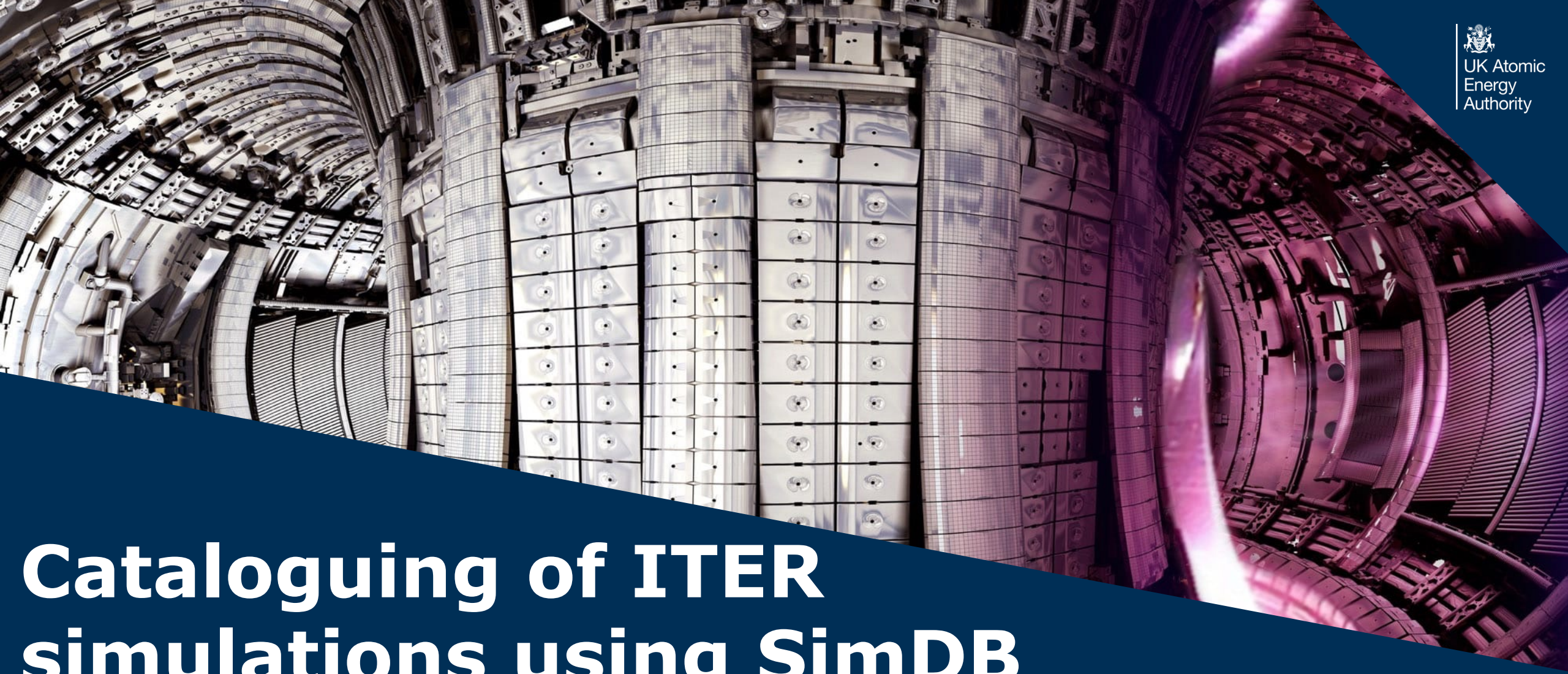


UK Atomic Energy Authority

Cataloguing of ITER simulations using SimDB

Jonathan Hollocombe (UKAEA), Simon Pinches (ITER), Olivier Hoenen (ITER), Andrew Lahiff (UKAEA) & Shaun de Witt (UKAEA)

4th IAEA Technical Meeting on Fusion Data Processing, Validation and Analysis 29/11/2021

iter

china eu india japan korea russia usa

Overview

UK Atomic Energy Authority

- Existing ITER simulation management
- Requirements for a new catalogue
- SimDB
  - Design
  - Features
  - Deployment
- SimDB dashboard
- Future developments

This work was done under the ITER contracts IO/16/RT/12202/JTR (JINTRAC) & IO/20/RFQ/20151/JLE (SimDB)

