



From the July WLCG MB

last GDB about data management activities within WLCG. The following comments were noted:

- Ian Bird observed that there are many activities ongoing within WLCG in the data management area. There are also opportunities for EU funding that are coming up in the near future and it would be useful to revisit activities like the Zephyr proposal in this context. He stated that he would strongly support having a group of people to coordinate these activities and suggested that it would be useful to prepare a draft mandate for this before the pre-GDB in September. Oliver Keeble agreed and stated that he would follow up. Ian Collier expressed his support for this idea and suggested that something like a twice-yearly pre-GDB on data management could be very useful. Torre Wenaus agreed that this is a critical area where communication and keeping an eye on open opportunities are very important. Stefan Roiser noted that this would especially be useful to follow up on the commonalities between the experiments, the move to SRM-less sites being one example. Frank Wuerthwein noted that there is one site in the US (Nebraska) that has no SRM anymore. Ian Bird concluded that the proposal was endorsed by the WLCG MB and should now be implemented.

Action: Oliver Keeble to prepare a draft mandate for a WLCG data management coordination team before the pre-GDB in September.

5. Information System

My take

- Plenty of bi-lateral communication
 - Exp<->site
 - Site<->storage
 - Storage<->Exp
- So what's missing?
 - Exp<->Exp (“let's ask for the same thing”)
 - Site<->Site (“let's evolve the infrastructure coherently”)
 - Storage<->Storage (“let's implement stuff compatibly”)
- i.e. this is needed for infrastructure transitions
 - “optional SRM” is a good example”

WLCG data coordination

- Would a more structured approach to WLCG communication about data be desirable?
- What form would that take?
 - mailing list (suggest converting wlcg-data-access-wg)
 - representation within WLCG operations
 - regular preGDB slot or other forum
 - Small steering group
 - Large Working Group

Proposal

- A data steering group is created
 - One person per experiment, storage provider, 2 or 3 for infrastructure
 - The representatives understand they are undertaking a WLCG responsibility
 - Would probably need a chair, or two co-chairs.
- This steering team
 - Tracks relevant technical themes
 - Meets every N months
- And on that basis oversees
 - A pre-GDB slot every 6 months, which should be confirmed 2 months before. Focussing on medium term developments.
 - Triggering of task forces if necessary
 - The WLCG Workshop
 - Creation of a `wlcg-data-management@cern.ch` list.
 - Reporting to WLCG ops meeting where requested.

WLCG Workshop S.F. - topics

- Current issues
 - Today's points
 - Caching solutions
 - Multi-site storage
 - Local federations
- Longer term / disruptive
 - Next-gen namespaces
 - SRMless tape
 - Event-level access
 - Object stores
 - In the cloud?