On demand data demo.
DAOD datasets and Lifetime Model

Alexei Klimentov

Brookhaven National Laboratory

WFM SW Weekly Meeting. August 11, 2022

Thanks

Tadashi, Xin, Torre, Andreu, Cedric, Mario, Martin, Sasha

General Considerations

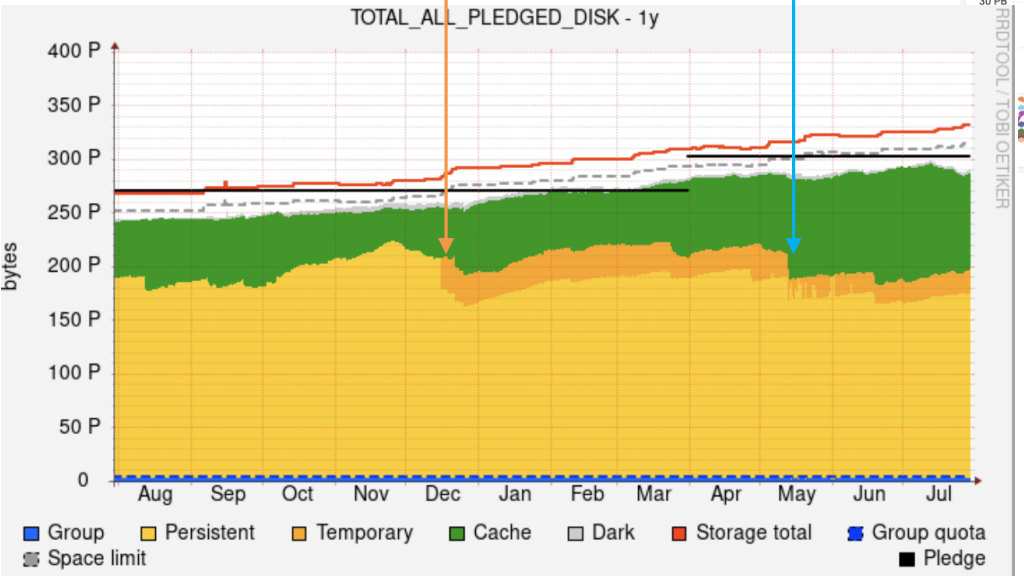
- Lifetime Model, data popularity, data placement, data deletion, and (may be) data carousel and data caching should be discussed on a coherent way

