Welcome!

Logistics

Everything is in this room (including lunch today), except dinner tonight and lunch tomorrow. Dinner is at the EVOO restaurant, lunch at Area 4. We will leave from here a bit before 6.30 tonight for diner (it is a short walk), since we've booked a private room for 30 at 6.30 pm.

Schedule:

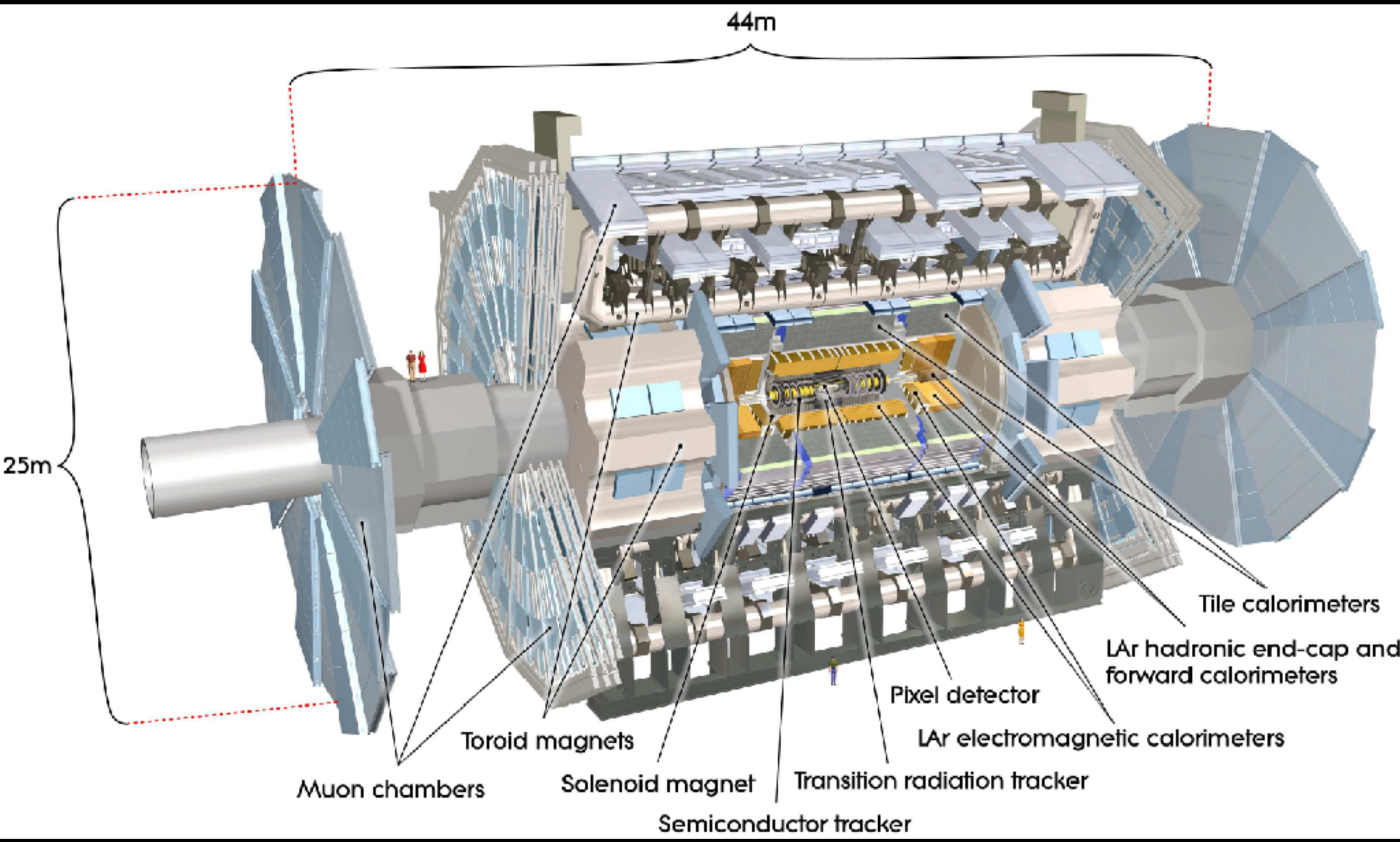
Th morning: accelerators (FPGAs, photonic crystals, etc);

Th afternoon: tracking + CS lightning talks;

F morning: clustering (calorimeters, jets) + CNNs (neutrinos);

