LEXIS PROJECT
LARGE-SCALE EXECUTION FOR INDUSTRY & SOCIETY

LEXIS Newsletter October 2021

The face-to-face progress meeting of the LEXIS project was a great success

The hybrid progress technical meeting of the LEXIS project took place from October 20 to October 22, 2021. Some project partners joined the meeting virtually and approximately 30 project colleagues attended the face-to-face meeting in Prague.

The main aim of the meeting was to present up-to-date results of the work, to plan activities for the final upcoming months and to engage in discussions that will help further successful cooperation.

All Work Package leaders presented the consortium and the External Advisory Board members key developments of their research with the particular focus on the LEXIS platform and the technical innovations resulted from the three project pilots – aeronautics, earthquake and tsunami and weather and climate. Last but not least, we dedicated the last half day to fruitful discussions aiming to sustain the exploitation of the LEXIS platform beyond the project lifetime.

LEXIS PROJECT
LARGE-SCALE EXECUTION FOR INDUSTRY & SOCIETY

LEXIS offers its Platform to SMEs in the 2021 Open Call

The LEXIS open call offers unique opportunities to third-party application experiments, apart from its three large-scale pilots, to further validate the LEXIS platform. Within this framework, the plan was to support more than five application experiments from the following key industrial sectors: healthcare, manufacturing and energy.

Projects with data intensive and computational-intensive requirements, representing various industrial sectors, research and academia were therefore encouraged to apply to the open call.

Selected applicants are able to test and scale their applications in the real HPC/Cloud/BD environments e.g. before entering the full-scale operation, using remote visualisation service if needed or to use the mixed CPU & GPUs architecture as required by their workflows.

The unified LEXIS web interface, **the LEXIS portal**, developed for workflow specification and execution, makes the use of HPC resources simple and possible to everyone, even to inexperienced HPC users.

During the entire open call period, user experience, including quantitative and qualitative data regarding implementing computational scenarios, technical functionalities and workflow efficiency etc. is collected and monitored.

As a consequence, valuable feedback and recommendations will be derived and used to fine-tune LEXIS technology to better meet user needs, leading towards better usability and sustainability of the LEXIS platform.

"In fact, application experiments leverage special workflows, knowledge and data gathered from three existing pilots. Additional application experiments selected in the open call are tested on the LEXIS Platform to confirm the innovative added value and general exploitability of the advanced LEXIS Hardware and Software co-design."
Marc DERQUENNES from EURAXENT

LEXIS OPEN CALL

In this issue, we are delighted to present you five innovative application experiments from selected SMEs participating in the LEXIS open call.

AIRMOBIS
Air mobility solutions

Application:
Intensive Multi-Phase Simulations – IMPS

Company:
AIRMOBIS, Czech Republic

Business sector:
Engineering, Aeronautics, Transportation

AIRMOBIS is an engineering company specialised in aerodynamics, hydrodynamics and thermal analyses. The company offers its applications mainly in the transportation domain (aircraft, boats, cars).

IMPS is an application for Computational Fluid Dynamics (CFD) simulation of the multi-phase system (air and water, air and solid) and thermal analysis, which will be tested on LEXIS. **Airmobis team is interested to use High Performance Computing resources through LEXIS platform in achieving higher result accuracy as well as making complex analyses with HPC requirements possible.**

In addition, possibility for using Paraview (remote visualisation functionality) to display model status during the computation will be explored. Example use cases are unsteady flow around maneuvering aircraft, multi-phase analysis of hulls and hydrofoils on boats, including wake deformation etc.

altrnativ

Application:
Altrnativ Data Networks – ADN

Company:
Altrnativ, France/Spain

Business sector: Cyber defence and Cybersecurity

Altrnativ is an operator and software publisher specialising in cyber defence, cybersecurity system and individual protection.

Altrnativ Data Networks, ADN, a data fusion platform solution, enables users in integrating, managing and securing all types of data. **Altrnativ will rely on LEXIS computing power to test, train and improve its solution, which has been developed based on Artificial Intelligence methodology.**

I-SEE
Internet Simulation Evaluation Envision

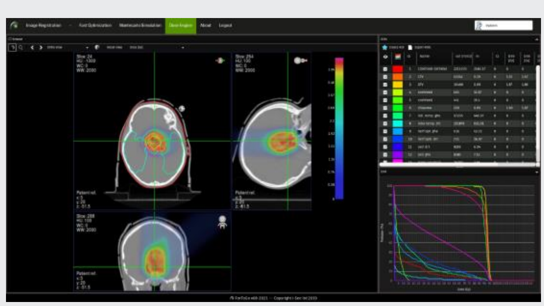
Application:
Full Monte Carlo treatment planning for particle therapy – 4SeePlan / SurviVe

Company:
I-SEE ltd., Italy

Business sector:
Healthcare

The goal of **4SeePlan** is to simulate the actual deposition of particle's energy into the patient in the advanced radiation therapy context (proton therapy). 4SeePlan evaluates the delivered dose given the irradiating field parameters using a Monte Carlo approach. The approach allows a greater accuracy in treatment planning dose evaluation at the expenses of the time of calculation.

4SeePlan is designed to run on a Multi-CPU platform, the workflow is based on a pure distributed computing. **I-SEE team will test its application via the LEXIS platform, aiming to achieve higher simulation results accuracy (dose evaluation) with significant runtime reduction.**



Application:
Radio-Frequency (RF) simulations in HPC

Company:
Open Engineering, Belgium

Business sector:
Engineering

Open Engineering has developed and validated an electromagnetic wave propagation coupled with temperature solver for **radio-frequency micro-electromechanical system** (RF MEMS) applications. Its core algorithm is based on the Finite-Difference Time-Domain (FDTD) method.

