



EUROfusion

Computing Hub: PSNC Core Software Integration ACH

Presenter: Marcin Płóciennik

Poznan Supercomputing and Networking Center

- PSNC is an **ICT research and development centre** affiliated to the Institute of Bioorganic Chemistry of the Polish Academy of Sciences
- One of the leading **operators of Polish e-Infrastructure**
 - National Research and Education Network – PIONIER (Polish Optical Internet)
 - Research Metropolitan Area Network - POZMAN
 - HPC Center
 - Data Center
 - Data repositories and Digital Libraries Federation
- **370+** employees, **15+** technical departments, **20** R&D laboratories
- Participates in many **European and national research projects (250+)**
 - coordination of 35+ projects
 - participation in 70+ projects in **Horizon2020**
 - Participation in 15+ fusion related projects (since 2005)



Poznan Supercomputing and Networking Center

- **Leading operator of Polish e-Infrastructure**
- HPC, Cloud-computing, Data Centre, NREN
- Currently PSNC runs the largest supercomputer in Poland.
- Eagle II is located in 85 position on Top500 Supercomputers list.
 - 7 Pflops; 90,000 cores
 - under upgrade to 15 Pflops
 - 0,5 Pflops GPU partition (to be upgraded)
- 2 Data Centres
- Cloud operator
- 80 PB of different class storage: tape to NVME under upgrade
- PSNC is an operator of national network for science (NREN)
 - Over 2.5 mln users, over 300 institutes



Location: Poznan, Poland

PSNC is located in Poznan, Poland, very well connected with major European cities.



Institute provides very good working conditions for hosting workshops:

- Several conference rooms, **co-working building**, equipped with Videoconference systems
- Very robust internet connections with high throughput
- PSNC's Poznan Living Lab, as effective member of ENOLL, for many years runs city challenge workshops, community hackathons and design thinking interdisciplinary courses.
- Many years of hosting **co-working space** for people who want to “hide” and focus on the development
- PSNC is located close to hotel facilities and is well connected with the centre of the city (public transport)

Experience of the team: Core Programming Team WP ISA EUROfusion

- ✓ Development, maintenance and support of the EU Integrated Modelling Platform
 - ✓ Support the development of the physics workflows (Kepler, Python, Matlab, Fortran)
 - ✓ Develop the tools and methods for running the physics workflows in HPC/Cloud environments
 - ✓ Develop the tools for monitoring, profiling, benchmarking, converting, storing results or the workflows
- ✓ Develop and maintain ITER Modelling and Analysis Suite (IMAS) infrastructure in a new community environment (multiple developer groups, multiple installation sites, ...)
 - ✓ Part of the development team of the IMAS Clusters/HPC: CINECA/ITER/JET/Gateway Continuous support to users/developers, trainings

Team composed of current CPT members, software engineers involved in fusion related projects (exascale, HPC, workflows), AAI initiatives, cloud and distributed data management, various expertise on board:

- ✓ Computer scientists, software engineers, initial IMAS developers, GUI developers, workflow system experts, HPC experts (optimisation, multiscale computing), containerisation experts, visualisation experts
- ✓ Python, C, Fortran, Java, React,, Data management, Databases/repositories, devOps, CI,
- ✓ Physicists

The Core Software Integration Advanced Computing Hub (CSI ACH)

The Core Software Integration Advanced Computing Hub (CSI ACH) focus mainly on addressing **category 2** objectives (Integrated Modeling and Control) in particular on the **support of the TSVV's requirements**

- **IMAS framework development**
- **Support for developers in terms of code adaptation to IMAS**
- **Integration of codes within IMAS based workflows**
- **Management of the repository for the community developed research software**
- **Production support of the integrated modeling tools**
- **User support and training**
- **Continued WPCD workflow support.**
- A set of transversal innovative activities is also proposed for benefit of overall EUROfusion, including:
 - support for integration of EUROfusion services (e.g. data, metadata, databases, web interfaces) with EUROfusion Federated Authentication and Authorization Infrastructure(AAI) and
 - support for exploitation of the Cloud technologies and infrastructures.

