessary to have as many baskets as events to store all the data effectively. For a mixed-content vector with one thousand entries, containing 500 floats and 500 integers (both are 4 bytes each), its size is approximately 4 kB. ROOT creates baskets of at least the size of the smallest branch entry, in this case the size of a single vector. So even though the basket size was set to 1 or 2 kB,

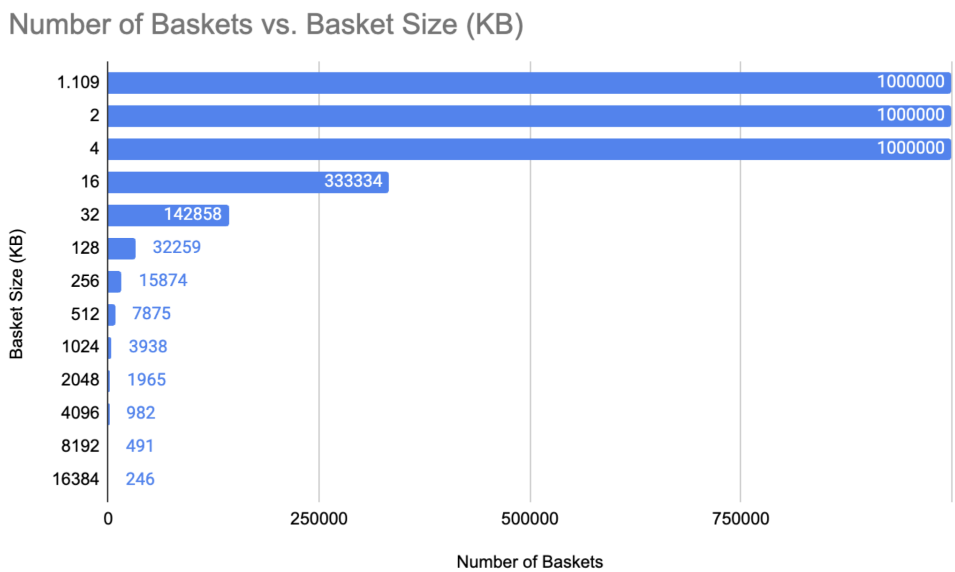

Figure 3.5: Number of Baskets vs Branch Size (1000 entries per vector, 1/2 Mixture $N = 10^6$ events)

ROOT created baskets of 4 kB. These baskets less than or equal to 4kB have a significantly worse compression than the baskets greater than 4kB in size, so the focus was shifted toward baskets. Once the basket size is larger than the size of a single vector, more than one vector can be stored in a single basket and the total number of baskets is reduced.

There were different types of configuration to the toy model investigated by this study. Looking further into the types of mixtures and how they would affect compression are shown in Figures 3.6 and 3.7. Here the same mixtures were used but the precision of the floating point numbers was decreased from the standard 32 floating-point precision to 16 and 8, making compression easier.

Each of these sets of tests indicate that after a certain basket size, i.e. 128 kB, there is no significant increase in compression. Having an effective compression at 128 kB, it's useful to stick to that basket size to keep memory usage down. Knowing that increasing the basket size beyond 128 kB yields diminishing returns, it's worth moving onto the next phase of testing with actual derivation production jobs.

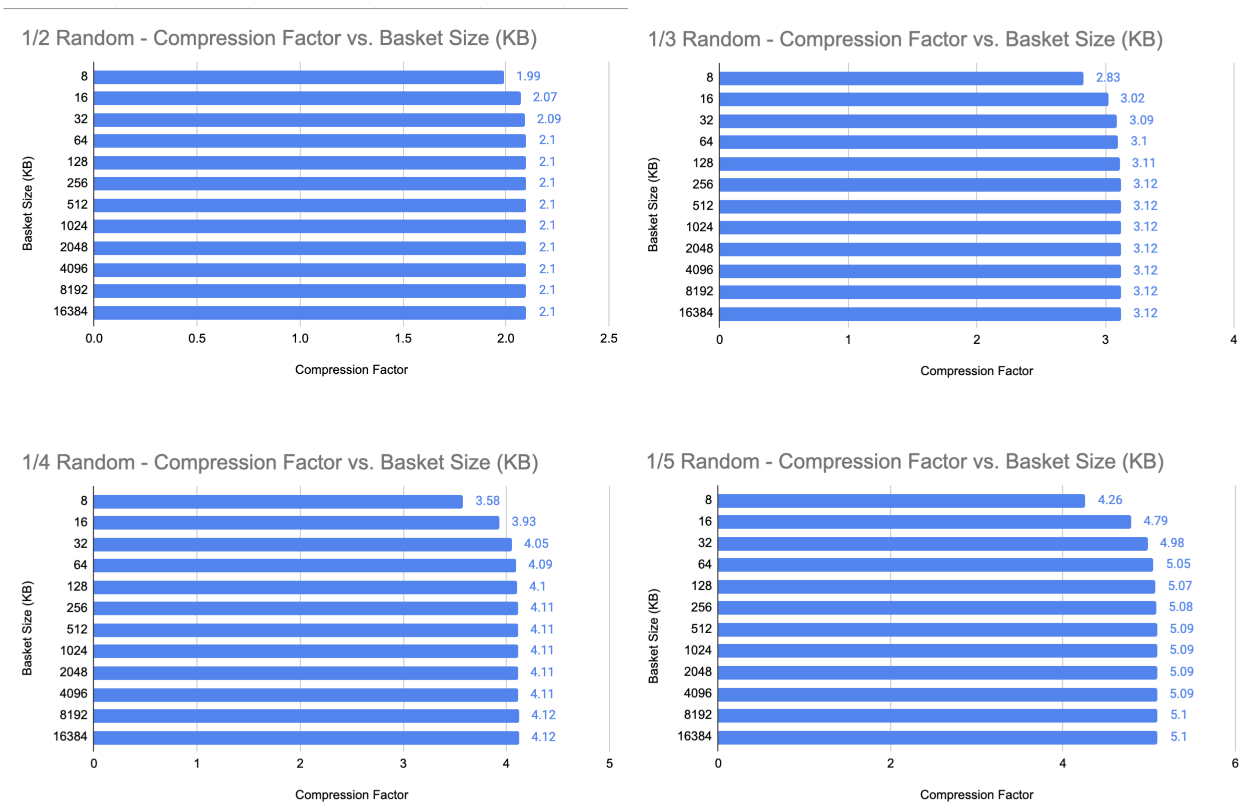

Figure 3.6: Varying Mixtures in 8 point precision - Number of Baskets vs Branch Size ($N = 10^6$ events)

