

Analyse-Software in ATLAS

Wo kommen meine Daten her?

Nikolai Hartmann

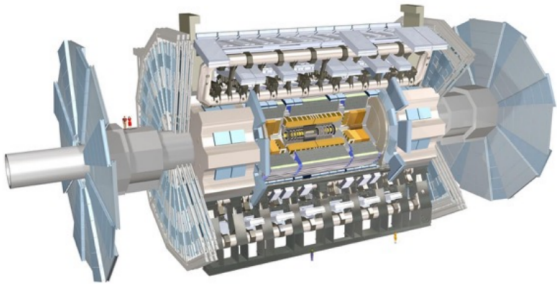
(Slides hauptsächlich von [Federica](#) und [ATLAS Software tutorial](#))

LMU Munich

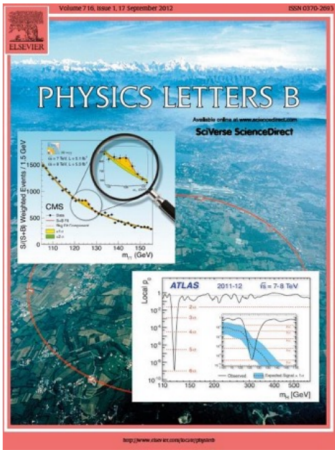
April 23, 2020



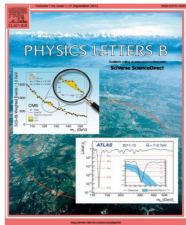
How do we get from this....



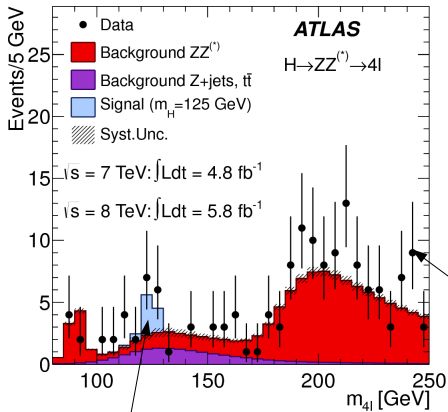
to this...?



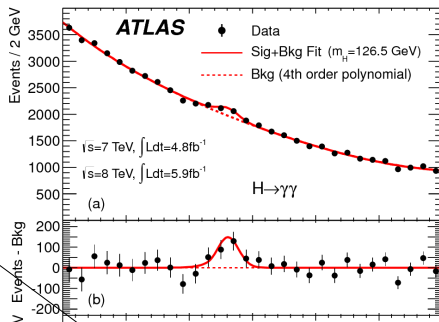
What's in



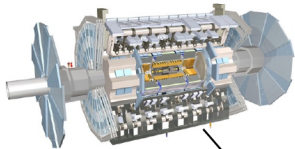
?



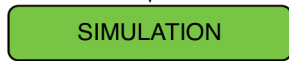
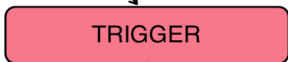
Monte Carlo data



LHC data



Data



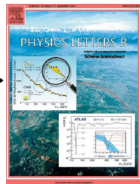
Monte Carlo

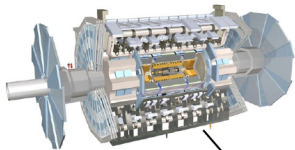
Online

Tier-0

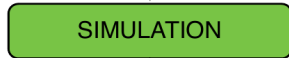
Grid

Local

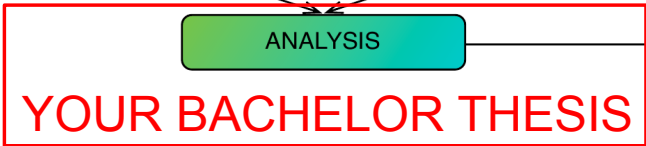




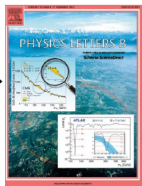
Data

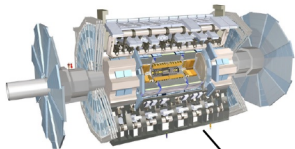


Monte Carlo

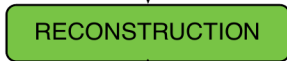
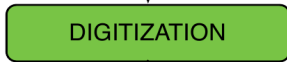
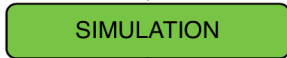


