User Analysis Workgroup Update

- **All four experiments gave input by mid December**
  - ALICE by document and links
    - Very independent from underlying infrastructure
    - Shares, allocation, access control by Alien framework
    - Production like analysis with "Train" concept
    - End User analysis by non grid means "PROOF" and via Alien task queues
  - LHCb by mail and links
    - Not very detailed at the moment
    - End user work requires very little storage (desktop scale)
  - CMS by a document with concrete requirements
    - In terms of missing functionality
    - In terms of storage requirements
    - In terms of split between production and analysis use of CPU
  - ATLAS provided a set of links
    - Two ways to access the resources.

General

- The analysis models of the experiments depend to a large degree on their own grid systems that are layered on top of the provided infrastructure
  - ALICE is in this respect very independent and flexible and has simple requirements
    - Since most complex aspects are handled in the Alien layer
  - The models look on a high level similar, but the different implementations impact the infrastructure in different ways
    - Calibration data access, data access control
  - In addition several different systems are used by the larger experiments
    - Pilot/WMS based submissions

- Which makes it difficult to give universally applicable advice to sites
  - And have a generic set of requirements for services

General

- **All experiments have exercised their frameworks**
    - It is not clear how close this is to the activity level when we have beam

- **For data access there is no clear metric to measure a T2s capability**
    - Data size, number of files accesses, parallel active users is too simple
    - Each system impacts the fabric a different way
        - The internal structure of files has shown to have a large impact on the I/O efficiency (which is still under investigation)

- **It might be instructive to take a look at some of CMSs requests and observed issues**

CMS

- **Strong concept of locality**
  - Each physicist is assigned to one or more T2s
  - User data is tied to these sites
    - Transfers with experiment's tools only after registering the files in a global catalogue
  - Jobs are sent by CRAB to the data
  - In addition official working groups manage their own data

- **However, this locality cannot currently be enforced by the underlying infrastructure**
  - It is enforced indirectly by the experiments tools

CMS T2 storage

- T2 storage
    - >20TB temporary space for production, controlled by prod team
    - 30TB space for centrally managed official data sets
    - N* 30TB for each official physics group
        - Each group can have multiple sites
        - Each site can host multiple groups
    - Regional (local) user spaces
        - Managed regionally
    - SRM based stage-out space (future)
    - T3 space at the site

- Complex ACLs and quotas are required
    - VOMS can express this
        - By using groups for locality and roles below that level
        - Complex mapping at the fabric level

CMS T2 CPU

- 50/50 % share between production and all other activities
  - With a timeframe for equalization of shares under load of less than a week

- Split and prioritization on a granularity of analysis groups and taking locality into account
  - Currently VOMS is only used for prod/analysis distinction
    - And at some places to express locality
  - The full implementation with the current infrastructure will be very cumbersome and put a significant load on the information system

- Prioritization: For individual users based on recent usage
  - Requires fixed mappings between users and local ids to work
    - For each and every VOMS identity of the user
    - Even more accounts….

- Related: lcgadmin mappings and priorities

CMS ISSUES

- Storage systems reliability
- Improved SRM APIs and tools (bulk ops)
- Quotas, accounting and ACLs for space
- SE and CE scalability
  - SRM operations
  - CE high load figures
    - Should have improved with recent mods (see GDB)
    - LCG will not handle the large number of groups and roles that are required to implement the desired behavior
- Lack of support for multi user pilot jobs (parallel to push model)
  - SCAS/glexec soon, GUMS/glexec in use
- Batch scheduler configuration appears to be not inline with agreements
- Lcgadmin mapping

- Except for the more divers resource allocation requests the reported issues apply as much to analysis as they do to production tasks.

Summary

- On an abstract level the different analysis models look similar
  - But the used tools are very different and by the way they are layered individually on the infrastructure create different requirements and constraints
    - Job flow, data transport and access, calibration data access
  - The analysis specific problems that are encountered are mostly specific to the way the infrastructure is used (stressed).
    - Data access, catalogue operations
    - Rollout experience supports this observation
- It is very difficult to estimate the true scale of activity after LHC start
  - In terms of users, "grid users" etc.
- There is a shift from Push to Pilot
  - Which will make a huge difference for the job management
- Is there a similar move for data access (xrootd)

Summary

- For resource configuration issues direct communication between the T2s and their experiment seems to be more efficient

- The common problem domain is storage specific
  - And solutions depend on the SE's implementation
  - And the details of the experiments access
  - There measurements of the systems have started
  - The analysis use cases are better handled within the scope of the SEs