A New Tool For Measuring Detector Performance in ATLAS

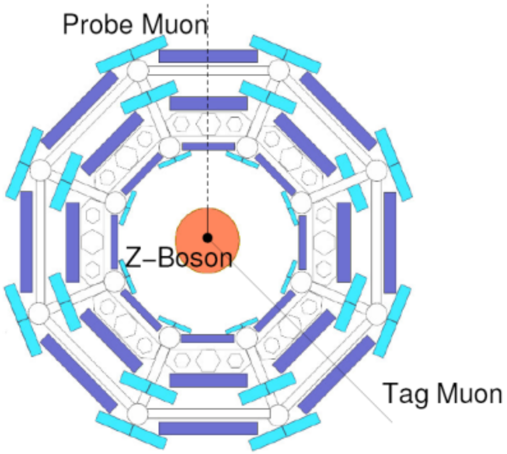
- Arno Straessner – TU Dresden
  Matthias Schott – CERN

on behalf of the ATLAS Collaboration

Computing in High Energy and Nuclear Physics
Prague
March 21-27, 2009

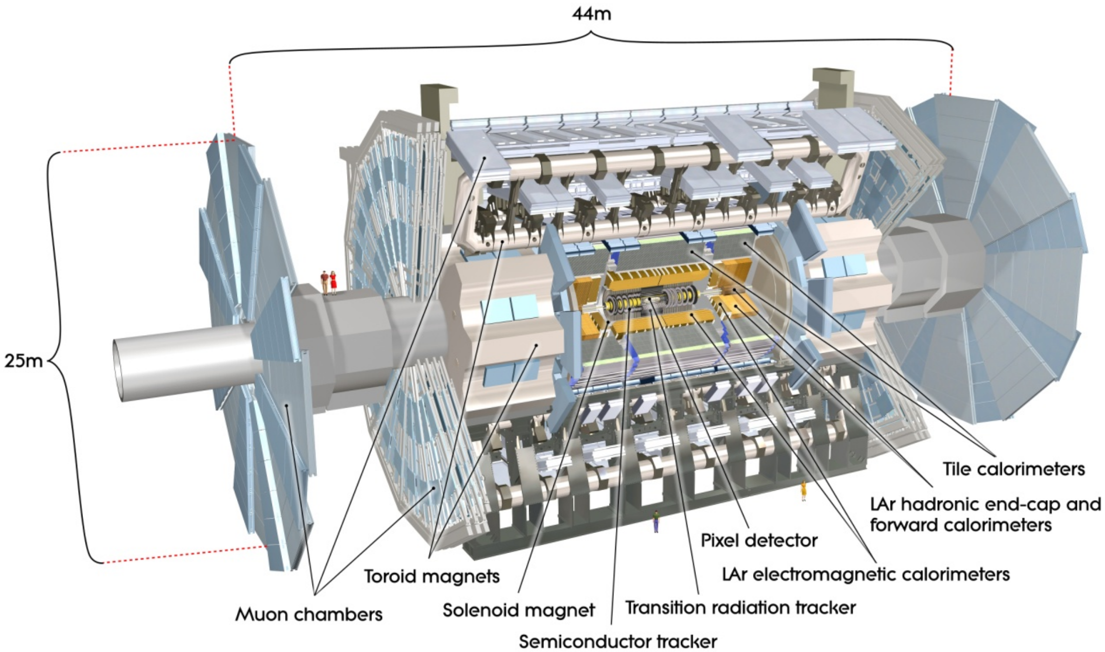
Outline

- The Physics: Detector Performance
- An Infrastructure For Performance Data
- Current Implementation in ATLAS
- Outlook



The Physics: Detector Performance

- Complex detector with main subsystems:
  - Inner tracking
  - Calorimetry
  - Muon tracking