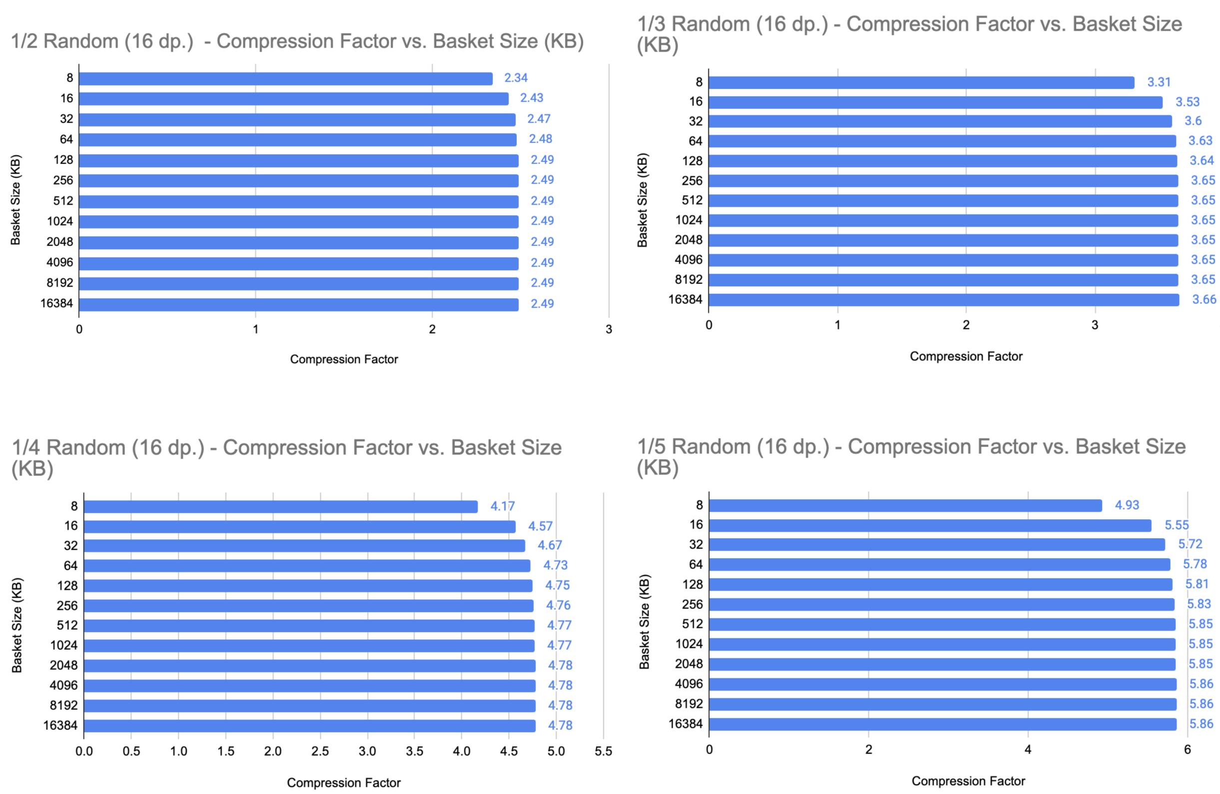

Figure 3.7: Varying Mixtures in 16 point precision - Number of Baskets vs Branch Size ($N = 10^6$ events)

CHAPTER 4

DATA AND MONTE CARLO DERIVATION PRODUCTION

Derivation production demands high memory usage, and DAODs make up a bulk of disk-space usage. DAODs are used in physics analyses and ought to be optimized to alleviate stress on the GRID and to lower disk-space usage. Optimizing both disk-space and memory usage is a tricky balance as they are typically at odds with one another. For example, increasing memory output memory buffers results in lower disk-space usage due to better compression but the memory usage will increase since the user will load a larger buffer into memory. This project opted to take is by optimizing for disk-space and memory by testing various basket limits and viewing the effects of the branches on both data and Monte Carlo (MC) simulated analysis object data (AODs).

4.1 Basket-size Configuration

As the toy model ruled out, the focus here was on optimizing Athena and not ROOTs contribution for optimization. The initial focus was on the inclusion of a minimum number of entries per buffer and the maximum basket buffer limit. The AthenaPOOL script directly involved with these buffer settings is the `PoolWriteConfig.py` found in the path `athena/Database/AthenaPOOL/AthenaPoolCnvSvc/python/`. As discussed in Section 4.2, further testing opted to keep the minimum number of entries set to its default setting, 10 entries per buffer.

Throughout the duration of this testing, the results of compression or file size are independent of any changes to the release or the nightly version of Athena. The data derivation job comes from a 2022 dataset with four input files and 160,327 events. The MC job comes from a 2023 $t\bar{t}$ standard sample simulation job with six input files and 140,000 events. The datasets are noted in Appendix A.1.

4.1.1 Derivation Job Configuration

To run a derivation job for testing purposes, AODs need to be downloaded by a data-management service, such as Rucio, to a machine dedicated to manually run tests.[35] Rucio is the data-management solution used for this project to procure the various AOD input files used for the derivation jobs. The machine running the Rucio client will need to have a valid proxy added for Rucio to run correctly.

A sample command would look like:

```
rucio download data22_13p6TeV:AOD.31407809._000898.pool.root.1
```

This downloads the AOD file from Rucio and saves it to the user's local directory.

The command used by Athena to run a derivation job takes the form of the following example:

```
ATHENA_CORE_NUMBER=4 Derivation_tf.py \
--CA True \
--inputAODFile mc23_13p6TeV.601229.PhPy8EG_A14_ttbar_hdamp258p75_SingleLep
    .merge.AOD.e8514_e8528_s4162_s4114_r14622_r14663/AOD.33799166._001224.
    pool.root.1 \
--outputDAODFile art.pool.root \
--formats PHYSLITE \
--maxEvents 2000 \
--sharedWriter True \
--multiprocess True ;
```

Where Athena allows one to specify the number of cores to use with the `ATHENA_CORE_NUMBER` environment variable. `Derivation_tf.py` is a script that runs the derivation job and is part of the Athena release. The `--inputAODFile` is the input file for the derivation job, in this case an AOD file. The user can specify multiple input files at a time by enclosing the input files in quotes and separating each file with a comma, like the following:

```
--inputAODFile="AOD.A.pool.root.1,AOD.B.pool.root.1,AOD.C.pool.root.1,
AOD.D.pool.root.1"
```

The `--outputDAODFile` is the output file for the derivation job, in this case a DAOD file. The `--formats PHYSLITE` flag allows the job to use the PHYSLITE format for the DAOD. Here is where the user may choose to include PHYS or PHYSLITE simply by inlusion of one or both. The `--maxEvents` flag allows one to specify the maximum number of events to run the job on. The `--multiprocess True` flag allows the job to use AthenaMP tools. AthenaMP is a mode of operation that allows for multi-process parallelism in many workflows since Run 2.[36] The `--sharedWriter True` flag allows the job to utilize SharedWriter, which is a memory allocation mechanism as part of AthenaMP which allows for multiple workers to share allocated memory in the writing process. The machine used to run these derivation tests was a CERN based machine, using an AMD EPYC 7302 16-Core Processor, supplied with 258 GB of memory, on version 9.4 of the AlmaLinux distribution.

