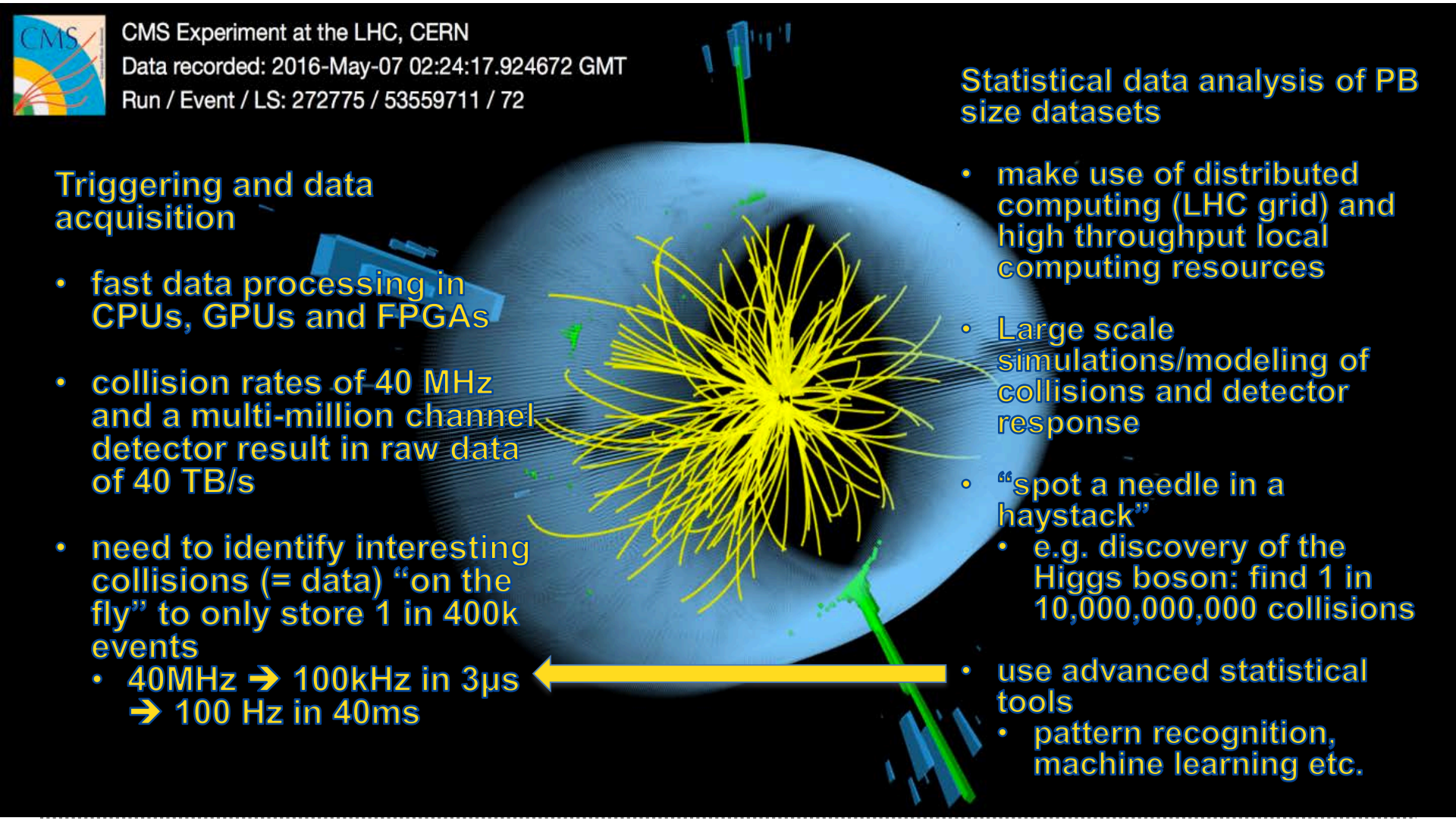
University of BRISTOL

Big Data challenges in Particle Physics

Jean Golding Institute
for data-intensive research



Triggering and data acquisition

- fast data processing in CPUs, GPUs and FPGAs
- collision rates of 40 MHz and a multi-million channel detector result in raw data of 40 TB/s
- need to identify interesting collisions (= data) "on the fly" to only store 1 in 400k events
  - 40MHz → 100kHz in 3µs → 100 Hz in 40ms

Statistical data analysis of PB size datasets

- make use of distributed computing (LHC grid) and high throughput local computing resources
- Large scale simulations/modeling of collisions and detector response
- "spot a needle in a haystack"
  - e.g. discovery of the Higgs boson: find 1 in 10,000,000,000 collisions
- use advanced statistical tools
  - pattern recognition, machine learning etc.