Two versions of the application exist, written in C++ and CUDA (CPU & multi-GPU). **Open Engineering aims to investigate software scalability using LEXIS capabilities and to explore technical, economical feasibility of the provision of RF solver in HPC environments to end users.**

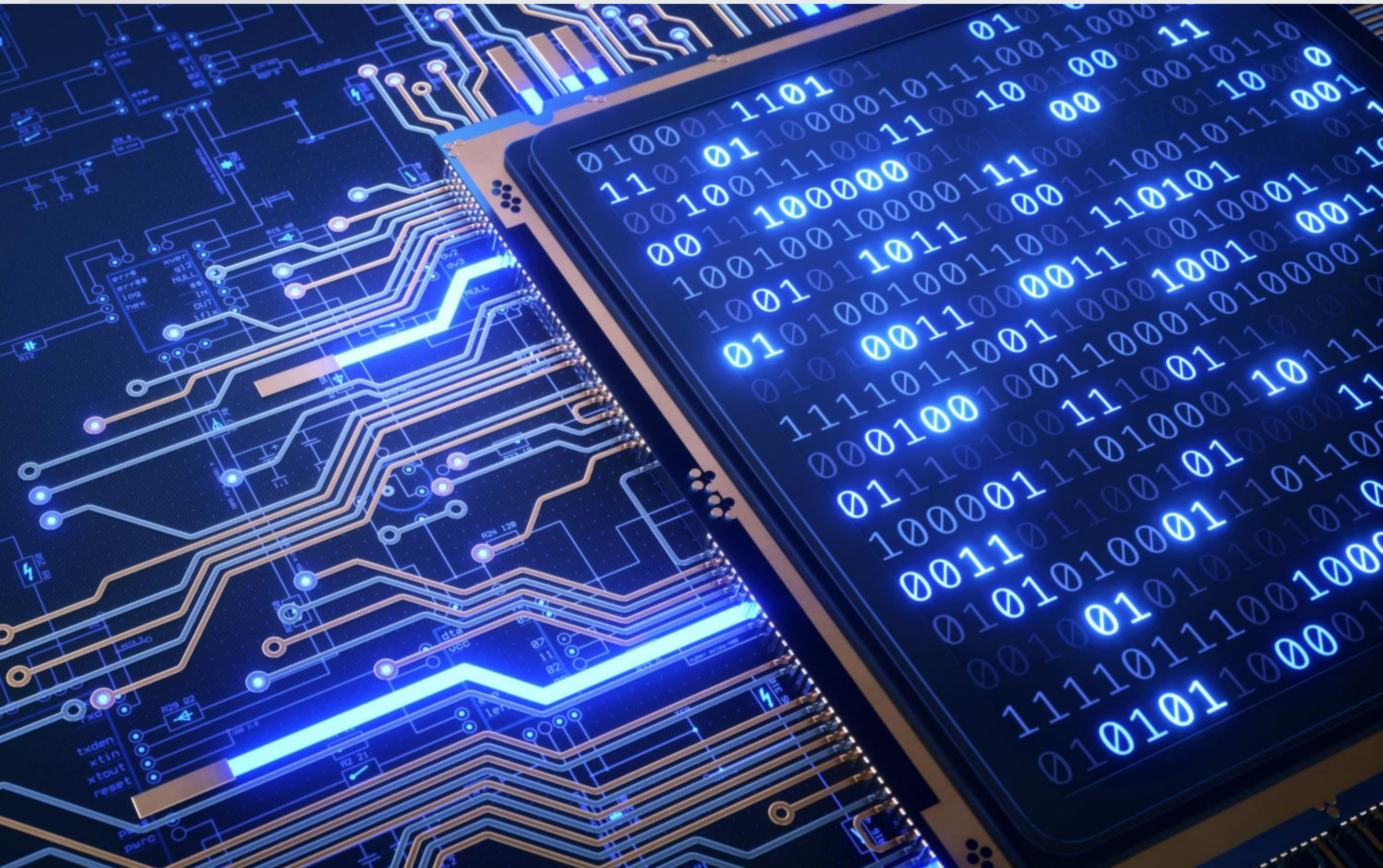
Application:
Virtual Screening for drug discovery – PHARMVS

Company:
Pharmacelera, Spain

Business sector:
Healthcare, Pharmacy

Pharmacelera has developed a software to help pharmaceutical companies in their R&D efforts. The tool that will be used in this experiment, **PharmScreen** is capable of analysing large libraries of millions of compounds to look for promising drug candidates in the preclinical stages of a drug discovery project.

Their goals when using the LEXIS platform are to study the scalability of the software in large computing systems, to learn how to integrate their software in other environments and to test easy access to large computing systems through the LEXIS portal.



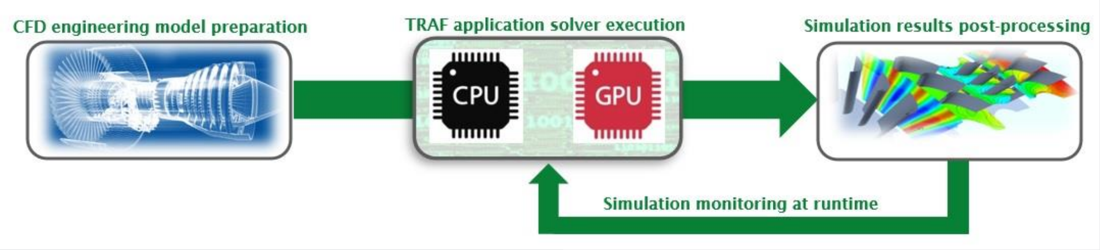
Aeronautics Pilot: Advanced CFD Simulations of Aeronautical Engine Critical Parts

The aeronautics large-scale pilot aims to significantly enhance the feasibility and exploitation of advanced Computational Fluid Dynamics (CFD) numerical modelling capabilities able to predict the fluid-dynamic behaviour of aircraft engine critical components. To meet this ambitious objective, the industrial applicability of the next-generation HPC/cloud/big data management technologies is assessed in LEXIS through two aeronautical engineering case studies, a turbomachinery application and a mechanical rotating part.

