DMP Implementation Status of IPP

2025-08-27

IMAS versions updated on the TOK and Citrix systems

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

- Entries in red have a problem after the cluster upgrade
  - Underlying python changed, together with numpy
    - Remove?

Worked out why horizon worked and the TOK cluster didn’t

- On the TOK cluster, I had done
  - module load IMAS-AL-Python/5.3.1-intel-2023b-DD-3.42.0
- since I didn’t need the rest of IMAS …
  - But this resulted in “UDA/2.7.5-intel-compilers-2023.2.1” being loaded
    - Which turned out to be problematic
- If instead, I loaded “IMAS/3.42.0-2024.09-intel-2023b”, this gave me “UDA/2.8.0-intel-compilers-2023.2.1”
  - which was able to access ITER without a problem (finally!)
- On horizon I had loaded “IMAS/3.42.0-2024.09-intel-2023b”

Worked out why horizon worked and the TOK cluster didn’t

- IMAS-AL-Python/5.3.1-intel-2023b-DD-3.42.0
  - IMAS-AL-Core/5.3.2-intel-2023b
    - UDA/2.7.5-intel-compilers-2023.2.1

- IMAS/3.42.0-2024.09-intel-2023b
  - IMAS-AL-Cpp/5.3.0-intel-2023b-DD-3.42.0
    - IMAS-AL-Core/5-intel-2023b
    - IMAS-AL-Core/5.4.3-foss-2023b
      - UDA/2.8.0-intel-compilers-2023.2.1

Continued going through notes from the UDA training

- Started looking at building “JSON_mappings/draft/summary/mappings.json”

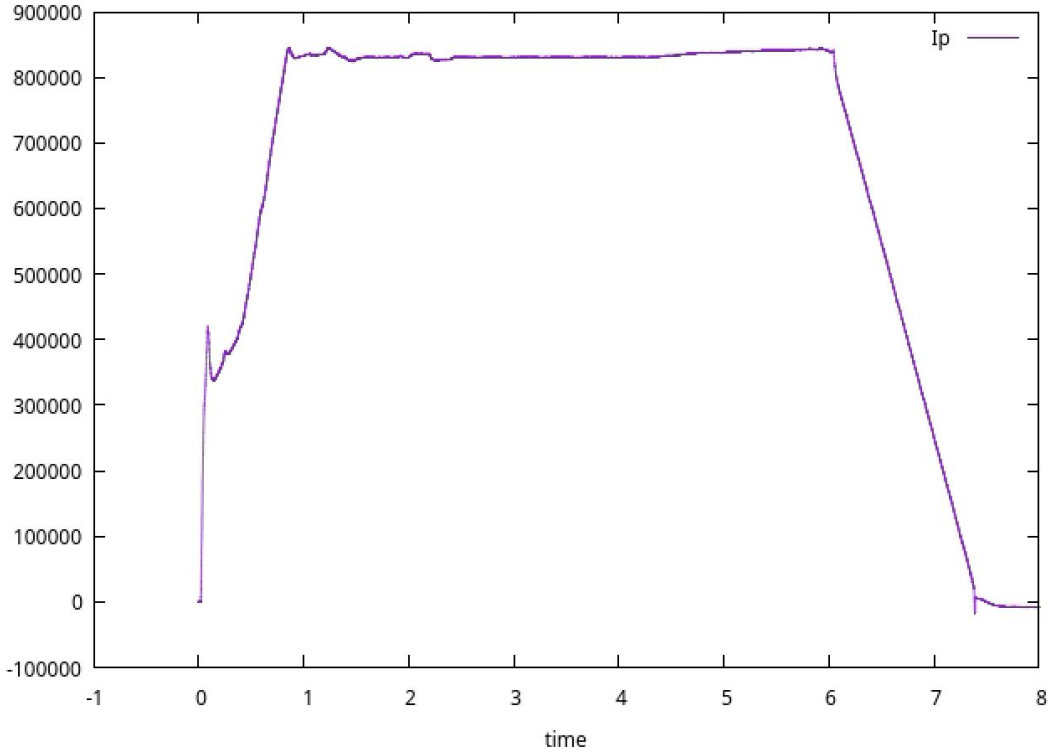
```
{
    "ids_properties/homogeneous_time": {
        "MAP_TYPE": "VALUE",
        "VALUE": 0
    },
    "global_quantities/r0/value": {
        "MAP_TYPE": "VALUE",
        "VALUE": 1.65
    },
    "global_quantities/r0/source": {
        "MAP_TYPE": "VALUE",
        "VALUE": "Official AUG value"
    },
    "time": {
        "MAP_TYPE": "VALUE",
        "VALUE": [0.0, 1.0, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0]
    }
}
```