The Core Software Integration Advanced Computing Hub (CSI ACH) in numbers

Beneficiary	2021 (PPY)	2022	2023 ->
PSNC/IPPLM	2.67	5	6 (7.1, 7.1)
VR	0.25	0.67	0.75
CEA	0.08	0.42	0.42

Beneficiary	2021 (PPY)	2022	2023 --->
VR	0.22	0.5	0.75
IST	0.11	0.25	0.33

Current level of **initial requests 8.43/year from TSVVs** (not yet finalised)

Support of the TSVV's requirements

- ACH role is to focus and adjust to the changing, emerging TSVVs requirements
 - All the mentioned objectives will be driven by the TSVV's requests just to mention few:
 - IMASification, support code adaptation to IMAS (TSVV 1,2,3,5, 7,8,12)
 - Visualisation, GUI development, results catalogues/databases (TSVV 8,10,11)
 - Workflows development in different frameworks/languages (TSVV11)
 - Automation of the processes, workflows (TSVV11)
 - IMAS related trainings, tutorials, documentation,(All TSVVs)
 - Containerisation, (TSVV11)
 - Workflow testing framework (TSVV11)
 - DevOPS platform (CI) (TSVV11)
 - Installation on other target systems(All TSVVs)
 -
 - We are aware that the initial requests are just starting point in reality, we expect lot of smaller requests

IMAS-based ecosystem development

- Periodical evaluation and extension of IMAS standard data model (addressing EF community user needs).
 - Extending IMAS Data Dictionary based interfaces in close and open collaboration with ITER Organisation. Requests coming from TSVVs and other ACHs will be treated as the main driving force for this activity.
- Integrating research software with IMAS libraries, support for data sources (e.g. HDF5, HDC, UDA)
- IMAS Access Layer maintenance and new developments, as needed by TSVVs
- Providing users with support related to preparation of the IMAS based workflows to be run in an HPC environment. It will be performed with close collaboration with ACH selected in the HPC area.
- Installation and maintenance of the software constituting IMAS eco-system at various target systems
- Improving means of accessing publicly available results by extending current approach (centrally located pulse files) with new Catalog QT implementation, integrating outcome of Fair4Fusion project (Dashboard).
- Applying containerization technologies (e.g Singularity, uDocker) based approach to improve portability of the framework (different platforms).

Code adaptation to IMAS

- Code adaptation to IMAS will address following points from call
 - As new reduced physics modules from ENR, TSVV, and other activities become available for integration and testing in integrated modelling frameworks support can be provided to the code owner for integration in IMAS and potentially staged for exploitation towards integrated workflows.
 - To meet this requirement the CSI ACH will provide relevant documentation on the IMAS data dictionary, examples on building interfaces and test programmes together with a test environment where IMAS infrastructure and relevant data are made available.
- The hub will provide:
 - Documentation introducing IMAS and subsystems
 - Detailed documentation on the data dictionary and standardized code interfaces.
 - Test environment(s) and data
 - Compound test workflows for integration validation
 - Hands on coaching to code owners to: Develop the physics actors and integrate them in workflows;
 - Integrate them in (compound) test workflows; Installation in repository, preparing for CI/CD

Integration of codes within IMAS based workflows and workflow support

- This task will be performed in close contact with the developers of the physics modules (TSVVs for example) and with the end users of the IMAS workflows (modelers).
 - Workflows will be developed/updated according to target modelling scenarios requested by the modelers also taking into account availability of relevant physics modules provided by the developers.
- The following activities are foreseen as a part of the objective:
 - Further develop and maintain the release procedures for the actors and workflows. Introducing of the continuous integration.
 - The release procedure should also be compatible with different workflow platforms (based on: Python workflow system, any other framework that might emerge in the future)
 - Provide the set of test cases based on the relevant experimental scenarios
 - Create and provide the workflow documentation describing the features/capabilities

