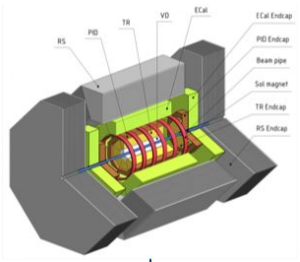
SPD Offline computing.

SPD offline computing

- Support of processing of ~ $2*10^{12}$ events per year (EPY):
  - *Reconstruction;*
  - *MC Simulation;*
  - *Reprocessing.*
- One event per one second - targeted processing rate
  - There are 31536000 seconds per year (SPY)
  - EPY/SPY = 63419 fully loaded CPUs cores ("supercomputer"?)
- Distributed computing system required (grid like)
  - **Based on resources from collaborators**
  - JINR - T0 (25%-30% of compute resources)
- Computing community are very welcome

HEP Data processing chain and (required) data formats



Online filter

Raw data

Reconstruction

Analysis object data

Derivation

Derived AOD

Generation

Event generator output

Simulation

Simulated interaction with detector

Digitization

Simulated detector output

Reconstruction

Analysis object data

Derivation

Derived AOD

Analysis

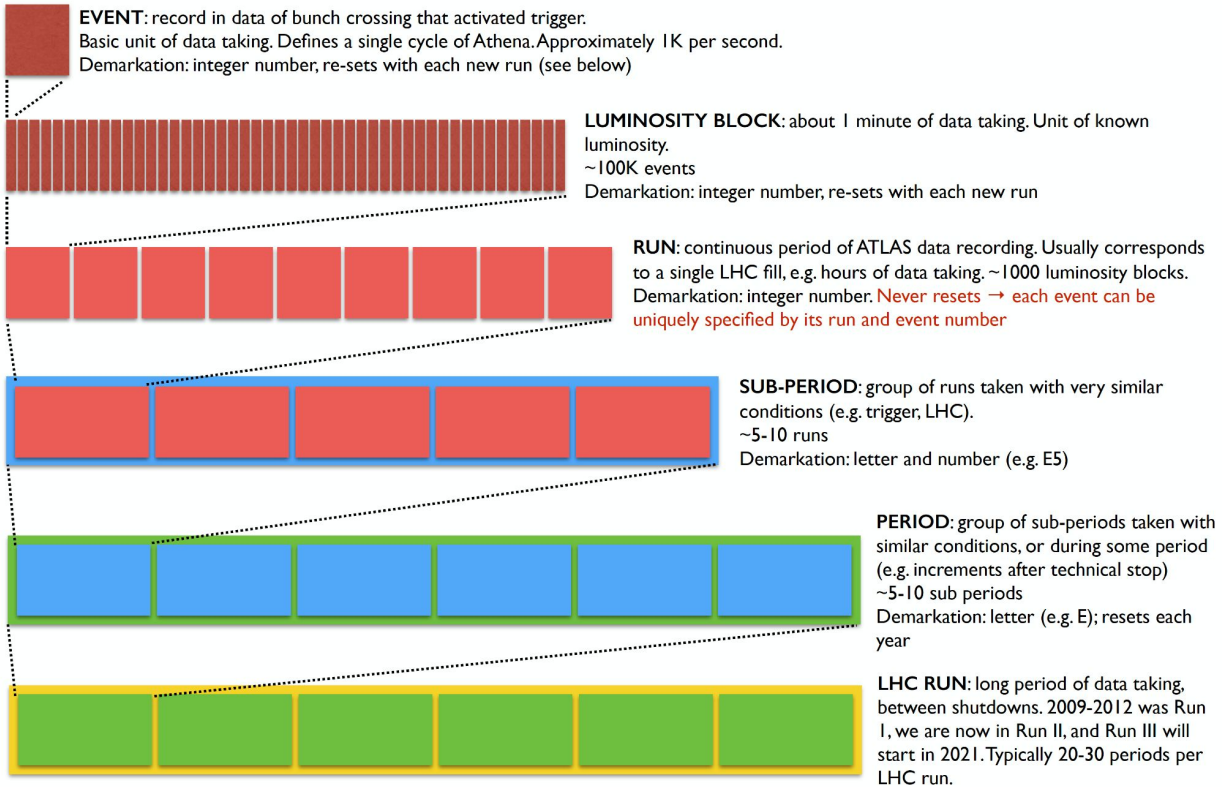
Online

Tier 0

GRID

Local

Data preparation (Reconstruction)



