



PanDA user analysis on Google PanDA queue

Google/ATLAS meeting
08 February 2023

Johannes Elmsheuser (BNL)

- Lots of debugging of user analysis on Google PanDA queue last week
- Mario/Cedric deployed patch for Rucio to add “`&multirange=false`” to signed TURL
 - `rucio list-file-replicas data22_13p6TeV:DAOD_PHYSLITE.30136824._000001.pool.root.1` returns now
GOOGLE-EU_DATADISK: `https://storage....3D&multirange=false&nconnections=10`
- Tested with prun jobs at `GOOGLE_EUW1_DYN` PanDA queue
 - Paul released new PanDA pilot version 3.4.8.6 to remove an old replacement of “&” with “&” and “8” in Google signed TURLs
 - New Pilot version hardcoded in CRIC - not yet released in production
- Fernando switched `GOOGLE_EUW1_DYN` and `GOOGLE_GPU` PanDA queues to direct I/O in CRIC
- AnalysisBase type analysis should hopefully work fine (see later “Caveats” slide)

- User access to Google storage:
 - Until today only selected users enabled in Rucio could write to GOOGLE-EU_SCRATCHDISK
 - Rucio team added today a quota for all users (20 TB) and added a 100TB watermark.
 - Allowed a general 2nd prun/pathena output copy on GOOGLE-EU_SCRATCHDISK
 - > monitor dataset replication and deletion on GOOGLE-EU_SCRATCHDISK
- Open GOOGLE_EUW1_DYN for user analysis:
 - GOOGLE production queue is fully used for TCO studies in the coming weeks
 - Enable user analysis in March (?)
 - Replicate of 0.5-1 PB of DAOD_* ?

Found 2 problems in Athena and ROOT:

- Athena code might not work with direct I/O https Google signed TURLs
 - Lines 62-96 in <https://acode-browser1.usatlas.bnl.gov/lxr/source/Gaudi/GaudiSvc/src/FileMgr/RootFileHandler.cpp> uses ROOTs TWebFile instead of TFile::Open with davix
 - Not sure if this old code is actually anywhere else - requires update of Gaudi
- ROOT TChain confuses/mangles Google signed TURLs:

```
import ROOT
filename="https://storage.googleapis.com:443/atlas-europe-west1-datadisk/rucio/data22\_13p6TeV/DAOD\_PHYSLITE.30136824\_000001.pool.root.1?GoogleAccessId=atlas-rucio-prod@atlas-rucio-prod.iam.gserviceaccount.com&Expires=1675437388&Signature=blabla&#multirange=false&nconnections=10"
chain = ROOT.TChain()
chain.AddFile(filename)
el=chain.GetListOfFiles()[0]
el.GetTitle()
'https://storage.googleapis.com:443/atlas-europe-west1-datadisk/rucio/data22\_13p6TeV/DAOD\_PHYSLITE.30136824\_000001.pool.root.1?GoogleAccessId=atlas-rucio-prod@atlas-rucio-prod.iam.gserviceaccount.com&Expires=1675437388&Signature=blabla&'
el.GetName()
'#multirange=false&nconnections=10'
```