Challenges in the ATLAS Monte Carlo Production during Run 1 and beyond

Wolfgang Ehrenfeld (Bonn)
On behalf of the ATLAS Collaboration

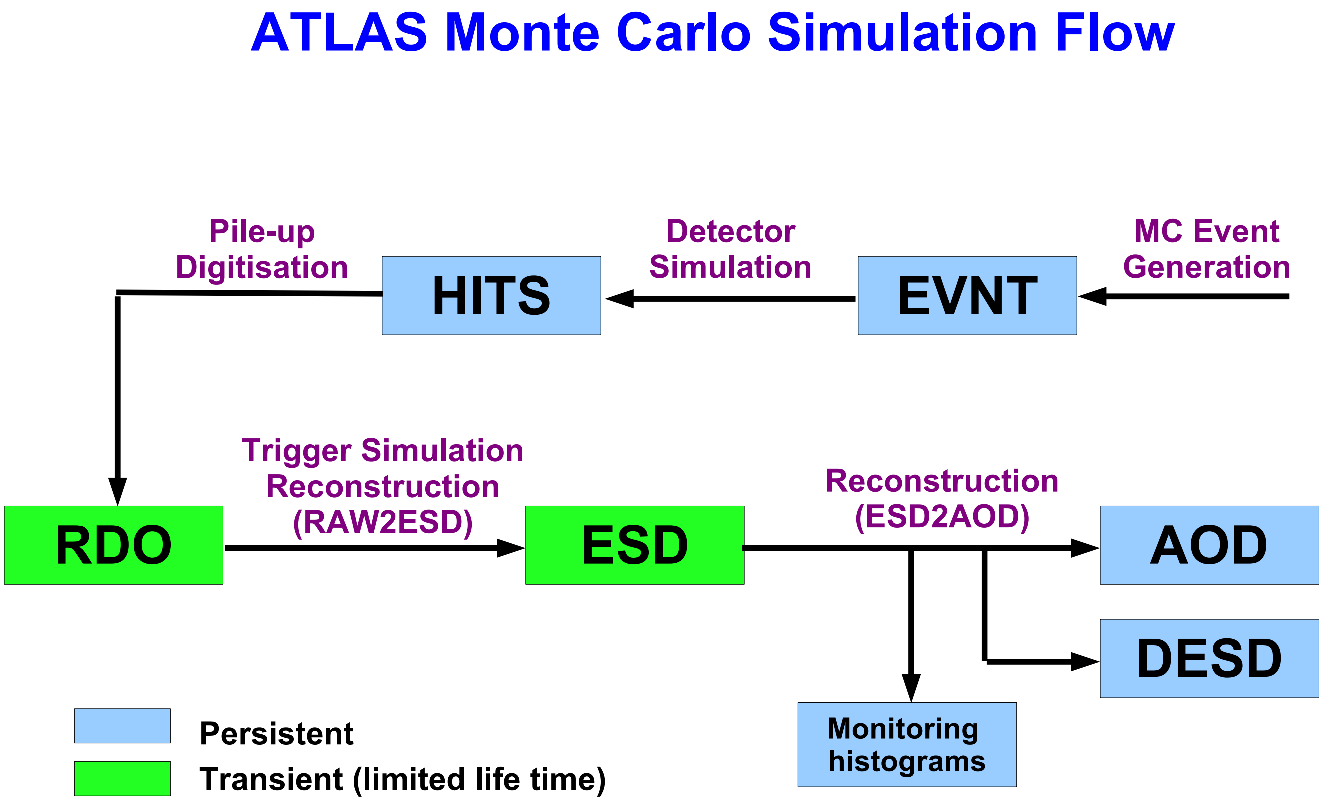
CHEP 2013
17 October 2013

Overview

- Monte Carlo production steps
  - event generation, simulation, digitisation, reconstruction
- Monte Carlo production campaigns
- grid resources and usage
- special production activities
  - validation, upgrade production, zero-bias data overlay
- summary

MC Production Steps

- event generation
- simulation
- digitisation
- reconstruction



Event Generation

- ~30 generators used in ATLAS
  - framework integrated generators
  - stand-alone generators
- event generation work flows
  - single step generation: Pythia6/8, Herwig(++), Sherpa
  - two-step generation: parton level generator coupled via LHEF files to framework generator for hadronisation (Pythia(6/8), Herwig(++))
    - default configuration: external, pre-made 4-vectors uploaded to the grid
    - on-the-fly configuration: run external generator before hadronisation in the same job
- distribution of job options for event generation
  - job options and generator configuration stored in SVN (versioning)
  - distribution independent of software release for frequent updates and fast turnaround
    - http based download of tar ball from CERN based web server
    - under development: distribute files via CVMFS to simplify software distribution and for fire-walled worker nodes

Event Generation - Performance

MC Campaign: setup corresponding to data taking period

- requested samples very diverse
  - 50 different generator combination in mc12 campaign
  - ~34 thousand different samples produced in mc12 campaign
- job characteristics
  - 5000 events per job → ~100 MB output file size
  - low memory requirements: < 0.5-1 GB
  - running time per job varies from
    - a few minutes for simple final states/hadronisation of external 4-vectors
    - hours or days for complex final states or low filter efficiencies
    - number of events needs to be adjusted for optimal running time of 8 hours
- performance improvements:
  - on-the-fly generator setups: avoid storing 4-vector input files on the grid
  - use pre-made integration files (Sherpa, Alpgen, MadGraph): reduce running time