- Online
- Tier-0
- Grid
- Local





Data



Monte Carlo

THIS LECTURE

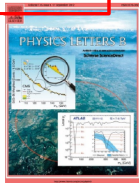
Online

Tier-0

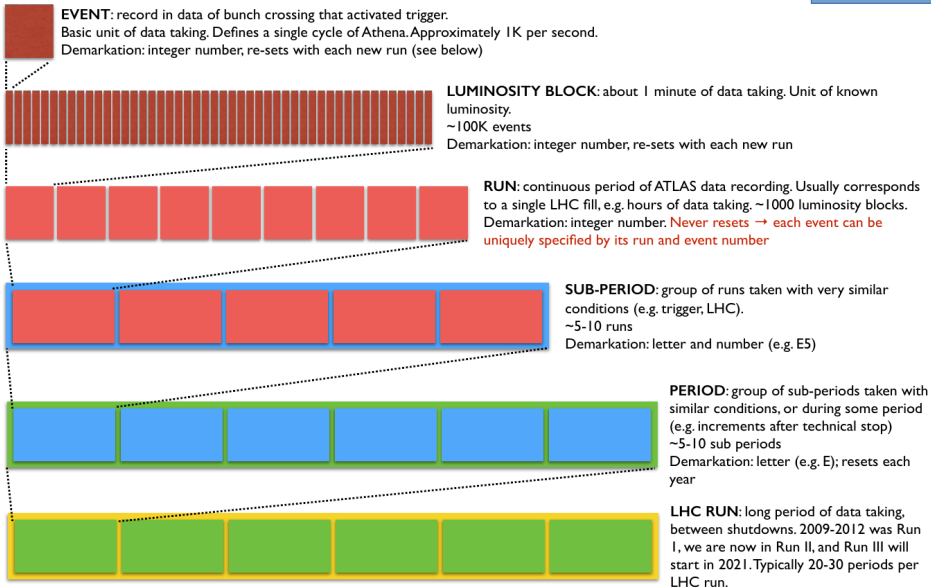
Grid

Local

If time allows



How data is divided up in ATLAS



Good runs lists

- Remember those luminosity blocks? ~ 1 minute of data taking
 - ▶ Several hundred lumi blocks per run; one or several lumi blocks per file
 - ▶ Some lumi block ranges will have *bad data quality* and should not be used for analysis

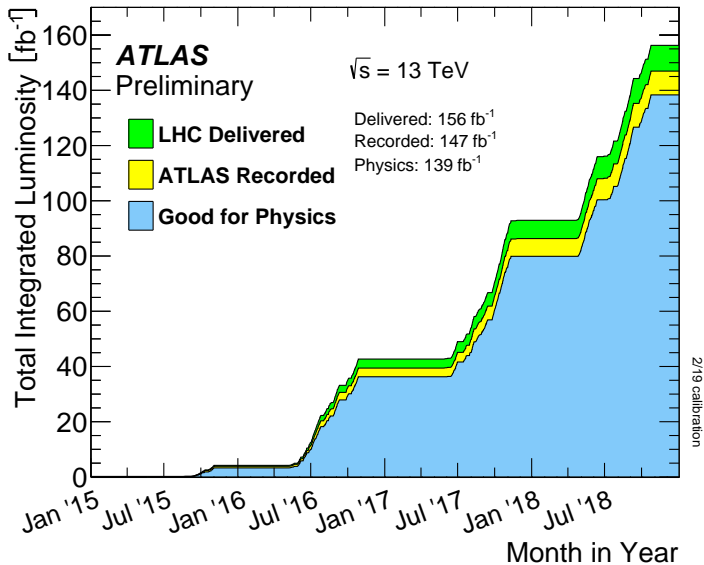


- The data quality flags are set in the data quality monitoring
- They are encoded in a database
- Different versions of the data quality flags get a different name, e.g. DetStatus-v6 l

