

ATLAS Workflows: Analysis

Alastair Dewhurst, Federica Legger



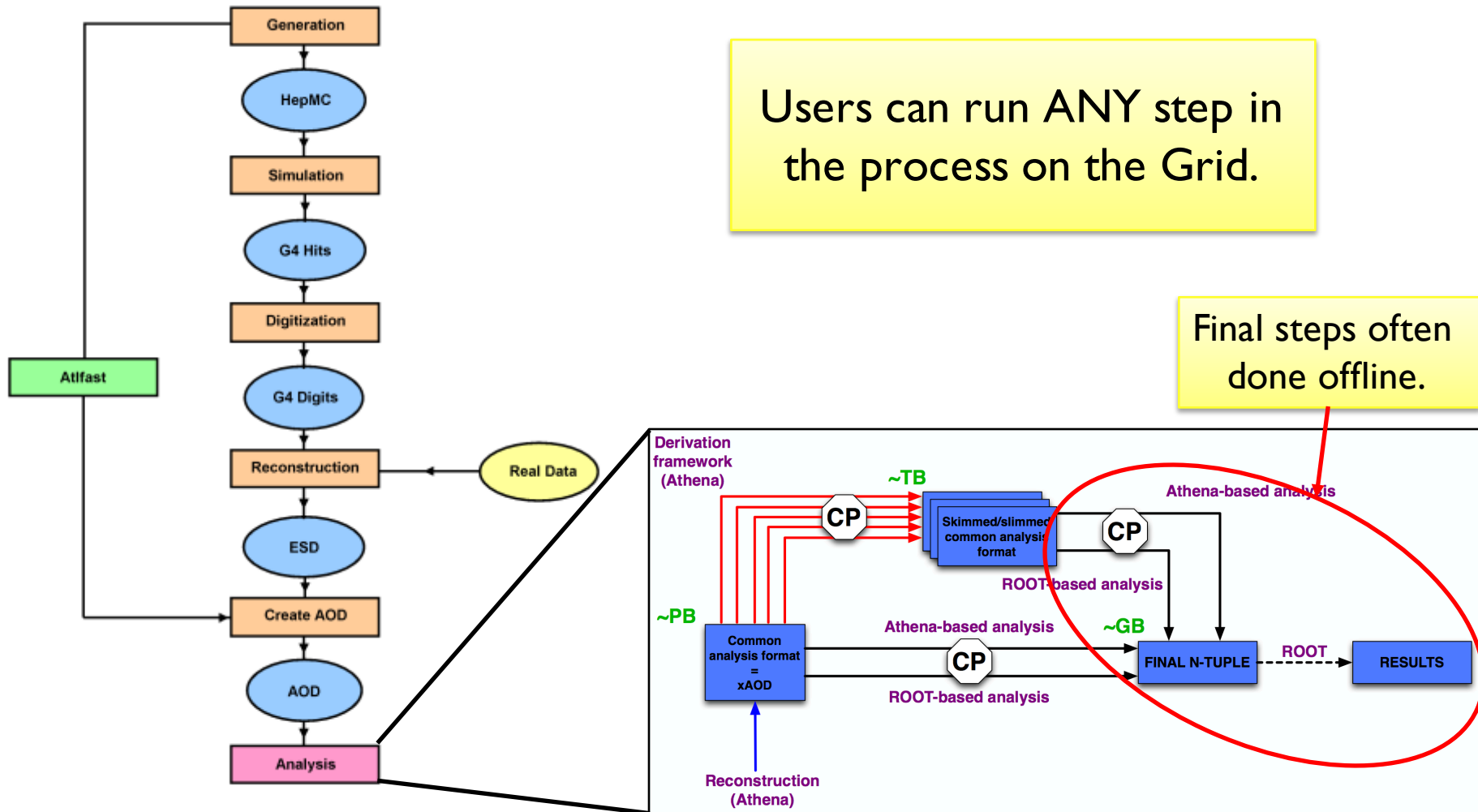
Distributed Analysis

*A Production user is just an Analysis user who has had a cup of coffee.
- Graeme Stewart circa 2009*

- Most users use PanDA to submit their jobs to the Grid:
 - P Athena – Runs Athena (with user modifications)
 - PRun – Allows users to execute any code!
- There are rules on what data can be used in published papers.
 - This prevents private reprocessing and MC generation.
- A lot of analysis work on the Grid is small scale testing/validation of production workflows.



RAW → Results



Users can run ANY step in the process on the Grid.

Final steps often done offline.



Multi-Core

- Users are not currently able to run Multi-Core jobs on the Grid.
 - Queues are technically available at RAL and BNL but very little demand so not expanding.
- There are two pieces of software that will spawn multiple threads if not configured properly.
 - MadGraph
 - PROOF lite
- Tests have shown cgroups limits the cores used without affecting job length (overall CPU usage).



Memory usage

- If users are running modifications on production code, the memory footprint will be very similar.
- If users are running on derivations, the memory footprint is likely to be smaller.
- Memory leaks are still a common mistake, although there are tools to help:
 - Add a new variable but forget to delete it.
 - Test on 100 events works. Fails when submitted to Grid in bulk.
- Sites should protect themselves from the pathological jobs but where possible allow flexibility.
 - E.g. Kill a job if it uses 2 - 3 times what it requests.



PanDA

- When a user submits a task to Panda, it will create 10 scout jobs.
 - If these fail (repeatedly), the task fails and the bulk creation of tasks stops.
- User priority starts at 1000 and fairly rapidly decreases.
 - Separate fair share at sites from production jobs.
- Average length of user jobs is a few hours.
 - Useful for backfilling batch farm. Sites should request a *_SHORT PanDA queue.



Questions?



Alastair Dewhurst, 27th January 2016