Faster and more accurate CAE analyses, that exploit newly deployed HW/SW resources in the LEXIS engineering platform, are carried out to enable the implementation of the improved or newly designed CFD-based engineering methodologies that both considered case studies rely on.

Aeronautics turbomachinery use case

Referring to the aeronautics turbomachinery use case, a high-level overview of **the deployed CFD simulation process is depicted *in the following figure*** that illustrates the three main stages (pre-processing, analysis solver execution, post-processing of results) of the application workflow.

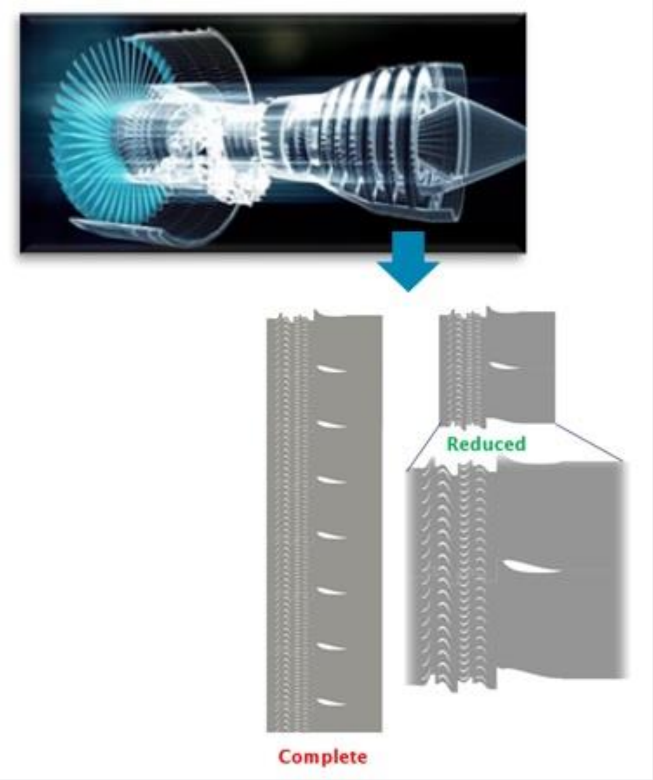


Three-stage application workflow for the CFD simulations in the aeronautics turbomachinery use case

Based on TRAF code, a CFD solver developed by University of Florence is used to investigate fluid dynamics phenomena with a special focus on turbomachinery simulations. The computational phase is here strongly HPC-demanding and time-consuming.

In order to drastically reduce the execution time of TRAF-based fluid dynamics analyses on low pressure turbines, a newly designed GPU-enabled solver release has been developed, showing promising results in terms of speed up.

The considered turbomachinery pilot test case is representative of a multi-stage turbine, shown *in the following figure*.



Low pressure turbine test case in the turbomachinery use case

Due to the large number of computational blocks, the underlying CFD model is very HW-intensive in terms of both the required memory resources and the execution time needed to perform the calculations.

Since the memory in the currently available GPU accelerated systems is not sufficient to manage this model, a reduced version of the test case has been used (***see the figure on the left***).

The shown representations of the test turbine are marked as “Complete” and “Reduced” depending upon whether a pure CPU-based system or a GPU-equipped system is used respectively.

Compared to the solely CPU-based system, **the CFD simulations executed through the GPU-ported release are currently running 5.22x faster on the Barbora HPC.**

Although the basic industrial target in terms of simulation speed up has been reached, the research team within the turbomachinery use case foresees further optimisations for the computational phase, looking to take advantage of latest compiler enhancements to minimise remaining performance bottlenecks and to exploit the best GPU-to-GPU communication channels.

Moreover, the opportunity to further reduce the execution time of the most HW-intensive subroutine of the CFD solver through FPGA-based HPC platforms is investigated.

The rotating parts use case

As regards the rotating parts use case, a high-level overview of the deployed CFD workflow is depicted ***in the following figure*** which illustrates its three main computational stages and the involved key software applications.

Based on ALTAIR nanoFluidXTM, a **smoothed particle-based hydrodynamics (SPH)** simulation tool conceived and optimised for use on clusters of GPUs to predict multi-phase flows around complex geometries under complicated motion, the computational phase here is also strongly HW-intensive and requires advanced HPC systems equipped with several GPUs due to the complexity of the input models.

With the aim of tuning a new CFD-based engineering practice to simulate, the flow field operates the multiphase flow inside gearboxes with increasing accuracy.