The input files for both data and MC jobs were ran with various configurations of Athena by modifying the basket buffer limit. The four configurations tested all kept minimum number of basket buffer entries at 10 and modified the basket limitation in the following ways:

1. "*default*" - Athena's default setting, and basket limit of 128 kB
2. "*256k*" - Limit basket buffer to 256 kB
3. "*512k*" - Limit basket buffer to 512 kB

4. "*no-lim*" - Removing the Athena basket limit, the ROOT imposed 1.3 MB limit still remains

4.2 Results

4.2.1 Presence of basket-cap and presence of minimum number of entries

The first batch testing was for data and MC simulation derivation production jobs with and without presence of an upper limit to the basket size and presence of the minimum number of basket buffer entries. PHYSLITE MC derivation production, from Table 4.2, sees a 9.9% increase in output file size when compared to the default Athena configuration. Since this configuration only differs by the omission of the "min-number-entries" requirement, we assume the minimum number of basket buffer entries should be kept at 10 and left alone.

Presence of features (Data)	Max PSS (MB) (Δ% default)	PHYS outFS (GB) (Δ%)	PHYSLITE outFS (GB) (Δ%)
basket-cap, min-num-entries (default)	27.1 (+ 0.0 %)	3.22 (+ 0.0 %)	1.03 (+ 0.0 %)
~~basket-cap~~ ~~min-num-entries~~	27.8 (+ 2.5 %)	3.22 (+ 0.2 %)	1.04 (+ 0.2 %)
~~basket-cap~~ min-num-entries	27.8 (+ 2.5 %)	3.22 (- 0.0 %)	1.03 (- 0.4 %)
basket-cap, ~~min-num-entries~~	27.3 (+ 0.7 %)	3.22 (+ 0.2 %)	1.04 (+ 0.7 %)

Table 4.1: Comparing the maximum proportional set size (PSS) and PHYS/PHYSLITE output file sizes (outFS) for data jobs while varying the presence of features in Athena PoolWriteConfig.py for 160327 entries.

Presence of features (MC)	Max PSS (MB) (Δ% default)	PHYS outFS (GB) (Δ%)	PHYSLITE outFS (GB) (Δ%)
basket-cap, min-num-entries (default)	14.1 (+ 0.0 %)	5.8 (+ 0.0 %)	2.6 (+ 0.0 %)
~~basket-cap~~ ~~min-num-entries~~	16.1 (+ 12.1 %)	6.0 (+ 2.9 %)	2.7 (+ 5.1 %)
~~basket-cap~~ min-num-entries	16.0 (+ 11.5 %)	5.7 (- 2.8 %)	2.5 (- 5.6 %)
basket-cap, ~~min-num-entries~~	14.2 (+ 0.4 %)	6.2 (+ 5.4 %)	2.9 (+ 9.9 %)

Table 4.2: Comparing the maximum proportional set size (PSS) and PHYS/PHYSLITE output file sizes (outFS) for MC jobs while varying the presence of features in Athena PoolWriteConfig.py for 140000 entries.

Table 4.2 also shows the potential for a PHYSLITE MC DAOD output file size reduction by eliminating our upper basket buffer limit altogether. However, since derivation production (or any job for that matter) is memory bound[1] neither case where basket buffer limits are removed are viable options for optimization.

4.2.2 Comparing different basket sizes

Pre-existing derivation jobs were ran for data and MC simulations to compare between configurations of differing basket sizes limits. The results for this set of testing are found from Table 4.3 through Table 4.4. The following tables are the DAOD output-file sizes of the various Athena configurations for PHYS/PHYSLITE over their respective data/MC AOD input files.

Athena Config (Data)	Max PSS (MB) (Δ% default)	PHYS outFS (GB) (Δ%)	PHYSLITE outFS (GB) (Δ%)
(default)	27.9 (+ 0.0 %)	3.3 (+ 0.0 %)	1.0 (+ 0.0 %)
256k_basket	28.2 (+ 1.3 %)	3.3 (- 0.1 %)	1.0 (- 0.3 %)
512k_basket	28.5 (+ 2.2 %)	3.3 (+ 0.0 %)	1.0 (- 0.3 %)
1.3 MB (ROOT MAX)	28.6 (+ 2.7 %)	3.3 (- 0.1 %)	1.0 (- 0.3 %)

Table 4.3: Comparing the maximum proportional set size (PSS) and PHYS/PHYSLITE output file sizes (outFS) for Data jobs over various Athena configurations for 160327 entries.

Athena Config (MC)	Max PSS (MB) (Δ% default)	PHYS outFS (GB) (Δ%)	PHYSLITE outFS (GB) (Δ%)
(default)	15.0 (+ 0.0 %)	5.9 (+ 0.0 %)	2.6 (+ 0.0 %)
256k_basket	15.3 (+ 1.9 %)	5.8 (- 1.4 %)	2.5 (- 3.1 %)
512k_basket	16.4 (+ 8.6 %)	5.7 (- 2.5 %)	2.5 (- 5.1 %)
1.3 MB (ROOT MAX)	16.9 (+ 11.3 %)	5.7 (- 2.8 %)	2.5 (- 5.6 %)

Table 4.4: Comparing the maximum proportional set size (PSS) and PHYS/PHYSLITE output file sizes (outFS) for MC jobs over various Athena configurations for 140000 entries.

"Max PSS" refers to the maximum proportional set size, which is the maximum memory usage of the job. Table 4.4 uses data from a 2022 dataset with four input files and shows

[1] Memory usage for the Grid is standardized at 2 GB per core on an 8-core configuration allowing any job to process on any Grid node.

there are marginal changes in both the memory usage for the job and the output file size of the DAODs. Whereas Table 4.4 shows a much more drastic change, with a 5.6% reduction in output file size for the MC PHYSLITE DAOD when compared to the default Athena configuration. While there's a 5.6% reduction in output file size for the MC PHYSLITE DAOD, there's also a 11.3% increase in memory usage.

4.2.3 Monte Carlo PHYSLITE branch comparison

Derivation production jobs work with initially large, memory-consuming branches, compressing them to a reduced size. These derivation jobs are memory intensive because they first have to load the uncompressed branches into readily-accessed memory. Once they're loaded, only then are they able to be compressed. The compression factor is the ratio of pre-derivation branch size (Total-file-size) to post-derivation branch size (Compressed-file-size). The compressed file size is the size of the branch that is permanently saved into the DAOD.

Branches with highly repetitive data are better compressed than non-repetitive data, leading to high compression factors–the initial size of the branch contains more data than it needs pre-derivation. If pre-derivation branches are larger than necessary, there should be an opportunity to save memory usage during the derivation job.

