



User Analysis - background

- Clarify roles of Tier 1, 2, 3 for user analysis for each experiment
 - Batch Analysis – the assumption is that this is grid based
 - End-user Analysis – what is this? – laptops, local analyses, ...?
 - → what are the distinctions?
- The presentations to the GDB on 8/10/08 are a good summary of the analysis models:
 - <http://indico.cern.ch/conferenceDisplay.py?confId=20234>
- It is important to understand what is missing to allow the experiments to implement and manage these models. This includes correct configuration of shares in batch systems, appropriate set ups of disk pools, etc. and a summary of which tools may be needed to implement the models. What is the set of services that are needed (in addition to what exists now) to support these analysis models?
 - NOT an open door for new service development requirements!
- What is a “standard model” for a Tier 2 supporting analysis?
 - E.g. File systems (Lustre, NFS4, gpfs, AFS, xrootd) with what interface, etc.?
 - How many users? Size per user, etc. How to scale the resources?



Mandate

- Clarify and confirm the roles of the Tier 1, 2, 3 sites for analysis activities for each experiment; categorising the various types of analysis: batch or interactive, centrally organised or individual users.
- List the minimal set of services, their characteristics (availability, scale, concurrent usage, users, etc.), or changes to existing services, essential to implement those models. These must be prioritised clearly between “essential” and “desirable”. The expectation is that this is addressing configuration of services, recommendations for service implementation, or new tools to assist managing or using the services rather than requests for new services. Such developments must be clearly justified.
- Deliverables:
 - Documented analysis models
 - Report on requirements for services and developments with priorities and timescales.



Scope

- The scope of the working group extends to analysis at any of the Tiers that are grid resources. It should also include the access to data residing at Tier 2 or 1 from Tier 3s, but if the Tier 3 is a local resource (not grid enabled) then it is out of the scope (apart from data access tools).
 - Tier 0, 1, 2 are grid enabled resources that include the formal resource commitments to WLCG via the MoU, while Tier 3 could be grid enabled or not and look like a Tier 2 or be as simple as individual laptops.



Membership & Timescale

- Membership
 - 1 -2 people from each experiment
 - 4-6 people from Tier 0/Tier 1 and Tier 2 sites
 - Representatives of Data Management development teams
 - Chairperson – Markus Schulz

- Timescale
 - Group should start immediately
 - Report by end of 2008, monthly reports to MB