H2020-INFRAEDI-2018-2020

The European Centre of Excellence for Engineering Applications

Project Number: 823691

D4.1
Application Fields and Competence Map

The EXCELLERAT project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 823691

Workpackage:	4	Enhanced services
Author(s):	Niclas Jansson	KTH
	Gordon Gibb	UEDIN
	Mark Sawyer	UEDIN
	Daniel Mira	BSC
	Adam Peplinski	KTH
	Thomas Gerhold	DLR
	Gabriel Staffelbach	CERFACS
	Dennis Grieger	USTUTT
	Christian Gscheidle	FRAUNHOFER
	Joachim Grimm	SSC
	Janik Schüssler	SSC
Approved by	Executive Centre Management	01/06/2019
Reviewer	Daniel Molka	DLR
Reviewer	Ivan Spisso	CINECA
Dissemination Level	PU	

Date	Author	Comments	Version	Status
2019-04-15	N. Jansson	Initial draft	V0.1	Draft
2019-04-30	All	First complete draft	V0.2	Draft
2019-05-08	N. Jansson	For QA review	V0.3	Draft
2019-05-27	All	Address reviewer's comments	V0.4	Draft
2019-05-28	All	Address reviewer's comments	V0.5	Draft
2019-05-28	N. Jansson	Version for final approval	V0.6	Draft
2019-06-01	N. Jansson	Document accepted for submission	V1.0	Final

List of abbreviations

API	*Application Programming Interface*
CAD	*Computer-Aided Design*
CFD	*Computational Fluid Dynamics*
CoE	*Centre of Excellence*
CPU	*Central Processing Unit*
DMD	*Dynamic Mode Decomposition*
DSL	*Domain Specific Language*
GPU	*Graphics Processing Unit*
GT	*Gas Trubine Engine*
HPC	*High Performance Computing*
IC	*Internal Combustion Engine*
I/O	*Input/Output*
IP	*Intellectual Property*
ML	*Machine Learning*
MPI	*Message Passing Interface*
PDE	*Partial Differential Equation*
POD	*Proper Orthogonal Decomposition*
SEM	*Spectral Element Method*
VR	*Virtual Reality*
WP	*Workpackage*

Executive Summary

The deliverable *D4.1 Application Fields and Competence Map* puts into context all relevant application fields and competences covered by the EXCELLERAT Centre of Excellence (CoE) partners, reference applications and offered services. Based on the tasks in WP4, an initial services portfolio is envisioned, building on the competence of the consortium and the common needs of the reference applications. We conclude with an outlook on how the service catalogue will evolve.

Table of Contents

Table of Figures

Table of Tables

1 Introduction

The strength of an engineering Centre of Excellence strongly depends on the application fields covered by the Centre, and the competence of the consortium behind it. An engineering CoE should support application fields relevant for both industry and academia to be able to offer a relevant services portfolio.

This document puts into context all relevant application fields and competences covered by the European Centre of Excellence for Engineering Applications through its offered services. The purpose of the document is to map competences within the consortium, the application fields covered by the Centre's reference applications and its offered services. First, we present the competences within EXCELLERAT followed by a presentation and motivation of the reference applications, their corresponding application fields and common challenges moving towards exascale. This is followed by a presentation of an initial services portfolio, and we conclude by giving an outlook on the evolution of this catalogue.

2 Application fields of EXCELLERAT

EXCELLERAT is built on top of a strong consortium, where each partner complements the Centre with key competences in several different application fields, various areas of HPC and in the field of computational science, covering the entire life cycle of an engineering simulation. Currently the Centre's core competences are:

- Pre-processing
  - Mesh generation (structured/unstructured)
  - Data transfer/handling
  - Data input (I/O)
- Simulation
  - Performance engineering
  - Porting to new architectures
  - Numerical methods (e.g. linear solvers, discretisation methods)
  - Intrinsic optimisation methods (e.g. adjoint methods)
  - Mesh adaption (structured/unstructured) and load balancing
- Post-processing
  - Data reduction algorithms
  - Data analytics
  - Data output (I/O)
  - Visualization methodologies (offline, remote or in-situ).

Besides the core competence of the consortium, EXCELLERAT is also supported by six reference applications (referred to as **CX**), accompanied by eleven use-cases (referred to as **UY**). These applications and related use-cases, presented in Table 1, have been selected to cover a wide range of engineering applications fields. Furthermore, the use-cases have also been selected such that each one exposes one or several common challenges related to the transition of engineering simulation towards exascale computing. Therefore, the services offered by EXCELLERAT will not be focused on a specific reference application or application field covered by a use-case. A more detailed description of the reference application, the use-cases and related challenges is given in the next section.