Existing ITER simulation management

UK Atomic Energy Authority

- ~2,000 ITER simulations
- Organised using IMAS shot/run file names
  - Shot number needs to be generated
- Additional metadata in simulation YAML files
  - Generated with scripts with manual input
- YAML watcher files

Requirements for a new catalogue

UK Atomic Energy Authority

- Make simulations listable and queryable
- Allow for validation of simulations & metadata
- Allow for acceptance and deprecation workflows
- To remove need to generate arbitrary shot numbers and organise simulations by shot/run
- Allow users to fetch found simulations using IMAS

- i.e. to make the simulations more FAIR

SimDB project

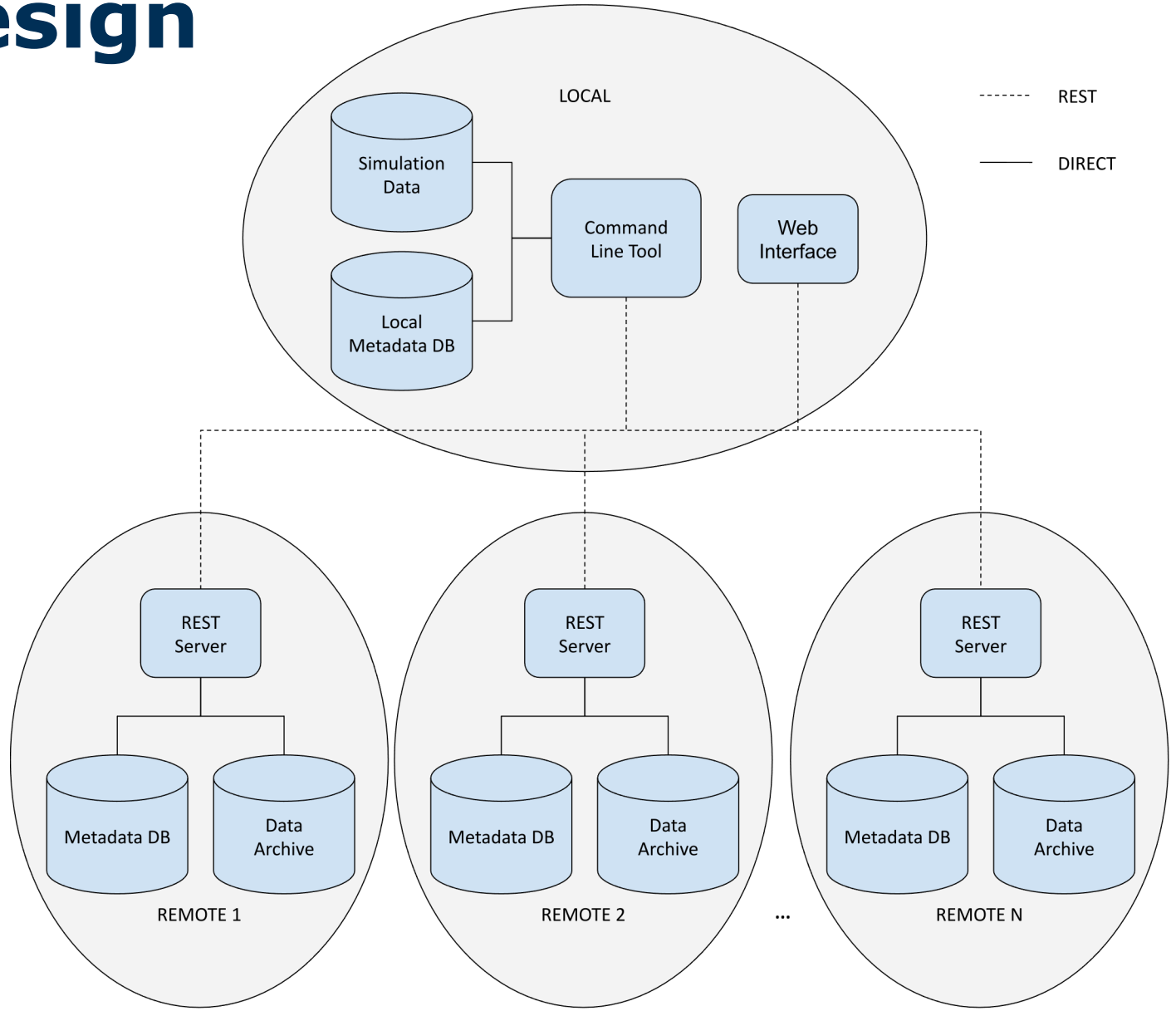
UK Atomic Energy Authority

- SimDB originally developed as part of ITER-JINTRAC project
  - Older version already used in managing STEP simulations

