

Workshop on workflow languages for HEP analysis Day 2

Matthew Feickert, Jamie Gooding, Lukas Heinrich, Clemens Lange

April 4th, 2024

Focus of today

Yesterday: general overview of (some) workflow languages/engines

Today: **actual experience with use of workflows in HEP** → **reality check**

- RECAST workflows in ATLAS SUSY
- Snakemake Workflows in LHCb
- Luigi/Law workflows in CMS

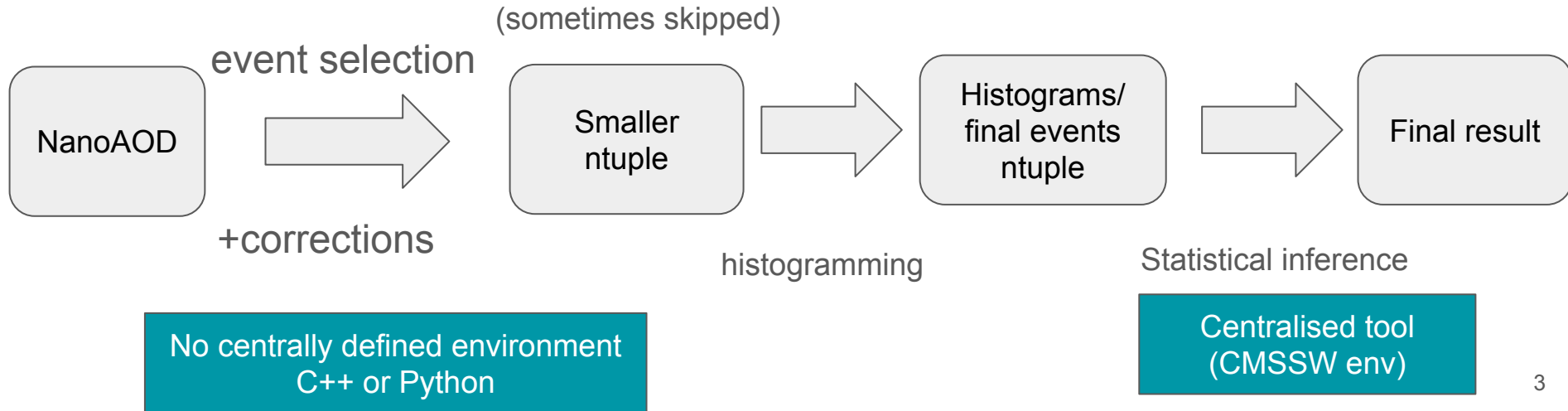
→ Try to answer what is working well and what isn't

Afternoon:

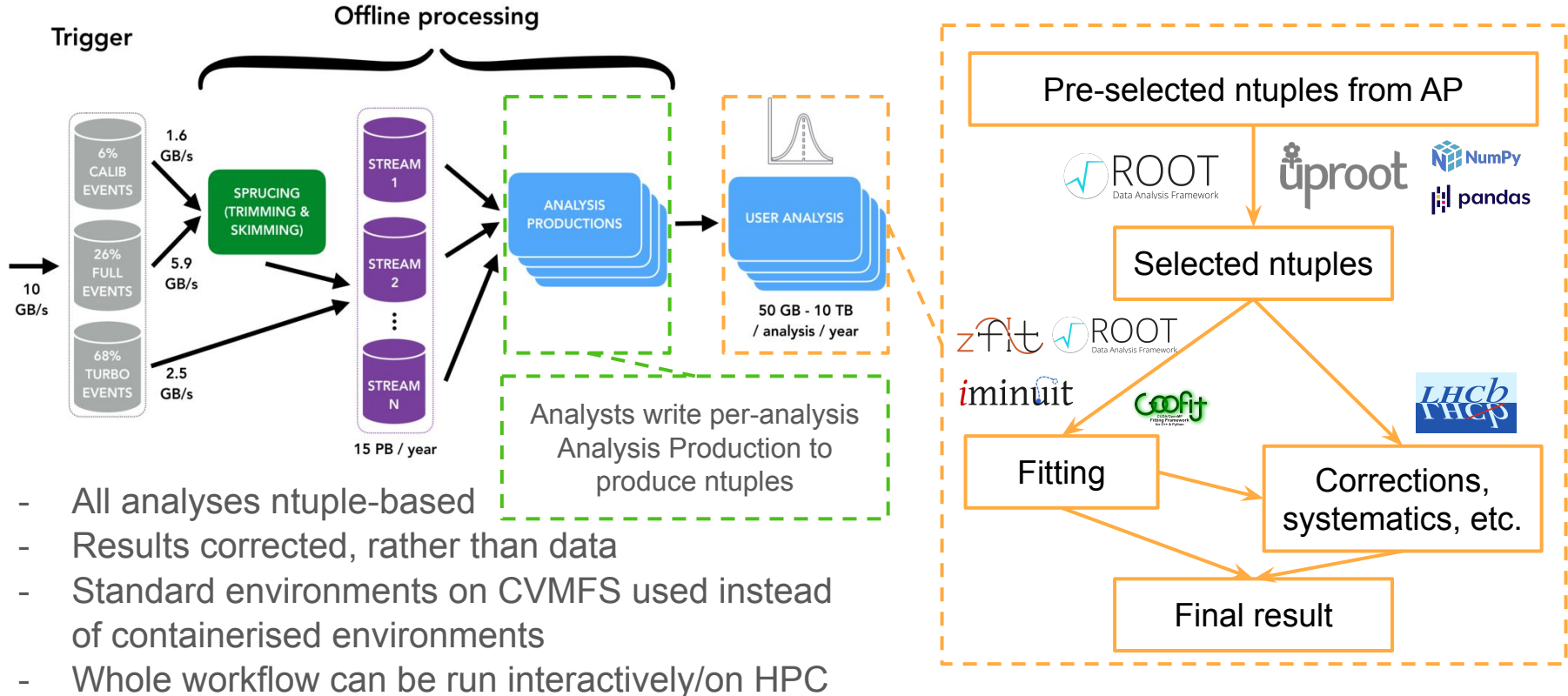
- HEP-CCE
- FAST-HEP: Confronting the challenges of developing a workflow language for HEP analysis

Analysis models of experiments: CMS / ATLAS

- >50% of analyses are using ntuple-like format for analysis
- While initial object corrections are applied, still requires additional scale corrections (for individual physics objects) and event weights
- Lots of iterations at each “arrow”, highly-parallelised batch jobs



Analysis models of experiments: LHCb



- All analyses ntuple-based
- Results corrected, rather than data
- Standard environments on CVMFS used instead of containerised environments
- Whole workflow can be run interactively/on HPC

Questions to be addressed

- Can we consolidate our experience and tools?
- Which tooling do we need to develop/share?
 - batch cluster profiles
 - Rucio and XRootD tools/scripts
 - Authentication

Try to use second half of afternoon to gain some hands-on experience and profit from presence of experts

Live notes

Please keep contributing/adding to live notes:

<https://docs.google.com/document/d/1DbQKHrhJsPo5Qx7QqzEu2e-Orvkpw2NSdc0A8iM3Q1s/edit#heading=h.oaythtr1m21n>

Workshop dinner

The workshop dinner will take place at [Luigia Academy Meyrin](#) on Thursday (today) at 19:00.

It's a 20-minute walk from IdeaSquare. If you want to walk there together, we'll **leave from CERN hostel building 39 at 18:35**. (so we'll be ending today at 18:00 sharp)

Food is covered by our sponsors – drinks will need to be paid for individually