

LUXE SAS meeting, 16th May 2022

Follow-up on proposal for dataset naming and plan for grid storage

Yee Chinn Yap, Federico
Meloni (DESY)

Introduction

- ❖ See last talk [here](#)
- ❖ Refined naming policy.
- ❖ Thanks for the suggestions and for looking into the files currently on DUST.

Dataset naming policy

`{Project}.{Process}.{Generator}.{Type}.{Tag}`

- ❖ Project: mc21 (no differentiation of mc campaign until we have uniform framework), dataXX (2-digit year)
- ❖ Process:
 - ❖ Signal: elaser_xi{}_{}_TW_{lp/cp}, glaser (brem-laser), ilaser (ics-laser)
 - ❖ Background/others: ebeam, gbeam, singlepositron, eneedle, etc.
- ❖ Generator: IPstrong, ptarmigan, G4gun, or runXXX (data run number).
- ❖ Type: GEN, SIM, DIGI, REC, NTUP_GEN, AUX (for logs and yaml/mac), RAW, ...
- ❖ Tag: 4 digit tags for generation, simulation, reconstruction, etc.
 - ❖ Separate sim tag into e-laser/g-laser setup: seXXXX, sgXXXX.
 - ❖ Separate reco tag into detector component (until we have complete framework): rtXXXX (tracker), rcXXX (calo), ...
 - ❖ One tag per g4 commit version or ptarmigan setting (apart from xi value and seed).

Files on group DUST

- ❖ Google spreadsheets with list of directories and plans for them (delete/keep/move to grid).
 - ❖ Please have a look and comment if you're using those files!
 - ❖ Proposed new names for those marked for moving.
 - ❖ In the same document, you can find also the tag database (which will be moved to confluence) and the naming policy.
- ❖ IPstrong generator-level and simulation samples, do we want to keep them on grid? If so, is everything there needed? Or can they be deleted?
- ❖ Test beam data also to be archived on grid.