- New project to build on this:
  - Phase 1: Develop and expand the SimDB tool and store the existing ITER simulation catalogue in SimDB
  - Phase 2: Create a dashboard to expose the ITER catalogue
  - Phase 3: Improve the remote data capabilities of IMAS to allow for fetching of simulations found using he dashboard or command line tools

- SimDB is not ITER specific but current development is driven by ITER project

SimDB design



SimDB features

UK Atomic Energy Authority

- Python command line tool with local metadata database and REST based communication with remote server(s)
- Zero data copy on local ingest with data checked and pushed to remote server on ”push”
- Metadata validation using Python Cerberus library
  - Configurable for each server
- Acceptance workflow with ability to update acceptance metadata
  - Can follow history of simulations
- Support of ingestion of IMAS data into metadata database for querying
  - Summary & dataset_description ingested for ITER

SimDB features – example metadata validation rules

```
device:
  required: true
  type: string
workflow:
  required: true
  type: dict
  schema:
    name:
      required: true
      type: string
      allowed:
        - CORSICA
        - something
description:
  required: true
  type: string
```

Validation can be as simple such as “required” …

```
equilibrium:
  required: true
  type: dict
  schema:
    time_slice:
      type: list
      schema:
        global_quantities:
          type: dict
          schema:
            ip:
              type: numpy
              coerce: numpy
              gt: -17000000
              le: 0
        profiles_2d:
          type: list
          schema:
            phi:
              type: float
            psi:
              type: float
            r:
              type: float
              gt: 0
            z:
              type: float
```

… or more complicated with array limits.

SimDB CLI

```
$ simdb --help
Usage: simdb [OPTIONS] COMMAND [ARGS]...

Options:
  --version                       Show the version and exit.
  -d, --debug                     Run in debug mode.
  -v, --verbose
  -c, --config-file FILE
  --help

Commands:
  alias       Query remo
  config      Query/upda
  database    Manage loc
  manifest    Create/che
  provenance  Create the
system.
  remote      Interact w
  sim         Alias for
  simulation  Manage ing
```

```
$ simdb remote --help
Usage: simdb remote [OPTIONS] [NAME] COMMAND [ARGS]...

  Interact with the remote SimDB service.

  If NAME is provided this determines which remote server to
communicate with,
  otherwise the server in the config file with default=True is
used.

Options:
  --username TEXT                 Username used to authenticate
with the
                                  remote.
  --password TEXT                 Password used to authenticate
with the
                                  remote.
  Remote options: [mutually_exclusive]
                                  Commands for managing remotes
    --set-default NAME            Set the remote as the default.
    --get-default                 Print the currently set
default remote.
```

SimDB CLI self documented with nested help

SimDB CLI

UK Atomic Energy Authority

Listing/querying simulations with additional metadata fields

```
$ simdb remote list -m workflow.name
UUID                                 alias     workflow.name
------------------------------------------------------------
29ca0344-3bf5-11ec-b660-000003b9fe80 100505/2  CORSICA
2b0ae62e-3bf5-11ec-b648-000003b9fe80 100506/1  CORSICA
2c604370-3bf5-11ec-9e13-000003b9fe80 100506/2  CORSICA
2ee22a78-3bf5-11ec-820b-000003b9fe80 100507/2  CORSICA
315c8ef6-3bf5-11ec-b781-000003b9fe80 101000/10 ASTRA
328671f2-3bf5-11ec-a2ba-000003b9fe80 101000/20 ASTRA
33b81cc4-3bf5-11ec-b777-000003b9fe80 101000/30 ASTRA
34ed4a7e-3bf5-11ec-8eea-000003b9fe80 101000/40 ASTRA
3634e59a-3bf5-11ec-b905-000003b9fe80 101001/0  ASTRA
3777f30c-3bf5-11ec-a9f7-000003b9fe80 101001/10 ASTRA
38a6098a-3bf5-11ec-b5df-000003b9fe80 101001/20 ASTRA
3b08539a-3bf5-11ec-9ac5-000003b9fe80 101001/40 ASTRA
3c3bc2c4-3bf5-11ec-8a83-000003b9fe80 101002/0  ASTRA
3d79190c-3bf5-11ec-b108-000003b9fe80 101002/10 ASTRA
3ea3c624-3bf5-11ec-b5db-000003b9fe80 101002/20 ASTRA
3fd337f0-3bf5-11ec-8857-000003b9fe80 101002/30 ASTRA
4106c34e-3bf5-11ec-869d-000003b9fe80 101002/40 ASTRA
423b9f50-3bf5-11ec-bd2a-000003b9fe80 101003/0  ASTRA
43730b4c-3bf5-11ec-9e4f-000003b9fe80 101003/10 ASTRA
44a26e0e-3bf5-11ec-a944-000003b9fe80 101003
```

