HiDALGO

CONTENTS

DON´T MISS!

Teratec Forum 2019
June 11-12
Palaiseau, France

International Supercomputing Conference 2019
June 16-20
Frankfurt, Germany

NEWSLETTER

June 6, 2019 - Issue No. 1

HiDALGO

Welcome from the Coordinator

Dear Colleagues,

Nowadays, it is very difficult to live isolated and solve certain challenges individually, as their size and impact are, probably, beyond our area of influence. This is the case in Global Challenges, which is the topic addressed by HiDALGO, understood as problems affecting large parts of the world in general (i.e. several countries, a whole continent, etc.) or local problems which are reproduced in different parts of the world (i.e. issues that any large city in the world may have). Our use cases are very good examples of such problems: migration of people from conflict zones, how messages are spread through social networks and urban air pollution affecting large cities.

These are particular problems that require the use of large amounts of data and computation, and there is where HiDALGO aims at providing solutions. The usage of HPC is important, due to the multiple simulations required, their complexity (usually involving millions of agents) and the large amount of data to analyze for supporting decision making. HiDALGO provides tools and services that are necessary to address such problems, enabling new solutions applicable in different domains, especially based on HPDA.

We know that defining the solutions and workflows for these problems is very time consuming and complicated. Therefore, HiDALGO proposes solutions based on AI in order to make stakeholders' life easier, and provides automatic support when defining and running such workflows.

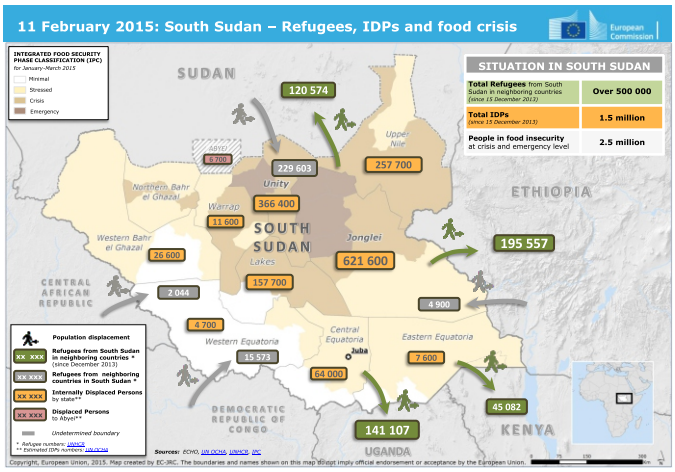
All of these tools will be supported by a Portal, which allows our stakeholders to configure and launch their simulations and applications through a user friendly GUI.

In this first issue, we present our three pilot applications of HiDALGO - migration, social networks and pollution - and give a brief description of the infrastructure used for our simulations. All in all, we expect that you will find HiDALGO solutions interesting and, of course, useful. Welcome to our project and I hope you will enjoy this newsletter.

Francisco Javier Nieto
ATOS Research and Innovation

HiDALGO

Migration Pilot



Forced migration involving South Sudan

There are more than 68 million people forcibly displaced worldwide, of which 24 million are refugees. These fleeing individuals are the unfortunate victims of civil wars and internal conflicts, who make decisions to migrate at the times of distress. To understand the causes of forced displacement, researchers establish three concerns faced by migrants, namely, the choice to stay or flee, the choice to flee internally or across borders, and the choice of destination.

As a baseline for the migration pilot in HiDALGO we have created a simulation development approach (SDA) that allows us to forecast population displacement in conflicts. This novel approach is important, because it could help to:

- Forecast population displacement when a conflict erupts, guiding decisions on where to provide food and infrastructure.
- Acquire approximate arrival estimates in regions where existing data is incomplete, to help prioritize resources to the most important areas.
- Investigate how border closures and other policy decisions are likely to affect the movements and destinations of migrants, to provide policy decision-makers with evidence that could support more effective policy and reduce unintended consequences.

In our approach we represent migrants as individual agents, combining simple rulesets for individuals to allow large scale and complex movement patterns to emerge. We have already successfully simulated population displacement in Burundi, Central African Republic, Mali and South Sudan using this prototype, and were able predict the destination of forced migrants with approximately 75% accuracy.