Athena v24.0.16 (default) MC branch	Branch size (kB)	Total-file-size (MB)	Compressed-file-size (MB)	Compression factor
PrimaryVerticesAuxDyn.trackParticleLinks	128	2146.2	24.0	89.4
HardScatterVerticesAuxDyn.incomingParticleLinks	128	118.5	1.7	71.6
HLTNav_Summary_DAODSlimmedAuxDyn.linkColNames	128	784.0	11.9	65.7
HardScatterVerticesAuxDyn.outgoingParticleLinks	128	108.6	1.9	58.7
TruthBosonsWithDecayVerticesAuxDyn.incomingParticleLinks	96	31.6	0.7	43.5
HLTNav_Summary_DAODSlimmedAuxDyn.linkColClids	128	390.6	10.7	36.6
AnalysisTauJetsAuxDyn.tauTrackLinks	128	75.0	2.0	36.6
HLTNav_Summary_DAODSlimmedAuxDyn.linkColKeys	128	390.6	11.7	33.4
AnalysisJetsAuxDyn.GhostTrack	128	413.8	13.1	31.5
TruthBosonsWithDecayVerticesAuxDyn.outgoingParticleLinks	83.5	27.3	0.9	31.0

Table 4.5: Top 10 branches sorted by compression factor, MC PHYSLITE [Athena v24.0.16 default configuration.]

Athena v24.0.16 (no-lim) MC branch	Branch size (kB)	Total-file-size (MB)	Compressed-file-size (MB)	Compression factor
PrimaryVerticesAuxDyn.trackParticleLinks	1293.5	2145.5	22.9	93.5
HardScatterVerticesAuxDyn.incomingParticleLinks	693.0	118.5	1.3	90.1
HardScatterVerticesAuxDyn.outgoingParticleLinks	635.5	108.5	1.5	74.0
HLTNav_Summary_DAODSlimmedAuxDyn.linkColNames	1293.5	783.5	11.9	65.8
TruthBosonsWithDecayVerticesAuxDyn.incomingParticleLinks	96.0	31.6	0.7	43.5
AnalysisTauJetsAuxDyn.tauTrackLinks	447.0	74.9	1.9	39.2
HLTNav_Summary_DAODSlimmedAuxDyn.linkColClids	1293.5	390.3	11.0	35.5
HLTNav_Summary_DAODSlimmedAuxDyn.linkColKeys	1293.5	390.3	11.3	34.5
AnalysisJetsAuxDyn.GhostTrack	1293.5	413.5	13.0	31.9
TruthBosonsWithDecayVerticesAuxDyn.outgoingParticleLinks	83.5	27.3	0.9	31.0

Table 4.6: Top 10 branches sorted by compression factor, MC PHYSLITE [Athena v24.0.16 without limit to the basket buffer.]

Athena v24.0.16 (default) MC branch	Branch size (kB)	Total-file-size (MB)	Compressed-file-size (MB)	Compression factor
PrimaryVerticesAuxDyn.trackParticleLinks	128	2146.2	24.0	89.4
HLTNav_Summary_DAODSlimmedAuxDyn.linkColNames	128	784.0	11.9	65.7
AnalysisJetsAuxDyn.GhostTrack	128	413.8	13.1	31.5
HLTNav_Summary_DAODSlimmedAuxDyn.linkColClids	128	390.6	10.7	36.6
HLTNav_Summary_DAODSlimmedAuxDyn.linkColKeys	128	390.6	11.7	33.4
AnalysisJetsAuxDyn.SumPtChargedPFOPt500	128	148.9	7.3	20.5
AnalysisJetsAuxDyn.NumTrkPt1000	128	148.8	8.7	17.2
AnalysisJetsAuxDyn.NumTrkPt500	128	148.8	11.9	12.5
HardScatterVerticesAuxDyn.incomingParticleLinks	128	118.5	1.7	71.6
AnalysisLargeRJetsAuxDyn.constituentLinks	128	111.5	7.1	15.8

Table 4.7: Top 10 branches sorted by total file size in bytes, MC PHYSLITE [Athena v24.0.16 default configuration.]

Athena v24.0.16 (no-lim) MC branch	Branch size (kB)	Total-file-size (MB)	Compressed-file-size (MB)	Compression factor
PrimaryVerticesAuxDyn.trackParticleLinks	1293.5	2145.5	22.9	93.6
HLTNav_Summary_DAODSlimmedAuxDyn.linkColNames	1293.5	783.5	11.9	65.8
AnalysisJetsAuxDyn.GhostTrack	1293.5	413.5	13.0	31.9
HLTNav_Summary_DAODSlimmedAuxDyn.linkColClids	1293.5	390.3	11.0	35.5
HLTNav_Summary_DAODSlimmedAuxDyn.linkColKeys	1293.5	390.3	11.3	34.5
AnalysisJetsAuxDyn.SumPtChargedPFOPt500	905.5	148.8	6.8	21.9
AnalysisJetsAuxDyn.NumTrkPt1000	905	148.8	8.5	17.6
AnalysisJetsAuxDyn.NumTrkPt500	905	148.8	11.8	12.6
HardScatterVerticesAuxDyn.incomingParticleLinks	693	118.5	1.3	90.2
AnalysisLargeRJetsAuxDyn.constituentLinks	950.5	111.4	6.4	17.4

Table 4.8: Top 10 branches sorted by total file size in bytes, MC PHYSLITE [Athena v24.0.16 without limit to the basket buffer.]

Athena v24.0.16 (default) MC branch	Branch size (kB)	Total-file-size (MB)	Compressed-file-size (MB)	Compression factor
PrimaryVerticesAuxDyn.trackParticleLinks	128	2146.2	24.0	89.4
AnalysisJetsAuxDyn.GhostTrack	128	413.8	13.1	31.5
AnalysisJetsAuxDyn.NumTrkPt500	128	148.8	11.9	12.5
HLTNav_Summary_DAODSlimmedAuxDyn.linkColNames	128	784.0	11.9	65.7
HLTNav_Summary_DAODSlimmedAuxDyn.linkColKeys	128	390.6	11.7	33.4
HLTNav_Summary_DAODSlimmedAuxDyn.linkColClids	128	390.6	10.7	36.6
AnalysisJetsAuxDyn.NumTrkPt1000	128	148.8	8.7	17.2
AnalysisJetsAuxDyn.SumPtChargedPFOPt500	128	148.9	7.3	20.5
AnalysisLargeRJetsAuxDyn.constituentLinks	128	111.5	7.1	15.8
HLTNav_Summary_DAODSlimmedAuxDyn.name	128	80.8	4.4	18.4

Table 4.9: Top 10 branches sorted by compressed file size in bytes, MC PHYSLITE [Athena v24.0.16 default configuration.]