Printing simulation metadata

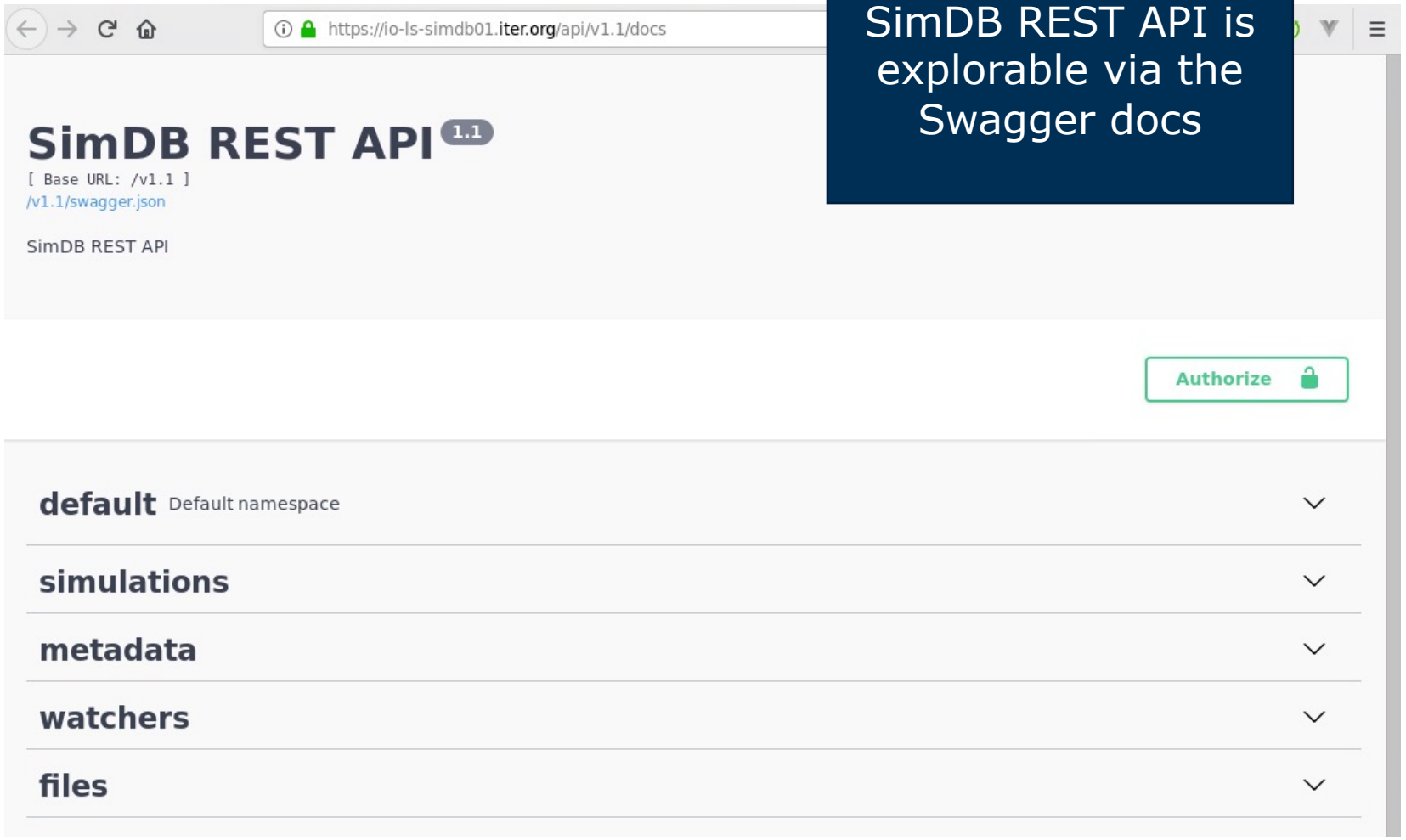
```
$ simdb remote info 100505/2
uuid:      29ca0344-3bf5-11ec-b660-000003b9fe80
alias:     100505/2
metadata:
  replaced_by: 856b3186-3bfb-11ec-93c8-000003b9fe80
  deprecated_on: 2021-11-02T23:10:42.667108
  summary.fusion.power.value: [0. 0. 0. ... 0. 0. 0.]
  summary.ids_properties.version_put.access_layer: 4.8.6
  hcd.p_hcd: 46.8
  dataset_description.simulation.time_begin: 2.0
  summary.global_quantities.li.source: CORSICA
  summary.local.magnetic_axis.q.value: [39.408     6.29465
4.89269  ...  0.965456  0.967934  0.947003]
  shot: 100505
  summary.global_quantities.current_bootstrap.source: CORSICA
  mary.time: [  2.        5.77416   6.7837  ... 301.973
  4   314.366  ]
  mary.global_quantities.fusion_gain.value: [0. 0. 0. ... 0.

  mary.global_quantities.beta_tor_norm.source: CORSICA
  kflow.type: predictive
```