Simulation - Improving Simulation Time

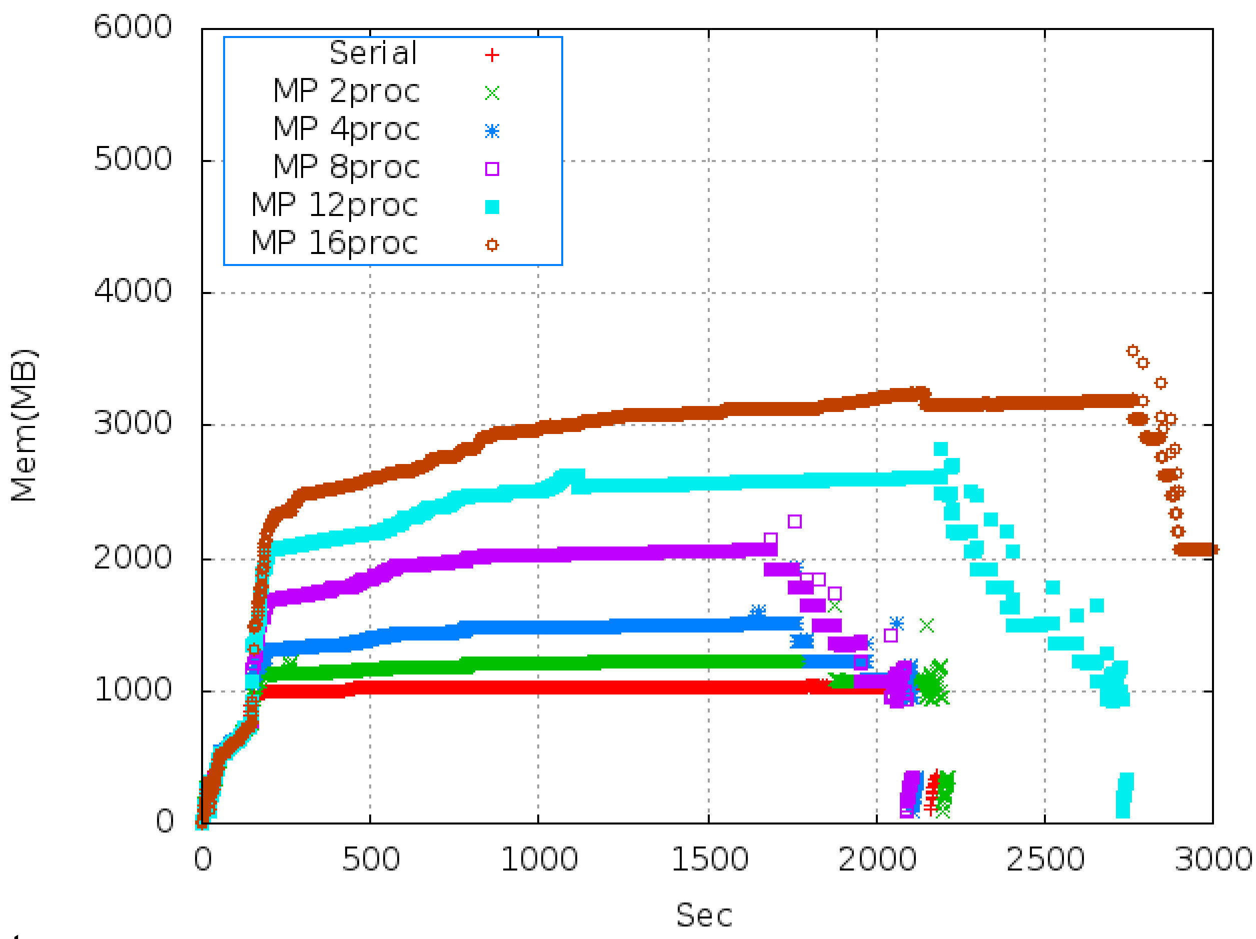
- G4 full simulation:
  - every stable particle is tracked through the ATLAS geometry
  - the list of possible interactions is defined by the physics list: QGSP_BERT as default
  - one event takes ~5 minutes → major simulation time spent in calorimeters
- G4 full simulation with Frozen Showers (FS) in calorimeters: 25% speed up in mc12
  - showers are tracked down to very low energy by G4 → stop showering at a threshold and substitute each end particle by a pre-made list of energy deposits
  - frozen showers in the forward calorimeters as default in mc11/mc12 including upgrade production
- AtlFast-II (AF-II): factor 10 speed up in mc12
  - parametrise all particles except muons in the calorimeters
    - do not simulate particles except muons in the calorimeter
    - parametrise non-simulated particles before the digitisation step
  - in production since late mc10
- Integrated Simulation Framework (ISF)
  - better integration of full and fast simulation based on sub-detectors and particles

Simulation - Performance

- job characteristics
  - full simulation: 100 events per job → ~80 MB output file size → merged up to 1000 events (0.8 GB file size) for better grid transfers and tape storage
  - fast simulation: 1000 events per job → ~0.5 GB output file size
  - low memory requirement: ~1 GB
  - run time per (averaged over grid cpus)
    - G4 full simulation: 335 s/evt
    - G4 full simulation with frozen showers: 250 s/evt
    - AtlFast-II: 20 s/evt
- performance improvements
  - run in 64 bit → better performance while slightly increasing the memory
  - Intel math library
  - modern random number generator: SIMD-oriented Fast Mersenne Twister
  - geometry and conditions DB access via frontier instead of pre-packed DB release → job only request needed data