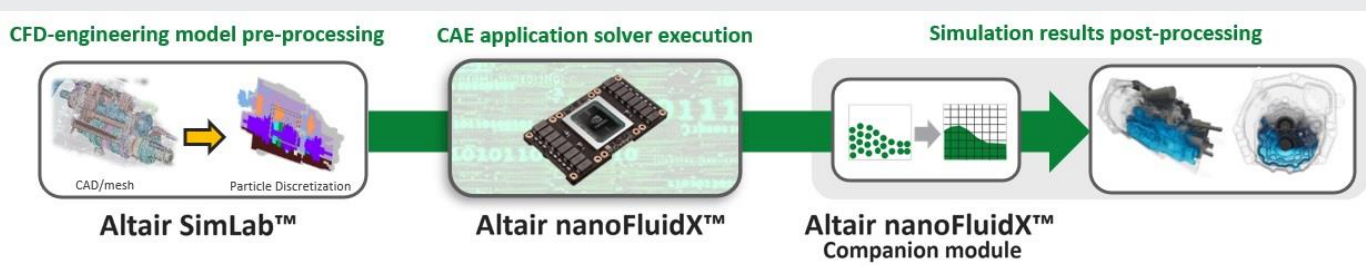
The presented application workflow is being deployed over the 3 engineering case studies, differing in their degree of complexity and sensitivity.

The three considered rotating parts test cases are:

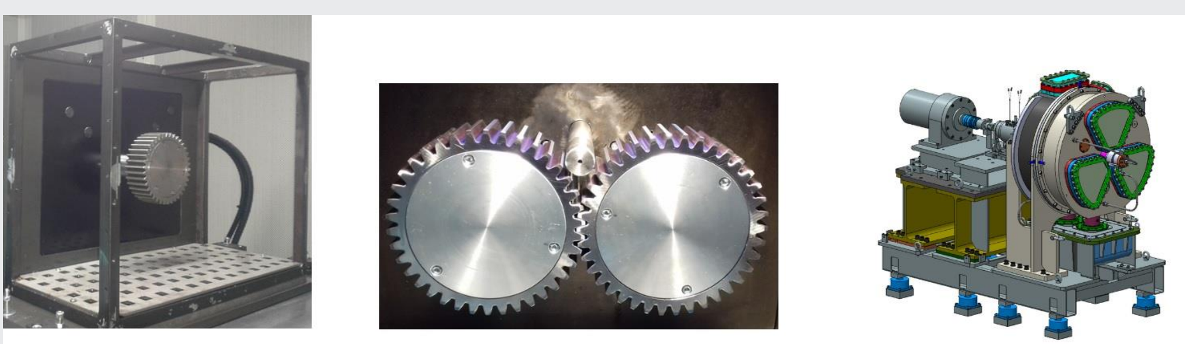
- **single high-speed wheel** used to run separately SPH Oil Phase simulations and air flow,
- **gear pair with oil jet lubrication** adopted with the target of applying multi-phase approach in simplest gearing geometries,
- **planetary gearbox** selected to carry out and validate industrial-like simulation predictions on a more complex environment, illustrated ***in the lowest figure of this page.***

Based on the preliminary outcomes from the research activity within the rotating parts use case, the first phase of investigation can be considered accomplished. SPH methodology is really promising with reference to the pre-processing phase, simple and very short, despite alternative CFD methods, like Finite Volume numerical which requires longer simulations (about 10x more time-consuming).

Finally, some best practices to carry out air/oil investigation have been defined with very good agreement with experimental data.



Three-stage application workflow for the CFD simulations in the aeronautics rotating parts use case



Selected test cases in the rotating parts use case

MPI Parallelisation enables Precise Tsunami Simulation with Inundation in Real Time

photo: nbcnews

LEXIS to serve real-time computing needs

The LEXIS project aims to enable demanding computational workflows, and to optimise their building blocks and components.

To better respond to the real-time computing requirements, we added parallelisation via the **Message Passing Interface (MPI)** to the tsunami simulation code TsunAWI and made it executable on the LEXIS platform.

The modified code shows substantial improvement in terms of quality of the simulation results (more precise) as well as the processing time.

Computing scenario in the past

In the past, our focus with TsunAWI has been on pre-computing tsunami scenarios in high resolution, based on the non-linear shallow water equations with flooding and drying. These simulations are of high quality.

Their results can serve not only as early warning products based on the wave height at the coast, but also as an input information for risk maps and evacuation planning.

Solutions to meet early-warning system requirements

However, without any further optimisation, high resolution simulations have been very time-consuming. Till now, real-time computations in case of a tsunamigenic earthquake had to be performed with simpler physical models, without inundation simulation, and at coarse resolution.

Precise simulations with TsunAWI at an acceptable processing time, meeting the requirements of an early warning system, has not yet been possible.

LEXIS enabled us to take a great leap!

Instead of using 4–8 processor cores of a personal computer or 20–40 cores of a server based on OpenMP instructions, by adding MPI parallelisation to the code, TsunAWI can now use parts of a supercomputer, consisting of many servers (so-called nodes). For example, in LEXIS, tests run on IT4I's Salomon cluster or on LRZ's Linux Cluster – typical HPC (High-Performance Computing) machines.

What is the computing workflow about?

By executing the TsunAWI code on these large machines, all processor cores work together to compute the tsunami propagation. We have now adopted the framework of the global ocean model, FESOM (https://fesom.de/), TsunAWI's big brother, which helps us to distribute the work among processors, dividing a large computational mesh into smaller submeshes.

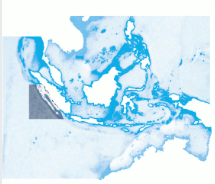
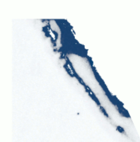
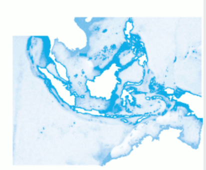
In order to compute the tsunami propagation of a large mesh parallely, values at the borders of submeshes have to be exchanged, for which different compute nodes communicate with each other using explicit MPI messages.