Athena v24.0.16 (no-lim) MC branch	Branch size (kB)	Total-file-size (MB)	Compressed-file-size (MB)	Compression factor
PrimaryVerticesAuxDyn.trackParticleLinks	1293.5	2145.5	22.9	93.5
AnalysisJetsAuxDyn.GhostTrack	1293.5	413.5	13.0	31.9
HLTNav_Summary_DAODSlimmedAuxDyn.linkColNames	1293.5	783.5	11.9	65.8
AnalysisJetsAuxDyn.NumTrkPt500	905	148.8	11.8	12.6
HLTNav_Summary_DAODSlimmedAuxDyn.linkColKeys	1293.5	390.3	11.3	34.5
HLTNav_Summary_DAODSlimmedAuxDyn.linkColClids	1293.5	390.3	11.0	35.5
AnalysisJetsAuxDyn.NumTrkPt1000	905	148.8	8.5	17.6
AnalysisJetsAuxDyn.SumPtChargedPFOPt500	905.5	148.8	6.8	21.9
AnalysisLargeRJetsAuxDyn.constituentLinks	950.5	111.4	6.4	17.4
HLTNav_Summary_DAODSlimmedAuxDyn.name	242	80.8	4.5	18.0

Table 4.10: Top 10 branches sorted by compressed file size in bytes, MC PHYSLITE [Athena v24.0.16 without limit to the basket buffer.]

Tables 4.5 - 4.10 look into some highly compressible branches that might lead to areas where simulation might save some space. An immediate observation: with the omission of the Athena basket limit (solely relying on ROOTs 1.3 MB basket limit), compression increases. *PrimaryVerticesAuxDyn.trackParticleLinks* is a branch where, among each configuration of Athena MC derivation, has the highest compression factor of any branch in this dataset. Some branches, like *HLTNav Summary DAODSlimmedAuxDyn.linkColNames* show highly compressible behavior and are consistent with the other job configurations (data, MC, PHYS, and PHYSLITE). Further work could investigate these branches for further areas of optimization for long term storage and better memory usage during derivation.

4.3 Conclusion to derivation job optimization

Initially, limiting the basket buffer size looked appealing; after the 128 kB basket buffer size limit was set, the compression ratio would begin to plateau, increasing the memory-usage without saving much in disk-usage. The optimal balance is met with the setting of 128 kB basket buffers for derivation production.

Instead, by removing the upper limit of the basket size, a greater decrease in DAOD output file size is achieved. The largest decrease in file size came from the PHYSLITE MC derivation jobs without setting an upper limit to the basket buffer size. While similar decreases in file size appear for derivation jobs using data, it is not as apparent for data as it is for MC jobs. With the removal of an upper-limit to the basket size, ATLAS stands to gain a 5% decrease for PHYSLITE MC DAOD output file sizes, but an $11 - 12\%$ increase in memory usage could prove a heavy burden (See Tables 4.2 and 4.4).

By looking at the branches per configuration, specifically in MC PHYSLITE output DAOD, highly compressible branches emerge. The branches inside the MC PHYSLITE DAOD are suboptimal as they do not conserve disk space; instead, they consume memory inefficiently. As seen from Table 4.5 through 4.10, we have plenty of branches in MC PHYSLITE that are full of seemingly duplicated data–as their compression factor is greater than $\mathcal{O}(10)$, showing the extent to which they are able to be compressed. Reviewing and optimizing the branch data could further reduce GRID load during DAOD production by reducing the increased memory-usage while keeping the effects of decreased disk-space.

CHAPTER 5

MODERNIZING I/O UNIT-TESTS

Athena uses a number of unit tests during the development lifecycle to ensure core I/O functionality does not break. Many of the I/O tests were originally created for the old EDM and haven't been updated to test the xAOD EDMs core I/O functions. The new software developed in this project takes in track information from a unit test using the T/P EDM, writes the data into an example xAOD object to file and reads it back.

5.1 xAOD Test Object

The object used to employ the new unit test is the `xAOD::ExampleElectron` object, where the `xAOD::` is a declaration of the namespace and simply identifies the object as an xAOD object. An individual `ExampleElectron` object only has a few parameters for sake of testing, its transvese momentum, `pt`, and its charge, `charge`. A collection of `ExampleElectron` objects are stored in the `ExampleElectronContainer` object, which is just a `DataVector` of `ExampleElectron` objects.[26] This `DataVector` acts similar to an `std::vector`.

The xAOD EDM utilizes a separation between between static and dynamic data stores.

The static data stores comprise variables directly attributed to the object associated with it, the dynamic counterpart stores data of variables added by the user. An example of a static variable might be an electrons transverse momenta or its charge, while an example of a dynamic attribute might be a link associating that object with a specific track.

Figure 5.1 illustrates how a simple setup of storing a `DataVector` of electrons that hold some specific parameters into one `IAuxStore` while also having a separate `IAuxStore` specifically for the dynamic attributes.

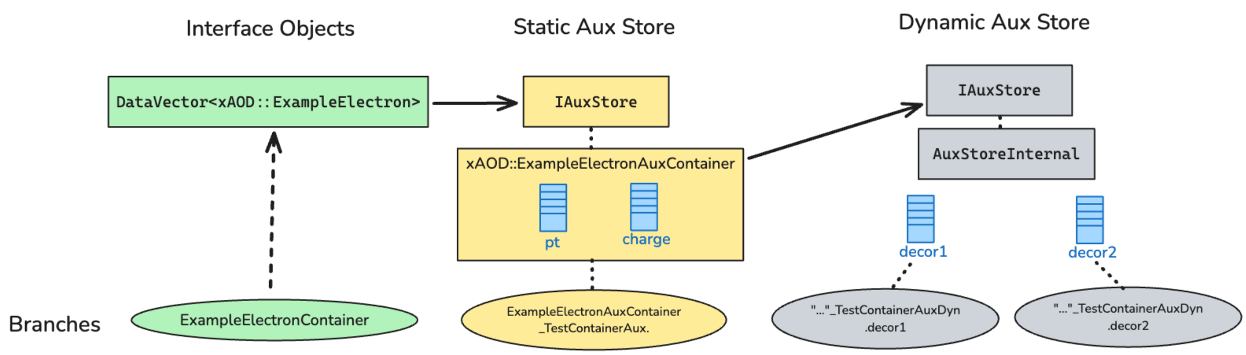

Figure 5.1: The framework between interface objects and the static/dynamic auxiliary data store for a collection of xAOD::ExampleElectrons.

5.2 Unit Tests

Unit tests are programs that act as a catch during the continuous integration of a codebase and test features that need to remain functional. Athena has a number of unit tests that check every merge request and nightly build for issues in the new code that could break core functionality, either at the level of Athena, ROOT, or any other software in the LCG stack. There were no unit tests in the appropriate packages to handle selection of dynamic attributes, or decorations, on xAOD objects created during writing and read back. To address this, a new xAOD test object needed to be created and written during a new unit test that fit into the existing unit tests. The list of AthenaPoolExample unit tests that are currently executed during a nightly build can be found in Table 5.1. These tests are executed in this order, as the objects created in one might be used in proceeding test.

