Physics Validation at ATLAS

HSF WLCG workshop

13/05/2020

J. Ferrando, I. Ochoa

Outline

- Overview of Physics validation at ATLAS
- General challenges for physics validation
- Challenges for Heterogeneous resources
- Challenges for full automation

Physics Validation

The role of physics validation within ATLAS is to check the effects of software changes in:

- Simulation software (full and fast simulation)
- Reconstruction software

Checks are then performed at the level of comparing the differences in distributions related to physics objects e.g. jets, electrons, photons, muons, etc. to verify that any changes are consistent with expectations.

Hierarchy of validation

1. Continuous integration:
    a. Can check O(10s) of events running locally
    b. Good for testing identical output from e.g. code restructuring/reconfiguration
2. ATLAS Release Tester (ART):
    a. Can run on grid and can check O(10k) events or shorter jobs locally
    b. Can test similarity of output with more statistical precision
    c. Possible to compare distributions and check for expected behavioural changes
3. Physics Validation:
    a. Run on many O(100k) samples of different physics processes
    b. Possible to check results are identical with yet better statistical precision
    c. Best for comparing distributions to look for expected behavioural changes
    d. When evaluating the effects of major changes (e.g. before a reprocessing) also run with physics analysis workflows (specific objects and regions of phase-space)

Description of the typical physics validation workflow

- Triggered by "major" update(s) to a development branch.
- A collection of MC samples of different physics processes are produced:
  - Several SM and Exotics processes, to test all aspects of object reconstruction.
- Samples are produced on the grid via central production system:
  - Large statistics (~100k events) to expose minor effects.
  - Production configured via the ATLAS Metadata Interface (AMI).
- Steps: generator level → simulation → digitization → reconstruction
  - For validation of reconstruction, start from same simulation samples.
- Output is collection of variables needed to verify quality of detector performance and single object reconstruction.
- The validation is performed by experts and against a stable (and previously validated) version of the software.
- A common framework is available for making histogram comparisons including a chi-square test and creating webpages for results.
- Outcome is documented and further development and bug-fixes tracked on JIRA.

Challenges for physics validation (I)

"Non-physics" sources of non-reproducibility:

- Different math libraries
- Different architectures
- Different random number seeds
- Re-digitization of samples (grid assigns different pile-up files)
- Different order of event processing (e.g. AthenaMP jobs)

Goal is to validate distributions, not event-by-event comparisons.

Challenges for physics validation (II)

Understanding results when non-identical output is expected:

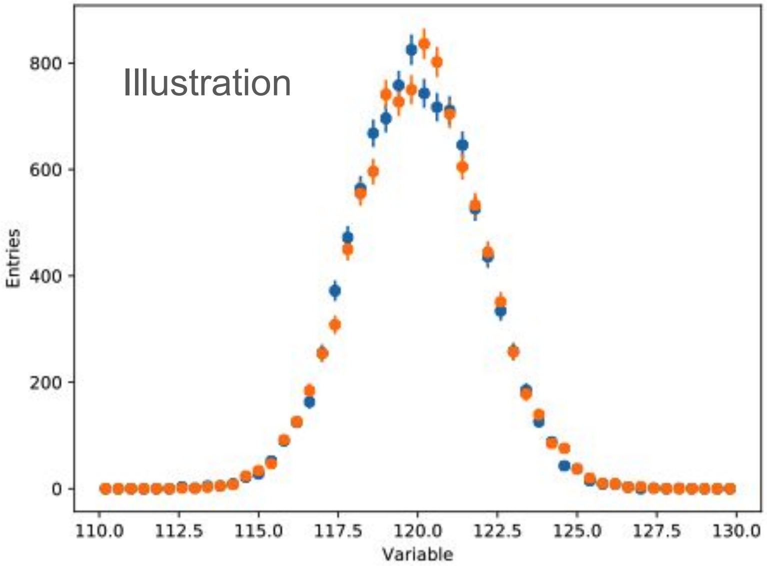
- When less precise output is used e.g. when introducing lossy compression.
- When non-identical but equivalent results are expected
    - e.g. some simulation change which alters random number seeds or is a close approximation of reference.
- When changes in low level objects propagate to higher-level ones (here Combined performance group input highly valuable)

Critical to have input from physics analyses: required precision levels will vary.

- For release changes/major reprocessing campaigns, a selection of analyses from all physics groups of ATLAS participate in the validation effort (aiming to cover all interesting phase-space for physics!).

Understanding Statistical Uncertainties

- Usually substantial correlations between events that have been put through simulation/reconstruction.
- Residual statistical variation from usage of random seeds through GEANT.
- Additional statistical uncertainty when including pile-up.
- We tested this using samples with different random seeds
- Working to quantify by generating ensemble of samples where only random seeds are changed.



Heterogeneous resources

- We can expect differences from different libraries, precision, ...
- Diverse configurations currently in use, e.g.:
    - X86_64-slc6-gcc62-opt, x86_64-centos7-gcc8-opt, x86_64-slc6-gcc48-opt x86_64-slc6-gcc49-opt
- HPC / MP running: different workflows require technical validation of the ATLAS Event Service.

We are working to quantify the variations from distinct workflows by generating ensemble of samples in different configurations.

Can we fully automate Physics validation?

For cases where fully identical results are expected, yes, e.g.:

- Moving algorithms to multi-threading

also for equivalent result cases, if effects of non-reproducibility are quantified/controlled for and there is a well-defined criteria for agreement.

Automation of validation when larger changes expected is difficult, e.g.:

- Optimization of full or fast simulation
- Changes to reconstruction of low-level objects

Summary and Outlook

- ATLAS software is continuously being improved and both technical and physics-related updates are expected.
- Physics validation ensures that rare effects are caught and that the impact of major updates is understood.
- Impact from sources of non-reproducibility, namely due to heterogeneous resources, is being evaluated.
- Important to contextualize in terms of precision with which physics analyses are carried out (which also depends on the physics).