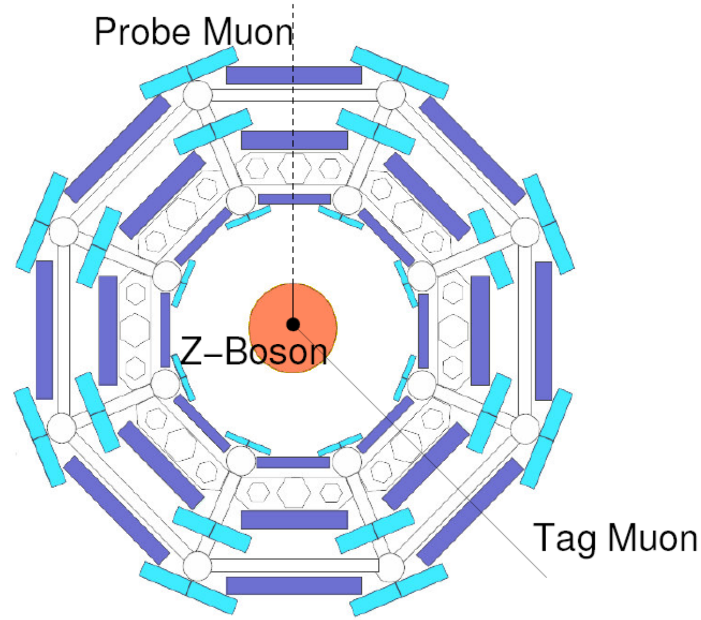
- 2500 physicists perform analyses

- All need detailed information about detector performance
  - General performance
    - Trigger, reconstruction and identification efficiencies
    - Resolutions of energy, momenta, angles, …
    - Energy and momentum scales

    } detector response
  - Time-dependent performance

Detector Performance Tool

- Commonly applied method to measure performance in data: → Tag & Probe
- Example: measure muon trigger efficiency in Z→μμ events
  - Identify triggered and well measured muon: → Tag
  - Use Z decay kinematics to find the 2nd muon: → Probe
  - Check if 2nd muon was triggered → efficiency