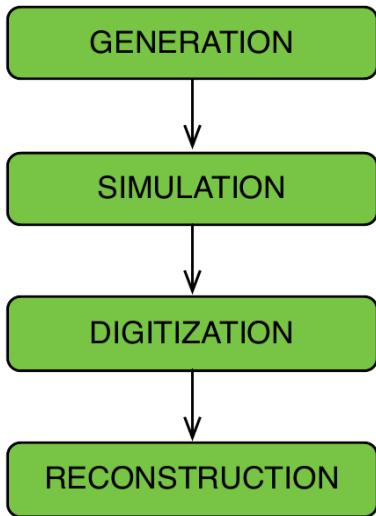
Lumi blocks and analysis

- When you do your analysis you need to filter out the bad lumi blocks
- This is done by means of a *good runs list (GRL)*
 - ▶ XML file listing each run and the range of good lumi blocks in each: read in by the analysis code
- The luminosity “seen” by your analysis depends on three things:
 - ▶ the luminosity blocks processed (from the GRL)
 - ▶ the trigger prescales applied
 - ▶ the level-1 trigger live-time (fraction of delivered luminosity that ATLAS recorded)

Delivered, recorded, good for physics



Monte Carlo Production

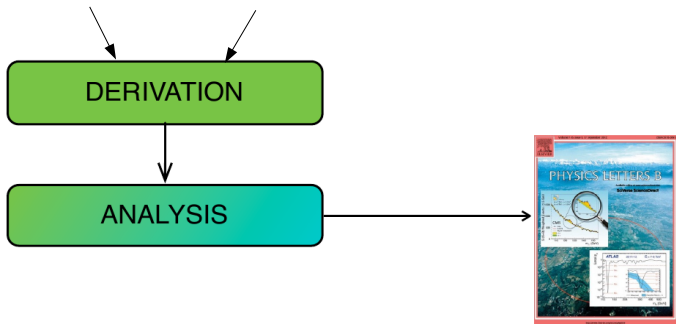


Monte Carlo production

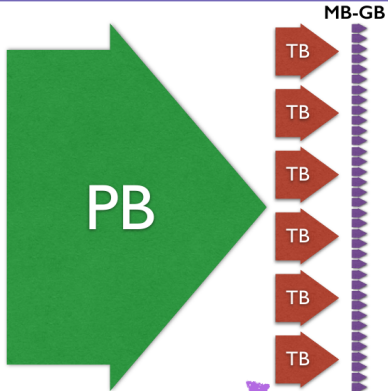
- All official Monte Carlo production is done on the Grid
- Several-step process
 - ▶ *Event generation*: simulation of the interaction between the quarks and gluons in the colliding protons, the subsequent parton showering and hadronization and decays into stable particles
 - ▶ *Detector simulation*: calculation of how the particles from the generator interact with the detector material; how they shower into secondaries; how much energy they deposit in each sensitive element
 - ▶ *Digitisation*: turning the simulated energy deposits into a detector response that “looks” like the raw data from the real detector
 - ▶ After this step, the process is the same as for real data
- The analysis data for MC and real data looks the same, *except* that in MC the original generated events (the “truth”) are available as well as the reconstructed objects
- Extra low momentum events must be injected into the chain to simulate the presence of multiple proton collisions (“pile-up”) in a given LHC bunch crossing
 - ▶ *Complication*: the average number of collisions per bunch crossing is a function of the LHC parameters, and we typically do not know this when we start the Monte Carlo production
 - ▶ Monte Carlo events must be weighted in analysis to account for discrepancies between the real and simulated pile-up.




