DMP Implementation Status of IPP

2025-02-19

IMAS versions updated on the TOK and Citrix systems

- No change

DD	AL	TOK cluster	Citrix system
3.38.1	4.11.4		2023-06-08
3.39.0	4.11.5	2023-07-26	2023-07-31
3.39.0	5.0.0	2023-10-10	
3.41.0	4.11.10	2024-04-17	
3.41.0	2024.07-foss-2023b		2024-08-11
3.42.0	2024.08-foss-2023b		2024-08-13
3.42.0	2024.08.1-foss-2023b	2024-08-23	2024-08-22
3.42.0	2024.08.1-intel-2023b	2024-08-23	2024-08-22
3.42.0	4.11.10		2024-10-09
3.42.0	2024.09-foss-2023b	2024-10-21	2024-10-21
3.42.0	2024.09-intel-2023b	2024-10-21	2024-10-21
4.0.0	2024.12-foss-2023b	2025-01-22	2025-01-22
4.0.0	2024.12-intel-2023b	2025-01-22	2025-01-22

Questions from Michal

- 1. Keycloak access and tests related to retrieving and decomposing Keycloak tokens
    - - did anyone apart from Cenk had time to do the tests?
        - https://confluence.eufus.psnc.pl/pages/viewpage.action?pageId=66256925
    - There is a whole discussion on Mattermost here:
        - https://chat.eufus.psnc.pl/dmp/pl/hphamm5hpfdo7xu5aetdwezziy
- Just done it ...
    - When can we expect this to work against the EUROfusion active directory?
        - Or can we configure it against the IPP IDP? Not that I know the details of this ...

Questions from Michal

- 2. Validator tests
  - - did anyone had a chance to give it a try to Validator?
  - There is a brief documentation here:
    - https://confluence.eufus.psnc.pl/display/DMP/IDS+Validator
  - and repository with a code and detailed description here:
    - https://gitlab.eufus.psnc.pl/dmp/importer
- Did this – validator worked but I wasn't able to get X working
  - Some years ago the CPT had a docker image that you could connect to via VNC
    - Could that be resurrected?

Questions from Michal

- 3. Future work
  - There are still some open points and open activities that we can tackle this year. However, we need to make priorities.
    - What should we start with in a first place:
      - Ability to upload data from UI (e.g. uploading / adding URI directly from UI level)
      - 2D / 3D arrays support (this is a heavy topic, doable, but will consume quite some resources)
      - HDF5 based data export (e.g. retrieving whole experiment and exposing it as a ZIP package with HDF5 files - quite time consuming, doable)
      - CSV export as HDF5 data (CSV like export but in HDF5 format - we need clarification here)
  - We need to know which of the listed topics brings most value to the product and which one should we tackle first.
- Any other ideas?

Questions from Michal

- Ability to upload data from UI (e.g. uploading / adding URI directly from UI level)
  - I think this is of strong interest and should be easy to implement (user must have the upload allowed flag enabled)
- 2D / 3D arrays support (this is a heavy topic, doable, but will consume quite some resources)
  - I think some brainstorming here might be useful.
    - This really only affects scenario B data
      - And only if we use the Catalog as a proxy for getting the data
        - Have we agreed on this?
        - What functionality do we want to expose?
          - Downloading data?
          - Plotting data?
- HDF5 based data export (e.g. retrieving whole experiment and exposing it as a ZIP package with HDF5 files - quite time consuming, doable)
  - Again, some brainstorming might be useful.
    - Would it make more sense to target netCDF from imaspy?
- CSV export as HDF5 data (CSV like export but in HDF5 format - we need clarification here)
  - I don't know what is meant here. The typical CSV data but exported in HDF5? If we build a pandas data frame with the data, both csv and hdf5 are supported output data options. This might be easier if we switch from java to python for the back-end.

Questions from Michal

- We need to know which of the listed topics brings most value to the product and which one should we tackle first.
  - My suggestions (highest priority first)
    - Adding URIs to other (non-summary IDSes)
      - Will not be ingested but added to info on summary page
    - Triggering data upload
    - CSV output as HDF5
    - netCDF based data export of IDSes
    - Plotting of profiles as a function of time
      - Leverage planned ITER developments?
- Any other ideas?
  - Would it make sense to brainstorm the preparation of data for ML?

