

# DQ2 'User' Experience

---

Borut Kersevan  
Jozef Stefan Inst.  
Univ. of Ljubljana



---

# Which perspective?

- Getting the CSC event generation to work
  - MC convener/CSC production manager.
- What we need to do is to pass input files of pre-prepared MC events to the Athena event generation loop.
  - Means registering hundreds of small files on the grid (one file/job).
  - The procedure evolved up to using DQ2 tools.
  - <https://twiki.cern.ch/twiki/bin/view/Atlas/PreparingLesHouchesEven>

---

# General experience

- The general experience when dealing with DQ2 and datasets registration is I believe quite positive:
  - Reasonably well documented.
  - Simplifies the procedure a lot.
- Now about the hiccups..

---

## Non-CERN DQ2

- The DQ2 tools installation seems to be somewhat non trivial...
  - Disclaimer: I haven't tried it myself.
- The symptom being that there were complaints that people cannot register the datasets properly from outside CERN/Ixplus:
  - Means one has to copy all the files to CERN/Ixplus and then register them from there: Difficult considering the mangy AFS quotas..

---

# CERN DQ2

- I know this is at least partially something that is out of DDM powers but still:
  - Quite often file registration is **very slow/failing** : more than 10 mins/1MB file (now multiply this by 100 files)..
  - When people go for lunch/short holiday while this goes on some datasets fail to register (e.g. #37 of 200 files) which people often miss till too late (CSC jobs failing).
    - In such cases the DQ2 warning might provide some enhanced visual effects.. (red flashing text?)

---

# CERN DQ2 cont'd

- When a file registration fails things get hairy:
  - Sometimes the file is physically present in CASTOR but not in LFC and DQ2
  - Sometimes the file entry is in DQ2 but not LFC and castor
  - All possible permutations of above..
- Removing a bad entry is not a one-line dq2 thing:
  - Finding the physical location of a file in CASTOR is not self-evident to everyone.
  - Removing LFC entries requires lower-level middleware commands.
  - Etc..
- Bottom line: **users get frustrated!** Possibly a simple tool/command would come handy? (**`dq2_verify_and_remove`**)

---

# CERN DQ2 cont'd

- A comment from my part:
  - When user's/responsible don't know how to fix missing/corrupt files this gets 'forwarded' to me..
    - dq2 states I am not the owner
    - I write a longish mail back what to check etc..
  - Can we have group permissions? No? Something like 'lower-level' admin?

---

# Documentation

- A lot of it is there, as I said, quite useful.
- Still...
  - Some more details would not be resented, like:
    - How do I check who owns a dataset?
    - What is the proper way to replace a corrupt/missing file?
    - Why does my dataset not like me/ what does 'closed' mean?
  - A FAQ?



---

## To sum up

- DQ2 tools are definitely a very welcome occurrence..
- Successfully used on a daily basis by a lot of people (from my chunk of CSC)
- There would be something wrong if we had no complaints...
- ... but keep up the good work!