Unit Test	**Employed Algorithms**	**Function**(Object Read) [Object Written]
Write	WriteData	[ExampleHit]
ReadWrite	ReadData, ReWriteData	(ExampleHit), [ExampleTrack]
Read	ReadData	(ExampleHit)
Copy	None	**Copies a file**
ReadWriteNext	ReadData, ReWriteData	(ExampleHit, EventInfo), [ExampleTrack]
WritexAODElectron	ReadData, WriteExampleElectron	(ExampleTrack), $[xAOD :: ExampleElectrons, decorations]$
ReadxAODElectron	ReadExampleElectron	(xAOD::ExampleElectrons, decorations)

Table 5.1: List of unit tests in the AthenaPoolExample package that are currently executed during a nightly build.

The mechanism for passing a unit test is done automatically by building the framework, running the unit tests, and comparing the diff of the output file to the unit test with a reference file associated with that particular unit test. If the unit test passes, then the diff, a product of the `git diff` command, will be empty and the unit test will be marked as passing. Conversely, if the unit test fails, then the diff will be non-empty and the unit test will be marked as failing.

5.2.1 WritexAODElectron.py

The two new tests added to the package were `WritexAODElectron` and `ReadxAODElectron`. During this first unit test, the first algorithm called is to `ReadData` which reads off all of the `ExampleTrack` objects stored in one of the files produced by the `ReadWrite` unit-test. Within the python script of the first unit test, the user is able to decide what decorations to have written to file. This is a part of the `OutputStreamCfg` parameter, `ItemList`, wherein the user specifies the object and its name in the format shown in Figure 5.2.

```
ItemList = [ "ExampleTrackContainer#MyTracks",
"xAOD::ExampleElectronContainer#TestContainer",
"xAOD::ExampleElectronAuxContainer#TestContainerAux.-decor2"] )
```

Figure 5.2: WritexAODElectron ItemList for the OutputStreamCfg parameter. Showing how to select dynamic attributes at the CA level.

The header file includes various packages needed by the algorithm, such as data objects, `Write/ReadHandleKeys`, base algorithms that give consistent structure to the algorithm, and whatever else is required. In the write-algorithm, there are `ReadHandleKeys` for `ExampleTrack` objects saved by a prior unit test. For the `WriteHandleKeys`, there is one for the `ExampleElectronContainer` and the name given to it is “TestContainer”. This “TestContainer” name will be needed for the `ReadExampleElectron` algorithm as the name is how it’s able to refer to the correct `ExampleElectronContainer` present in the input file. Additionally, a `WriteHandleDecorKey` for the decoration objects is needed for appending each decoration onto each `ExampleElectron` object. Figure 5.3 shows the syntax for how these keys would be presently defined.

```
// Read key ExampleTracks
SG::ReadHandleKey<ExampleTrackContainer> m_exampleTrackKey{
    this, "ExampleTrackKey", "MyTracks"};

// Write key for the ExampleElectronContainer
SG::WriteHandleKey<xAOD::ExampleElectronContainer>
    m_exampleElectronContainerKey{this, "ExampleElectronContainerName",
                                   "TestContainer"};

// Decoration keys
SG::WriteDecorHandleKey<xAOD::ExampleElectronContainer> m_decor1Key{
    this, "ExampleElectronContainerDecorKey1", "TestContainer.decor1",
    "decorator1 key"};
SG::WriteDecorHandleKey<xAOD::ExampleElectronContainer> m_decor2Key{
    this, "ExampleElectronContainerDecorKey2", "TestContainer.decor2",
    "decorator2 key"};
```

Figure 5.3: WriteExampleElectronheader file setup

Then the `WriteExampleElectron` algorithm is called and takes `ExampleTracks`, creates an `ExampleElectron` object and sets the electrons `pt` to the tracks `pt`. As shown in Figure 5.4, the `ExampleElectronContainer` and `ExampleElectronAuxContainer` are created and set to the `elecCont` and `elecStore` respectively. The `elecCont` has an associated aux store, so the `setStore` function is called with the `elecStore` pointer. The track container is

```
auto elecCont = std::make_unique<xAOD::ExampleElectronContainer>();
auto elecStore = std::make_unique<xAOD::ExampleElectronAuxContainer>();
elecCont->setStore(elecStore.get());

SG::ReadHandle<ExampleTrackContainer> trackCont(m_exampleTrackKey, ctx);
elecCont->push_back(std::make_unique<xAOD::ExampleElectron>());

for (const ExampleTrack* track : *trackCont) {
    // Take on the pT of the track
    elecCont->back()->setPt(track->getPT());
}

SG::WriteHandle<xAOD::ExampleElectronContainer> objs(
    m_exampleElectronContainerKey, ctx);
ATH_CHECK(objs.record(std::move(elecCont), std::move(elecStore)));
```

Figure 5.4: Algorithm to initialize and write T/P data (ExampleTracks) to an xAOD object container (ExampleElectronContainer).

accessed by using StoreGate's `ReadHandle`, which associates the `m_exampleTrackKey` with the `ExampleTrackContainer` specified in the header file. This is then looped over all elements in the container and the `pt` of each track is set to the `pt` of the electron. A `WriteHandle`, called `objs`, is then created for the container of `ExampleElectrons` which is then recorded.

Within the same algorithm, the next step is to loop over each of the newly produced `ExampleElectrons`, accessing the decorations `decor1` and `decor2`, and setting each to an arbitrary float value that are easily identifiable later. Figure 5.5 shows how this is done using two handles for each decoration. Note the difference here using the `WriteDecorHandle`, where the prior handle type was `WriteHandle`.

5.2.2 ReadxAODElectron.py

The only algorithm called in this test is `ReadExampleElectron`. The header file for the `ReadExampleElectron` only creates `ReadHandleKey` for the container of ExampleElectrons, with the same name from the header of the `WriteExampleElectron` algorithm header,

```
SG::WriteDecorHandle<xAOD::ExampleElectronContainer, float> hdl1(
    m_decor1Key,ctx);
SG::WriteDecorHandle<xAOD::ExampleElectronContainer, float> hdl2(
    m_decor2Key,ctx);

for (const xAOD::ExampleElectron* obj : *objs) {
     hdl1(objs) = 123.;
     hdl2(objs) = 456.;
}
```

Figure 5.5: Writing of dynamic variables for each of the ExampleElectron objects.

syntax shown in Figure 5.6. From the source file, we can initialize the `ReadHandleKey`

```
SG::ReadHandleKey<xAOD::ExampleElectronContainer>
m_exampleElectronContainerKey{this, "ExampleElectronContainerName",
                                    "TestContainer"};
```

Figure 5.6: ReadHandleKey for the container of ExampleElectrons

object by a simple `ATH_CHECK(m_exampleElectronContainerKey.initialize());` in the `initialize()` method. This allows for, when defining the `ReadHandle` in execute, identifying the correct container defined in the header file. The same can be done for the decoration key, which needs a separate read handle, `ReadDecorHandle`. Once this is setup, all the read algorithm needs to do is to loop over all the `ExampleElectrons` in the "TestContainer" and access their p_T and charge.