DDM Global Accounting

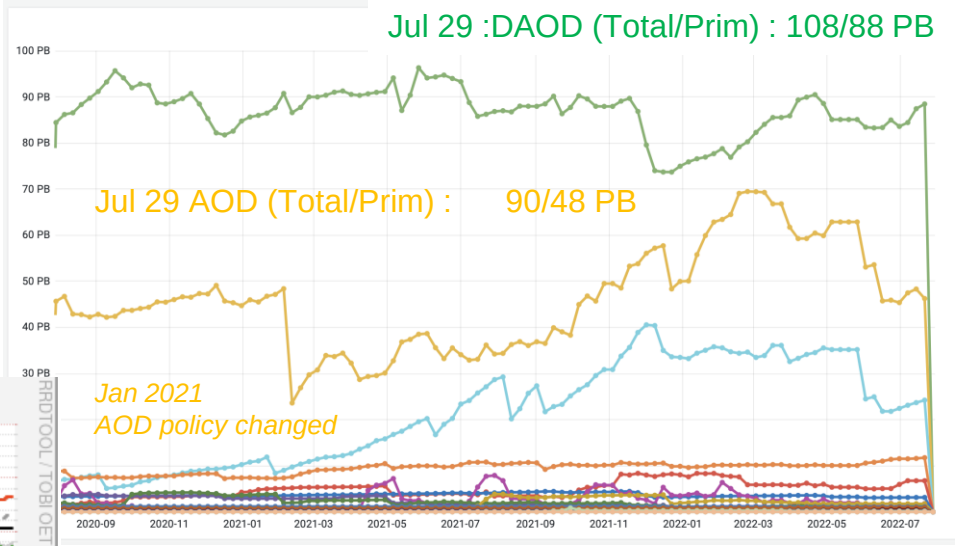
Pledged disks Aug 2021 – Aug 2022

May 2022 :
“catmore rule” execution frequency increased

Dec 2021 :
New data class Temporary

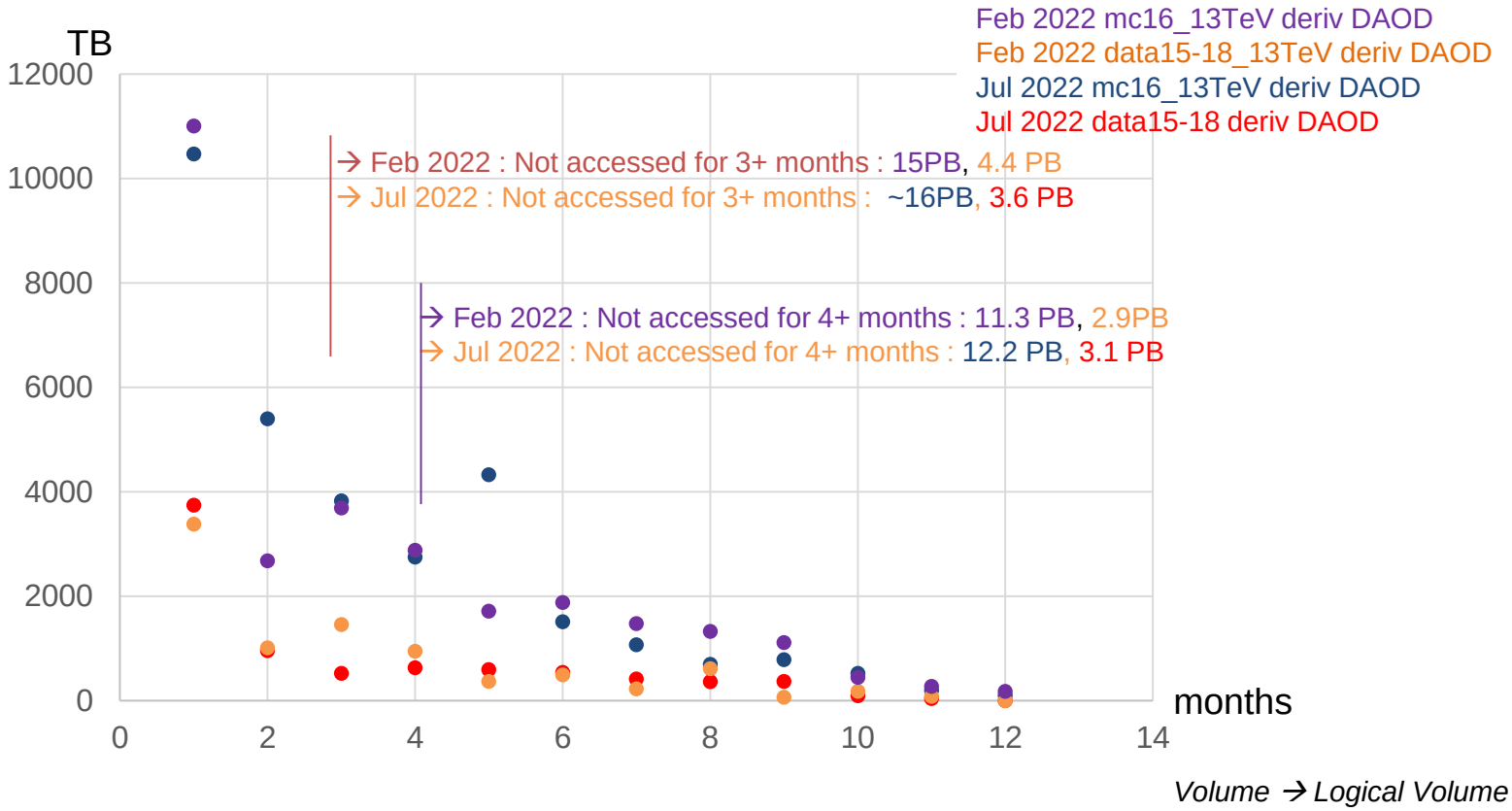


Jul 29 :DAOD (Total/Prim) : 108/88 PB



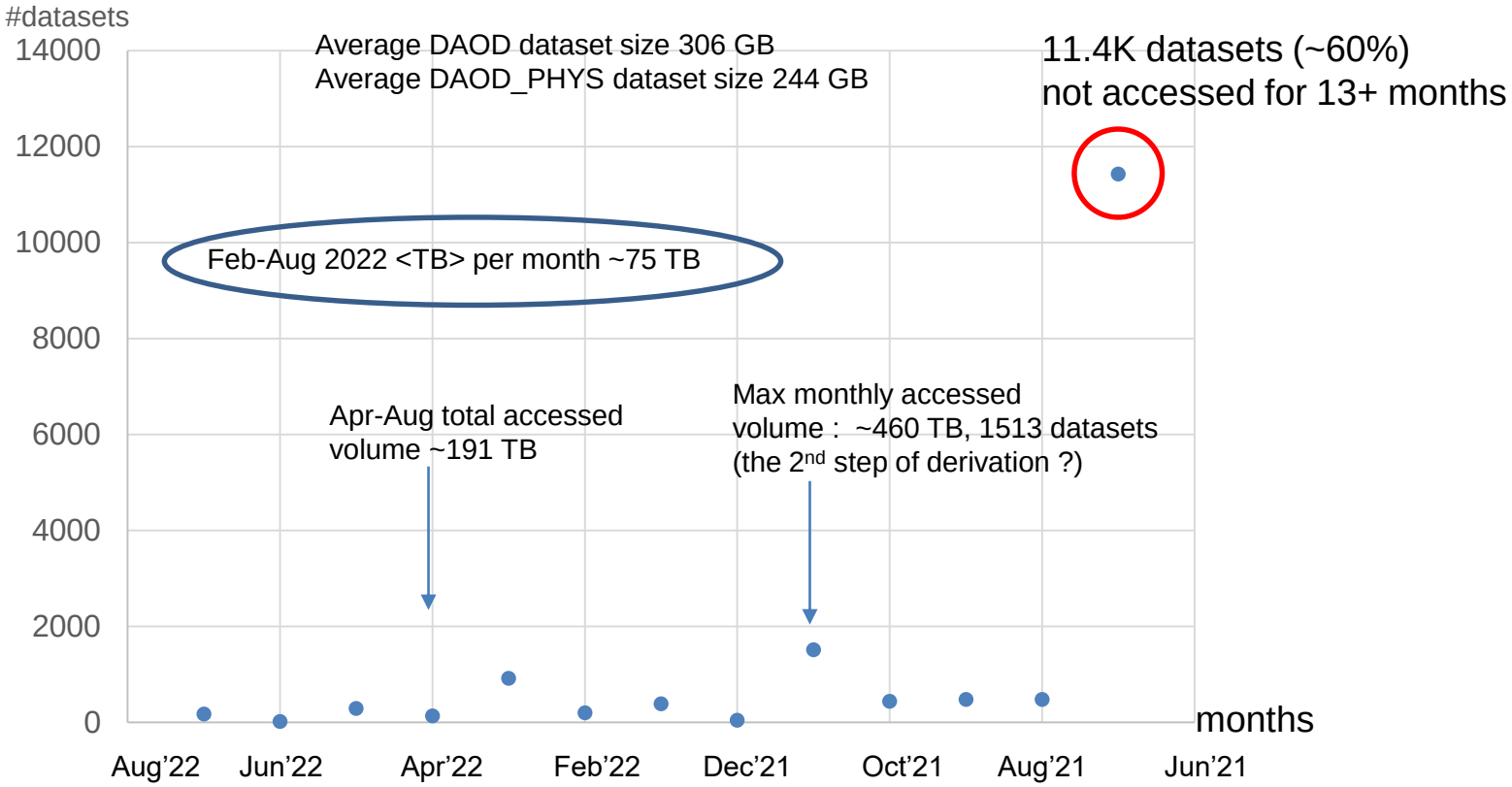
Disk size Aug 2020 – Aug 2022
Primary datasets volume vs Production step

DAOD disk volume accessed in the last 12 months



Volume → Logical Volume

Jun 2021 lifetime model campaign. DAOD datasets from the lifetime exception list
Total number of datasets 24.6K, known to Rucio (Aug'22) : 19.1K, total <logical> volume 2.6 PB



DAOD datasets handling

- DAOD datasets in exception requests (or "unpopular" datasets in the future)
  - *we can estimate annual data volume (~8-10 PB)*
  - *we know access frequency*
  - *we do not need to invent a sophisticated algorithms to move them to tape*

…for this data sample

1. we can copy (archive) datasets to tapes and to release disk space for secondary copies of popular data – *demoDC will be presented at DDM meeting (XZ)*

   …at the same time, we have a possibility to work on data grouping scenarios together with Tier-1s using the above datasets to debug our future smart writing algorithms

2. Delete DAOD datasets from the lifetime model exception list and reproduce them on demand if needed – *demoODD, today's discussion*

On demand data demo I

Step 1. PanDA demonstrator

- –Data sample : 2-4 DAOD datasets from September’22 lifetime model list
- –Check that corresponding AOD datasets are available (on tape at different T1s and there are no AOD disk replicas)
- –Delete DAOD datasets from disks
