ROOT, Analysis, Experiments, LHCC, and all that

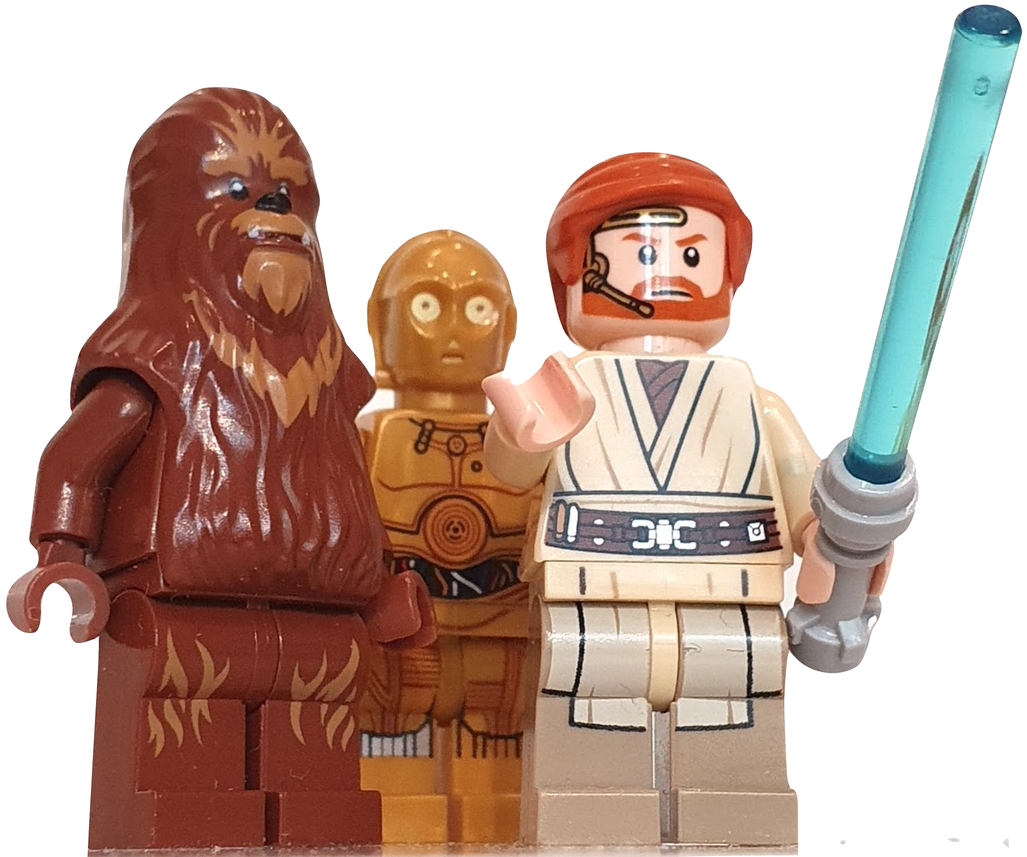
Axel for the ROOT Team, 2021-05-04

Topics

- Analysis
- ROOT's job
- HL-LHC Challenges for Analysis
- Challenges for ROOT

Topics

- Analysis
- ROOT's job
- HL-LHC Challenges for Analysis
- Challenges for ROOT



Analysis

- Traditionally no hard requirements (unlike e.g. online or reco)
  - If software / computing part of analysis takes longer, publication will just be later
  - OTOH analysis performance is significant competitive edge: publish *this* year! React to review *in time* for conference!
- Traditionally deemed "parasitic": use resources as available, all of them
  - Having dedicated, tweaked analysis set-up can be competitive edge

Analysis (2)

- Turn-around time: slide before stays true for doubling
- Factor 10 is a different story:
  - Analysis code that needs 2 week for processing input is still realistic
  - 0.5 years unfeasible for PhD students
- Turn-around time is crucial factor for how HL-LHC data can be used!

ROOT's Job

- Simple, sturdy interfaces for HEP analysis; as efficient (physicist hours and CPU hours) as possible
  - Follow and pre-empt needs of analyses
- "Pick and choose": ROOT is a set of libraries that people interface with
- ROOT's core in C++: consciously designed to enable optimized interplay, such as efficient data movement
- Analyses written in Python or C++, without having to deal with ROOT's internal, optimized parts (similar to other Python libraries)

HL-LHC Challenges

- More data doesn't just mean more analysis input data but also more stat power, means:
  - Can probe more complex models: more parameters / higher dimensions; sys uncertainty dominates: more studies / correlations / higher dimensions
  - Differential analysis / ML-optimization of parameters is expensive; end-to-end optimization even more
- Scaling of computing resources < scaling of data rate; growth of analysis needs >= growth of data rate?