The ATLAS Analysis Model and the xAOD

LHC data MC data



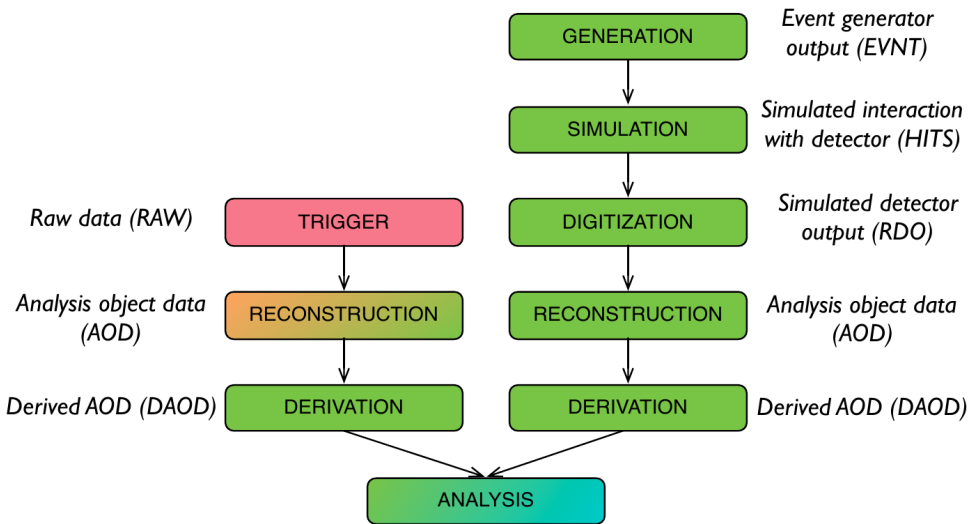
Derivation step



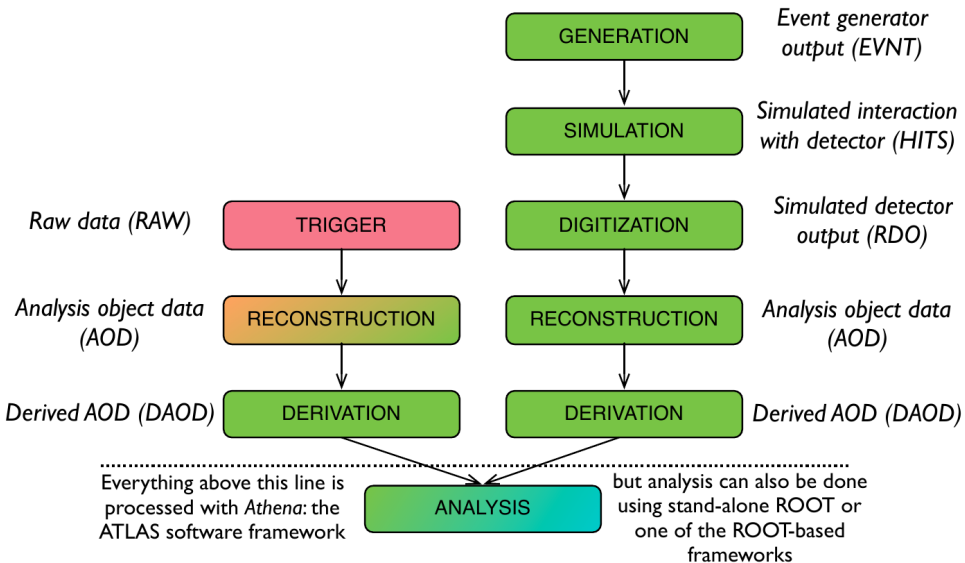
	Full output of reconstruction, ~PB size	One format
	Intermediate analysis format ~TB size	~100 formats
	Final n-tuple ~MB-GB size	~1000 formats

- 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

Data formats in ATLAS

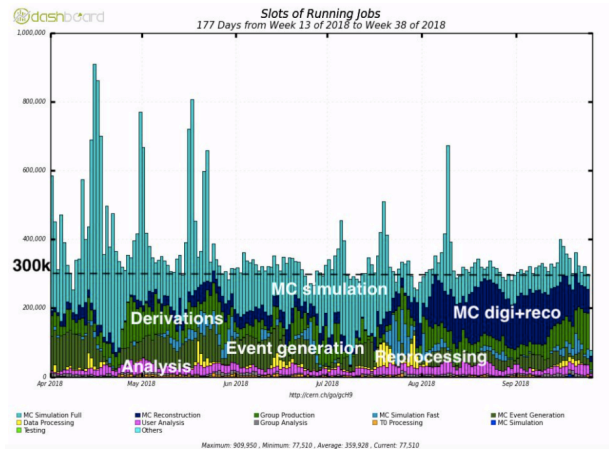
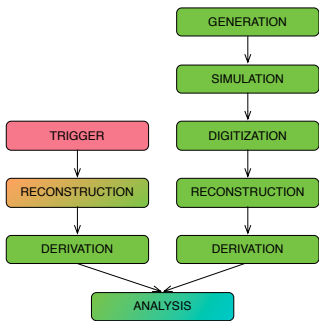


