# HPC-as-a-Service for Global Challenges

In recent years, an increasing interest in evidence-based computer-aided approaches to address global challenges (GC) including climate modeling and prediction of migration flows arose. This has led to a rapid development of computational global system sciences (GSS), a new research field that deals specifically with the design and evaluation of policies through digital twinning of global challenges. Accurate numerical handling of these global challenges inevitably leads to computationally expensive and costly simulations, though. These simulations are not only composed of complex models to solve different physical phenomena, but also require massive data coming from heterogneous sources. Then again, these simulations produce vast amounts of data outputs, which need to be further analysed to produce policies. Last but not least, new stakeholders are addressed that have often no experience with HPC. We therefore need to set-up mechanisms to lower the hurdle of using HPC resources, and to simplify the overall workflow in order to let non-experts execute complex simulations on HPC. All of this makes GSS applications particularly suitable for the implementation with an HPC-as-a-Service concept.

The concept of HPC-as-a-Service and the compute requirements imposed by global challenges are addressed with the EU-funded project HiDALGO Centre of Excellence (CoE). The mission of HiDALGO is to support global system sciences with high-performance computing and data analytics via coupling of simulations to produce more accurate results and to improve decision making. Thus, supporting non-experts in HPC usage is one key requirement, because GSS workflows are complex: They often make use of Artificial Intelligence (AI) and Machine Learning (ML) methods, require support for multiple cluster infrastructures (HPC, HPDA, Cloud), and data management systems (e.g., Lustre, HDFS, GridFTP). Both, the application developers and users, want to abstract all complexities and run their application smoothly by leveraging HPC infrastructure as a Service. Here, HiDALGO relies on Cloudify as a means for orchestrating workflows.



Cloudify is used to abstract different (remote) infrastructures and data management systems to define application and data flows. The workflow description language of choice is Cloudify-TOSCA. Cloudify-TOSCA is the defacto standard when it comes to define application execution flows. These textual descriptions are then used by Cloudify and its HPC plugin, Croupier, to submit and control the execution of applications on target infrastructures. Croupier extends out-of-the-box Cloudify functionalities to provide also support for HPC (interfacing both Slurm and Torque), HPDA (via Apache Spark) clusters, and different data management systems (e.g., CKAN and GridFTP), which allows altogether to implement the

complex GSS workflows. Furthermore, HiDALGO has set up Cloudify to interconnect with all the infrastructure and data management system to execute the applications by using the Cloudify-TOSCA based application model. Users can control the applications remotely by using the Cloudify REST interface instead of typical Unix Command Line Interface (CLI). These abstractions help different kinds of HPC workloads to run smoothly by using the multiple clusters and exchanging the data between them easily.

HiDALGO offers support to stakeholders by tackling today's and future global challenges through the power of high-performance computing, e.g. by facilitating contacts within the global challenges community, providing real-world solutions, and through consultancy, training, and infrastructure access. HiDALGO is also open for collaborations through its associate partners program, which SODALITE is an active member of.

For more information, please visit the HiDALGO portal at www.hidalgo-project.eu, or get in touch with the team via contact@hidalgo-project.eu.