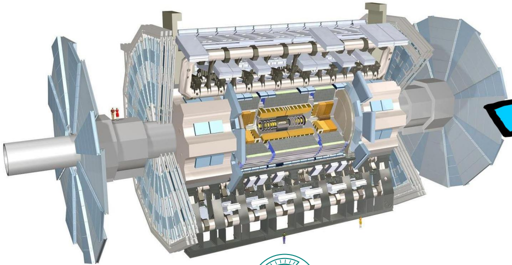
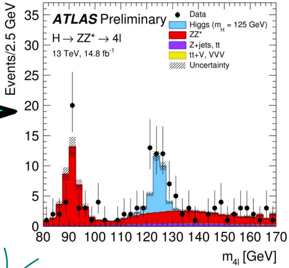
Making use of experimental data: Computing and analysis

Ludovic Scyboz

Max-Planck-Institut für Physik





MAX-PLANCK-GESELLSCHAFT

Max-Planck-Institut für Physik

From the detector output to the physics

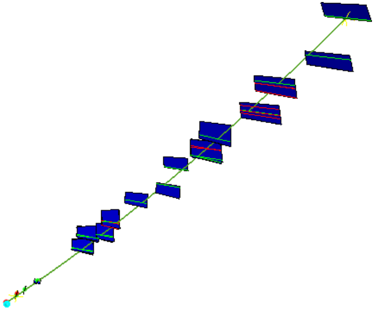
- Goal: **store/manage** the data

- **reconstruct** objects
- and **extract** the physics

From the detector output to the physics

- Goal: **store/manage** the data

- **reconstruct** objects
- and **extract** the physics

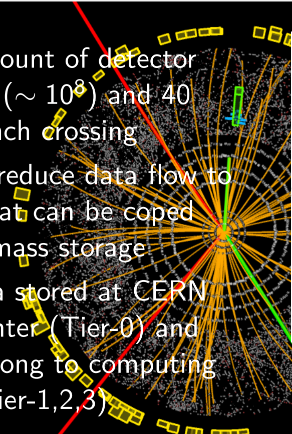


Outline

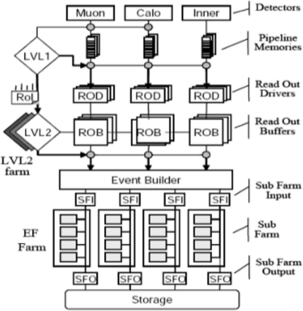
- The ATLAS Computing Model
  - What happens with the raw data?
- The Event Data Model
  - Or how to make everything readily accessible
  - The Athena framework
- Physics analysis and the production chain
  - Monte Carlo and the Grid
  - The full treatment: from generation to reconstruction
- Conclusion

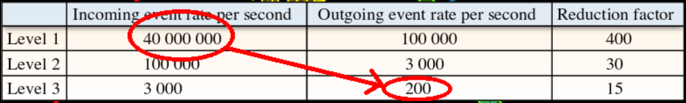
Trigger and Data Acquisition

ATLAS EXPERIMENT
http://atlas.ch



- Huge amount of detector channels ($\sim 10^8$) and 40 MHz bunch crossing
- Need to reduce data flow to values that can be coped with by mass storage
- Raw data stored at CERN Data Center (Tier-0) and passed along to computing farms (Tier-1,2,3)



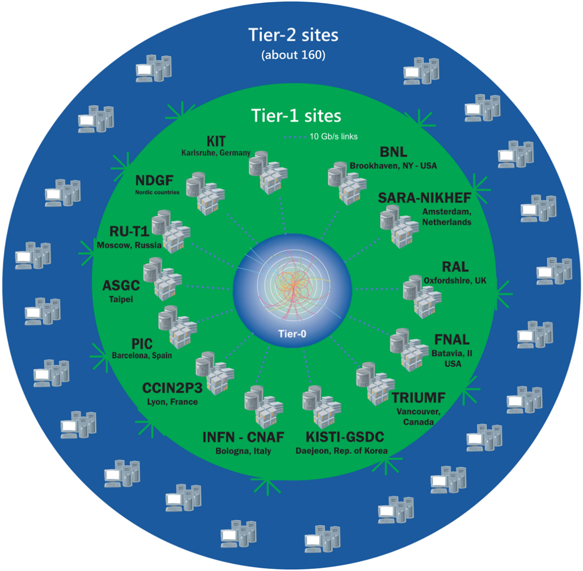


	Incoming event rate per second	Outgoing event rate per second	Reduction factor
Level 1	40 000 000	100 000	400
Level 2	100 000	3 000	30
Level 3	3 000	200	15