EVENT: record in data of bunch crossing that activated trigger.
Basic unit of data taking. Defines a single cycle of Athena. Approximately 1K per second.
Demarkation: integer number, re-sets with each new run (see below)

LUMINOSITY BLOCK: about 1 minute of data taking. Unit of known luminosity.
~100K events
Demarkation: integer number, re-sets with each new run

RUN: continuous period of ATLAS data recording. Usually corresponds to a single LHC fill, e.g. hours of data taking. ~1000 luminosity blocks.
Demarkation: integer number. Never resets → each event can be uniquely specified by its run and event number

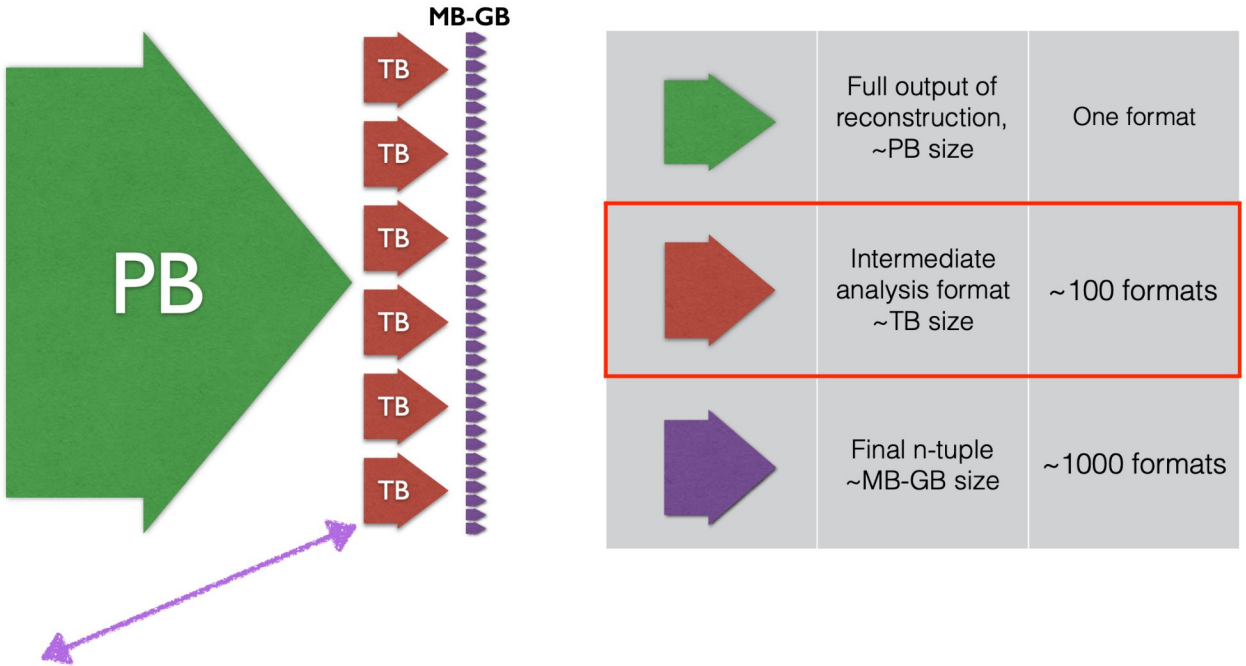
SUB-PERIOD: group of runs taken with very similar conditions (e.g. trigger, LHC).
~5-10 runs
Demarkation: letter and number (e.g. E5)

PERIOD: group of sub-periods taken with similar conditions, or during some period (e.g. increments after technical stop)
~5-10 sub periods
Demarkation: letter (e.g. E); resets each year

LHC RUN: long period of data taking, between shutdowns. 2009-2012 was Run 1, we are now in Run II, and Run III will start in 2021. Typically 20-30 periods per LHC run.