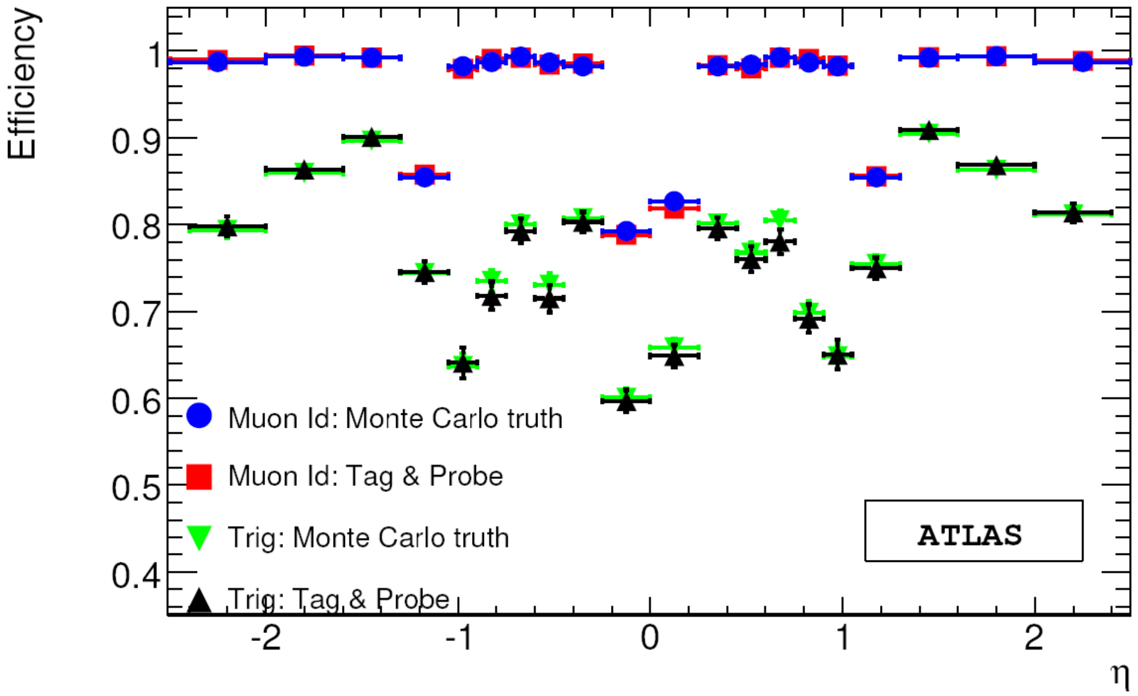


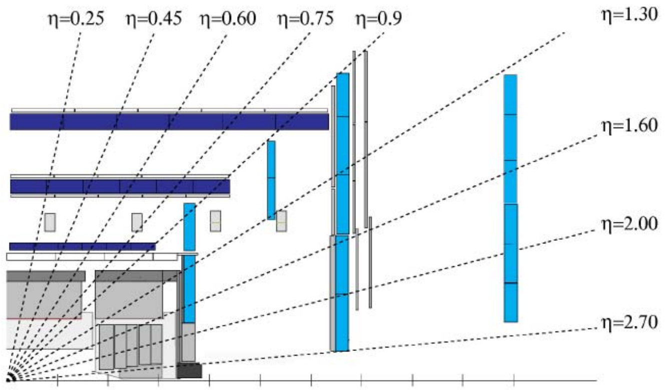
- Many more examples:
  - Z→ ee, J/psi → ee, Z+jets for tau fake rates, …
- Common infrastructure to
  - Implement the object tag and event selection
  - Store the collection of probe objects
  - Analyse the probes → efficiencies and detector response
  - Store matrices with efficiency, resolution and scale information

Clients:
- Performance groups
- Fast Monte Carlo simulation
- Individual Physicists

Application: Performance Groups

- Early version of the package was used in ATLAS physics book
- Estimation of muon identification and trigger efficiencies:





- Performance Tool is the ATLAS solution for:
  - Direct use in Standard Model physics → closely related to performance groups
  - Benchmark comparisons for reconstructed objects used in searches etc.

The Procedure

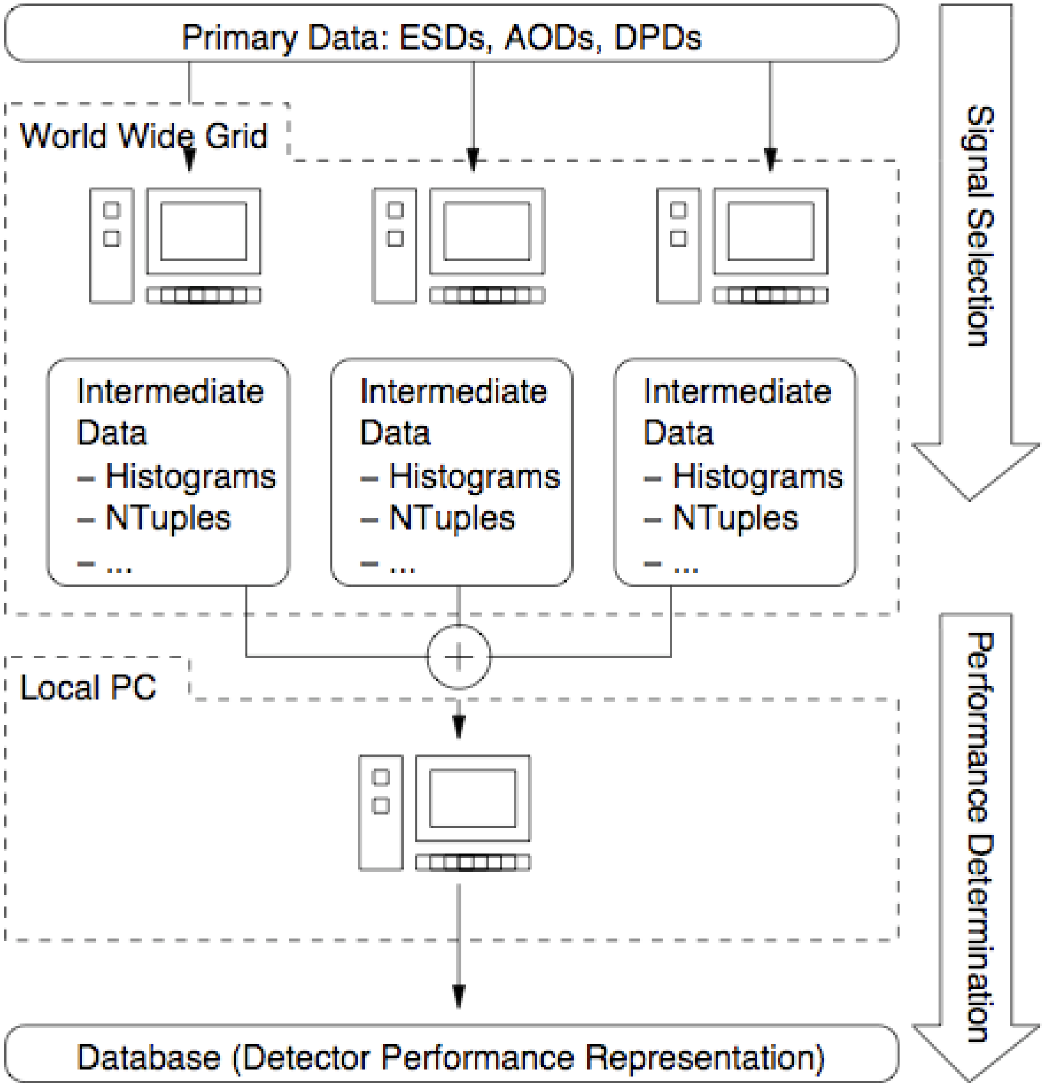
- ATLAS data storage
  - Event Summary Data (ESD)
  - Analysis Object Data (AOD)
  - Derived Physics Data (DPD)

