Ian Bird

WLCG MB, 15th July 2014

Topics for the future

Introduction

- As well as work that is ongoing in WLCG and the experiments, there are a number of topics that need to be addressed

- Some of these may need some coordination and effort

- Today:
  - List some of the most urgent topics for discussion
  - Understand how they should be addressed – working groups, reporting/discussion of existing activities, or ???

Optimisation and performance

- Extend the work started in the TEG I/O benchmark group (Dirk)
  - Need to develop performance metrics and how they should be measured and reported
  - These would be the basis for tuning of overall performance
  - Needs some analysis of parameters and their related costs
    - I/O, memory, cpu, etc etc
  - Core use cases: MC, reconstruction, analysis etc (obviously by experiment)
- This will be essential for the work of the HEP sw foundation, as well as new CPU benchmarking etc
- Related: Techlab exploration of CPU architectures as part of sw foundation, but also related to the performance

Operations costs

- Analyse where the main costs are in terms of effort (for sites and experiments)

- Understand where we could potentially reduce the costs and how to do it

Optimising use of small or opportunistic sites/resources

- Lots of individual explorations already ongoing in several experiments and IT
- Should we have a general
- BOINC (or similar) service to support this?
- More a strategic discussion than technical
  - May need some consideration of chunks of processing (groups of events rather than datasets, for example)
- Could help to make better use of small or occasional resources

YEARS / ANS CERN

Cloud strategy

- Useful to clarify what direction would be most useful here
  - What is the most useful way to support access to resources including clouds?
    - Via a CE as now;
    - Via pilot factories?
    - Direct to a cloud interface?
    - Something else?
    - Is there still a need for (complex) batch systems (will need scheduling at some level)
  - i.e. is there a commonly agreed architecture for workflow in a virtualised environment
- Also need to agree on how to update accounting in this environment
  - Wall vs CPU accounting, how to manage quotas, etc

Other topics

- Networking – do we need something in this area?

- Disk usage reporting (required by RRB/RSG)

- Monitoring activities – what is common infrastructure vs experiment-specific?

YEARS / ANS CERN

Run 3 and Run 4

- How do we start the planning for the 5-10 year timescale
- Activities at various levels in the experiments
- Expectation of the FA's that this includes as much commonality as possible
  - Also anticipated by the LHCC
- Probably should have some brainstorming activity around the possible needs and solutions on that timescale