Event rate after each trigger level (Level-1, Level-2, Event Filter)

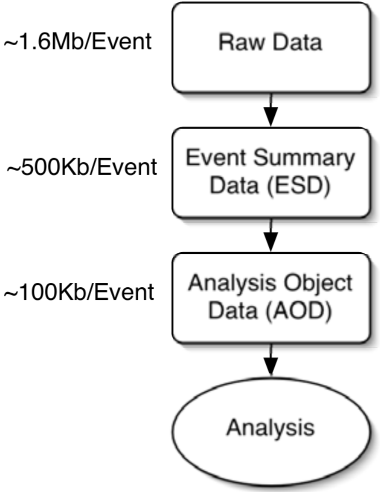
Computing model

- Tier-0: CERN Data Center
- Tier-1: Support for Tier-0
- Tier-2: Universities/institutes
- Tier-3: Local clusters/individuals



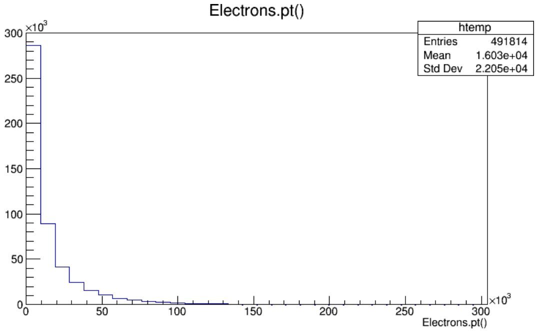
The Event Data Model: data formats

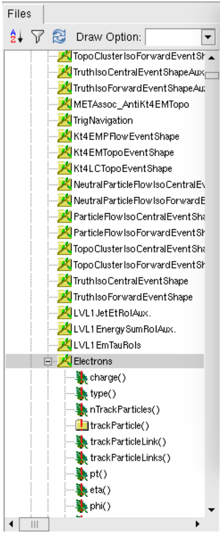
- **RAW**
- **ESD** (Event Summary Data): reconstructed detector output → information used for particle identification, track fitting, jet calibration...
- **AOD** (Analysis Object Data): summary of event reconstruction with physics objects (electrons/muons, jets, ...)
  → see next slide!
- **TAG**: general features of the event, used to quickly select interesting events in ESDs or AODs



xAODs: analysis-oriented, derived data sets

- New format introduced for Run 2
- Combines AODs from Run 1 and the concept of derivation (skimmed/slimmed events)




Files
Draw Option:
TopoClusterIsoForwardEventSh
TruthIsoCentralEventShapeAux
TruthIsoForwardEventShapeAu
METAssoc_AntiKt4EMTopo
TrigNavigation
Kt4EMPFlowEventShape
Kt4EMTopoEventShape
Kt4LCTopoEventShape
NeutralParticleFlowIsoCentralEv
NeutralParticleFlowIsoForwardE
ParticleFlowIsoCentralEventSha
ParticleFlowIsoForwardEventSh
TopoClusterIsoCentralEventSha
TopoClusterIsoForwardEventSh
TruthIsoCentralEventShape
TruthIsoForwardEventShape
LVL1JetEtRoIAux.
LVL1EnergySumRoIAux.
LVL1EmTauRoIs
Electrons
charge()
type()
nTrackParticles()
trackParticle()
trackParticleLink()
trackParticleLinks()
pt()
eta()
phi()

Reconstructed physics objects can be accessed and their properties used for plots, cuts, etc.

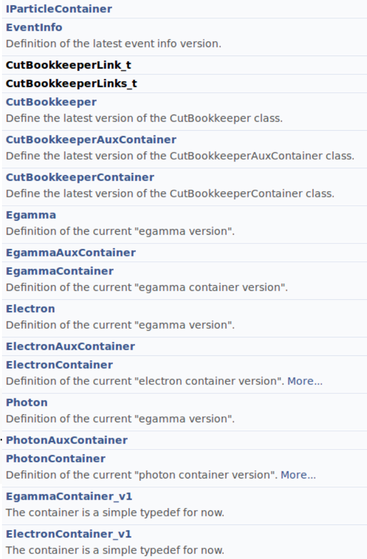
xAODs: why and how use them?