- 1st step:
  - Signal selection and object tag
  - Input: ESD, AOD, DPD
  - Grid task

→ Performance-DPD with probe objects

- 2nd step:
  - Performance determination
  - Executed on local cluster
  - Overall fits might be necessary (background shape, …)

→ Performance database



Probe Object Data

- Probe objects are usually: tracks, reconstructed leptons, jets
- Only parameters necessary for further analysis are stored → can be freely defined
- Example for electron calorimeter identification efficiency:

Electron reconstruction category
Electron quality cut
Electron trigger information
Reconstructed Z mass from tag+probe electron

} To calculate number of trials and successes for efficiency

Z boson p_T
Angular difference (ΔR) to next jet
Jet multiplicity
Sum of jet p_T
Max of jet p_T
Electron isolation variables

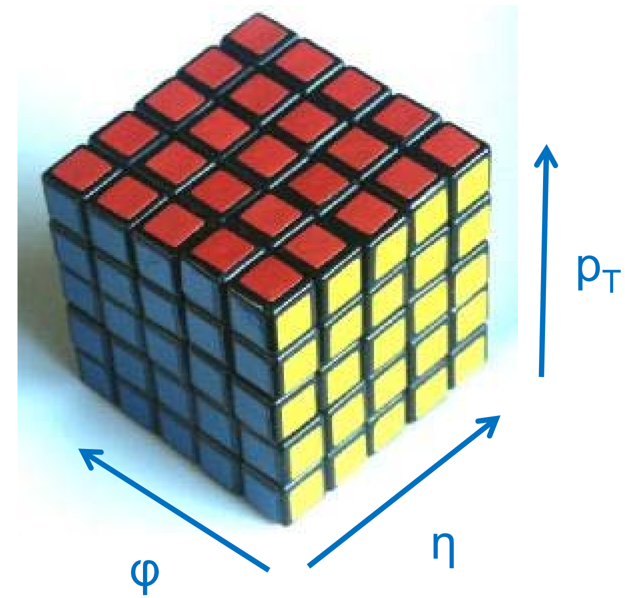
} Interesting variables for differential efficiencies

Matching angle to generator truth electron

} Optional for Monte Carlo verification of the method

- Map<tag,float> with string tags in meta-data → flexible, user-friendly structure
- Small data size: ~ 0.33 kB per event in DPD file