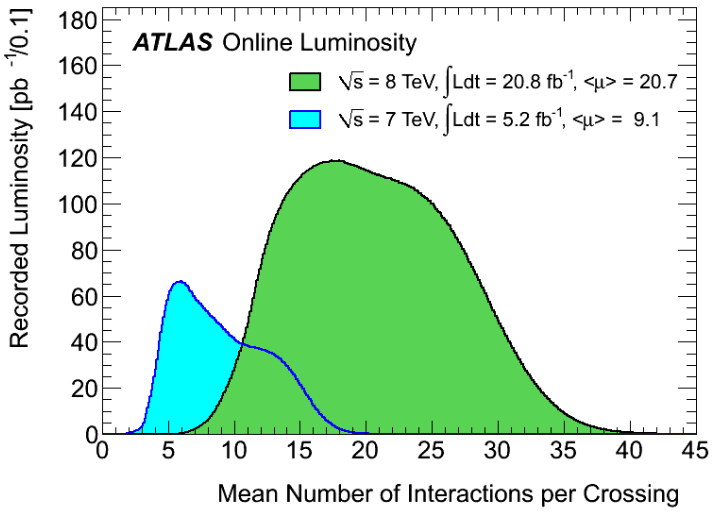
Simulation - Multi Core Utilisation with athenaMP

- single core: 1.0GB
- double core: 1.2GB
- …
- 8 cores: 2.6GB
- 0.8GB + 0.16GB/core
- athenaMP validated for simulation
- production scenarios
  - reducing number of job
  - back filling multi-core slots
  - high performance computing resources



Digitisation

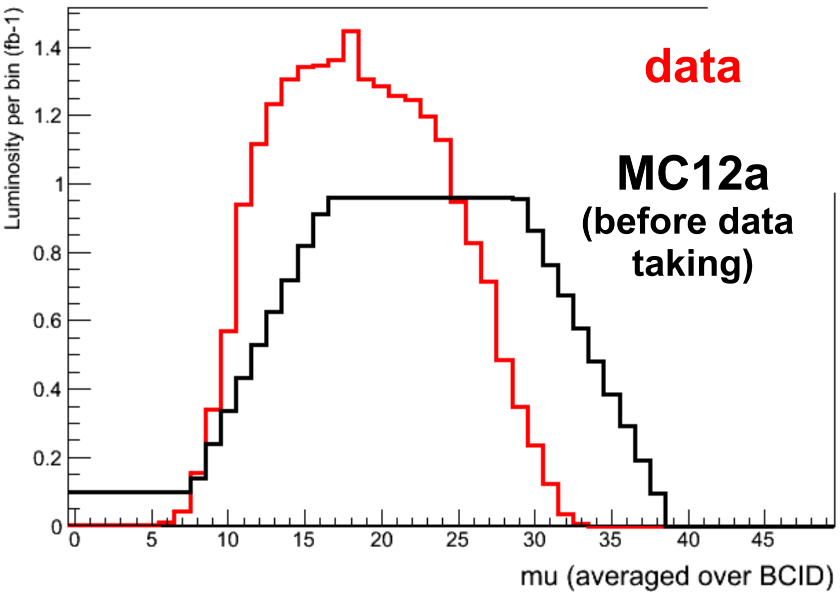
- simulate detector readout
- simulate pile-up contributions (multiple pp interactions on top of hard scatter event)
- overlay a number of pre-simulated minimum bias events on each signal event
  - <μ> average number of additional pp collisions
  - fixed <μ> (for performance studies)
  - pre-defined <μ> profile (default for physics samples)
    - sample given <μ> profile over 5000 events
      → small samples should be multiple of 5000 events

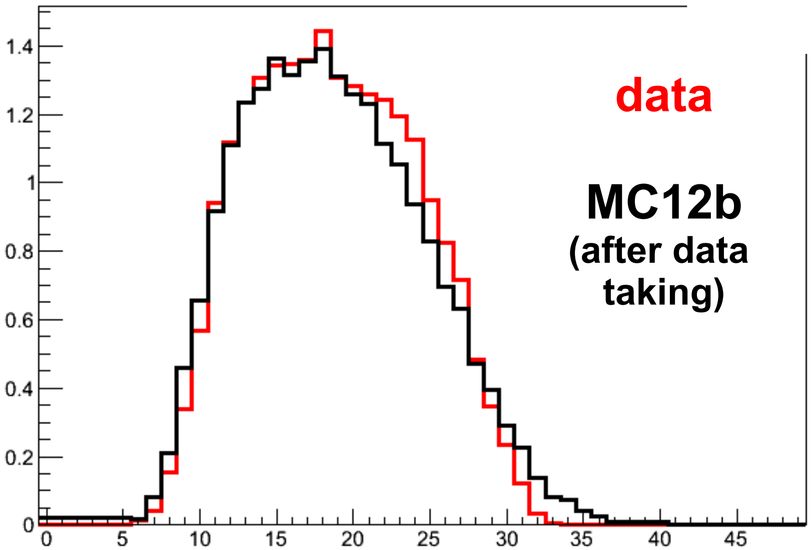


- optimise pile-up event storage and access
  - cache pile-up events in memory → memory intensive
  - flush memory early and re-load from disk on demand → I/O and CPU intensive
- minimum bias pile-up samples
  - separate into low-Q and high-Q (Q=35GeV) samples to allow for frequent re-use of low-Q events per job and limit re-use of within one sample

MC12 Pileup Simulation

- pile-up profile in MC matched to observed distribution in data if possible





- mc12 pile-up sample configuration
  - <μ> profile samples from 0 to 40, with a mean of <μ>=20
  - 10M low/high-Q (1.5/4.8 TB = 6.3 TB) → 5000/500 events per file
  - 500 events per job: one signal file, 5 low/high-Q files → 4.8 GB of input files per job (100 events per job: one signal file, 1 low/high-Q file → 1.1 GB of input files per job)
  - distribute minimum bias pile-up sample to T1 and larger T2 sites → 0.3-0.4 PB total