Implementing this in an optimised way involves quite a programming effort, but it has two advantages.

First, real-time applications can fully exploit the high-performance computing resources. **Second**, explicit MPI parallelisation allows the programmer to understand speed limitations caused by passing messages.

Thus, the overhead time due to the parallelisation can be analysed and reduced, improving the parallelism efficiency, also for runs on a single server system.

We have obtained the best performance based on a hybrid-parallel approach, where 4 parallel processing "threads" are started via OpenMP, and the rest of the parallelisation occurs via MPI.

	regional coarse	regional fine	Indonesia
			
Number of vertices	230,000	1,240,000	11,110,000
Resolution	200m-15km	20m-5km	150m-20km
Time step dt	1.5s	0.15s	1.0s
Model time	2h	2h	12h
Compute time for time stepping: 2x Intel Xeon Cascade Lake 48core			
1 node	4s	171s	707s
2 nodes	3s	96s	378s
4 nodes	–	58s	350s
10 nodes	–	34s	152s
20 nodes	–	–	89s
40 nodes	–	–	61s

The table above shows the run-time improvement of the MPI parallel TsunAWI when more cores are used (strong scaling), or when number of cores and problem size are modified (weak scaling). 2–4,000 cores can be used by TsunAWI with good efficiency.

Faster & near real-time, higher result accuracy, larger processed area – the computing capabilities enhanced by LEXIS

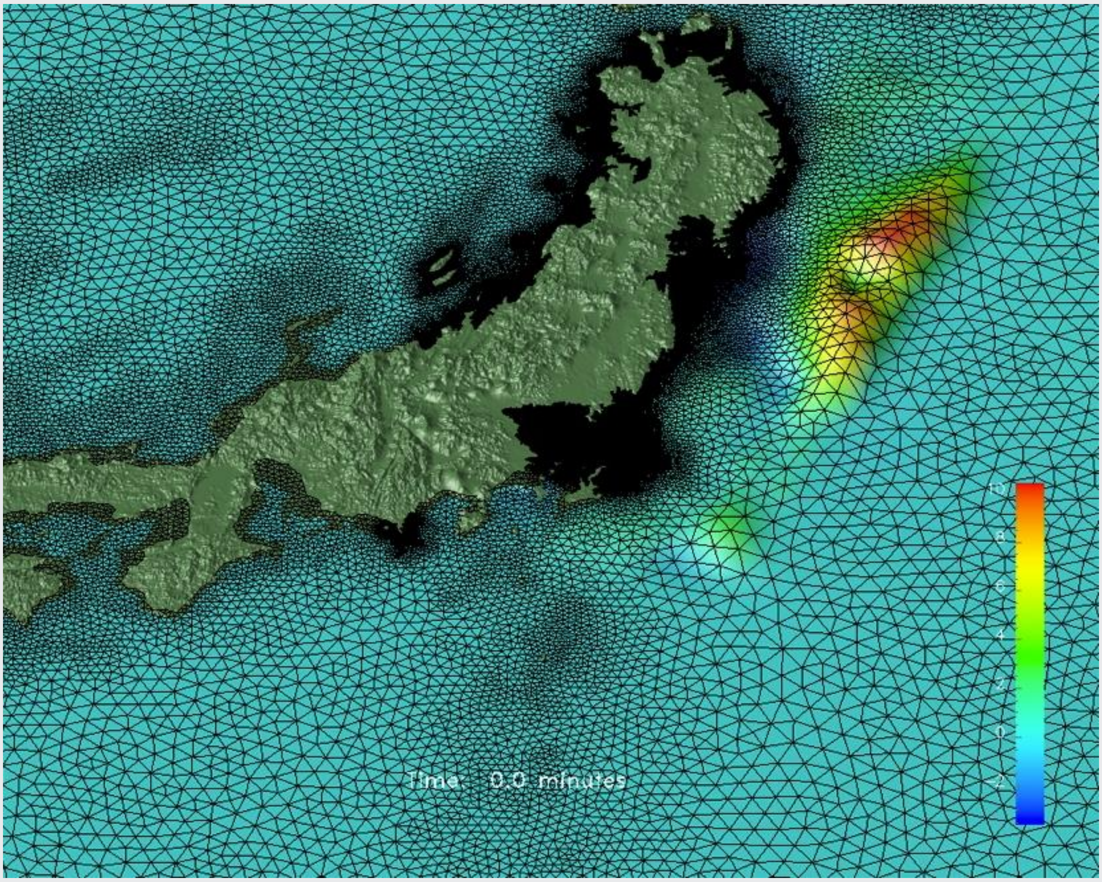
With sufficient compute resources at hand, simulation areas can be extended without additional time needed for the simulation. In the past without MPI parallelisation, we had to restrict the TsunAWI simulation to one region of interest, e.g. the city of Padang (Sumatra island), at coarse resolution for fast estimates.

By executing the enhanced code on the LEXIS infrastructure, we can now compute the tsunami propagation and inundation for the whole Indonesian Archipelago within just one minute.

A simulation for this same area with its 5,000 km extent from Sumatra in the West to New Guinea in the East, would have cost us about three hours some years ago.

In addition to the LEXIS project, this effort has been made possible by more people and entities, in particular the North German Supercomputing Alliance (HLRN) and Atos/Bull.

We thank Sebastian Krey (GWDG) for optimising the compute node setup, the Scientific Board of HLRN for granting a consulting project, and Alesja Dammer and John Donner (Atos/Bull) for many hints on code optimisation in a hybrid-parallel context.