Have the uda_cli producing the expected results

- uda_cli --host localhost --port 56565 --request "IMAS_JSON_MAP::read(path=summary/ids_properties/homogeneous_time, mapping=draft, datatype=int, rank=0)" --source ""
  - request: IMAS_JSON_MAP::read(path=summary/ids_properties/homogeneous_time, mapping=draft, datatype=int, rank=0)
  - 0
- uda_cli --host localhost --port 56565 --request "IMAS_JSON_MAP::read(path=summary/global_quantities/r0/value, mapping=draft, datatype=float, rank=0)" --source ""
  - request: IMAS_JSON_MAP::read(path=summary/global_quantities/r0/value, mapping=draft, datatype=float, rank=0)
  - 1.65
- uda_cli --host localhost --port 56565 --request "IMAS_JSON_MAP::read(path=summary/global_quantities/r0/source, mapping=draft, datatype=str, rank=0)" --source ""
  - request: IMAS_JSON_MAP::read(path=summary/global_quantities/r0/source, mapping=draft, datatype=str, rank=0)
  - Official AUG value
- uda_cli --host localhost --port 56565 --request "IMAS_JSON_MAP::read(path=summary/time, mapping=draft, datatype=float, rank=1)" --source ""
  - request: IMAS_JSON_MAP::read(path=summary/time, mapping=draft, datatype=float, rank=1)
  - [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, ... (11 elements)]

And have dug up a C example for accessing AUG data

- SUMMARY IDS might be a bit of a challenge since we need to devise an operator to perform the appropriate mean and standard deviation calculation to move the data onto a standard time-base (100ms resolution) for all signals
  - Of course the example was out-of-date with regard to IPP computer systems!



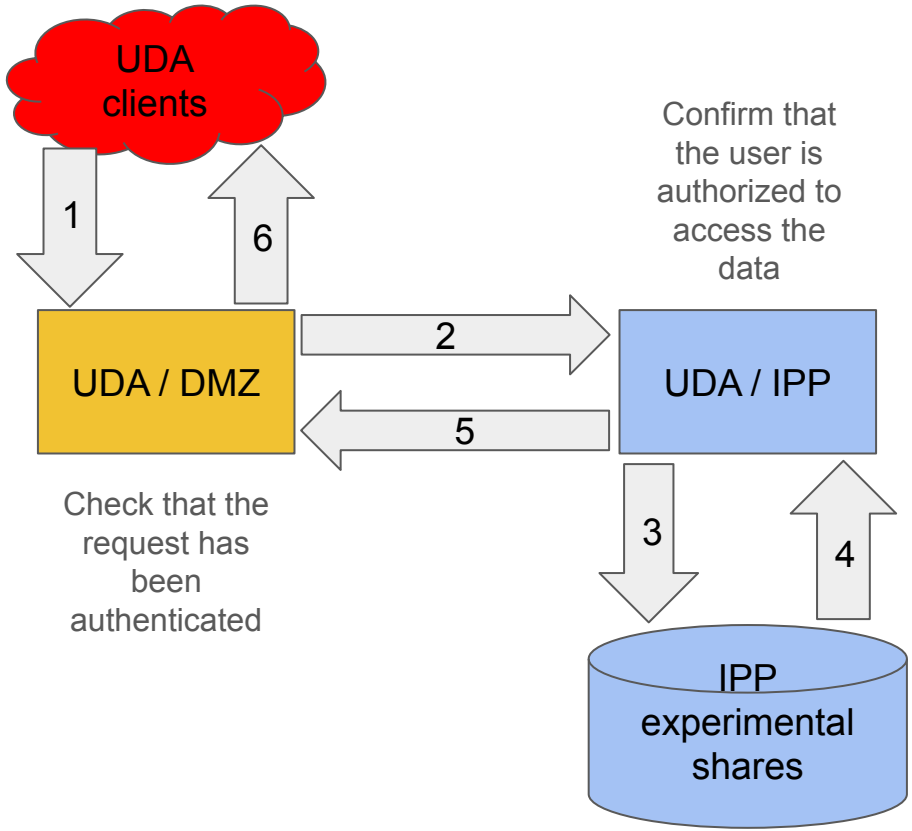
Not sure how to move from uda_cli to UDA/get

- Still digging through documentation to make the move from uda_cli access working on particular fields
  - To being able to being able to do an IMAS get via UDA to get the entire SUMMARY IDS
    - ~/GIT/uda-test/uda.py -u 'imas://localhost:56565/uda?mapping=DRAFT&verbose=10&shot=30417&batch_size=20' -c summary
    - Accessing data from imas://localhost:56565/uda?mapping=DRAFT&verbose=10&shot=30417&batch_size=20
    - ERROR:root:b'al_begin_dataentry_action: [ALException = [read_data]: Error locating data plugin\n]'
    - DBEntry failed after 1.103 seconds
- So I’m missing a plugin in but which one?
- Too many repositories!