A collection of classes and types: to ensure commonality across the detector subsystems and subgroups such as trigger, test beam reconstruction, combined event reconstruction and physics analysis.

- **xAOD::EventInfo**: what's the pileup? What's the run and event number?
- **xAOD::IParticle**: interface for all particle types, clustered energy deposits and tracks

```
root [0] gROOT->Macro( "$ROOTCOREDIR/scripts/load_packages.C" );
root [1] xAOD::Init();
root [2] f = TFile::Open( "/afs/cern.ch/atlas/project/PAT/xAODs/.
root [3] t = xAOD::MakeTransientTree( f )
root [4] t->Draw( "ElectronCollection.pt() - ElectronCollection.
```

be directly handled in Athena (see next slide)!



IParticleContainer

EventInfo
Definition of the latest event info version.

CutBookkeeperLink_t

CutBookkeeperLinks_t

CutBookkeeper
Define the latest version of the CutBookkeeper class.

CutBookkeeperAuxContainer
Define the latest version of the CutBookkeeperAuxContainer class.

CutBookkeeperContainer
Define the latest version of the CutBookkeeperContainer class.

Egamma
Definition of the current "egamma version".

EgammaAuxContainer

EgammaContainer
Definition of the current "egamma container version".

Electron
Definition of the current "egamma version".

ElectronAuxContainer

ElectronContainer
Definition of the current "electron container version". More...

Photon
Definition of the current "egamma version".

PhotonAuxContainer

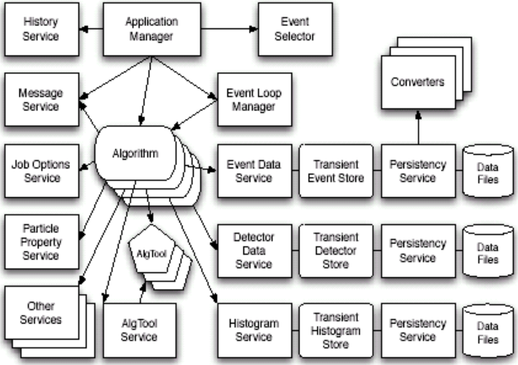
PhotonContainer
Definition of the current "photon container version". More...

EgammaContainer_v1
The container is a simple typedef for now.

ElectronContainer_v1
The container is a simple typedef for now.

The Athena Framework

- Basically, after Run I, most of the analysis code had grown naturally by itself

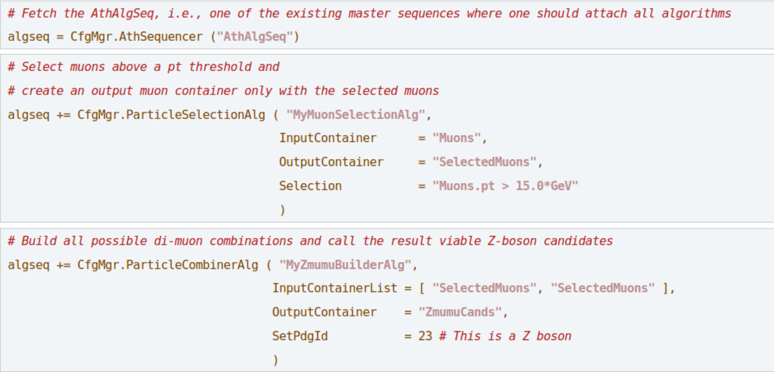


```
              AtlasProduction
                    |
               AtlasOffline
                    |
        +-----------+----------+
        |                      |
   AtlasAnalysis               |
        |                      |
    AtlasTrigger               |
        |                      |
AtlasReconstruction     AtlasSimulation
        |                      |
        +-----------+----------+
                    |
                AtlasEvent
                    |
             AtlasConditions
                    |
                AtlasCore
                    |
                  Gaudi
                    |
                  LCGCMT
```

- Need for a harmonized and modularized analysis framework

The Athena Framework: algorithm sequencing

- Physics analysis implemented sequentially
  - Calibration of the muons, jets, ...
  - Selection cuts
  - Histogramming