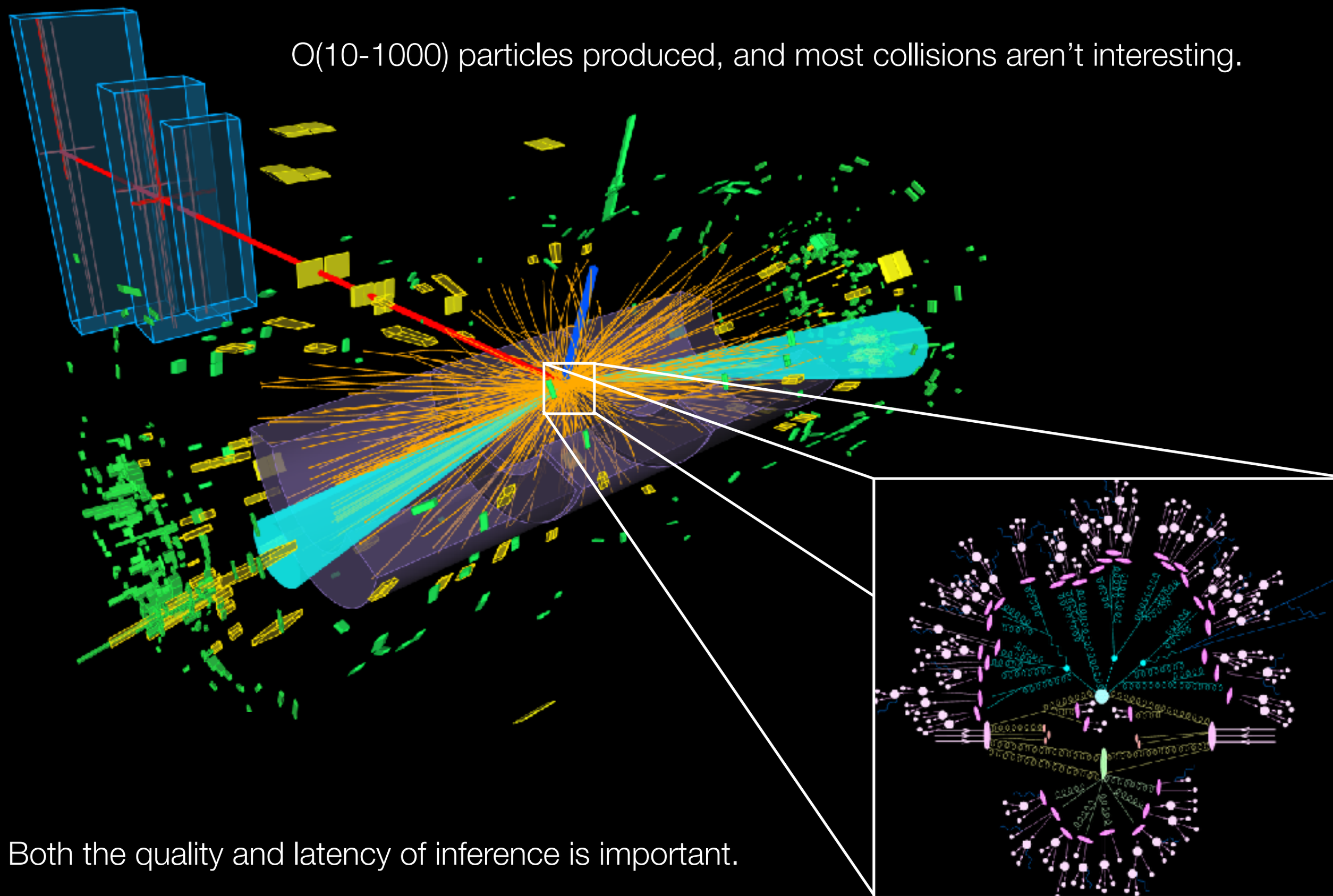
F afternoon: compression (autoencoders), pruning, theory perspective.

There should (hopefully) be plenty of time for discussion during and after the talks — and especially at lunches and dinners. More on some focus items for discussion in a few slides.


44m
25m
Tile calorimeters
LAr hadronic end-cap and
forward calorimeters
Pixel detector
LAr electromagnetic calorimeters
Toroid magnets
Muon chambers
Solenoid magnet
Transition radiation tracker
Semiconductor tracker

Example “Event”