Status of data mappings: LINK

IDS name	Type	AUG	COMPASS	COMPASS-U	JET	MAST	MAST-U	TCV	WEST
amns_data	Data	N/A						N/A	
barometry	Experimental								
b_field_non_axisymmetric	Experimental								
bolometer	Experimental							gdat	
bremsstrahlung_visible	Experimental							N/A	
calorimetry	Experimental								
camera_ir	Experimental								
camera_visible	Experimental								
camera_x_rays	Experimental								
charge_exchange	Experimental								
coils_non_axisymmetric	Experimental								
controllers	Control								
core_instant_changes	Derived								
core_profiles	Derived	trview						gdat	
core_sources	Derived	trview						gdat	
core_transport	Derived							gdat	

Proposed topology



Just received: do we want to provide a community response

改进的邮件模板
Dear David Coster,

I am Xueyi Li, a Ph.D. Candidate at Wuhan University. I know you through your scientific dataset entitled "Experimental investigation and SOLPS-ITER modeling of Ne-seeded radiative divertor H-modes plasma on EAST" from ScienceDB. And I am writing to kindly solicit your participation in a survey on researchers' use of data search tools for discovering open data to find out your general views.

Data is a fundamental resource for research. The open data movement provides opportunities for sharing, obtaining, integrating, and reusing existing datasets. However, researchers face a new challenge: before data collection and reuse, how do we effectively find relevant datasets published on the Web? The open dataset is as useful as its discovery mechanisms and dataset search tools provide a promising solution to overcome this challenge. These tools allow researchers to search for the datasets relevant to their data needs and obtain up the description of datasets to evaluate datasets. It contributes to discovering the existence, characteristics, and accessibility of datasets. And I am carrying out a research project aiming to explore researchers' discovery of open data and the role of data search tools, especially generative AI tools in this process, which provides support for optimizing the practice of open data retrieval. This survey is a vital component of this project.

Your participation is greatly important to us and the questionnaire will take you around 5 minutes to complete. We guarantee that all responses will be treated as confidential and used only for this research project, in which the results of this study will be included in an article discussing researchers' dataset discovery. We will strictly follow the law including the International Covenant on Civil and Political Rights of the United Nations, General Data Protection Regulation and the Cybersecurity Law of the People's Republic of China and ethical research guidelines in ensuring the confidentiality of any private information that you have disclosed. All results will be analyzed in aggregate only and the answer will never be associated with individual participants. The e-mail contact details will be deleted from the survey results data file.

If you have used the Science Data Bank's Data Discovery Chat (beta) or any other data retrieval systems integrated generative AI technologies, please follow this link:

https://forms.office.com/r/fXkcLtGjt7

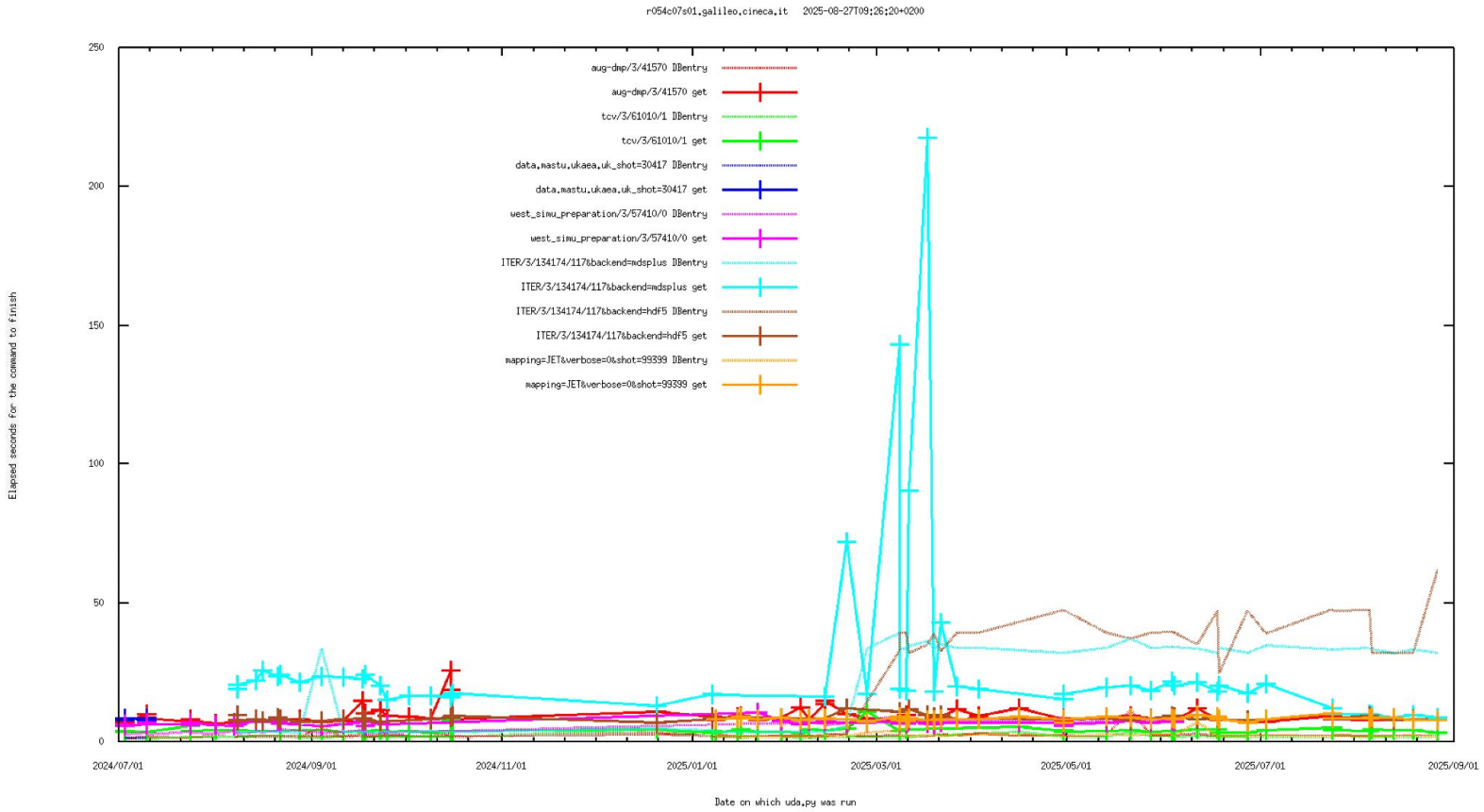
By submitting your responses are indicating your consent to participate in this survey. Your participation in this research is voluntary. You have the right to withdraw at any point during this survey, for any reason, and without any prejudice. If you have any questions about this survey, please feel free to contact me at lixueyi_sim@whu.edu.cn.

Thank you for your consideration and participation. It is very valuable to me.