Reconstruction

- reconstruct simulated events in the same way as data
- trigger simulation
- two step process:
    - RAWtoESD: main reconstruction → output is Event Summary Data (ESD)
    - ESDtoAOD: fast slimming process → output is Analysis Object Data (AOD)

  - for MC ESD are transient files → can be stored on request (in group space)
  - ntuples and derived formats from ESD or AOD are produced by group production
  - some work flows have different output formats
    - for heavy ion ESD and ntuples are produced and stored
- job characteristics
  - 500 events per job → ~220 MB output file size → merged up to 5000 events (~2.2 GB file size) for better grid transfer, processing and tape storage
  - high memory usage:
    - 3.6 - 3.8 GB in 32 bit
    - 64 bit would exceed the 4 GB (grid queue limits)

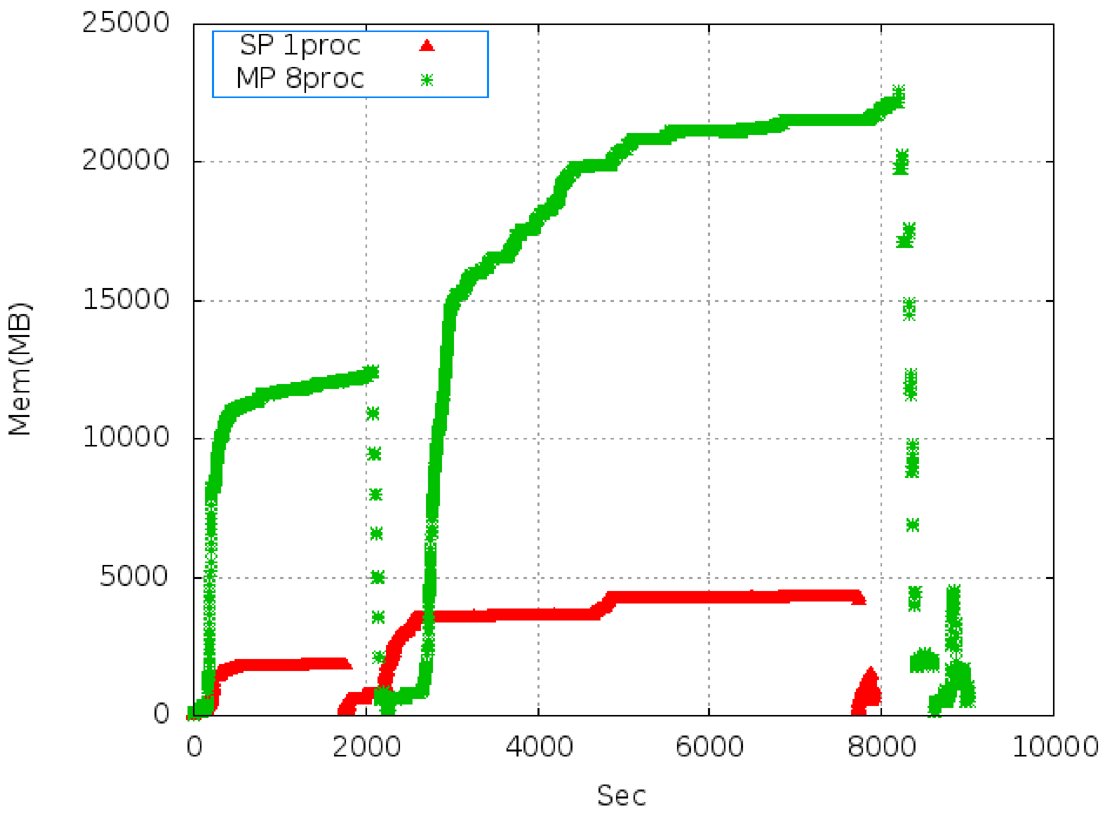
athenaMP - Memory Sharing in Digitisation+Reconstruction

- running in 64 bit
- single core: 4.3GB
- 8 cores: 22.6GB

- 2.8GB/core

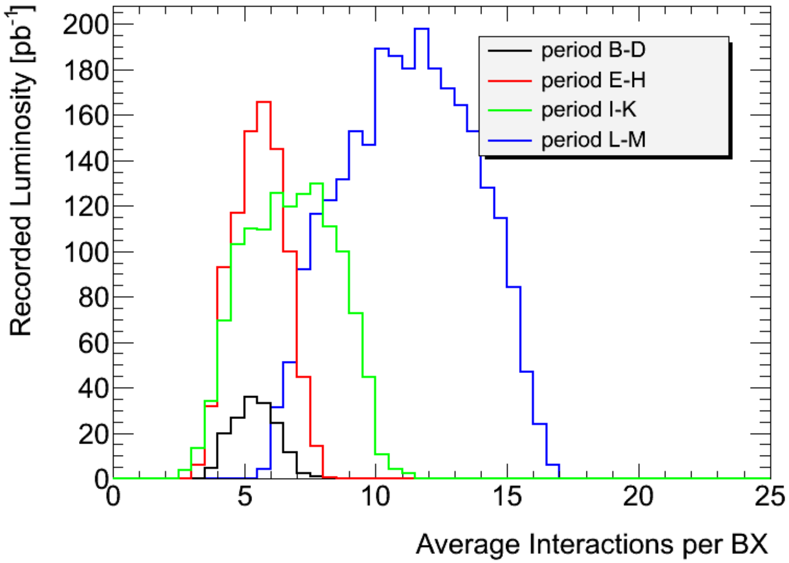
- better than 4GB/core
  but aim is 2GB/core