Initial condition of tsunami simulation result of Sendai (Japan)

LEXIS PROJECT
LARGE-SCALE EXECUTION FOR INDUSTRY & SOCIETY

The Effective Management of Workflows through a Dynamic Job Allocation Strategy

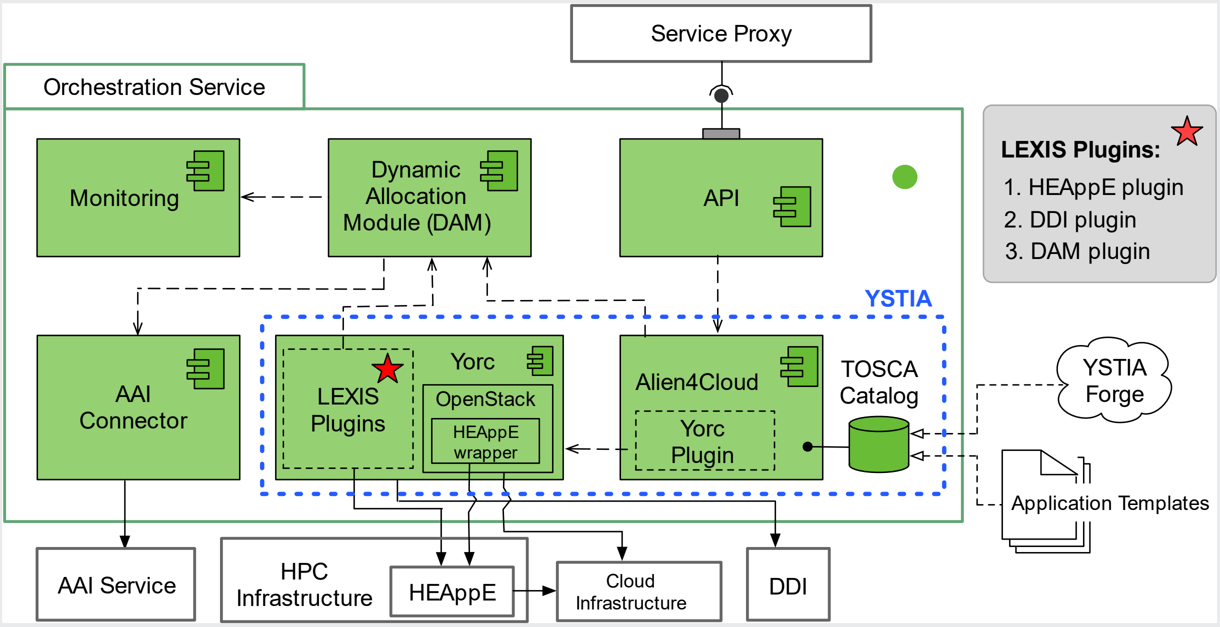
Application workflows have grown in complexity over the last years by extending their execution requirements by far and including even more AI and big data analytics tasks. Supporting these complex application workflows requires to distribute their execution on computing and storage resources made available by a federation of HPC centres.

Reasons behind this approach are many for instance; accessing to different types of resources (hardware accelerators – GPUs, FPGAs etc. – and virtual instances), providing the resources for executing **urgent computing (UC)** tasks, overcoming both planned and unplanned unavailability of underlying infrastructures.

LEXIS Orchestration Service (OS) building blocks are designed and integrated together to form a flexible and effective solution for managing computing and storage resources i.e. the execution resources.

This solution allows workflow jobs or tasks to seamlessly access the execution resources on geographically distributed HPC centres in a dynamic fashion.

In its final form, the LEXIS OS will perform a flexible and dynamic allocation of execution resources across the federated HPC centres to optimise workflows efficiency or resource-usage balance.



Architectural diagram showing the main building blocks of the LEXIS Orchestration Service and their relationship

The LEXIS Orchestration Service features a workflow execution engine (**Yorc**) and a user-friendly frontend (**Alien4Cloud**), which are collectively called **YSTIA orchestration suite.**

This software stack, which exposes an API for facilitating its integration with the remainder of the LEXIS platform, is well suited for the execution of complex workflows composed of several jobs/tasks while checking their status and dependencies.

It also allows external modules to perform specific functions. Last but not least, the TOSCA catalogue stores the set of TOSCA components used to define application templates (i.e. a description of the resources, software elements and their dependencies needed to derive a running workflow), as well as to connect to a public repository (YSTIA Forge).

In this context, the LEXIS consortium identified an important feature in the dynamic job placement, which is needed for an efficient and robust workflow execution in a federated and multi-resource environment. As a result, the **Dynamic Allocation Module (DAM)** has been developed and deployed as a function-specific delegate for YSTIA.

The idea behind the DAM is **to have a resource selection strategy that is able to select the best execution location among all the available computing centres, based on the analysis of different evaluation criteria** e.g. occupancy of the clusters, data transfer speed among locations etc.

To this end, relevant information is collected from the HEAppE Middleware instances and the LEXIS Distributed Data Infrastructure (DDI). The DAM then computes and returns the "best" location, where to execute a given job based on the status and availability of the monitored resources to YSTIA.

Thus, a set of specific plugins for the DAM and other LEXIS platform components i.e. DDI, HEAppE have been integrated in the YSTIA stack to connect to HPC/cloud resources, as well as to authorise operations through the LEXIS AAI service ***(see architectural diagram above)***.

A simulator was used to verify the performance of this approach that should not be worse than purely **random** and **round robin** strategies. We refer the DAM selection strategy as **"greedy"**. Since there is no awareness about list of the submitted jobs, the strategy decides job by job independently.