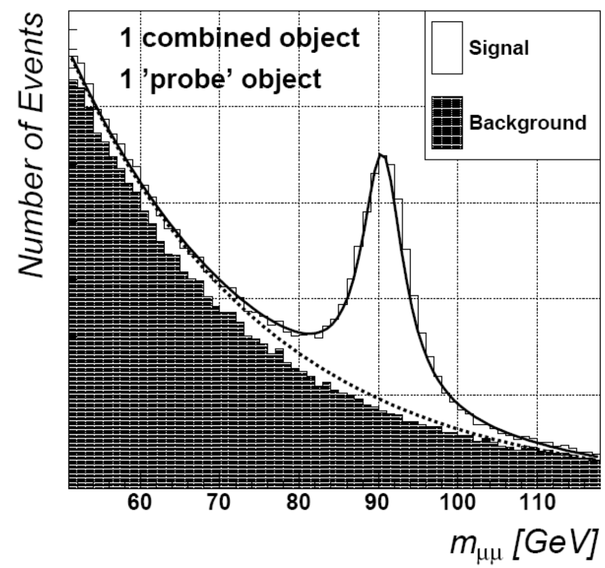
Performance Database



- Objects stored:
  - N-dimensional matrices to map detector areas and physics
- Typically 4-dimensional, not larger because of
  → data statistics per matrix entry
  → storage space
- Matrix defined by N axes objects with free binning
- Matrix entries:
- For "simple" efficiency calculations:

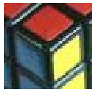
  - Number of trial and success counts
- With side-band subtraction of background:

  - Data and background histograms
  - Pre-defined fitting functions
  - Background subtraction assuming predicted S/B ratio

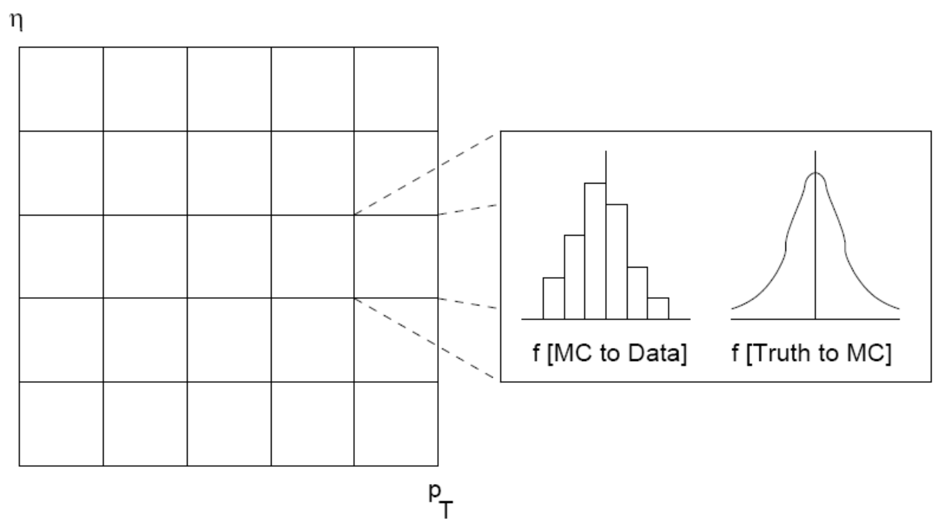


Performance Database

- Matrix entries for detector response:

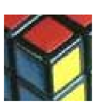


  - Resolution/scale histogram for Monte Carlo → data reference
  - Smearing functions for generator level → fully simulated Monte Carlo



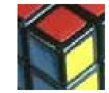
  - Once ATLAS data is available:
    - Smearing functions for generator level → data projections
    - To be used in fast detector simulations

- Matrix entries for resolution and scaling:

