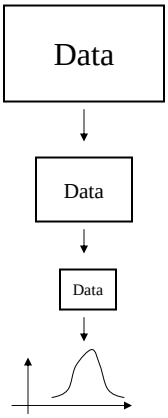
A web-services based data/analysis indexing framework for LHC?

Szabolcs Hernath

1.Physical data analysis: reducing the amount of data to find interesting information



Raw detector data

↓

Events, tracks

↓

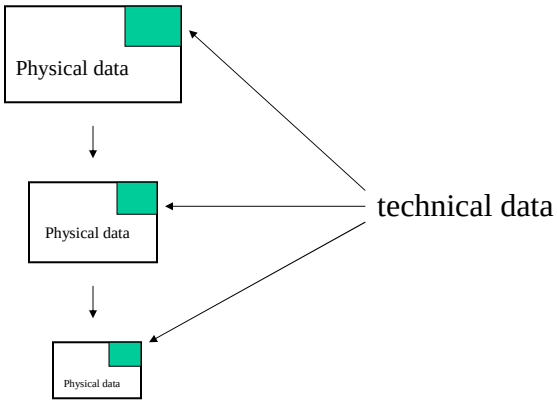
Cross sections etc.

↓

Result: new physics

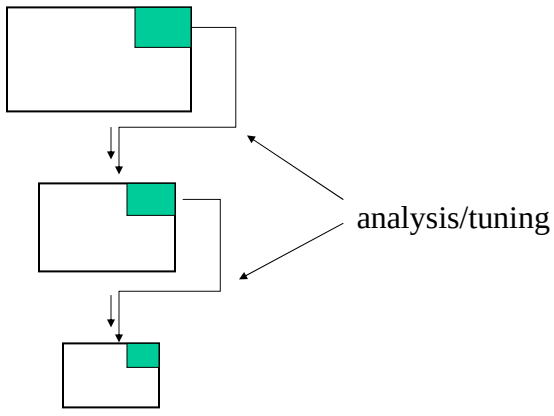
2. Two types of data:

- physically interesting data (kinematics, charges etc.)
- technical data (detector geometry, beam parameters etc.)



3. Analysis/tuning of technical parameters:

will set the way of data reduction (selection criteria)



4. Availability of data:

- physical data: indexed, searchable, widely published
- technical data: often internal to experiments; recorded, but not necessarily indexed/searchable/published

This is not a problem within an experiment!

5. Problems for ‘outsiders’:

- borrowing of ideas: need to understand technical details to adapt solutions.

  This may be difficult in ‘cross-experiment’ cases.

- avoiding duplicate work: how do you know whether some specific analysis has already been performed by others (‘dead branches’).

  This may again be difficult in ‘cross-experiment’ cases.

6. Possible improvement: finding useful ways of indexing technical data/analysis/dead branches

- web based infrastructure
- flexible/extensible
- integration with the GRID
- integration with analysis tools
- integration with the workflows/tools of experiments
- ...

7. Questions:

- is it a problem at all?
- could it be a problem?
- is it already addressed?
- has it been tried and failed?
- is it technically infeasable?
- ...