5.3 Results

This project sought to replace existing unit tests that created `ExampleHits`, T/P EDM objects, to be written and read back. An independent xAOD object, `ExampleElectron`, was created and implemented into two new unit tests that write and read `ExampleElectron` objects along with their chosen dynamic attributes. A merge request was created, approved, and merged into the Athena software framework. Future work can be done to fully modernize the package these unit tests reside, `AthenaPoolExampleAlgorithms`, including unit tests that test core functionality of AthenaMT/AthenaMP, and newer storage formats like RNTuple.

CHAPTER 6

CONCLUSION

The work done for this thesis was primarily motivated to find avenues to optimize resource usage for GRID I/O operations. The toy model testing allowed us to create branches with data similar compression ratios to real and simulated data, allowing to investigate the hypothesis that modifying the basket buffer limit had an effect on disk and memory usage. It led to the conclusion that, upon investigating with real data and real MC simulation, that there might be an avenue to look at both ROOT and Athena to limit basket sizes.

Modifying the basket buffer sizes at the Athena level shows there was a balance was struck when using the Athena basket buffer size limited to 128 kB between memory-usage and the size of the DAOD to be saved long-term. Removing the basket buffer size limit, the 5.5% saving in PHYSLITE MC disk-usage at the expense of an 11% increase in memory-usage could be a trade-off worth making in some scenarios. A class of potentially unoptimized AOD branches in MC simulated data was also brought to light during this study. The leading indicator to potential optimization is the highly compressible nature of these branches post-derivation. Further work could be done to look into these AOD branches to identify areas where further work can be done to reduce the overall AOD footprint.

The xAOD EDM comes with a number of new additions to bring about optimization the future of analysis work at the ATLAS experiment. Integrating the new features into a few comprehensive unit tests allow for the nightly CI builds to catch any issues that break core I/O functionality as it pertains to the xAOD EDM, which has not been done before. These new unit-tests exercise reading and writing select decorations ontop of the already existing data structures attached to an example object called `ExampleElectron`.

BIBLIOGRAPHY

[1] Jean-Luc Caron for CERN. *LHC Illustration showing underground locations of detectors*. 1998. URL: `https://research.princeton.edu/news/princeton-led-group-prepares-large-hadron-collider-bright-future` (cit. on p. 2).

[2] Oliver Sim Bruning et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: `10.5170/CERN-2004-003-V-1`. URL: `https://cds.cern.ch/record/782076` (cit. on p. 1).

[3] *ATLAS: technical proposal for a general-purpose pp experiment at the Large Hadron Collider at CERN*. LHC technical proposal. Geneva: CERN, 1994. DOI: `10.17181/CERN.NR4P.BG9K`. URL: `https://cds.cern.ch/record/290968` (cit. on pp. 1, 5).

[4] Nir Amram. "Hough Transform Track Reconstruction in the Cathode Strip Chambers in ATLAS". Presented on 19 Mar 2008. Tel Aviv, Tel Aviv U., 2008. URL: `https://cds.cern.ch/record/1118033` (cit. on p. 3).

[5] Beniamino Di Girolamo and Marzio Nessi. *ATLAS undergoes some delicate gymnastics*. 2013. URL: `https://cerncourier.com/a/atlas-undergoes-some-delicate-gymnastics/` (cit. on p. 4).

[6] G Aad et al. "ATLAS pixel detector electronics and sensors". In: *Journal of Instrumentation* 3.07 (July 2008), P07007. DOI: `10.1088/1748-0221/3/07/P07007`. URL: `https://dx.doi.org/10.1088/1748-0221/3/07/P07007` (cit. on p. 2).

[7] Glenn F. Knoll. *Radiation Detection and Measurement*. New York: John Wiley & Sons, Inc., 2010 (cit. on p. 3).

[8] A. Abdesselam et al. “The barrel modules of the ATLAS semiconductor tracker”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 568.2 (2006), pp. 642–671. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2006.08.036. URL: https://www.sciencedirect.com/science/article/pii/S016890020601388X (cit. on p. 3).

[9] A Andreazza. *The ATLAS Pixel Detector operation and performance*. Tech. rep. Geneva: CERN, 2010. URL: https://cds.cern.ch/record/1287089 (cit. on p. 4).

[10] The ATLAS TRT collaboration et al. “The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance”. In: *Journal of Instrumentation* 3.02 (Feb. 2008), P02013. DOI: 10.1088/1748-0221/3/02/P02013. URL: https://dx.doi.org/10.1088/1748-0221/3/02/P02013 (cit. on p. 4).

[11] Bartosz Mindur. *ATLAS Transition Radiation Tracker (TRT): Straw tubes for tracking and particle identification at the Large Hadron Collider*. Geneva, 2017. DOI: 10.1016/j.nima.2016.04.026. URL: https://cds.cern.ch/record/2139567 (cit. on p. 4).

[12] *ATLAS muon spectrometer: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331068 (cit. on p. 5).

[13] ATLAS Experiment at CERN. *Trigger and Data Acquisition*. URL: https://atlas.cern/Discover/Detector/Trigger-DAQ (cit. on p. 7).

[14] ATLAS Outreach. “ATLAS Fact Sheet : To raise awareness of the ATLAS detector and collaboration on the LHC”. 2010. DOI: 10.17181/CERN.1LN2.J772. URL: https://cds.cern.ch/record/1457044 (cit. on p. 7).

[15] K. Bos et al. *LHC computing Grid: Technical Design Report. Version 1.06 (20 Jun 2005)*. Technical design report. LCG. Geneva: CERN, 2005. URL: https://cds.cern.ch/record/840543 (cit. on pp. 7, 13).

[16] E Martelli and S Stancu. "LHCOPN and LHCONE: Status and Future Evolution". In: *Journal of Physics: Conference Series* 664.5 (Dec. 2015), p. 052025. DOI: 10.1088/1742-6596/664/5/052025. URL: https://dx.doi.org/10.1088/1742-6596/664/5/052025 (cit. on p. 7).

[17] ATLAS software group. *Athena Software Documentation*. URL: https://atlassoftwaredocs.web.cern.ch/athena/ (cit. on p. 7).

[18] Georges Aad et al. *Software and computing for Run 3 of the ATLAS experiment at the LHC*. Tech. rep. Geneva: CERN, 2024. arXiv: 2404.06335. URL: https://cds.cern.ch/record/2895022 (cit. on pp. 8, 13, 14).

[19] J. Catmore. "The ATLAS data processing chain: from collision to paper". Joint Oslo/Bergen/NBI ATLAS Software Tutorial. University of Oslo, 2016. URL: https://indico.cern.ch/event/472469/contributions/1982677/attachments/1220934/1785823/intro_slides.pdf (cit. on p. 8).