Data formats in ATLAS



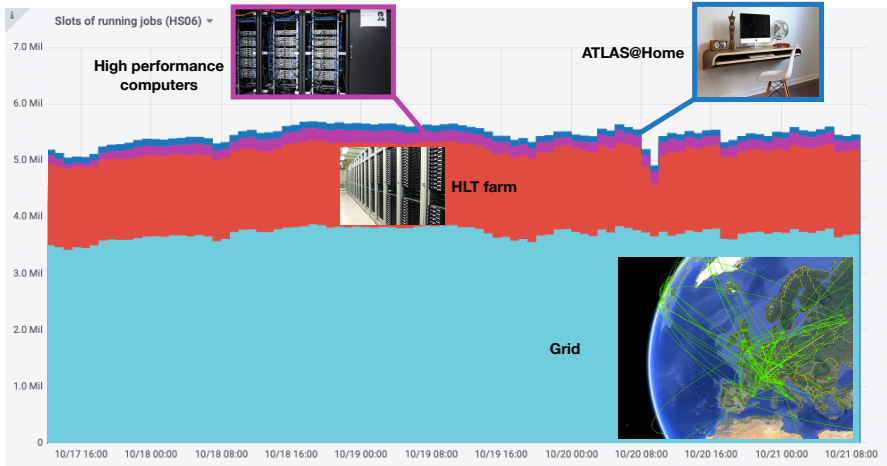
What are we running?

29



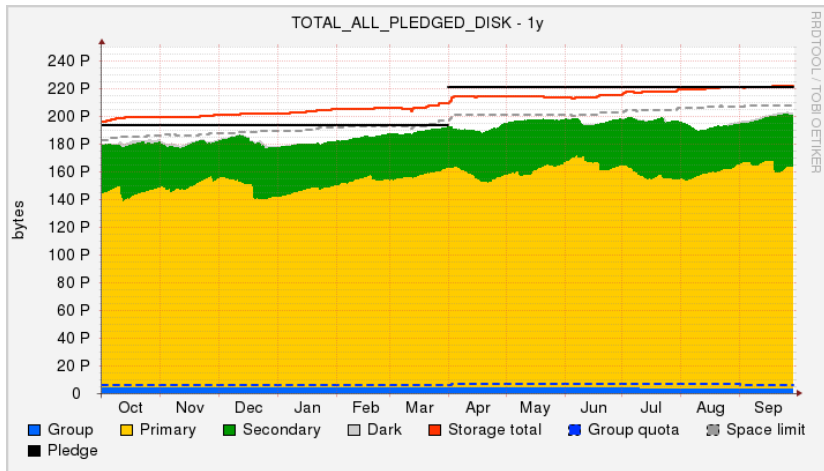
Where are we running it?

23

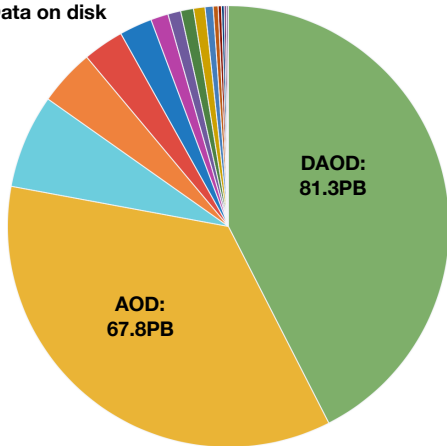


Our computing resources aren't infinite...

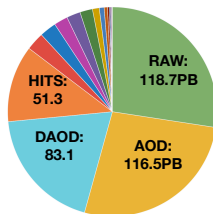
24



Data on disk



Data on disk +tape



Message: don't assume that computing resources are infinite. Keep disk space usage in mind when planning your work.

Computing is a shared resource.

Dataset nomenclature

DATA

Project Stream Format
data18_13TeV.00349051.physics_Main.deriv.DAOD_EGAM1.f926_m1955_p3544
Run number Step AMI tag

or in Physics Containers:

Period Format
data18_13TeV.periodM.physics_Main.PhysCont.DAOD_EXOT0.grp18_v01_p3583
Project Stream Compound tag

