

WLCG Technical Forum - Proposal

Ian Bird

MB

CERN, 15th September 2015



Background

- As discussed and agreed in June
 - Set up a technical forum to discuss/agree and publish the technical strategy and directions of the distributed computing infrastructure for WLCG
 - Needed for ourselves, but also to address concerns of LHCC and funding agencies that we have a process to ensure that HL-LHC computing can be managed and afforded
 - Also need a way to make sure that any available funding fits with *our* strategy and that we are not driven by external influences

- Proposal for discussion today:
 - Scope, topics to address (starting set), membership

Scope

- ❑ Prepare the long-term future of the WLCG infrastructure, complementary to, and supporting, the evolution of the computing models of the experiments.
- ❑ The timescale of Run 3 and Run 4 are the focus, although improvement should be ongoing.
- ❑ Understand how the overall distributed computing infrastructure must evolve in preparation for the upgrades, and to adapt to the changing environment and technology.
- ❑ In particular investigate where reductions in necessary effort and resources may be achieved, through consolidation, automation, and novel techniques.
- ❑ Review all aspects of the infrastructure and its implementations
- ❑ Key points:
 - ❑ Must show that where it is feasible we can have common solutions and implementations – this will be essential to convince FA's
 - ❑ Keep in mind the desire to have an infrastructure/tools useful for HEP globally

Members & Reporting

- ❑ Experiment nominees (1+1) + experts as needed
- ❑ Site representatives (1/Tier 1 + 2-3 reps of Tier 2's)
- ❑ Technical providers as needed and invited
- ❑ Chair:
 - Nominated by PL, 2-year term?
- ❑ Reports to:
 - MB
 - Uses GDB (& workshops) for broader discussion of issues

Topics

- ❑ Propose a list of key topics that should be reviewed for the long term evolution
- ❑ However, MB should use the group as a core of expertise to address other technical concerns that don't necessarily fit with operations coordination or that have longer term implications

Topics to be discussed

- Security and trust model
 - Trust VO's or users?
 - Traceability and reactivity vs “control”
- Software performance
 - Supporting HSF – but needs to be clear it is a key concern
- Compute:
 - Simple access to any resources: grid, cloud, HPC, volunteer, etc – what is minimal need?
 - Need for “CE”
 - Need for batch systems?
 - Volunteer computing service as mechanism for “small” sites
 - Cloud provisioning: interfaces and how to provision; commercial clouds as extensions of sites, stand-alone?
- Data:
 - Common set of data management tools?
 - What are key functionalities/use cases
 - A&A vs performance
 - Better use of disk – roles of Tiers
 - Consolidation of functions
 - Evolution of federated storage mechanisms
 - “Dropbox” functionality etc
- Networking:
 - Evolution of LHCOPN, LHCOne
 - Deployment of IPV6
 - Monitoring/performance
 - Evolving requirements
 - Need for advanced tools? Define why.

Topics ... (2)

- Monitoring and information systems
 - What are use cases now
 - Service discovery vs monitoring
 - Include reporting needs, e.g. accounting
 - SAM-like functionality still needed? What?
 - Dashboards – what?
- Accounting
 - Job vs client (need for clouds to validate billing)
 - Understand limitations and define strategy
- Benchmarking and performance monitoring
 - Benchmarking
 - In-job performance check
- Long term distributed computing models
 - Long-term thinking – e.g. virtual data centre ideas, different analysis models
 - How best to use offered sites – large vs “small” – “data” sites vs “compute” sites
- Define anticipated running conditions for Run 3 and Run 4; resource models
- Simulation/modelling of the overall system?
- Need for prototyping long term ideas?

Discussion & next steps??

Concept and scope?

Proposal of chair person

If we agree:

- Should start immediately
- Draft outline whitepaper with main ideas for the future
- MB should prioritize most urgent topics