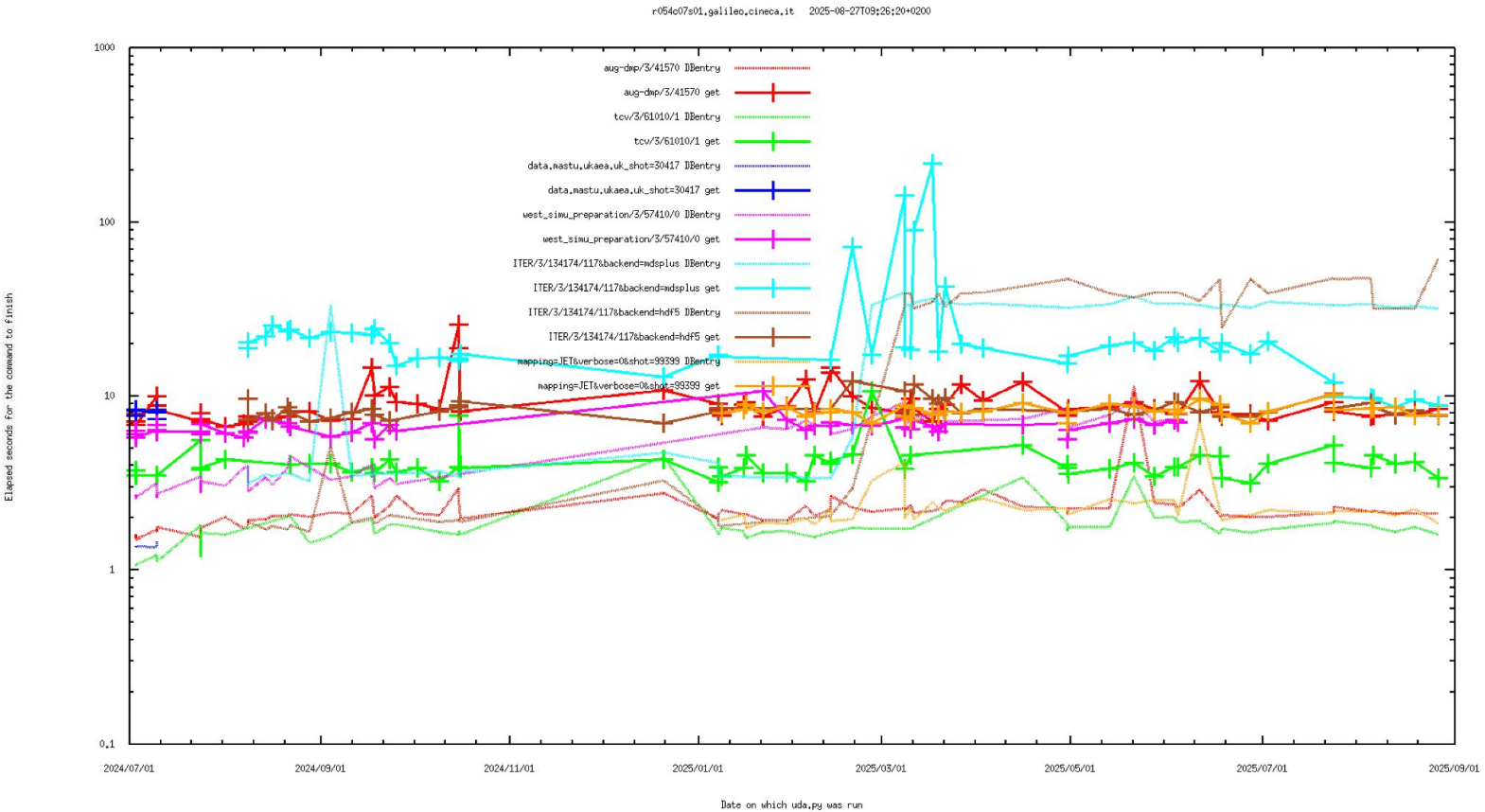
With regards,

Xueyi Li
Ph.D. Candidate
School of Information Management Wuhan University
Email: lixueyi_sim@whu.edu.cn