- athenaMP validated for digitisation+reconstruction
- production scenarios
  - reducing memory consumption
  - reducing number of jobs



Joining Steps in one Job

- joining two or more steps from the simulation can be useful
  - digitisation+reconstruction (default)
    - avoid storing large digitisation output on the grid
    - easier for multi-period pile-up and trigger simulation (mc11)
  - fast simulation
    - avoid storing intermediate outputs → simplify data management
    - fast simulation → small loose in CPU in case of re-reconstruction



Campaigns

- MC production campaigns correspond to data taking periods with same conditions
  - centre-of-mass energy, detector configuration, conditions, …

- Major MC production campaigns
  - mc11: simulation configuration for 7 TeV in 2011
    - mc11a: digitisation+reconstruction configuration with Pythia 8 pile-up sample, estimated beam spot and pile-up profile based on three run periods
    - mc11b: same as mc11a with updated pile-up profile/conditions based on four run periods and two trigger menus
    - mc11c: same as mc11b with Pythia 6 pile-up sample
  - mc12: simulation configuration for 8 TeV in 2012
    - mc12a: digitisation+reconstruction configuration with Pythia 8 pile-up sample, estimated pile-up profile and beam spot based on 2011 data
    - mc12b: same as mc12a with beam spot and pile-up profile from data
    - mc12c: improved geometry description for precision measurements: simulation based on mc12 and digitisation+reconstruction based on mc12b

Produced MC Events

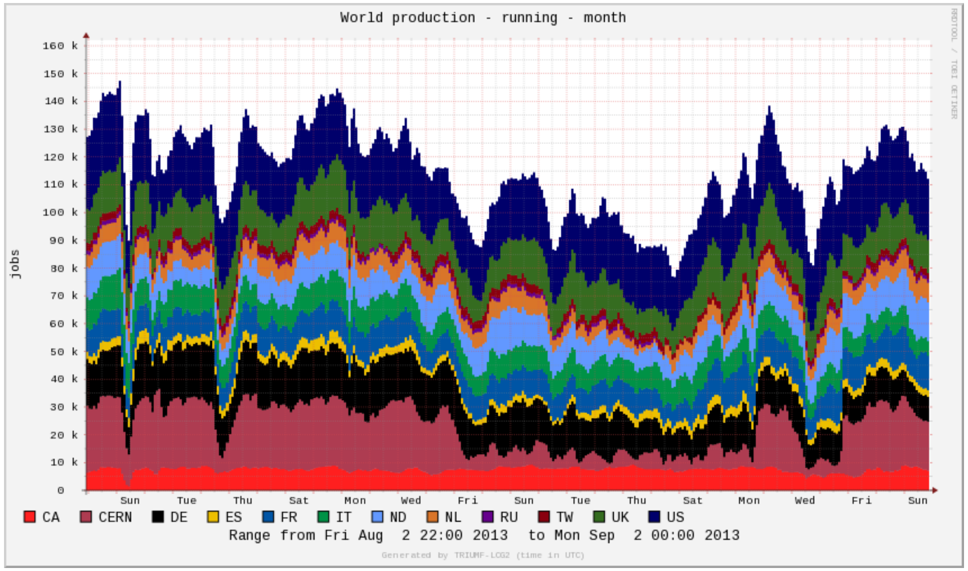
- mc11: 2.4 x 10^9 full and 2.1 x 10^9 fast simulation events
  - mc11a: 0.8 x 10^9 events
  - mc11b: 1.0 x 10^9 events (super seeds mc11a)
  - mc11c: 4.8 x 10^9 events (super seeds mc11b) → total: 4.8 x 10^9 events
- mc12: 3.8 x 10^9 full and 3.0 x 10^9 fast simulation events
  - mc12a: 5.9 x 10^9 events
  - mc12b: 0.5 x 10^9 events
  - mc12c: 0.2 x 10^9 events → total: 6.6 x 10^9 events

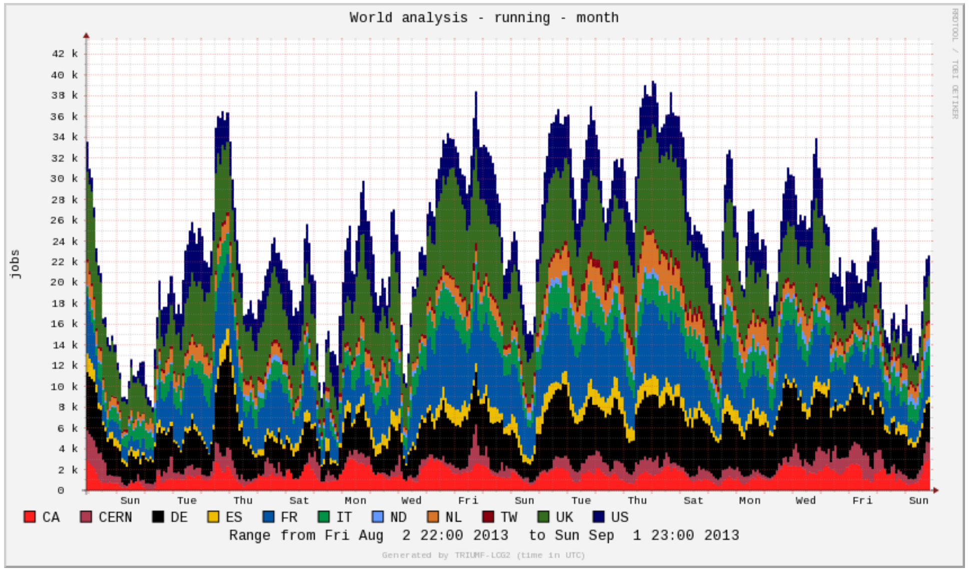
→ total of 6.2 x 10^9 full and 5.1 x 10^9 fast simulation events