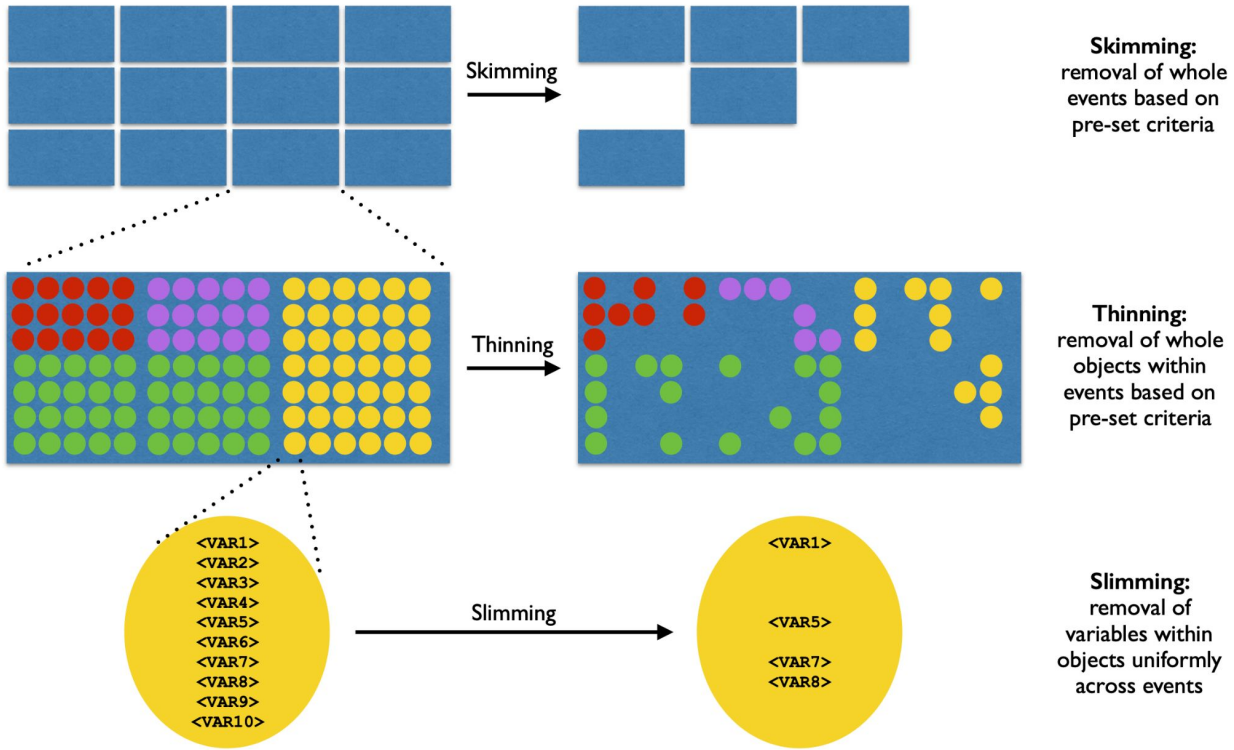
ATLAS Example

Derivation (ATLAS example)



- These formats tend to be specific to a single analysis or group of analyses
- Calibrations and common object selections are often applied as they are made
- They generally need to contain all variables needed for calculating systematics

Derivation step: data reduction



Skimming: removal of whole events based on pre-set criteria

Thinning: removal of whole objects within events based on pre-set criteria

Slimming: removal of variables within objects uniformly across events

Derivation step: augmentation

- New information (augmentation) is typically done in two ways:
  - Adding new reconstructed object containers: typically jets made with a modified algorithm.
  - Decorating existing objects with extra variables: typically the results of object selection by combined performance tools (e.g. "this is a good muon")