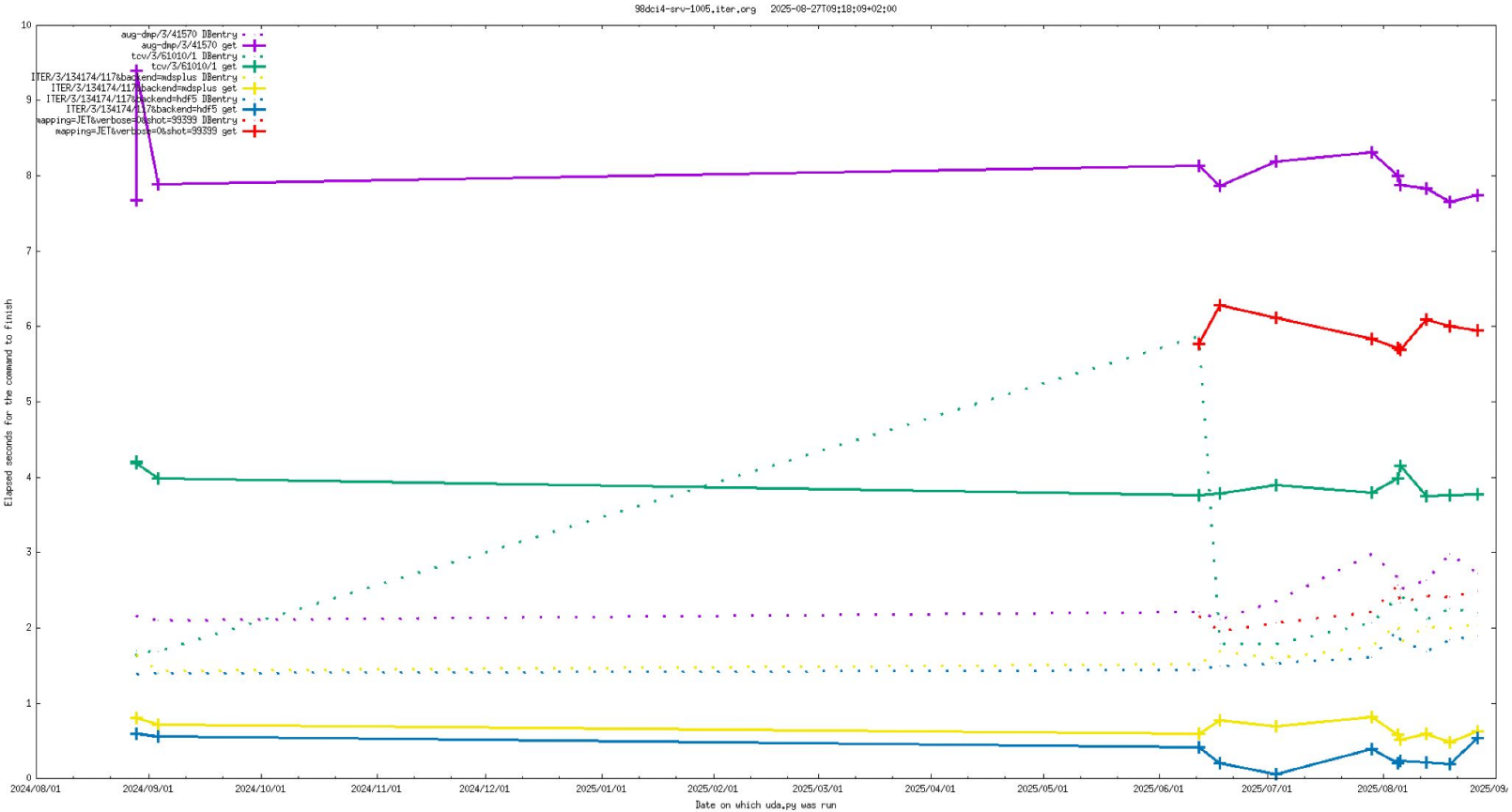
Data access times: run on Gateway



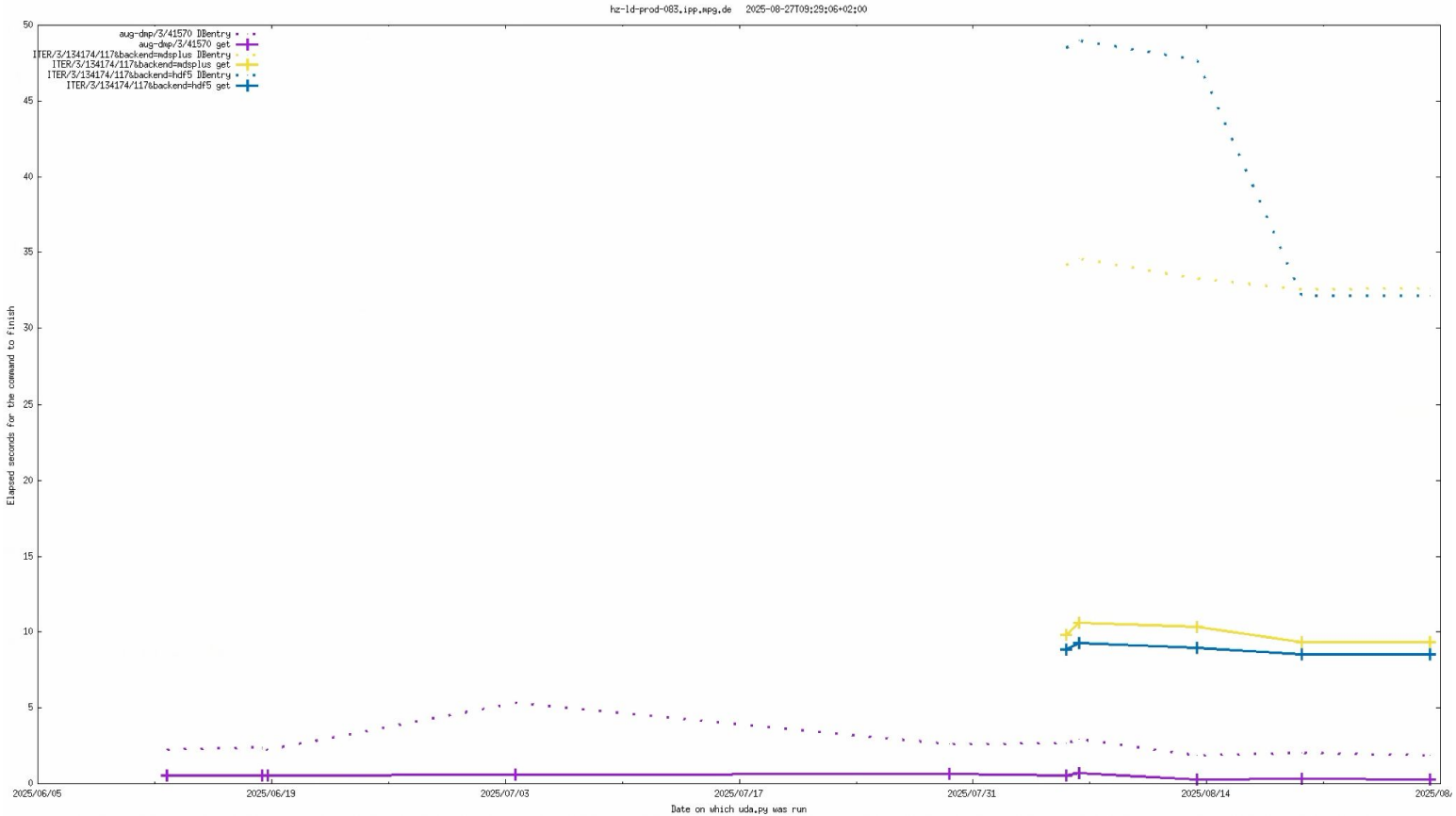
Data access times: run on Gateway