ATLAS Grid Resources

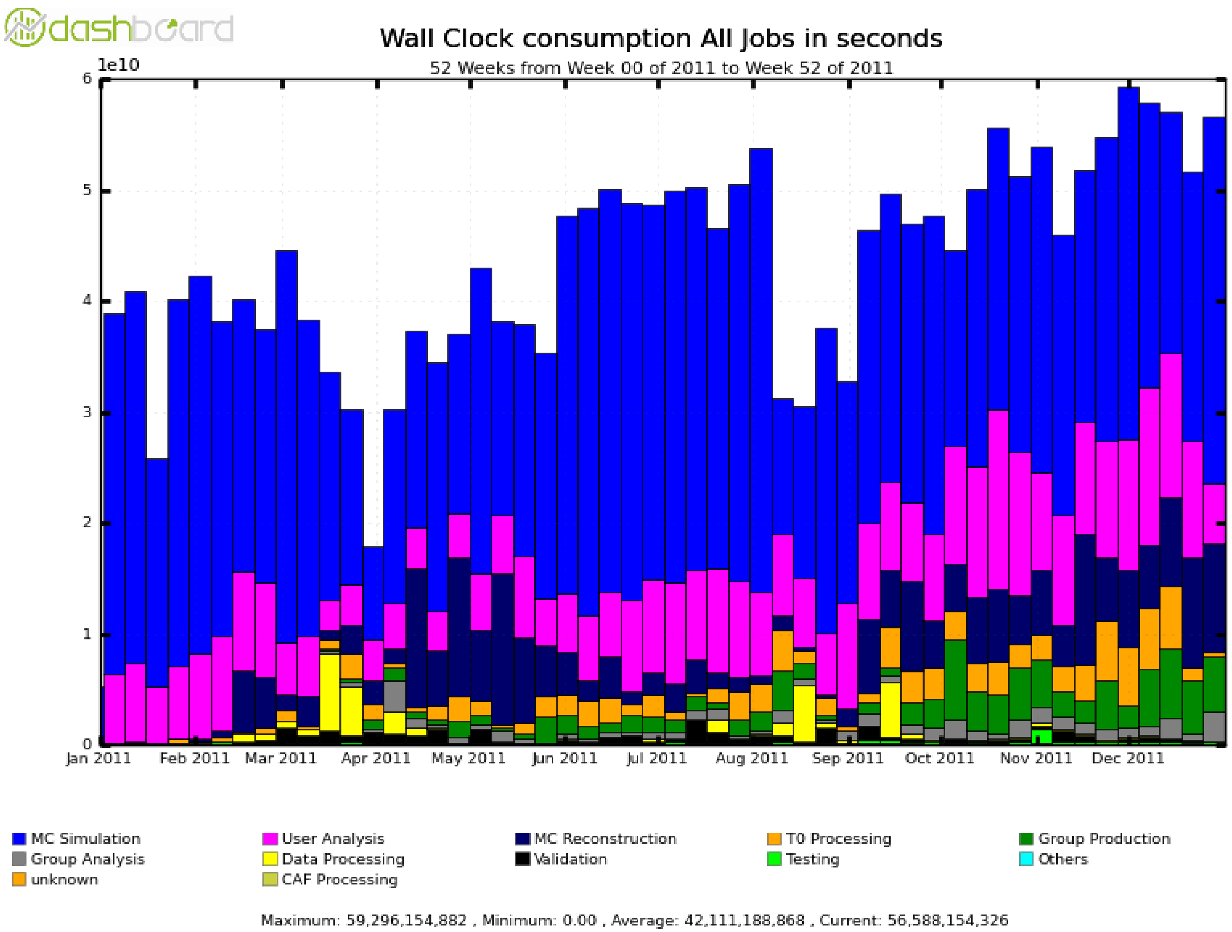
- grid resources
  - Tier0: CERN
  - Tier1: 10 (11) sites
  - Tier2: ~70 sites
  - Tier3: ~20 sites