Job placement requires a fast and effective decision-making procedure. To this end, the DAM implements a greedy-based job placement algorithm which allows selecting the “best” execution location, by combining information on the federated resources availability and status, where datasets are located, and (in its final form) the number of core-hours/credits available.

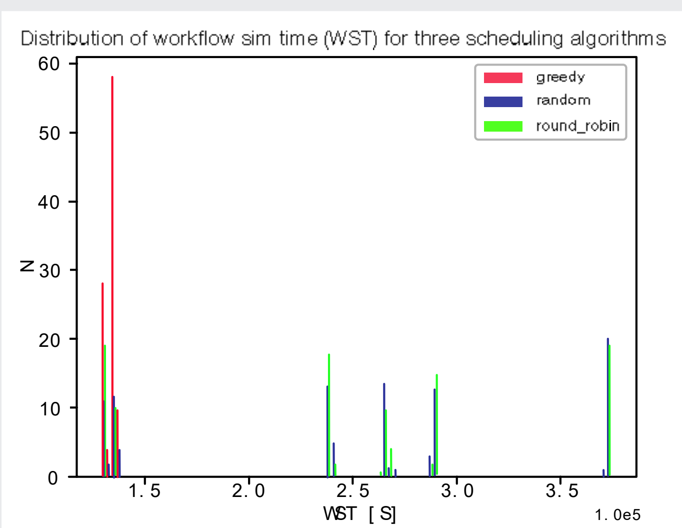
Testing of this algorithm has been done through a dedicated simulation framework, which has been developed in the context of the LEXIS project based on SimGrid library.

The framework (written in C++) allows for modelling the resource federation by defining HPC/cloud resources and their network connections, more specifically, by specifying a set of parameters i.e., number of nodes/cores, computing capability of cores, connection speed etc.

Similarly, input workflows can be defined beforehand by setting their own parameters (i.e. number of jobs, required resources for each job, size in terms of number of operations to perform etc.). The software simulates the placement and execution of jobs over the time through a scheduling and placement policy defined by the user.

To our testing purposes, we defined the greedy job allocation policy along with other two policies, round robin and random.

As a result of the simulation framework, several files, containing relevant simulated values e.g. execution time, location, cluster occupation as a function of simulated time, are generated (*see the following figure*).



Allocation Algorithm	Avg (s)	Std (s)
Random	248927	85807
Round Robin	246124	84648
Greedy	**134294**	**2472**

An example of simulation outputs: the distribution of (simulated) execution times for a simple workflow. Three different job placement strategies are compared. The average and the standard deviation values over 100 executions are provided.

The LEXIS Dynamic Allocation Module

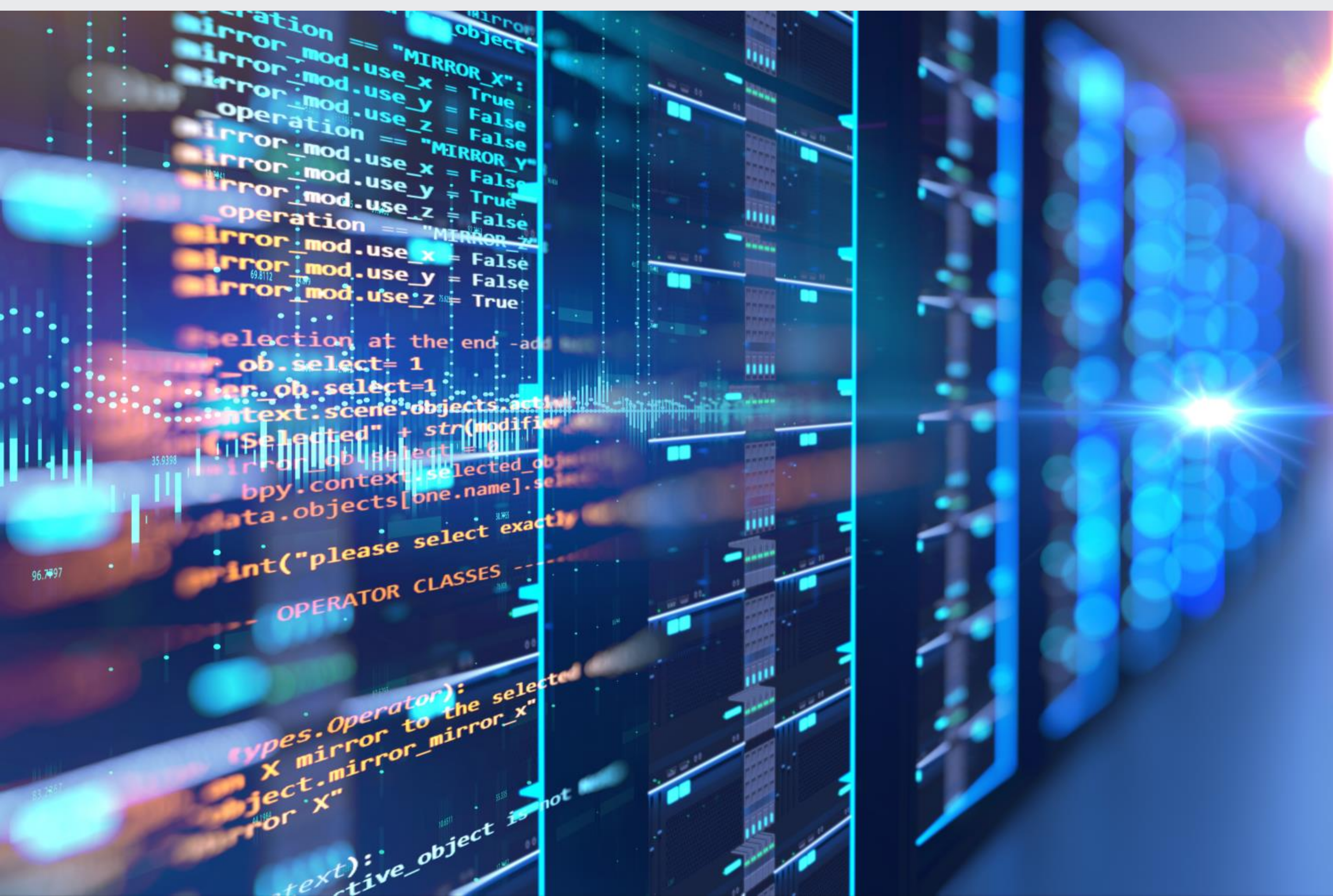
The “best” execution location for a given job is computed through the aforementioned greedy algorithm, looping over all the available computing locations to check the compatibility of hardware/software features with the job’s strict requirements and to evaluate a set of metrics used to identify the “best” solution.