Production support of the integrated modeling tools

- Support the development of the EUROfusion Integrated modelling workflows in IMAS.
- Provide adequate release cycle of the maintained workflow following the modifications of the infrastructure or updates of the underlying physics models
- Provide flexible visualization tools to satisfy a large number of use cases for grouping and plotting relevant multi-dimensional data arrays as defined in the IMAS data dictionary, according to physicists needs. IMASViz plugins based python tool already installed at ITER is just one example of such a solution for which new plugins will be developed (but other tools will be also evaluated/considered/supported).
- Develop user friendly tools (GUI/CLI) for automatic creation of API adapters between physics based codes and data access layer, post processing tools to ease the application of software tools
- Promoting the concept of technical responsible officers for the workflows

Continued WPCD workflow support

- The ACH should provide continued support, maintenance and upgrades for WP CD established workflows
- ETS, EQSTAB and EQRECONSTRUCT (with sub workflows)
- (Negotiating AMNS and EDGE support with IPP for future voluntary contribution)
- Resources:

Beneficiary	2021 (PM)	2022	2023 ---->
VR	2.6	6	8
IST	1.3	3	4

- Requests of workflow support from TSVV9, TSVV10, WP SA, WP TE, (ENR)
- Request for components TSVV11, [TSVV8 (transport)?]
- Limited support for physics modules within ACH (distributed smaller scale support deemed not consistent with ACH [nor TSVV?] remit) - other mechanisms for support need to be found in addition.

Management of software repositories, CI, ticketing systems - considerations

- Source code hosting
- CI/CD platform for all codes will be available
- We are considering different options
 - Atlassian (Confluence + JIRA + Bamboo)
 - JetBrains (Space + YouTrack + TeamCity)
 - various tools from various vendors (e.g. Bugzilla, GitLab, Jenkins, etc.)
 - cheap but without support and requires manual integration from ACH side
- Source code can be hosted different ways
 - our custom installation of GIT - we can't benefit from features like pull requests
 - GitLab based installation (free) - we can benefit from various features but with some limitations
 - GitLab based installation (paid) - we can benefit from all the features available in GitLab
- Collaboration with Data Management ACH is foreseen

Communication, management

- The team's daily work will be organized according to agile development principles
 - Regular coordination meetings with selected stakeholders to follow the status of activities and discuss clients/stakeholders requirements - TBD
 - Strategy meetings on the topics/activities mentioned above including one or several parties - TBD
- Technical coordination with ITER IM technical representatives (following current CPT-ITER cord. meetings)
- Ticketing systems will be used for gathering the technical aspects, and to track the communication of the feature requests and the bug reports. Tracking of the activities, related to product development, will take place in Issue & Project Tracking Software.
- In post-pandemic times - F2F meetings/scientists visits and code camp style sessions (organised in Poznan) are foreseen
- We will issue also cyclic bulletins, and in addition other community channels will be used to inform users about latest changes, releases and plans for the developments (like slack??TBD, etc)

Open points

- The current level of requests **8.43/year from TSVVs** (that is most probably underestimated), is far more than the hub manpower is currently, taking into account the hub we have also booked PMs for
 - IMAS ecosystem support and development
 - IMAS installation support on different systems
 - Management of the repositories
 - other activities e.g. related to AAI, etc
- Usage of Gateway
 - software repositories availabilities/tools
 - resources/devops environment availability/contenerisation env.
- The need for prioritisation to effectively use limited ACH resources (especially in 2021)
 - After analysis of all requirements/priorities to focus re-adjust
- Few entries, that are requested to PSNC IMAS related hub, but probably should rather be addressed by HPC hubs,like
 - TSVV-05 Algorithmic development: implementation and optimization of code
 - TSVV-05 HPC-conform interfaces
 - TSVV-10 Speeding up transport models for burning plasmas in presence of Alfvénic waves and energetic particles once these models are established.
- TSVV11 - python actors and workflows:
 - iWrap will be available starting from 2022
 - The need for recommendation concerning WF design from IO

Contacts:

PI: Marcin Plociennik marcinp@man.poznan.pl

Questions?