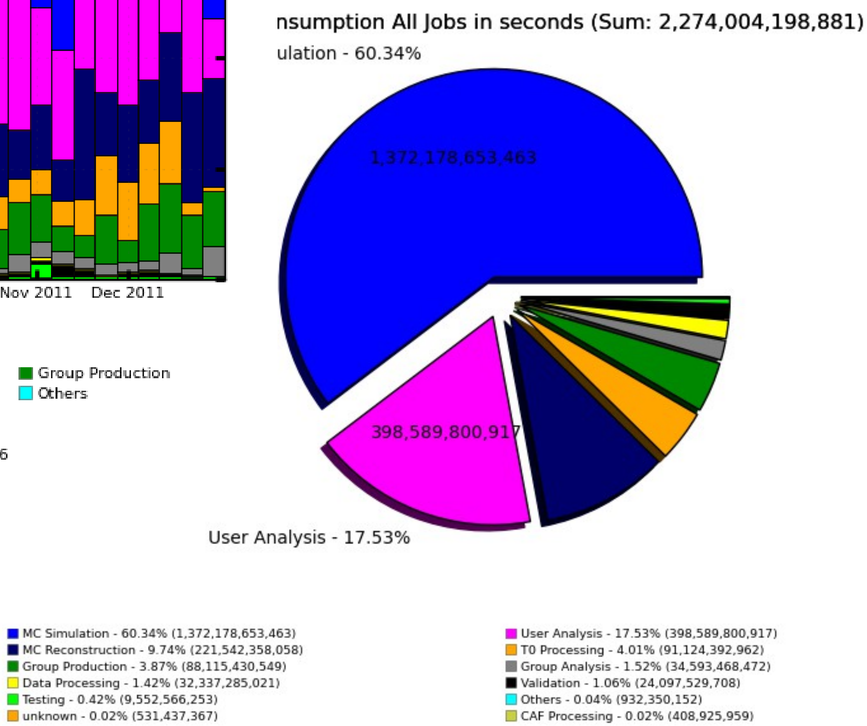
  → ~90 000 single core slots for MC production
- clouds
  - Amazon EZ2 cloud
  - Google Computing Engine cloud
  - Open Clouds
- opportunistic sites
  - online trigger farm (16 000 slots)
  - High Performance Computing