In fact, the algorithm computes the value of a set of functions, modelling specific evaluation criteria for each location e.g., current load, cost of data movement, then these values are combined (averaged) to determine the final (best) location. Evaluation criteria allow to select the execution location in such a way the resources usage of federated clusters remains balanced over the time.

In the final version of the DAM, real-time billing information for different HPC centres, gathered and elaborated by the LEXIS Account and Billing system, will be used to exclude locations with insufficient core-hours/credits.

The DAM has been implemented in Python using the Flask library, which eases the creation of a REST-API and the connection with the other building blocks of the LEXIS Orchestration Service. All the requests are authorised by means of the integration with the LEXIS AAI system, in order to adhere to the zero-trust security concept.

An integrated database solution (based on the InfluxDB technology) is used to log all the decisions made (i.e., locations selected along with the list of values of the evaluation criteria) and to store information related to planned maintenances of the HPC/cloud resources as well as the data transfer speed among them.



LEXIS PROJECT
LARGE-SCALE EXECUTION FOR INDUSTRY & SOCIETY

LEXIS
Activities in brief

AUGUST 2021

ETP4HPC released European High-Performance Computing Projects Handbook 2021

The ETP4HPC association released the final version of the **"European High-Performance Computing Projects Handbook 2021"** available on its web page. The handbook gathers selected ongoing European projects in the context of high-performance computing. Undoubtedly, the LEXIS project is presented there too.

Handbook:
https://www.etp4hpc.eu/european-hpc-handbook.html

SEPTEMBER 2021

European Meteorological Society 2021 – EMS 2021

September 3–10, 2021, digital

The LEXIS team from Weather and climate pilot, **Paola Mazzoglio, Antonio Parodi, Paolo Pasquali and Andrea Parodi** contributed a talk at the EMS 2021 event. Their work **entitled "Improving weather forecasts by means of HPC solutions: the LEXIS approach in the 2020 Bitti flood event"** represents a good example, in which data-intensive workflow of a weather forecast system can be benefited from the powerful HPC computing resources like LEXIS.

Abstract: https://doi.org/10.5194/ems2021-125

19th Workshop on high performance computing in meteorology 2021

September 20–24, 2021, digital

LEXIS partner, ECMWF, organised its **19th Workshop on HPC in meteorology**, this time fully digital.

On September 22, **Tiago Quintino** (ECMWF) gave a talk **"On the Convergence of HPC, Cloud and Data Analytics for Exascale Weather Forecasting – ECMWF Present and Future"** in which innovative data access algorithm researched in the context of LEXIS project was introduced.

Workshop:
https://events.ecmwf.int/event/169/timetable/

Researchers' Night 2021 in the Czech Republic

September 24, 2021, Ostrava, Czech Republic

The National Supercomputing Center IT4Innovations joined the Researchers´ Night again. The evening event, during **which hundreds of scientific workplaces in Europe open their doors to showcase researches and its impacts on the people daily lives**, took place on September 24 this year. The theme of this year's event was TIME. The LEXIS project was present at Researchers´ Night and successfully introduced its innovative outcomes thanks to its video, quizzes, competitions, and promotional items which won the favour of all visitors.

That evening, IT4Innovations welcomed 590 visitors, successfully brought researches closer to the public by offering them a range of activities: excursions, lectures, competitions and most importantly, fun and entertainment.

Event:
https://www.it4i.cz/en/events/researchers-night-2021

OCTOBER 2021

The F2F progress meeting of the LEXIS project

October 20–22, 2021, Prague, Czech Republic, hybrid (*see page 1*)

NOVEMBER 2021

SC21

November 14–19, 2021, St. Louis, USA

The LEXIS project together with IT4Innovations will be present at Supercomputing Conference 2021 in St. Louis, USA during November 14–18. The LEXIS project will have a wonderful booth and it will promote its activities and also other projects. You will have chance to get LEXIS goodies and we have great power banks for three lucky winners. Stop by our booth 1417 to find out more!

https://lexis-project.eu/web/event/sc21/

Tech Forum "TRAF code GPU porting: final release"

November 2021, organised by Avio Aero, digital

The tech forum will mainly discuss key results and achievements obtained up to now in the development of TRAF code in the context of LEXIS project. New perspectives in modelling capabilities, aimed at improving knowledge of the flow evolving in turbo-machinery passages will be also presented.

This work was supported by the LEXIS project funded by the EU's Horizon 2020 research and innovation programme (2014-2020) under grant agreement No 825532.