- –Submit User tasks

Metrics :

- – Total time to execute ODD DAOD task vs DAOD disk resident task
- –Time to stage-in DAOD datasets (for different T1s) - demoDC
  - •This will be a good comparison. It also means we need to write these DAOD sample to tape, and then rerun the user analysis tasks with staging DAODs from tape directly.
- –“Derivation step” time

On demand data demo II

Step 2. repeat step1 for ~150TB sample (sample size is TBD)

Metrics :

- Load on Tape systems
  - a) How much extra tape bandwidth will this workflow introduce? May be we should add it to the overall Run3/Run4 delivered tape throughput estimates;
  - b) How efficient ADC/Tier-1s can stage the AODs out of tape ? If the efficiency is 50%, that means site will buy more tape drives to meet the extra bandwidth requirements that's calculated from (a)
    - Note that, the overall efficiency of staging AOD from tape can be estimated by looking into the general derivation campaigns (plus info from site). But if this particular workflow has a timing requirement, which is very likely, we will have to treat its staging efficiency differently in the estimate.
- Time to reproduce DAOD datasets

Step 3. PanDA and Rucio integration (see ggdoc discussion)

Evaluation for other data types or/and unpopular DAOD datasets

Summary

- This discussion was triggered by the US ATLAS technical meeting in August (although the idea has been circulating for a long time)
- Positive initial feedback from PanDA and Data Carousel
- Ongoing discussion with Rucio et al
- This demo should be considered in a global context
- *Many implementation details are still TBD after the completion of step 1*