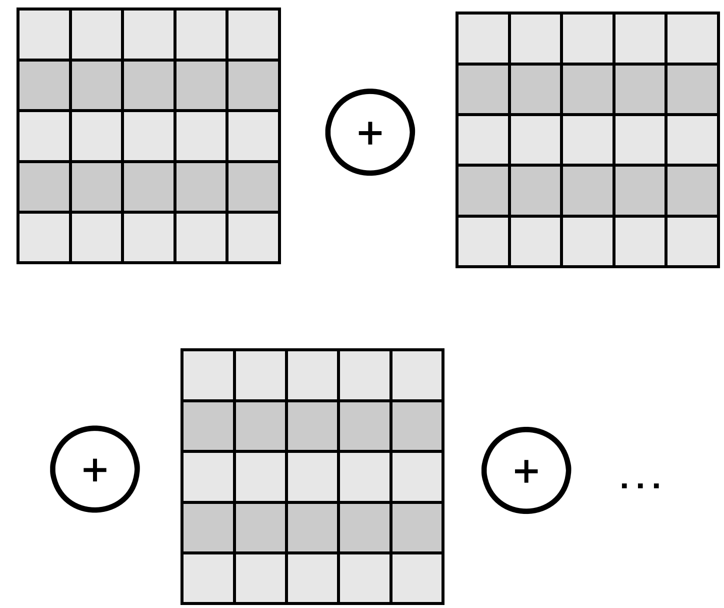


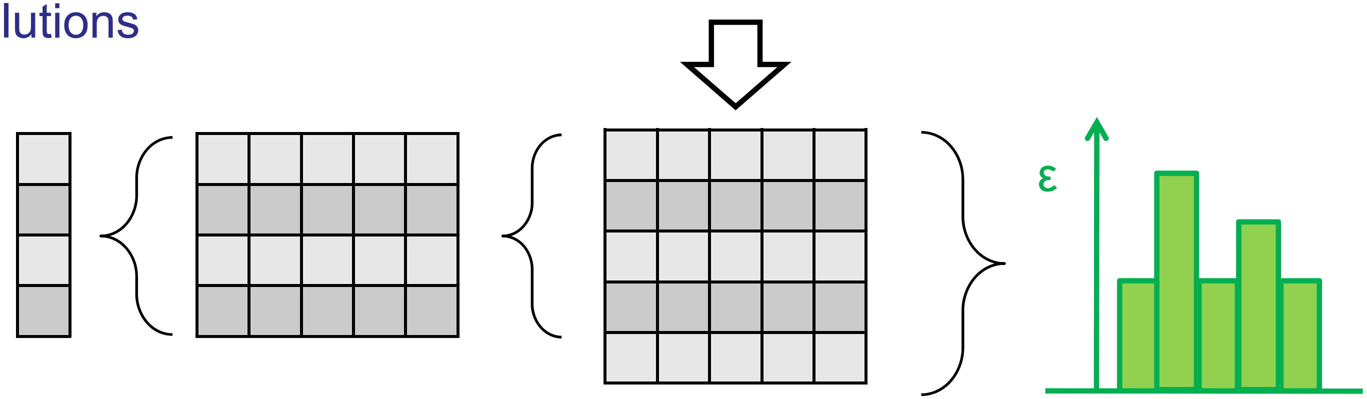
  - ROOT Histograms → in future: RooWorkspace of RooFit

Matrix Objects



- Methods to calculate efficiency and resolution from stored data
  - Classical and Bayesian efficiencies and uncertainties
  - Efficiencies and detector response using fits to data and background
  - Caching: fit result is stored and then directly accessed
- All matrix entry objects are additive:
  - Distributed analysis
  - Averaging over different run periods
- Projections to any number of axes is supported:
  - Projections in form of matrices
  - Projections into histograms
  - Slices of matrices → cuts
  - Not trivial for resolutions
- Hide complexity from user