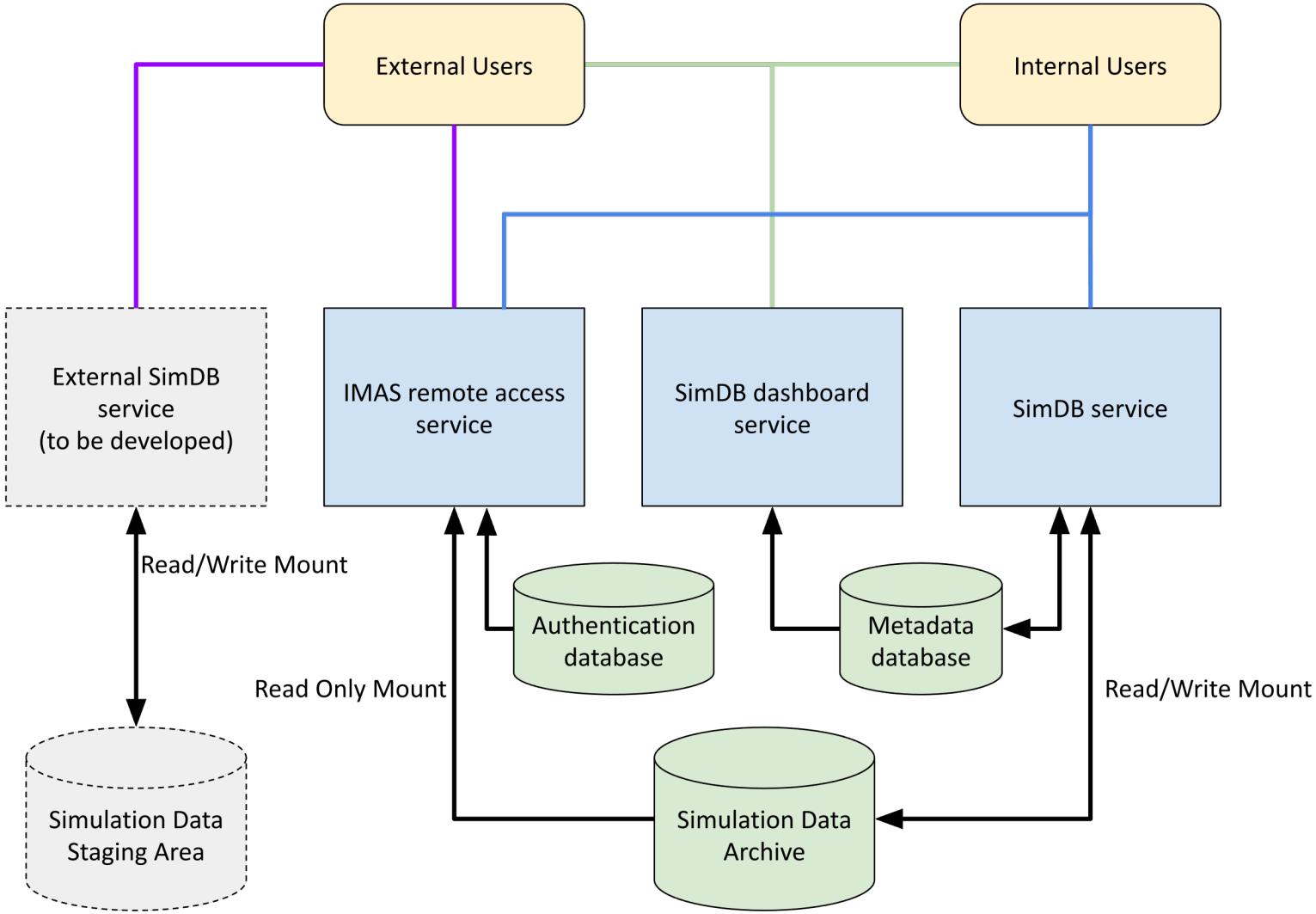
SimDB API

UK Atomic Energy Authority

SimDB REST API is explorable via the Swagger docs



SimDB deployment at ITER



SimDB deployment

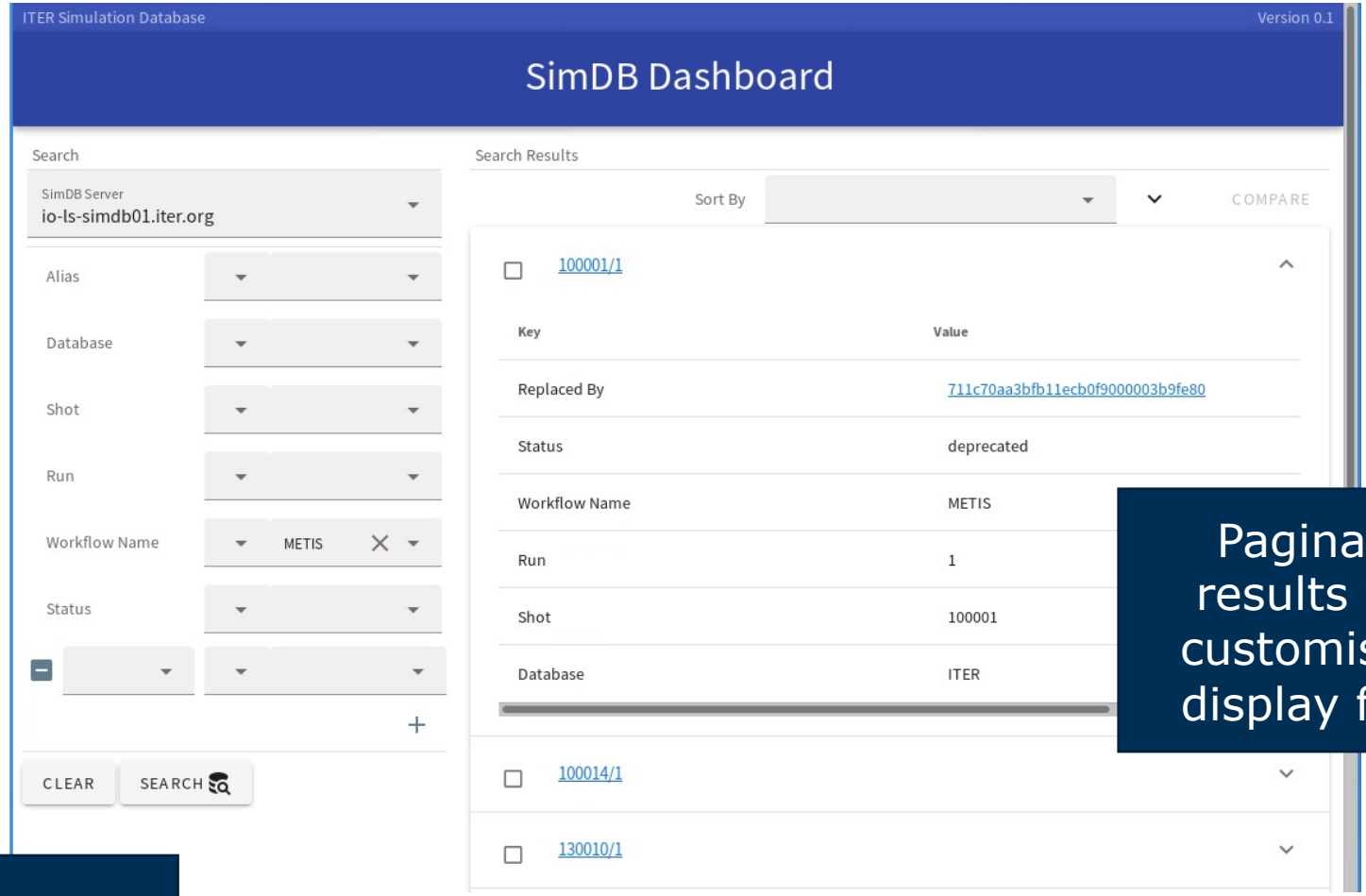
UK Atomic Energy Authority

Site specific configuration:

- One or more SimDB servers
- Database connection
- Location of simulation data archive
- Validation rules
  - Auto validate option
  - Error on fail option

- Dashboard configuration (default data to view etc.)