MONTE CARLO

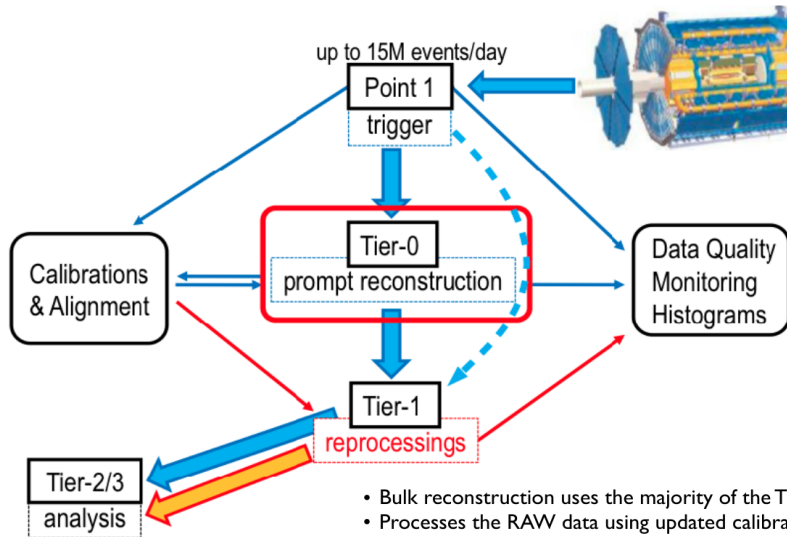
Project Physics description Format
mc16_13TeV.364250.Sherpa_222_NNPDF30NNLO_1111.deriv.DAOD_EGAM1.e5894_s3126_r9364_p3526
Dataset ID Step AMI tag

Zusammenfassung

- Die ersten Schritte der Datenanalyse laufen zentral
 - Datenanalyse läuft in mehreren Schritten - auf gleiche Weise für simulierte Daten wie Monte Carlo
 - Datenmenge zu groß für lokale Speicherung/Prozessierung
→ Weltweit verteilt im Grid
 - Rekonstruktion bzw. Detektorsimulation für Monte Carlo zu einem Objektorientierten Datenformat (AOD)
 - Reduktion der Datenmenge durch "Derivations" für verschiedene Analysekatgorien (DAOD)
- Prozessieren der DAOD Formate pro Analyse (mehr dazu später)

Backup

Bulk reconstruction



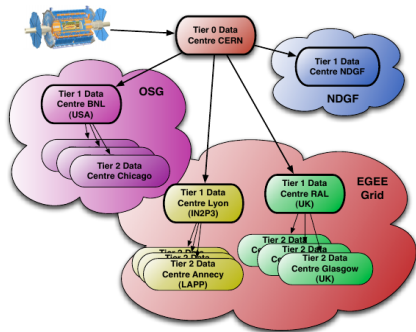
- Bulk reconstruction uses the majority of the Tier0 resources
- Processes the RAW data using updated calibrations determined in the Prompt Calibration
- Produces many outputs used for a variety of purposes, the most important being the AOD

What's all this grid nonsense then?

- Recognition in the early planning of LHC computing that a single centre would not supply enough resources for the LHC experiments
 - Mainly because no one would fund a central computing facility

- In fact ATLAS uses three grids:
 - EGI in Europe, Asia and Canada (UMD middleware)
 - OSG in USA (OSG stack)
 - NorduGrid in Nordic countries (ARC middleware)
 - All badged as “Worldwide LHC Computing Grid compatible”
- 

Computing Model Cartoon



- 1 Tier-0: CERN
- 10 Tier-1: National Computing Centres (BNL, RAL, IN2P3, ...)
- 40 Tier-2: Regional Computing Centres (ScotGrid, Frascati, Toronto, ...)
- Composed of multiple individual sites
 - ~100 Analysis queues in PanDA

- T0: prompt reconstruction
- T1s/T2s: reprocessing, MC production, derivation production, analysis (mainly at T2s)
- T3s: analysis only

ATLAS Grid Architecture

- Distributed Data Management System (DDM): Data movement and catalog system, policy engines
- Other databases and data book keeping (AMI, GRL, Coma for metadata; conditions data in distributed Oracle + frontier; Event Index; etc.)
- Production System (PanDA)
- Distributed Analysis Interfaces (ganga + pathena)



New version
for Run 2:
Rucio

New for for
Run 2: JEDI
and DeFT



DDM

Rucio

- Data management is at the core of all ATLAS grid activities
 - All data organised as datasets
 - Which contain multiple files
 - Datasets are the 'work units' on the grid
 - So the DDM central catalog records the location of datasets on the grid, as well as the content of each dataset
 - Datasets are also the basic input to an analysis
- DDM also moves data between sites on the grid



Send in the Jobs!

- To run ATLAS jobs on the grid you need an distributed analysis framework which
 - Understands the ATLAS data model
 - Understands the grid
 - Understands the analysis code
- Distributed Analysis Interfaces (ganga + pathena)