Although we have been able to construct basic simulations of population displacement, many major scientific challenges lie ahead of us. For instance, we wish to improve the accuracy and resolution of our model such that we not only can make predictions for the forced migrant destinations, but also reproduce the routes that they are likely to take in their travel. We also seek to clarify the role of weather and telecommunication in our simulations. The former could affect the journeys that migrants are likely to undertake, while the latter may help us identify the typical routes that migrants take in their journey towards their destination.

Last but not least, this simulation infrastructure will be constructed such that it can exploit large supercomputers, and accounts for sources of uncertainty (through the VECMA project, www.vecma.eu), as well as sensitivity to key assumptions.

HiDALGO

Social Networks Pilot

Social networks are omnipresent in our everyday life and have a significant influence on our behaviour. Political decision makers as well as economic players utilise social media to reach out and attract the attention of a substantial number of individuals. Also, people of different age use social networks for communication and interaction with each other. Examples of social networks with enormous influence are Facebook with over 1.4 billion daily users or Twitter with more than 500 million daily tweets .

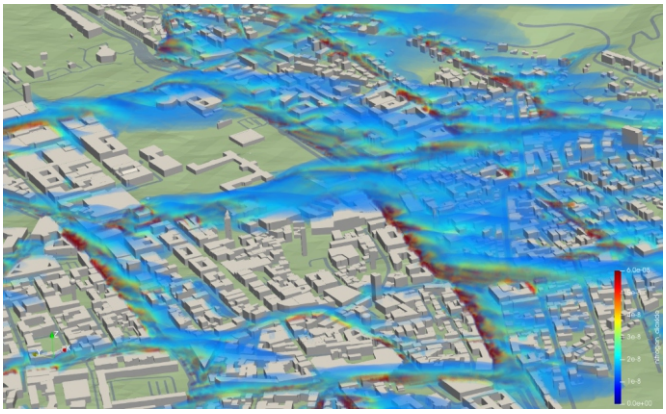
The analysis and simulation of the spread of messages in social networks is one of the three major use cases in the HiDALGO project. Our main goal is to understand the characteristics of the corresponding stochastic processes in such networks. We believe that this understanding will also allow us to identify malicious messages that intend to change the behaviour of a substantial number of users in these networks. This, in turn, will probably lead to algorithmic approaches on how to cope with such messages and allow decision makers to adopt countermeasures, in order to prevent the spread of such messages on a large scale. As part of the HiDALGO project, we plan to develop a highly scalable and accurate simulation framework, which is able to mimic the “life” of different real world social networks.

To achieve our main objectives, several scientific problems have to be addressed. Examples are the analysis of the structural and algorithmic properties of social networks, the stochastic behaviour of information diffusion, as well as the interplay between these aspects. Furthermore, several programming related issues must also be considered, e.g. the efficient simulation of the processes mentioned before.

First of all, we analyze existing social networks to understand their characteristics. Apart from classical properties, such as the degree distribution, diameter, or the clustering coefficient, we aim to understand the structure of the communities in these networks. Then, the structural characteristics are utilized to design synthetic graph models, which accurately describe these networks. Next, we design stochastic models for the spreading behaviour of messages in social media. Clearly, all these models have to be validated against their real world counterparts. Finally, we combine these models, and build an agent based simulation framework, which encompasses node level performance, parallel scalability and algorithmic efficiency.

HiDALGO

Pollution Pilot



Simulated NO2 concentrations
for Stuttgart (preliminary results)

Bad air quality in cities leads to 3 millions deaths per year according to WHO reports. One of the most severe pollutant is NO2 due to vehicular traffic. The vision of HiDALGO's urban air pollution pilot is to make the air cleaner in cities by using HPC and mathematical technologies.

The project will provide policy makers and our stakeholders a computational tool as service that accurately and quickly forecasts air pollution in cities with very high resolution. Furthermore, a traffic control system will be developed as well to minimize air pollution while considering traffic flow constraints.

First, the pilot will develop a HPC-framework for simulating the air flow in cities by taking into account real 3D geographical information data of the city, applying highly accurate computational fluid dynamics simulation on a highly resolved mesh and using weather forecasts data as boundary conditions. Dispersion of the vehicular traffic emitted pollutants in the wind field is computed with strong coupling to the air flow computation. Here the emission is computed from weakly coupled traffic simulations. The initial phase of the pilot will be based on the 3DAirQualityPrediction use case of the MSO4SC-project (see https://mso4sc.eu).

