Brainstorming HL-LHC era computing

Ian Bird
Version 1.0
17 December 2015

At the next WLCG workshop we intend to have a brainstorming day on how computing for the HL-LHC era might evolve. In order to set the scene, I suggest we take the following assumptions as guidelines around which to start the discussions. These are simply suggestions and of course can also be discussed. It is also true that this aims mainly at ATLAS and CMS, since the planning for LHCb and ALICE for the phase 1 upgrades are more advanced. However, we should really aim to be as inclusive as possible and indeed to have an eye on the broader HEP and global science communities as we think about this.

Some assumptions and boundary conditions

Budgets

The current guidance for Run 2 and the foreseeable future is that we should plan around flat budgets for computing. There is presently no indication that this situation is likely to change on any timescale, but with some potential concern that it might decrease. For the moment therefore, we should continue to assume flat budgets.

Data volumes

Anticipated data volumes seem to be on the order of 0.5-1 Exabyte of raw data per year combined across 4 experiments, without major changes in philosophy of data selection/filtering. In addition the compute load with the naïve extrapolations so far made, have computing needs between 5-10 times in excess of what may be affordable as anticipated by technology evolution (which of course itself is very uncertain).

Given these two conditions, we must plan for innovative changes in the computing models from DAQ to final analysis, in order to optimise the physics output.

Cost drivers

We can see three broad areas where changes in model that will drive the cost (see Figure):

1. The amount of raw data to be processed. Can this be more intelligent – e.g. by moving the reconstruction into the "online" (which does not necessarily have to be at the pit – but make a decision based on accurate reconstruction before recording)?
    a. Actually is it even realistic now to assume that we can blindly write raw data at a maximum rate and assume it can be sorted out later?
    b. Storage cost reduction with compression (lossless and lossy) – this will not reduce the offline load, but may help manage the flow.
    c. Clearly most of this implies in-line automated calibrations and alignments, and single pass high quality reconstruction.
    d. Ultimately data volume management might require physics choices.
    e. Consideration of computing costs (esp for simulation) during the design of detectors for the upgrades. (e.g. ATLAS calorimeter is expensive to simulate).
2. New algorithms and re-thinking of how reconstruction and simulation are performed.

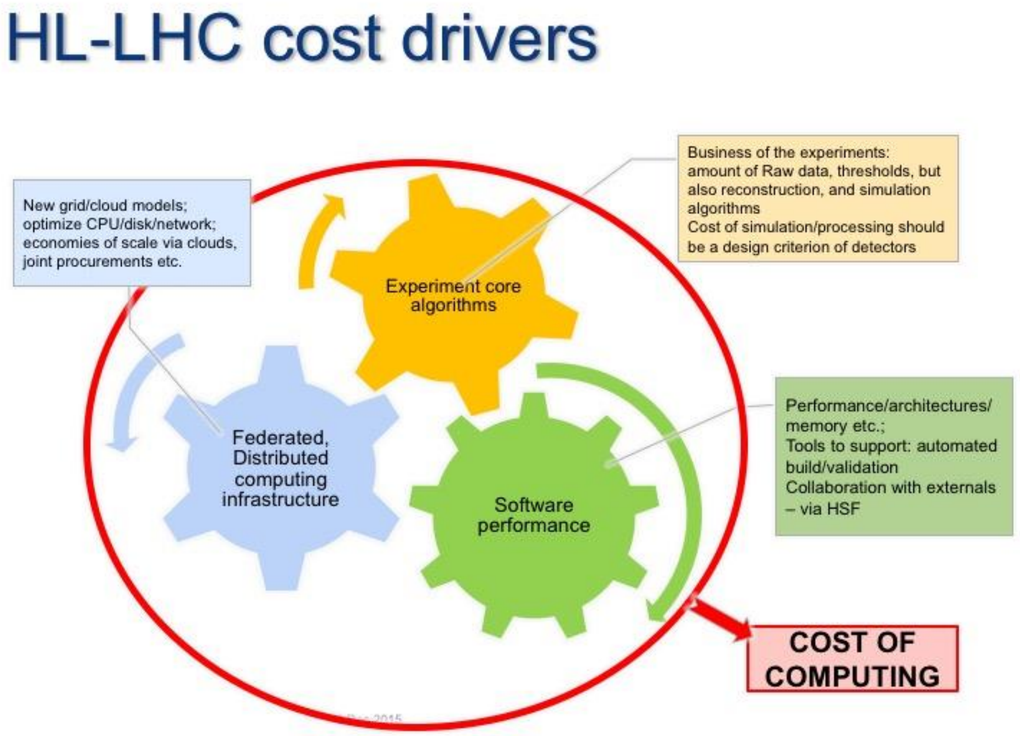
   a. This area needs some inspired thinking and innovation around algorithms.
   b. Adaptation and optimisation of code for different architectures which will continually evolve.
   c. Rethinking of how data & I/O is managed during processing to ensure optimal use of the available processing power.
   d. All of this probably requires a continual process of automation of build systems and physics validation, since it is highly likely that key algorithms must be continually optimised across architectures and compilers. The current situation of single executable for all architectures will not be optimal.
3. Optimisation of infrastructure cost. Many parameters that could be investigated here.
   a. Reduce the amount of disk we need, to be able to acquire more CPU/tape. Today disk is the largest single cost, and is mostly driven by the number of replicas of data.
   b. Can we get economies of scale through joint procurements, large scale cloud providers, etc. On this timescale we may expect commercial computing costs (of CPU and cache) may be much cheaper that our internal costs, if we can reach a large enough scale.
   c. Likely we will have significant network bandwidths (1-10 Tb/s) at reasonable commercial costs.
   d. Should think about very different distributed computing models – for example a large scale (virtual) data centre integrating our large centres and commercially procured resources. Data would not be moved out of that "centre" but rather processed in-situ and available remotely to the physics community. Such a model would minimise storage costs, and enable potentially different analysis models since all of the data could be (virtually) co-located. Simulation load would be maintained at remote centres.

It is clear that 3) alone will not give the order of magnitude improvement in performance, but all of these areas need to be addressed. Point 1) is clearly the responsibility of the experiments, while 2) and 3) can be the subject of collaborative efforts via WLCG. It is probably still the case that the LHCC will want to review computing in the broadest sense which would cover all these areas.

Other considerations

In this timescale, computing (and thus physics) is going to be limited by cost, thus we must make every effort to maximise what can be done, in terms of infrastructure, but also where it costs real skilled effort (such as algorithms and performance optimisation). This requires cross-boundary collaboration (experiments, countries, etc.)

There is significant scope here for common projects and developments, common libraries, frameworks, etc. A significant level of commonality is going to be required by our reviewers (scientific and funding). Four solutions are not going to be sustainable, but we must justify where differences need to be made.



The HSF is also a mechanism through which to address some of these concerns.

Many of these areas also will need infrastructure support at the level of intelligent automation (continuous validation etc.).

The WLCG and the experiments are asked to produce a timeline of checkpoints for planning the costs of computing for HL-LHC.