[20] I. Bejar Alonso et al. "High-Luminosity Large Hadron Collider (HL-LHC): Technical design report". In: 10 (2020), p. 390. DOI: https://doi.org/10.23731/CYRM-2020-0010. URL: https://e-publishing.cern.ch/index.php/CYRM/issue/view/127 (cit. on p. 9).

[21] J Elmsheuser et al. "Evolution of the ATLAS analysis model for Run-3 and prospects for HL-LHC". In: *EPJ Web Conf.* 245 (2020), p. 06014. DOI: 10.1051/epjconf/202024506014. URL: https://doi.org/10.1051/epjconf/202024506014 (cit. on p. 9).

[22] *ATLAS HL-LHC Computing Conceptual Design Report*. Tech. rep. Geneva: CERN, 2020. URL: https://cds.cern.ch/record/2729668 (cit. on p. 9).

[23] Javier Lopez-Gomez and Jakob Blomer. "RNTuple performance: Status and Outlook". In: *Journal of Physics: Conference Series* 2438.1 (Feb. 2023), p. 012118. DOI: 10.1088/1742-6596/2438/1/012118. URL: https://dx.doi.org/10.1088/1742-6596/2438/1/012118 (cit. on p. 10).

[24] Blomer, Jakob et al. "ROOT's RNTuple I/O Subsystem: The Path to Production". In: *EPJ Web of Conf.* 295 (2024), p. 06020. DOI: 10.1051/epjconf/202429506020. URL: https://doi.org/10.1051/epjconf/202429506020 (cit. on p. 10).

[25] A. Buckley et al. *Report of the xAOD Design Group*. 2013. URL: https://cds.cern.ch/record/1598793/files/ATL-COM-SOFT-2013-022.pdf (cit. on p. 12).

[26] A. Buckley et al. "Implementation of the ATLAS Run 2 event data model". In: *Journal of Physics: Conference Series* 664.7 (Dec. 2015), p. 072045. DOI: 10.1088/1742-6596/664/7/072045. URL: https://dx.doi.org/10.1088/1742-6596/664/7/072045 (cit. on pp. 12, 42).

[27] ATLAS software group. *Athena*. URL: https://doi.org/10.5281/zenodo.2641997 (cit. on p. 13).

[28] ROOT Team. *ROOT, About*. URL: https://root.cern/about/ (cit. on p. 15).

[29] ROOT Team. *ROOT, TTree Class*. 2024. URL: https://root.cern.ch/doc/master/classTTree.html (cit. on p. 15).

[30] R Trentadue et al. "LCG Persistency Framework (CORAL, COOL, POOL): Status and Outlook in 2012". In: *Journal of Physics: Conference Series* 396.5 (Dec. 2012), p. 052067. DOI: 10.1088/1742-6596/396/5/052067. URL: https://dx.doi.org/10.1088/1742-6596/396/5/052067 (cit. on p. 16).

[31] *Athena gitlab repository*. URL: https://gitlab.cern.ch/atlas/athena (cit. on p. 17).

[32] *Jenkins*. URL: https://www.jenkins.io (cit. on p. 17).

[33] Schaarschmidt, Jana et al. "PHYSLITE - A new reduced common data format for ATLAS". In: *EPJ Web of Conf.* 295 (2024), p. 06017. DOI: 10.1051/epjconf/202429506017. URL: https://doi.org/10.1051/epjconf/202429506017 (cit. on p. 17).

[34] P. J. Laycock et al. "Derived Physics Data Production in ATLAS: Experience with Run 1 and Looking Ahead". In: *Journal of Physics: Conference Series* 513.3 (June 2014), p. 032052. DOI: 10.1088/1742-6596/513/3/032052. URL: https://dx.doi.org/10.1088/1742-6596/513/3/032052 (cit. on p. 18).

[35] Martin Barisits et al. "Rucio: Scientific Data Management". In: *Computing and Software for Big Science* 3.1 (2019), p. 11. DOI: 10.1007/s41781-019-0026-3. URL: https://doi.org/10.1007/s41781-019-0026-3 (cit. on p. 34).

[36] Alaettin Serhan Mete and Peter van Gemmeren. "Shared I/O Developments for Run 3 in the ATLAS Experiment". In: *PoS* ICHEP2022 (2022), p. 219. DOI: 10.22323/1.414.0219 (cit. on p. 35).

APPENDIX A

DERIVATION PRODUCTION DATA

A.1 Derivation production datasets

For both the nightly and the release testing, the data derivation job, which comes from the dataset

```
data22_13p6TeV:data22_13p6TeV.00428855.physics_Main.merge.AOD.
    r14190_p5449_tid31407809_00
```

was ran with the input files

```
AOD.31407809._000894.pool.root.1
AOD.31407809._000895.pool.root.1
AOD.31407809._000896.pool.root.1
AOD.31407809._000898.pool.root.1
```

Similarly, the MC derivation job, comes from the dataset

```
mc23_13p6TeV:mc23_13p6TeV.601229.PhPy8EG_A14_ttbar_hdamp258p75_
    SingleLep.merge.AOD.e8514_e8528_s4162_s4114_r14622_r14663_
    tid33799166_00
```

was ran with input files

```
AOD.33799166._000303.pool.root.1
AOD.33799166._000304.pool.root.1
AOD.33799166._000305.pool.root.1
AOD.33799166._000306.pool.root.1
AOD.33799166._000307.pool.root.1
AOD.33799166._000308.pool.root.1
```

APPENDIX B

ATHENA CONFIGURATION JOB

B.1 Athena job configuration example

An Athena job configuration is a script that allows the user to steer a specific program in the framework. Steering allows one to, at a high-level, configure low-level behavior of any kind of production job. A general Athena application using `ComponentAccumulator` written in pseudocode would take the form:

```
# Import Packages
from AthenaConfiguration.AllConfigFlags import initConfigFlags
from AthenaConfiguration.ComponentFactory import CompFactory
from OutputStreamAthenaPool.OutputStreamConfig import OutputStreamCfg ,
outputStreamName

# Configure Output
outputStreamName = "StreamA"
outputFileName = "output.root"

# Setup flags
flags = initConfigFlags()
flags.Input.Files = ["input.root"]
flags.addFlag(f"Output.{streamName}FileName", outputFileName)
flags.lock()

# Main Service(s)
from AthenaConfiguration.MainServicesConfig import MainServicesCfg
acc = MainServicesCfg( flags )

# Add algorithms to the accumulator
acc.addEventAlgo( CompFactory.MyAlgorithm(MyParameters) )

```

```python
    # Run
    import sys
    sc = acc.run(flags.Exec.MaxEvents)
```

The `acc` is the `ComponentAccumulator`, so here the user might have more than one Algorithm it needs to call, but each one would have a separate `.addEventAlgo` call. When `flag.lock()` is called, any previously established flags will be set in place and unable to be changed.