Within the HiDALGO-project, a traffic monitoring network will also be developed. The sensor network will consist of plate recognition camera system that covers all traffic at the main junctions of the demonstration area, the city of Győr, Hungary, thus providing real time traffic information including origin-destination and trip data. The traffic monitoring system will be completed with affordable cost air quality sensors for validation purposes.

Using the accurate HPC-simulation, the project will apply results of modern mathematical technologies, the model order reduction. This will result, based on several offline HPC-simulations and data analytics methods, in a new simulation framework of the air flow and pollutants dispersion that is of almost the same accuracy as the original HPC-simulations but requires hopefully much less computing resources.

The results will be visualised with the tools provided by the HiDALGO infrastructure. The urban air pollution simulation framework will be provided as a service via the HiDALGO portal in an easy-to-use web interface.

HiDALGO

Our Infrastructure

Photo © Boris Lehner for HLRS

HLRS' Hazel Hen is one of the HPC systems used for our simulations

HiDALGO relies on a tripartite concept in order to facilitate coupled simulations in the domain of Global Challenges. It is based on traditional High Performance Computing, High Performance Data Analytics, but also Cloud Computing resources, which are provided to the project by the three supercomputing centres HLRS, PSNC and ECMWF. Large-scale simulations are executed on the traditional High Performance Computing infrastructures, whereas Big Data analytics and Machine Learning based pre- and post-processing workflows are served by specific High Performance Data Analytics environments. These infrastructures are connected to remote Cloud services in order to foster local processing capabilities. Consequently, a spatially distributed compute service is built up within the frame of the project that follows an innovative approach for leveraging modern data-centric computing.

The HiDALGO Portal interconnects these three independent computing infrastructures. It is hosted in the Cloud and unites infrastructure services like data management, parallel data transfer, simulation governance or remote visualisation with community services such as user management, training, support and discovery. Hence, it represents a One-Stop-Shop that allows its users to control entire simulation and data processing workflows for around 500.000 HiDALGO compute cores and to build up the Centre of Excellence community at the same time.

All the infrastructures are tightly integrated and managed by a Continuous Integration and Deployment (CI/CD) pipeline that guarantees a maximum of stability and maturity for the HiDALGO services and application packages. Additionaly it will be utilised to support manifold computing architectures through pre-configured software compilation and testing as well as to install the tools on the various PRACE systems in an efficient manner.

HiDALGO

Partner Institutions

Events HiDALGO participated in

May 28-29, 2019 in Stuttgart, Germany **HiDALGO Plenary Meeting** The project partners had ample discussions about the current progress of the project and the future work.	**May 28-29, 2019 in Stuttgart, Germany** **URBAN SYSTEMS / GLOBAL CHALLENGES / DIGITAL TOOLS** Michael Gienger and Zoltán Horváth from HiDALGO gave presentations at the symposium.
May 9-10, 2019 in Amsterdam, the Netherlands **VECMA All-Hands Meeting** John Hanley from HiDALGO gave a presentation (Delivering ECMWF Data and Services via the Cloud within HiDALGO) at the VECMA (www.vecma.eu) meeting.	**April 16-18, 2019 in Edinburgh, Scotland** **CSW spring** Martin Maritsch and Kevin Winter (Area Knowledge Discovery) participated in a Student Challenge organized by the HiPEAC consortium.
April 08, 2019 in Brussels, Belgium **Digital4Med Conference** Derek Groen from the HiDALGO consortium was one of the participants in the panel discussion "Capacity building: how to create a data economy?".	**April 2019** **HiDALGO featured in HiPEACinfo** An article about HiDALGO appeared in HiPEACinfo, vol. 57, pp. 15-16.
February 21, 2019 in Frankfurt, Germany **FocusCoE Meeting** Different collaboration strategies between the funded Clusters of Excellence were discussed.	**December 12-13, 2018 in Madrid, Spain** **HiDALGO Kick-off Meeting** We shared our views, information, and knowledge, to boost the usage of HPC in Europe.

Project #824115 funded by the Horizon 2020 Framework Programme of the European Union