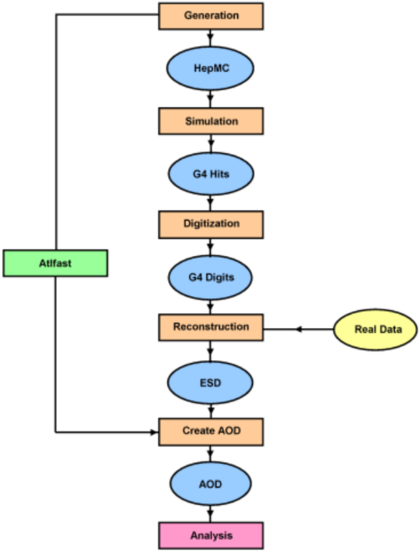


```
# Fetch the AthAlgSeq, i.e., one of the existing master sequences where one should attach all algorithms
algseq = CfgMgr.AthSequencer ("AthAlgSeq")
```

```
# Select muons above a pt threshold and
# create an output muon container only with the selected muons
algseq += CfgMgr.ParticleSelectionAlg ( "MyMuonSelectionAlg",
                                       InputContainer      = "Muons",
                                       OutputContainer     = "SelectedMuons",
                                       Selection           = "Muons.pt > 15.0*GeV"
                                       )
```

```
# Build all possible di-muon combinations and call the result viable Z-boson candidates
algseq += CfgMgr.ParticleCombinerAlg ( "MyZmumuBuilderAlg",
                                      InputContainerList = [ "SelectedMuons", "SelectedMuons" ],
                                      OutputContainer    = "ZmumuCands",
                                      SetPdgId           = 23 # This is a Z boson
                                      )
```

Monte Carlo production and comparison to data

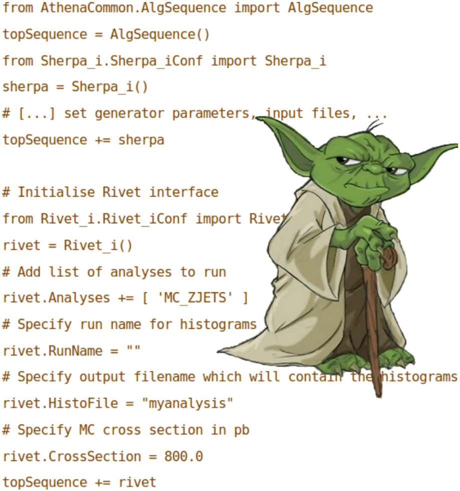


Schematic representation of the Full Chain Monte Carlo production.

- To account for detector inefficiencies, geometric acceptance, etc..., Monte Carlo-produced samples have to be simulated, digitized and reconstructed
- All steps can be run in parallel on the ATLAS Grid
- Also done in Athena!
- AODs can then be constructed and analyzed

Analyses: datasets and MC samples

- Lots of possible tools and custom analyses (C++, Python, ROOT...)
- Rivet is directly implemented in Athena as well
- Histogramming observables in YODA format: data and MC directly comparable

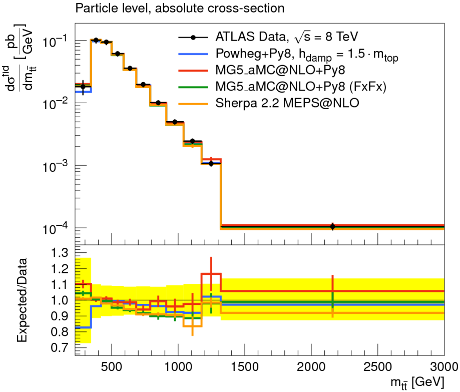


```
from AthenaCommon.AlgSequence import AlgSequence
topSequence = AlgSequence()
from Sherpa_i.Sherpa_iConf import Sherpa_i
sherpa = Sherpa_i()
# [...] set generator parameters, input files, ...
topSequence += sherpa

# Initialise Rivet interface
from Rivet_i.Rivet_iConf import Rivet
rivet = Rivet_i()
# Add list of analyses to run
rivet.Analyses += [ 'MC_ZJETS' ]
# Specify run name for histograms
rivet.RunName = ""
# Specify output filename which will contain the histograms
rivet.HistoFile = "myanalysis"
# Specify MC cross section in pb
rivet.CrossSection = 800.0
topSequence += rivet
```

Analyses: RIVET

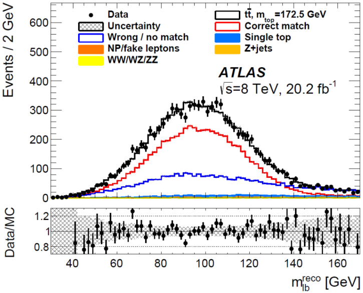
- Library of predefined functions for jets, event shapes, ...
- Based on physical objects with the help of projections:
  - Dressed electrons/muons
  - Jets (FastJet)
  - Final state hadrons
  - Reconstructed bosons
- Validated analyses with datasets available for download
- Plugin to write your own analyses