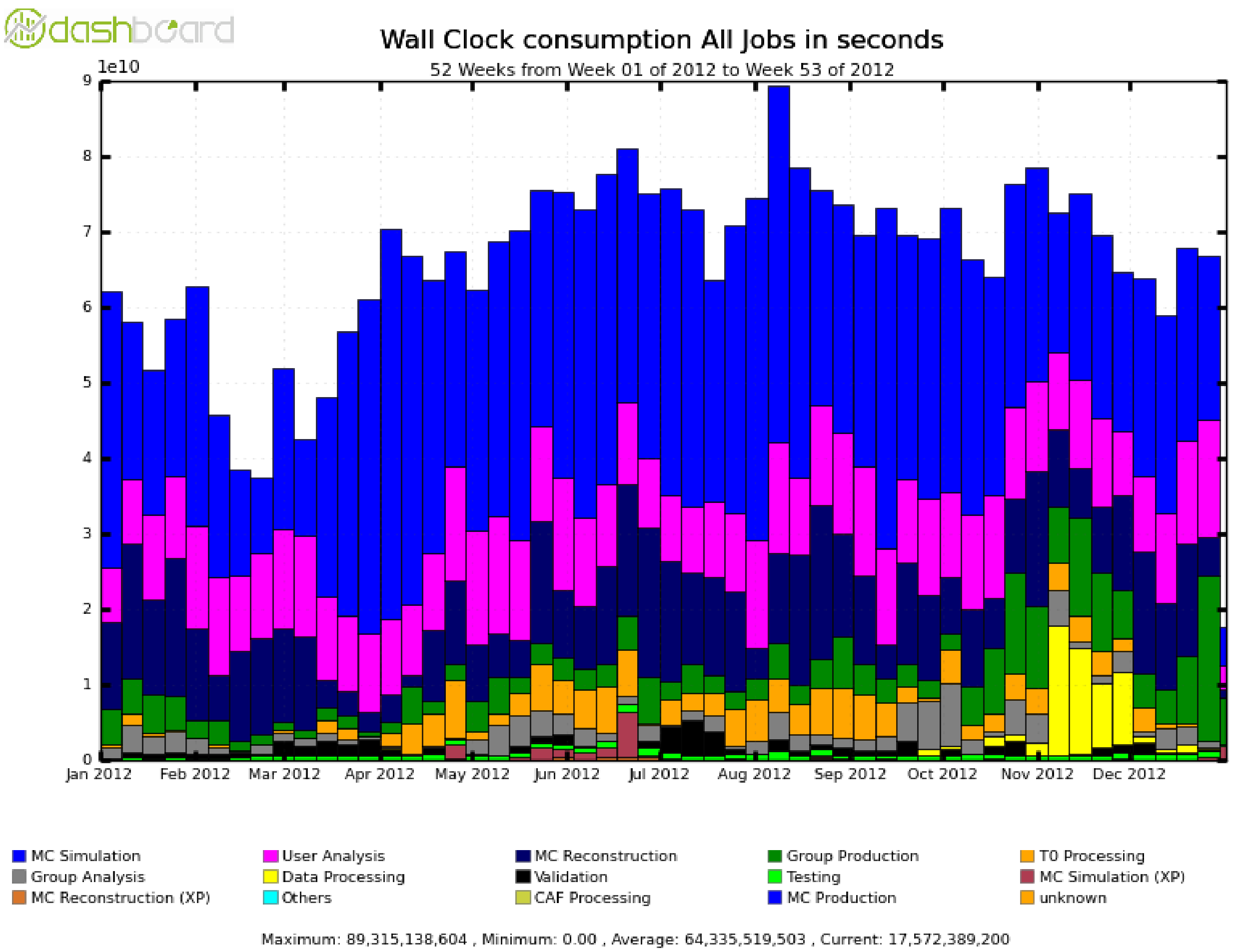


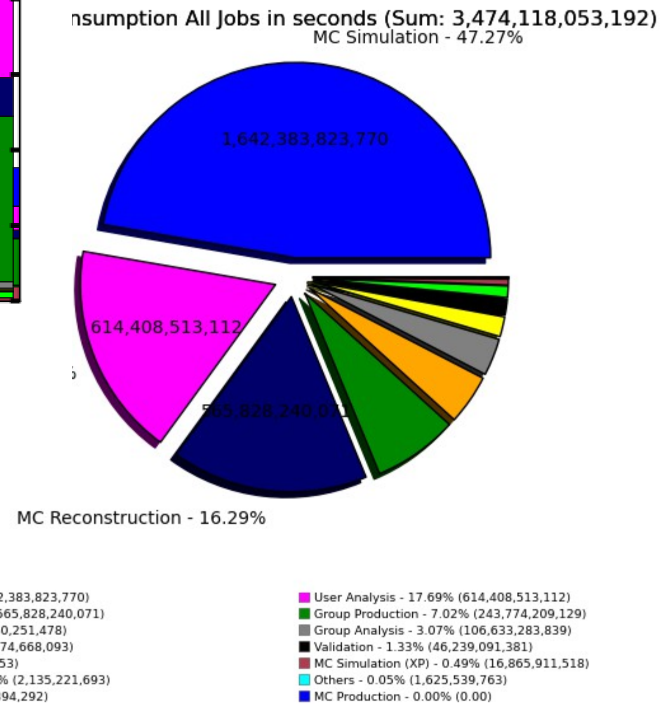
Grid Usage: 2011





Grid Usage: 2012





Validation

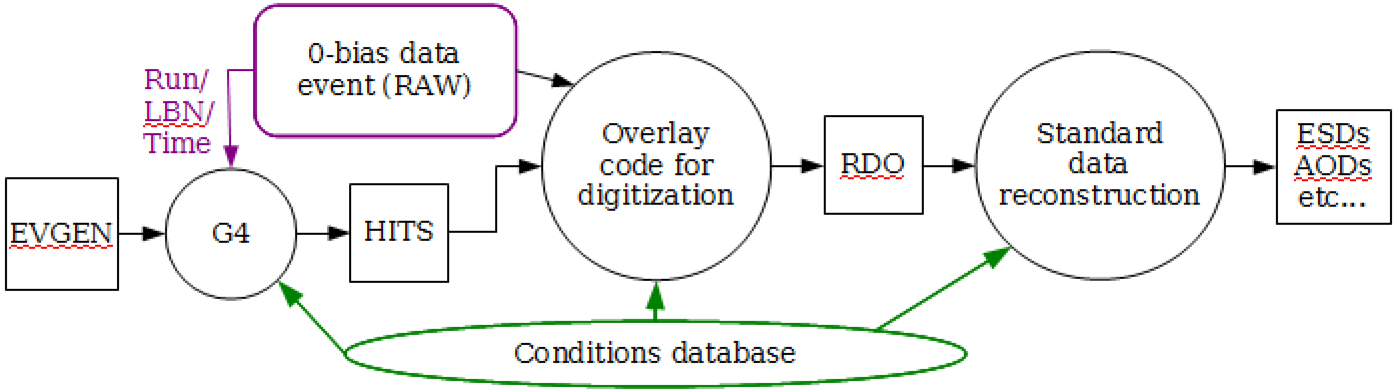
- physics validation
  - ~1 million events split over different performance and physics samples processed on the grid
  - checked and compared to previous validation runs by performance and physics contacts concentrating on relevant physics quantities
  - samples need to be proceeded at highest priority for a quick turn around (< 1 week)
- aims for production
  - large scale validation → detecting rare run-time problems
  - testing job in grid environment
- validation tasks
  - regular validation of new software release for simulation and digitisation+reconstruction
  - validation of performance and technical improvements
  - validation of new simulation related features
    - improved geometry description
    - Geant4 physics lists
    - Geant4 bug fix patches

Upgrade Production

- preparations for Run 2, Phase 2 and beyond
  - planed detector upgrades
    - ATLAS+IBL (Insertable b-Layer (pixel detector extension) for Run 2)
    - ATLAS+ITK (silicon only inner tracker upgrade for Phase 2)
  - machine constraints: 50ns or 25ns bunch spacing and pile-up level
- ATLAS+IBL configuration: 25/50ns and $\langle\mu\rangle$=20, 40, 60, 80
  - simulation time increases due to higher centre-of-mass energy/more particles per event
  - higher pile-up level increases memory usage, especially in reconstruction and trigger simulation
    - reduce trigger menu (<4 GB for 60@25ns and 80@50ns)
    - run on dedicated high memory queues (<6GB for 80@25ns)
    - for Run 2 simulate trigger between digitisation and reconstruction
    - running time: 100 s/evt for mu=20; mu=40 → x 2.2; mu=60 → x 1.8
- ATLAS+ITK configurations: 25ns and $\langle\mu\rangle$=80, 140, 200
  - reconstruction stays well below 4GB
    (trigger simulation not yet supported and no transition radiation tracker)

Zero-Bias Data Overlay

- improve pile-up simulation by using zero-bias data events



- conditions and beam spot need to be adjusted for each signal event to the corresponding zero bias event → run simulation and overlay in one job
- mc12 overlay configuration
  - ensure a representative pile-up sampling in sets of 50 000 events
  - 100 events per job: one signal file, 1 overlay file → 0.4 GB of input files per job
  - 2012 pp 8 TeV zero-bias sample contains 50 million events (160 TB)
    → grid distribution needs to be improved
- overlay heavily used by heavy ion analysis for PbPb and pPb collisions as simulating heavy ion collisions is difficult and very CPU intensive

Summary

- review of ATLAS Monte Carlo production setup
  - event generation, simulation, digitisation, reconstruction
  - configuration
  - resource usage
  - performance improvements
- special production activities are fit into the Monte Carlo production
  - validation, upgrade, zero-bias data overlay
- preparations ongoing for large scale data challenge in 2014
  - test improved software and anticipated LHC conditions
  - get prepared for the next data taking