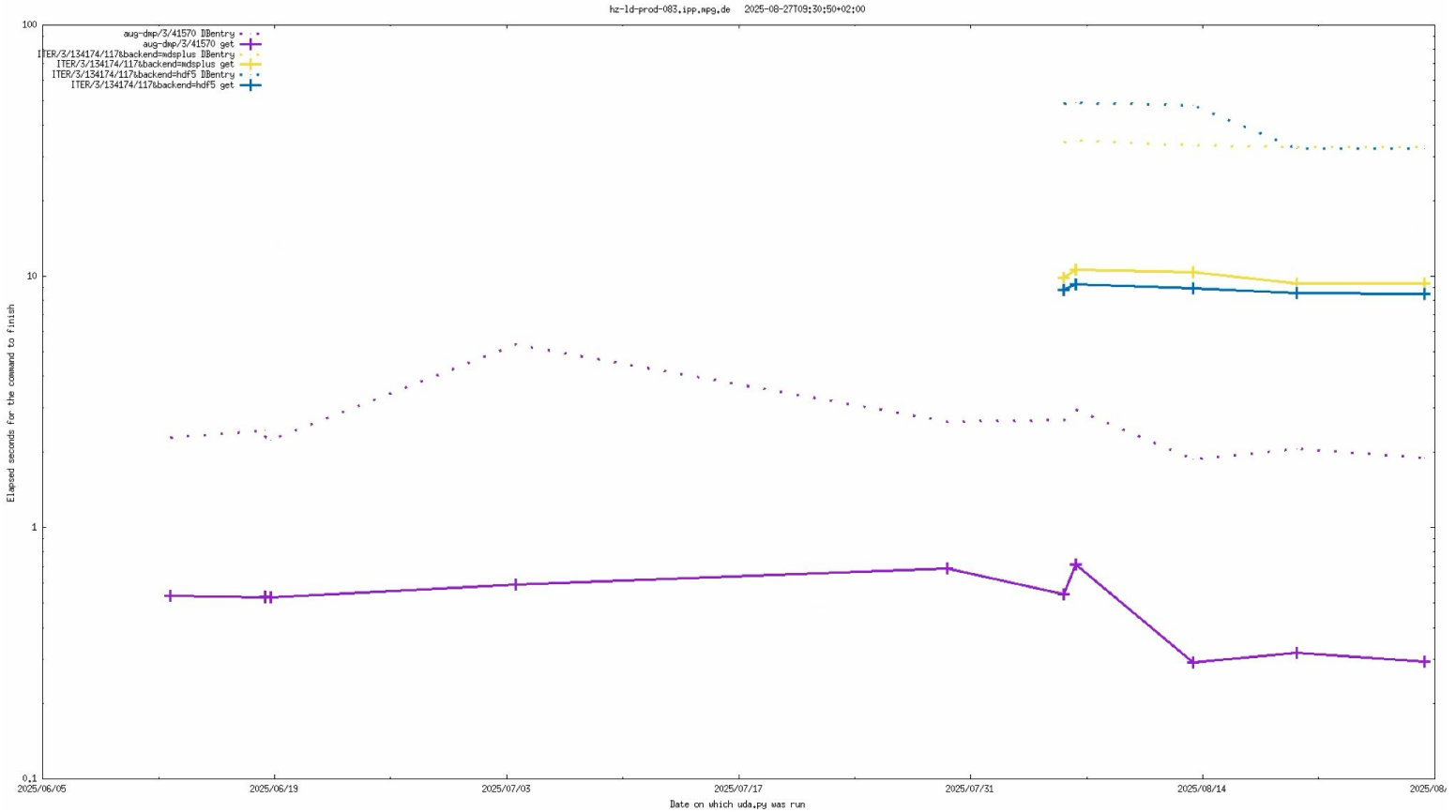
Data access times: run at ITER



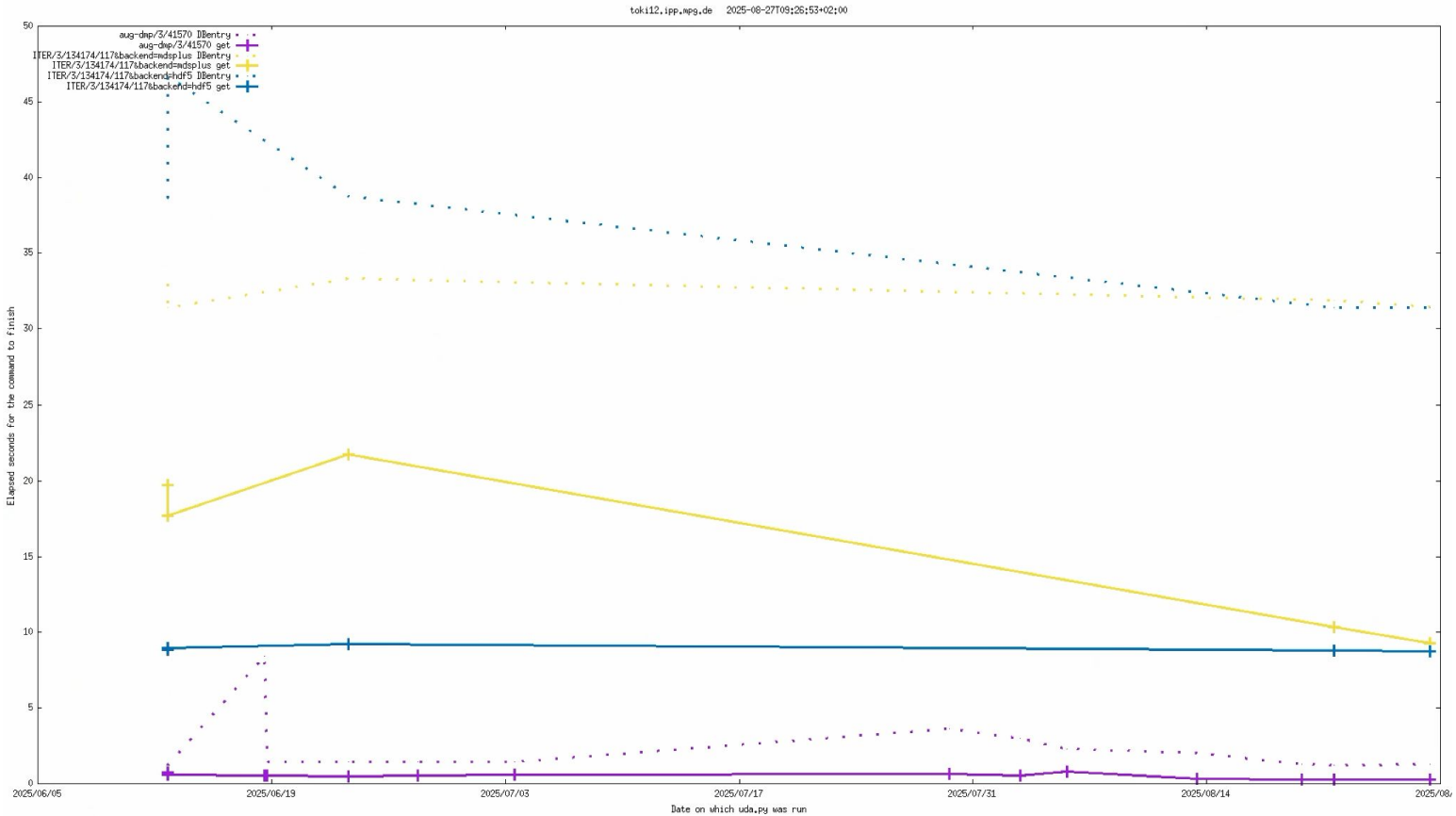
Data access times: run at IPP (horizon)



Data access times: run at IPP (horizon)



Data access times: run at IPP (TOK cluster)



Data access times: run at IPP (TOK cluster)