Automating exascale workflows for fusion engineers

W. Smith[1], O. Woolland[1], A. Turner[2], L. Margetts[1]

[1] *University of Manchester, Manchester, UK*

[2] *EPCC, University of Edinburgh, Edinburgh, UK*

The productivity of teams of engineers involved in the design of fusion power plants is hindered by poor data exchange across technical specialist and organisational silos. To improve collaboration, we propose developing digital twins in the industrial metaverse that are powered by workflows. Workflows accelerate design and expose the opportunity to fully leverage automation, the first step towards introducing AI across the design process. However, when complex workflows involve simulation on exascale systems, traditional systems set-up using command line login and 2-factor authentication for a specialist human user become a barrier to automation. We explore using the Galaxy workflow engine and Pulsar to remove the human from the loop. A fusion engineering case study is presented that uses workflow automation on the HPE Cray EX ARCHER2 supercomputer hosted at EPCC. In the server-client model adopted by Galaxy, simulation tool installation and configuration becomes the job of system administrators, not engineers. Galaxy has been used to create tools that wrap existing fusion simulation, pre- and post-processing steps and allow them to be used in FAIR workflows, where tools can be arranged in many configurations via a GUI. Pulsar endpoints have then been configured on both local and remote compute resources. Pulsar is a Galaxy extension which allows jobs to be transferred from the Galaxy server to external compute resources, enabling jobs to be load balanced and endpoints selected based on the characteristics of a job (GPU needed, job size, etc.). Running on the endpoint, Pulsar performs the simulations and returns the results for further processing and continuation of the workflow. This means that scientists and engineers across all domains in fusion can use complex, computationally intensive tools with specific requirements on exascale or lower tier computer systems, from a digital twin environment hosted in the industrial metaverse.