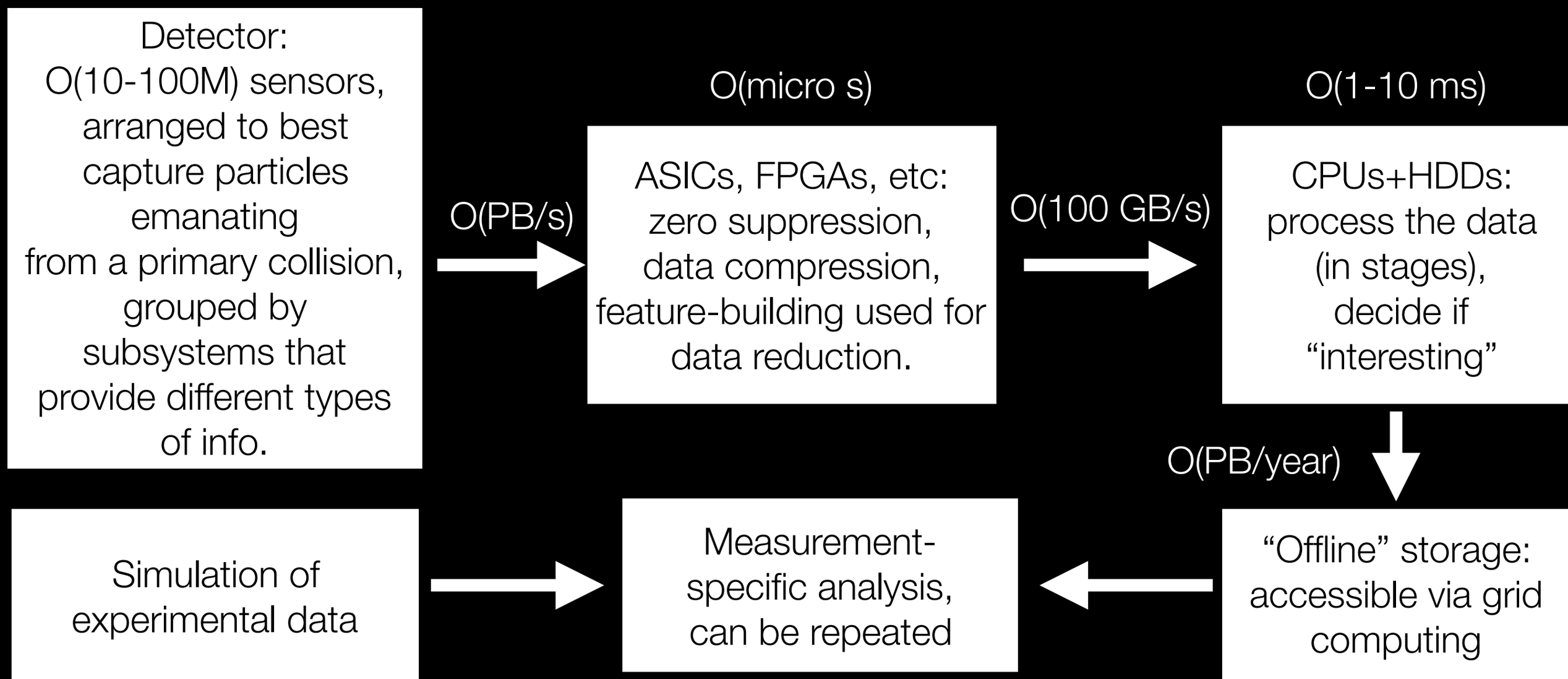
O(10-1000) particles produced, and most collisions aren’t interesting.



Both the quality and latency of inference is important.

LHC Data

The LHC collides bunches of protons every 25 ns (40 MHz), this clock places hard constraints on how we process the data.



Interesting features: sparse, irregular geometries, heterogeneous information, physical symmetries and conservation laws, etc.

N.b., these numbers go up substantially in the next decades, but only minimal advances in CPUs are expected.

HEP + ML

A non-exhaustive list:

- classification of object types, e.g. identifying what type of particle was detected;
- determining particle properties, e.g. regression to improve energy resolution;
- classification of reaction/event types, e.g. selecting signal candidates;
- optimizations
- etc.

We've also provided data for several "Challenges", including one starting now on Tracking (more on this later).

Recently deep learning is taking off:

- replacing shallow algorithms in classification tasks;
- CNNs for image-like outputs, e.g. neutrinos;
- RNNs for jet tagging;
- GANs to replace expensive microphysics (GEANT) simulations;
- etc.

Systematics!!! We care a lot about uncertainties (transfer learning).

Software Institute

How — and even should? (I think we should) — we collaborate?

Past:

Problem-specific feature engineering & shallow algorithms = not much motivation to collaborate across experiments.

Future:

Move to deep algorithms (custom architecture?) from low-level info (similar representations across experiments) = good reason to collaborate.

How should we collaborate? What does “collaboration” even mean here?

Do we need a better platform than Kaggle? (More from Andrey later today.)

What are we going to produce? Workflows, packages ???

Workshops, hackathons, ramps ???