HL-LHC Challenges (2)

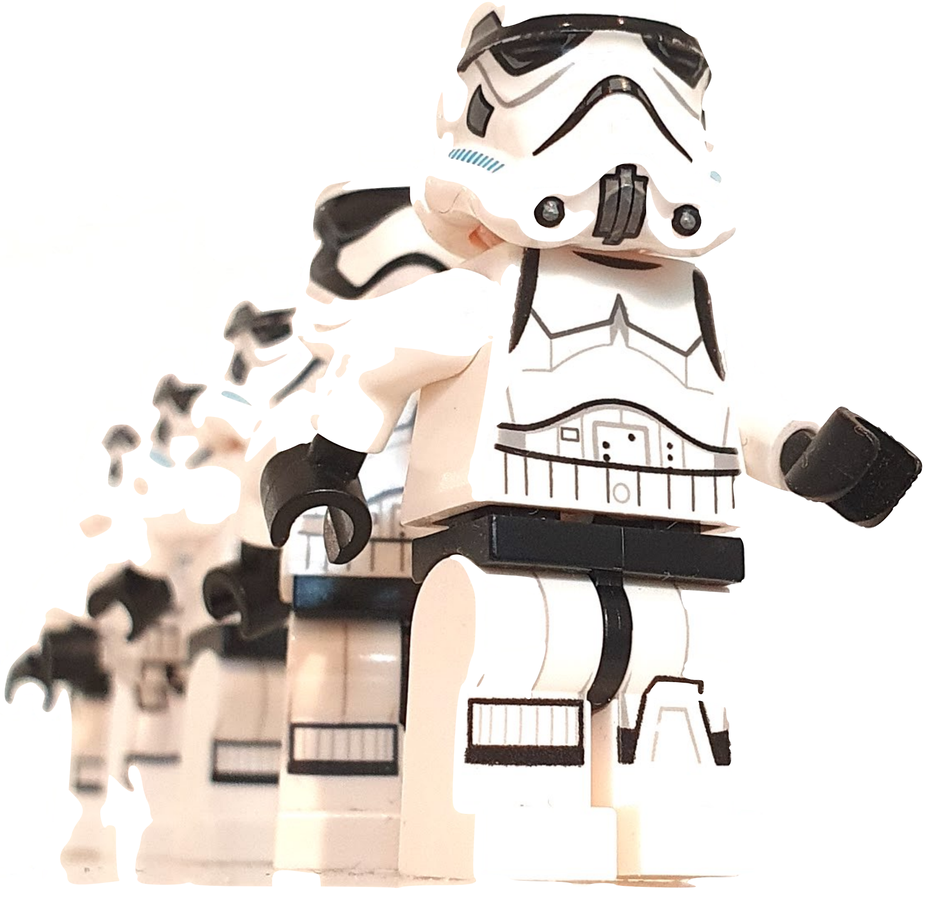
- Expect higher number of samples per analysis
- Expect investment in dedicated analysis facilities:
  - GPUs, effective caching
  - Some analyses require significantly more memory (e.g. high-dimensional calibration/correction/weight maps): beefy machines
- Apart from grid, expect heterogeneous analysis facilities

HL-LHC Challenges for ROOT

- HL-LHC R&D already started years ago. Developments in the works, because it takes a long time to do it right. We assume ROOT data, and pave the analysis path
- Bulk data handling: RDF, ML, RooFit; GPUs. Accelerate wrt ROOT 6.16 by factor 100: threads, SIMD, GPU, better implementations, **accelerated by default!** (I.e. it shouldn't take a ROOT Padawan to have this.)
- Scalable analysis setup: laptop, HPC, analysis facility, uni cluster, grid. Caching, distributed, parallel, heterogenous
- Ergonomics: interfaces, interplay, documentation, debuggability, training

HL-LHC Challenges for ROOT

- Several higher-order challenges
  - Sustainability: use "off the street" knowledge wherever possible; sharing of expertise
  - Benchmark: make sure the community knows why it's using ROOT
  - Community awareness: ROOT 25yo and reliable, yet innovative + evolving, source and sink of R&D also with community
  - Team's generational transition: need to grow succession, both regarding competencies and community trust



Conclusion

Conclusion

- HL-LHC analysis challenges are to a large extent known
- While we cannot know analysis in 15 years (ML^2 in Rust?), we can make sure we have done the groundwork to accommodate it
- I.e. important to not paint ourselves, as a community, into a corner. Don't assume today's solutions will solve everything tomorrow
- Invest in growth and retention of expertise, R&D through new features, sustainability of code and projects
- Goals: ergonomics, efficiency, interoperability, scalability, sustainability