Data access times
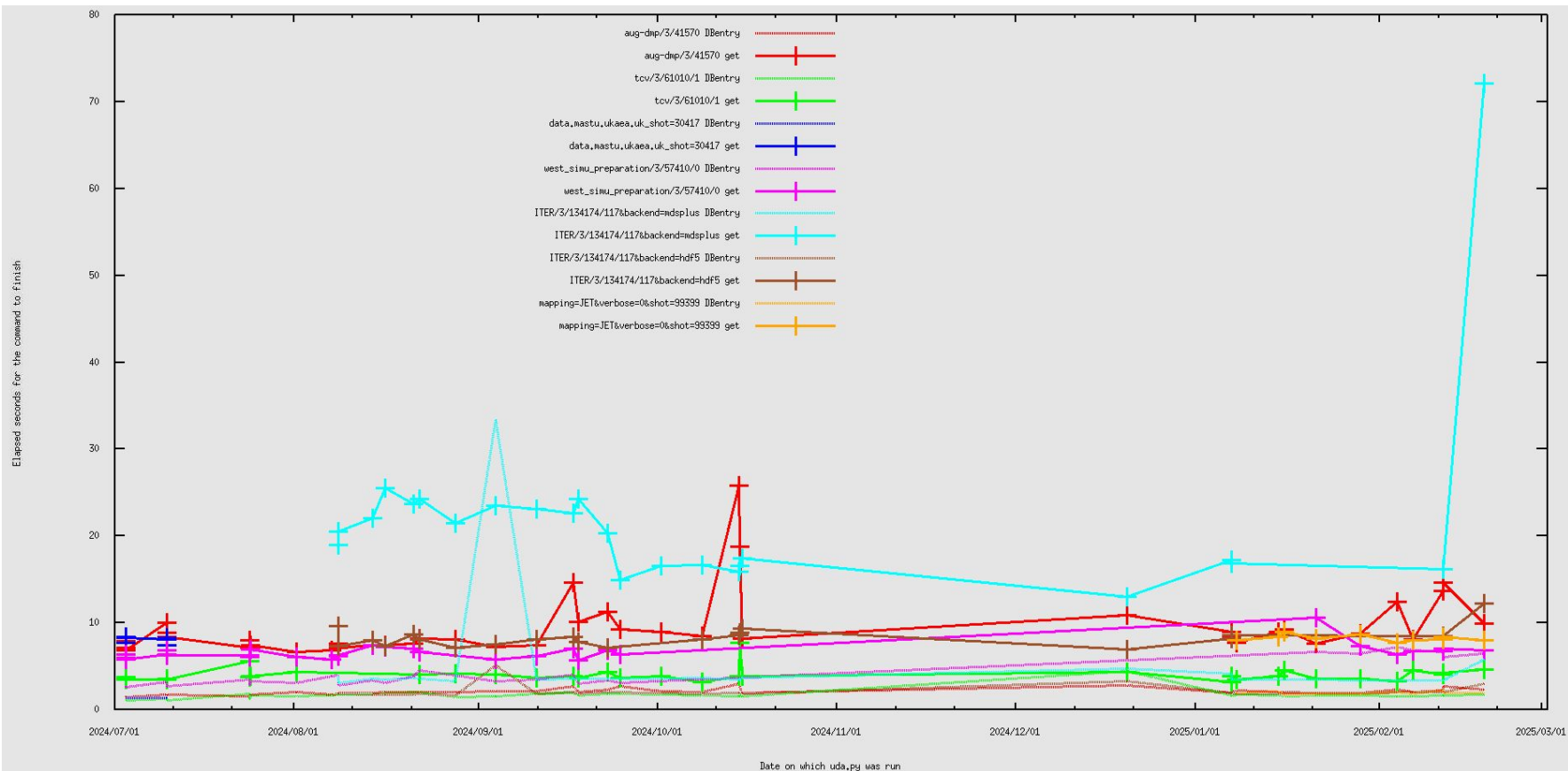
Data access times (log scale)
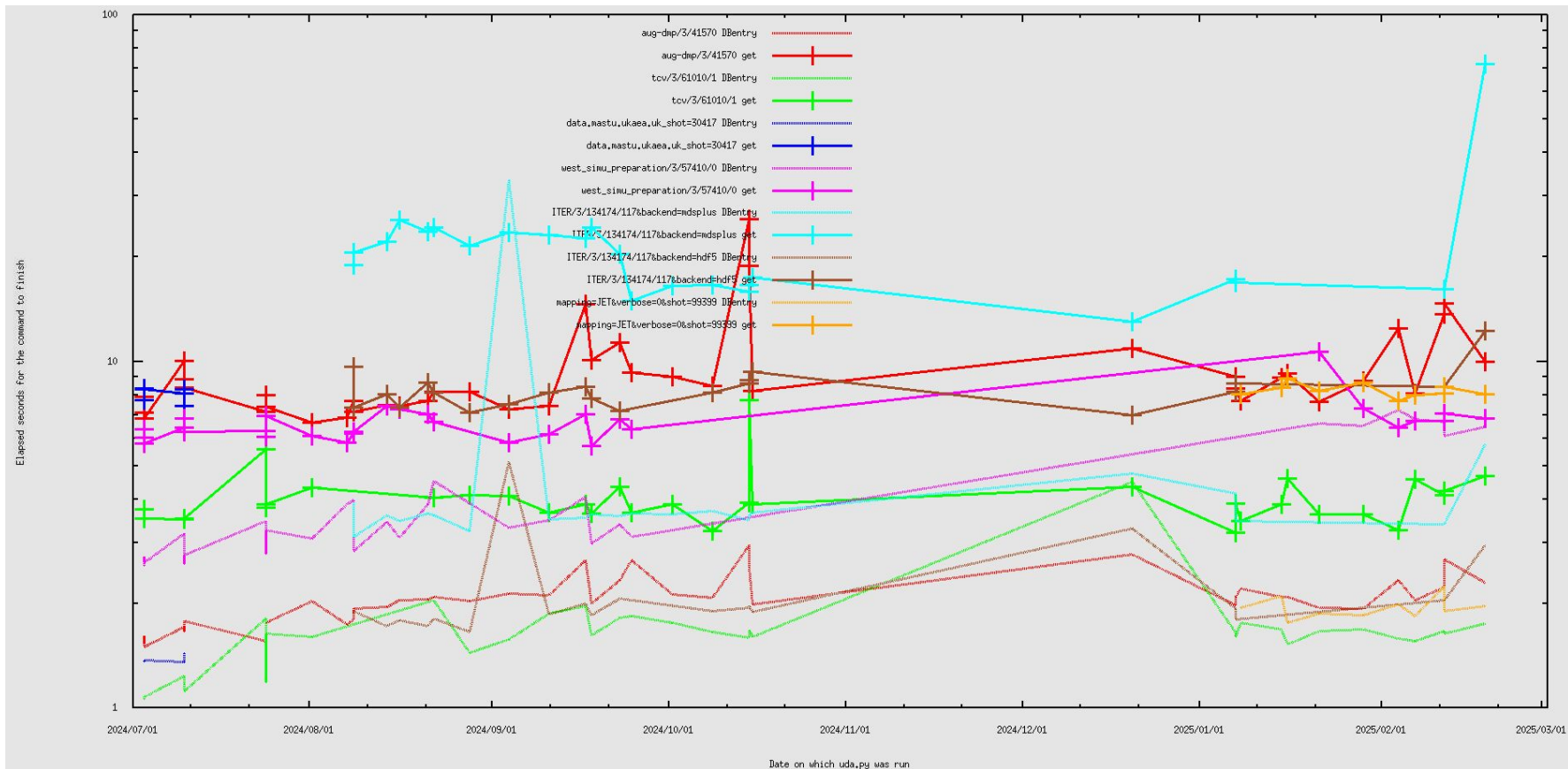
Possible issue with uda-keycloak-test

- Anybody know why I see this?
  - dpc@srv-uda-1 ~/GIT/uda-keycloak-test (git)-[uda-keycloak] % git status
  - On branch uda-keycloak
  - Your branch and 'origin/uda-keycloak' have diverged,
  - and have **23** and **66** different commits each, respectively.
  - (use "git pull" to merge the remote branch into yours)
  -
  - nothing to commit, working tree clean
- Am I going to see a large number of problems if I do a "git pull"?