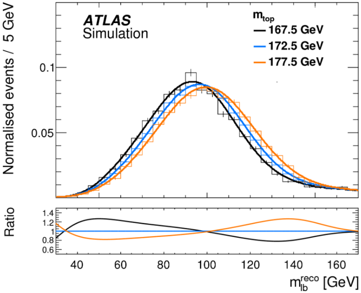
[ATL-PHYS-PUB-2016-020]

MC/data example: top mass determination in the dilepton channel

- Uses the template method: varying the top mass in Monte-Carlo



[arXiv:1606.02179]



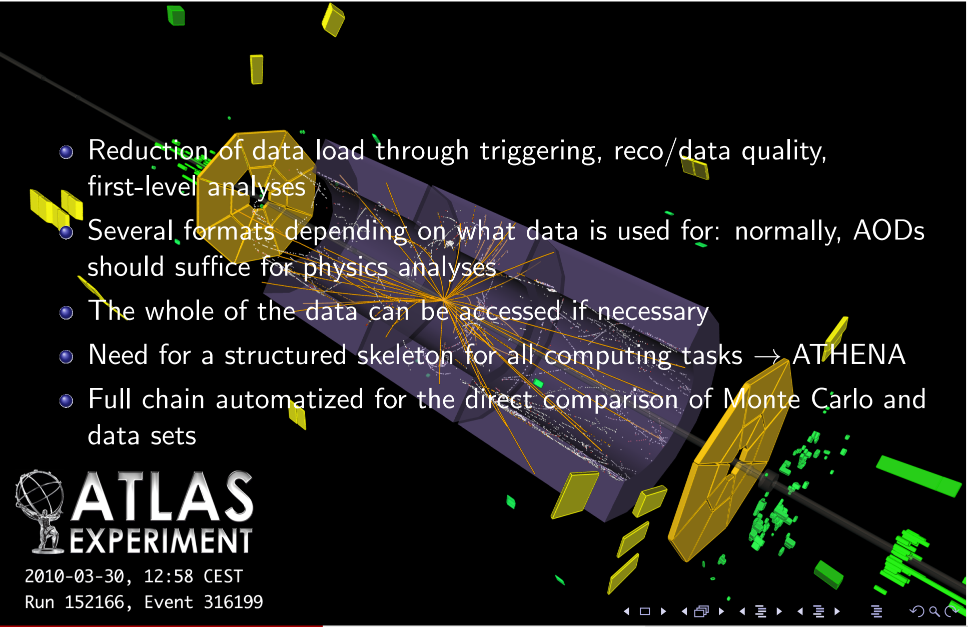
[arXiv:1606.02179]

- And fitting the template to the data

$$m_{\text{top}} = 172.99 \pm 0.41 \ (\text{stat.}) \pm 0.74 \ (\text{syst.}) \ \text{GeV}$$

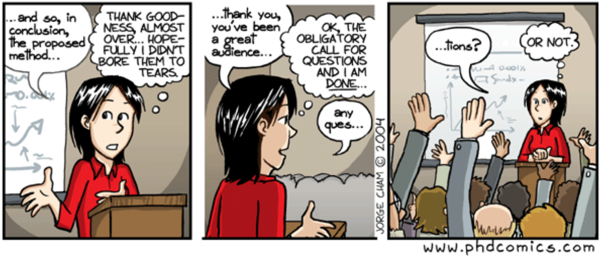
Conclusion

- Reduction of data load through triggering, reco/data quality, first-level analyses
- Several formats depending on what data is used for: normally, AODs should suffice for physics analyses
- The whole of the data can be accessed if necessary
- Need for a structured skeleton for all computing tasks → ATHENA
- Full chain automatized for the direct comparison of Monte Carlo and data sets



2010-03-30, 12:58 CEST
Run 152166, Event 316199

Ques...tions?


...and so, in conclusion, the proposed method...
THANK GOODNESS, ALMOST OVER... HOPEFULLY I DIDN'T BORE THEM TO TEARS.
...thank you, you've been a great audience...
OK, THE OBLIGATORY CALL FOR QUESTIONS AND I AM DONE...
any ques...
JORGE CHAM © 2004
...tions?
OR NOT.
www.phdcomics.com