Implementation in ATHENA

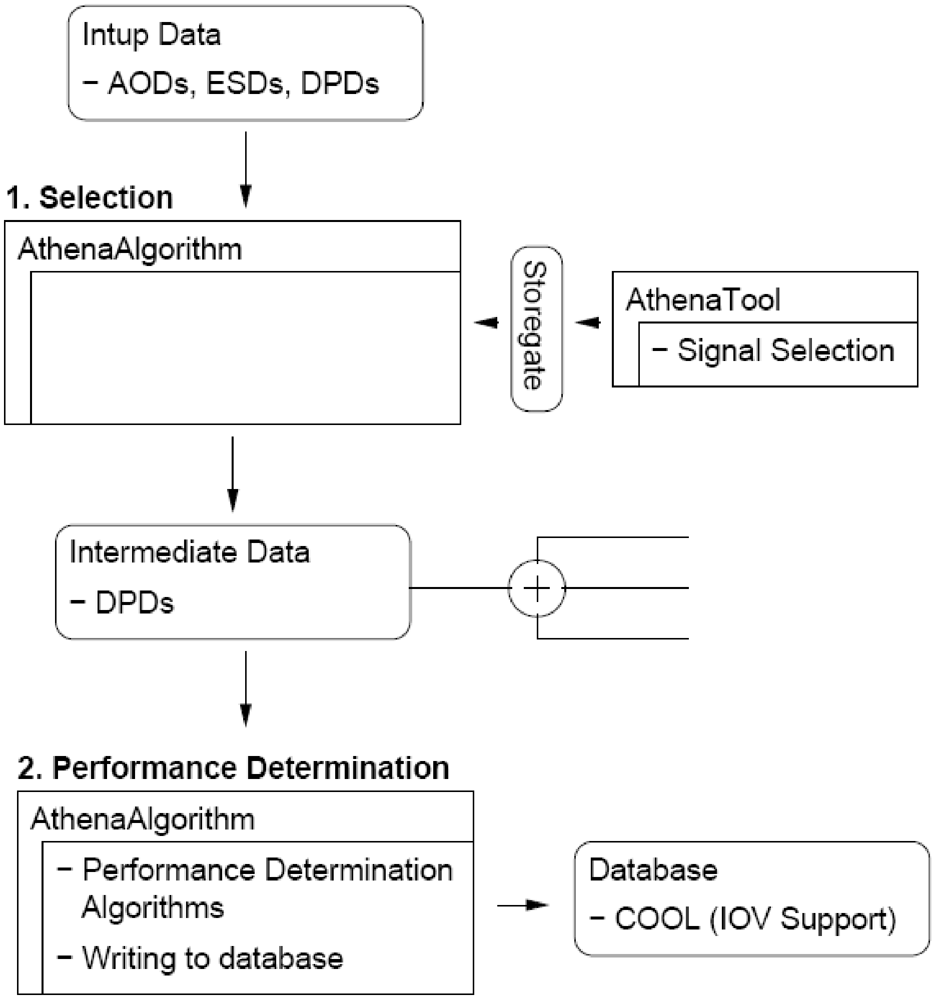
- ATHENA = ATLAS software framework

- “Algorithm” is executed at each event
- Uses “Tool” to perform signal selection

- Intermediate data is stored in
  - DPD
  - Transient memory “Storegate”

- Performance determination

- Output to ATLAS official conditions database:
  - LCG product: COOL database
  - With Interval-Of-Validity (IOV)



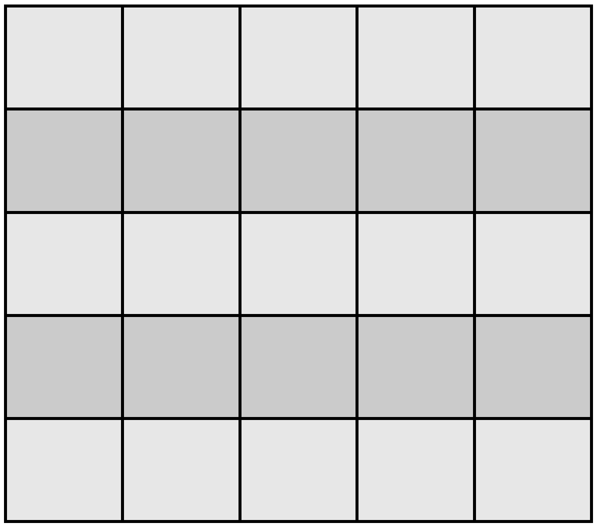
Supported Database Structures

- ATLAS official conditions database: LCG Pool/COOL → for collaborative use
- Plain ROOT file → for online Tier-0 applications, development, debugging and individual use
- Unique database identifier:

Database ID	Description	Examples
Object	Physics Parameter	Muon, Electron, Tau, Jet, ….
Type	Performance Parameter	Efficiency, Scale/Resolution, Fake-Rate
Channel	Physics Channel	$Z \rightarrow \mu\mu$, $J/\psi \rightarrow ee$, tt, …
Author	Author's name	MuonPerformanceGroup, PJenni, …
RecoSWV	Software version (ATLAS Metadata Interface tag)	14.5.2.1
IOV	Interval of validity (runs), For MC: simulation software release	Run 1000-2000 13.0.1

Database Object Structure

Matrix



Unique ID
Database ID

Payload:

- Vectors of integers and floats
- Data vectors
  - Matrix entries
- Info vectors
  - Matrix dimensions and axes definitions