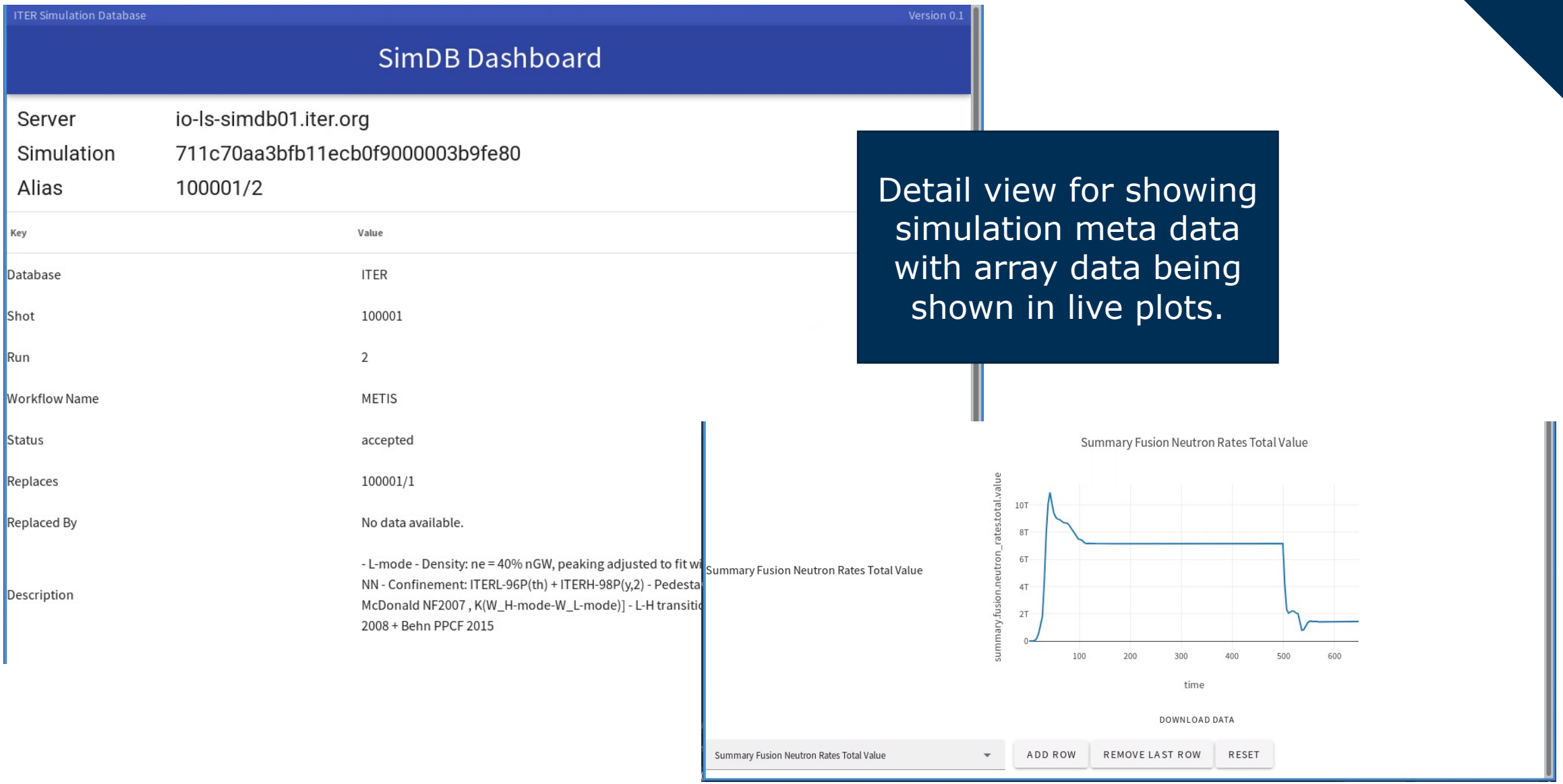
SimDB dashboard



Paginated results with customisable display fields

Query Functionality

SimDB dashboard



Detail view for showing simulation meta data with array data being shown in live plots.

SimDB dashboard

UK Atomic Energy Authority

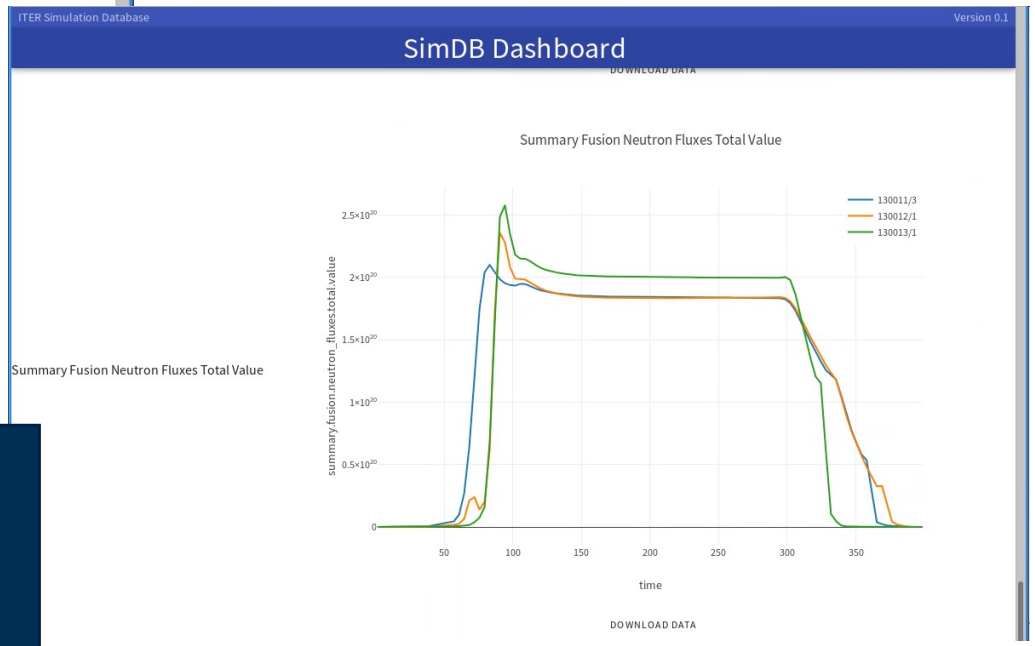
Comparison view showing tabulated view for scalar data

ITER Simulation Database

SimDB Dashboard

Server io-ls-simdb01.iter.org

Key	59c344ee3bf911ecabd7000003b9fe80	5b3ff9d43bf911ecb4ce000003b9fe80	5cc2a9003bf911ecbf94000003b9fe80
Database	ITER	ITER	ITER
Shot	130011	130012	130013
Run	3	1	1
Workflow Name	METIS	METIS	METIS
Status	deprecated	deprecated	deprecated
Replaces	130011/1	130011/1	No data available.



& comparison plots for array data

Future developments

UK Atomic Energy Authority

- Complete the development on the SimDB dashboard
- Phase 3: Development of IMAS remote data access
  - URI based file descriptors
  - Integrated with SimDB
  - Improving performance

- Code is available at:
  - https://git.iter.org/projects/IMEX/repos/simdb/browse
  - https://git.iter.org/projects/IMEX/repos/simdb-dashboard/browse

- Email: jonathan.hollocombe@ukaea.uk