Pool / COOL / ROOT

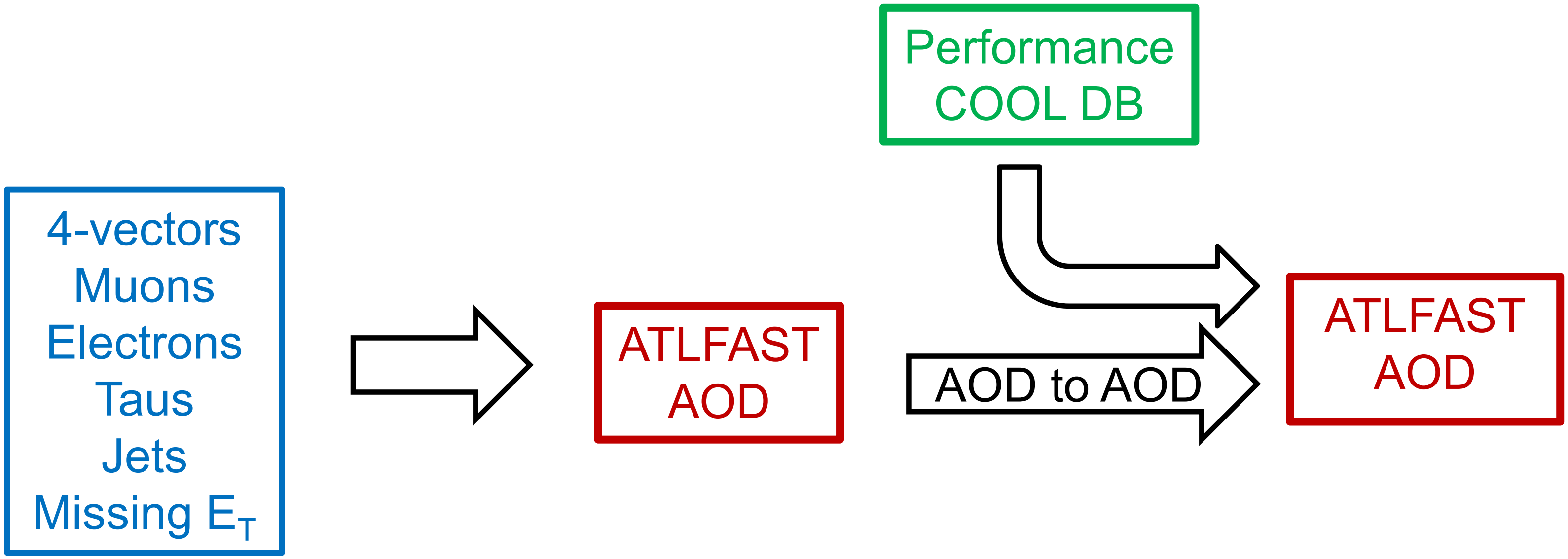
COOL – Database

DATA	IOV
Data [float, float, int, int , float, ...]	10–23
Data [float, int , int]	12–35
...	
Data [Reference to ROOT–File]	11–18
...	

Application: Fast Monte Carlo Simulation

- Fast Monte Carlo simulation
  - Correction of the generated 4-vectors
  - Smearing and efficiency correction

- Ideal place for input from performance tool → currently “hand-made” input tables
- Communication via COOL database → to be implemented



→ A better Monte Carlo simulation

→ Reference numbers from data for systematic detector studies

→ towards a “realistic” Monte Carlo: possibility to map on run-time effects

File Sizes

- Probe DPD files are much smaller than typical AOD:
  - 200 kB/event in AOD → 0.3-0.5 kB/event in DPD
- DPD files produced on grid and collected on local storage

- Matrix files eventually stored in COOL
  - Depends on number of dimensions
  - Potentially large – for a database – if full histograms are stored
  - For 2x50 bin histograms in 20x20x20 matrix ~ 3 MB

Current Status

- All underlying functionalities are implemented
  - Tagging framework
  - DPD creation
  - Matrix representations and operations
  - COOL and ROOT database operations
    → converters

- Working use-cases:
  - Electron reconstruction and identification efficiency
  - Muon reconstruction and identification efficiency
  - Muon trigger efficiency
  - Inner detector tracking efficiency with Z and J/psi tags

Summary and Outlook

- Performance tool provides useful and standardized service to the collaboration
- Performance data can be distributed via ATLAS central database
- Interesting use cases:
  - Individual physics analysis
  - Performance groups
  - Realistic Monte Carlo simulation
- Full working analyses planned to be available in a month's time
- "Online" exercise on Tier-0 center to be done
- Many more performance analyses to be integrated
- Currently evaluating structures for dealing with systematic uncertainties

Backup

Backup

Current Analysis Structure