Nek5000 (C1)	**C1U1**: Aerospace – Wing with three-dimensional wing tip
	C1U2: Aerospace - High fidelity simulation of rotating parts
Alya (C2)	**C2U1**: Automotive/ Aerospace - Emission prediction of internal combustion (IC) and gas turbine (GT) engines
	C2U2: Aerospace - Active flow control of aircraft aerodynamics including synthetic jet actuators
	C2U3: Transport systems - Coupled simulation of fluid and structure mechanics for fatigue and fracture
AVBP (C3)	**C3U1**: Aerospace and energy - Combustion instabilities and emission prediction
	C3U2: Safety applications - Explosion in confined spaces
Fluidity (C4)	**C4U1**: Renewables - Tidal Energy Generation Modelling of drag and sedimentation effects of tidal turbines
FEniCS (C5)	**C5U1**: Aerospace and Automotive - Adjoint optimization in external aerodynamics shape optimization
FLUCS (C6)	**C6U1**: Aerospace - Design process and simulation of full equipped aeroplanes
	C6U2: Aerospace - CFD coupling with computational structural mechanics including elastic effects

Table 1: Reference applications and their application fields

2.1 The reference applications

The first reference application **Nek5000 (C1)** [1] is a fluid and heat-transfer solver based on the Spectral Element Method (SEM) with a long development history. In the mid-1990s it was the first available code for distributed memory computers, which received the Gordon Bell prize for algorithmic quality and sustained parallel performance in 1999. The good scaling properties are granted by use of SEM-based domain decomposition into a set of disjoint spectral subdomains, which allows to split global operators into a set of locally performed dense matrix-matrix multiplications combined with the communication step through direct-stiffness summation. This domain decomposition can be used as well to increase simulation reliability, as domain decomposition can be dynamically modified during the run to minimize estimated computational error. Within EXCELLERAT, KTH is going to work on Nek5000 development focusing on number of aspects corresponding to number of WP4 services e.g.: adaptive mesh refinement using adjoint algorithms (Intrinsic Optimization Methods), uncertainty quantification (Numerical Methods, Data reduction algorithms), use of accelerators (Porting to new Architectures, Node-level Performance Engineering) or post-processing data reduction (In-situ Visualization). It spans the whole simulation cycle starting from the pre-processing stage, where the coarse, hex-based mesh for relatively complex geometry have to be generated (Meshing Algorithms). At the simulation stage, we are going to concentrate on the pressure preconditioners for nonconforming meshes (Numerical Methods) and the communication kernels (System-level Performance Engineering).

To test and demonstrate new code abilities we are going to run two applications. The first one (**C1U1**) is a flow around three-dimensional wing tip, which is a relatively complex test case relevant for number of industrial users like aeronautics, automotive or bio engineering. It allows to study a number of physical phenomena responsible for e.g. generation of the drag on the airplane, which increases fuel consumption in aviation. Despite its environmental and economic importance, there is lack of high-fidelity numerical data for high-Reynolds-number turbulent flows. This case is a good platform for testing a number of different aspects of numerical modelling including mesh adaptivity, uncertainty quantification or data reduction.

It is a potential exascale test case as well. We aim for simulation of flow around NACA0012 wing tip and validation against existing literature data. The second application (**C1U2**) is the high-fidelity simulation of rotating helicopter blades, which adds the additional complexity of rotating parts compared to **C1U1**.

Alya (C2) [2] is a parallel multi-physics / multi-scale simulation code featuring hybrid parallelization and has been developed to run efficiently in supercomputing environments. Its parallelization includes distributed memory (MPI), shared memory (OpenMP, OmpSs [3]) and stream processing (OpenACC, CUDA). Alya is one of the two engineering codes (CFD) of the PRACE Benchmark Suite and has been tested showing excellent parallel performance on most European supercomputer platforms. Alya is developed by around 40 developers in BSC, who has joined the EXCELLERAT consortium as full member, and many users worldwide.

The first use-case **C2U1** deals with the prediction of pollutants e.g. NOx and soot using advanced numerical simulations with detailed chemical kinetics in combustion applications. The second use-case **C2U2** targets the investigation of active flow control for full aircraft aerodynamics, which is a very relevant topic for the development of the new ultra-high bypass ratio (UHBR) engines. The third use-case **C2U3** is focused on the study of damage monitoring and detection for civil aircraft composite fuselage using piezoelectric sensors as impacts from external objects on the fuselage structure may provoke a local damage of the material and reduce the safety of civil aircraft.

The three use-cases have been selected as they correspond to fundamental challenges in the aeronautical and aerospace sector and share common limitations from the technical point of view. In detail, the core development team will work together with the EXCELLERAT consortium on key challenges of large scale HPC applications - the development of a fully parallel workflow, adaptive mesh refinement and enable parallel multiphysics coupling at extreme scale. Additionally, further porting Alya to new architectures as well as node level performance engineering with a focus on efficient vectorization and the efficient usage of hierarchical memory systems, will be another key challenge on which Alya's core developers will work with EXCELLERAT partners.

AVBP (C3) [4, 5] is a 3D Navier-Stokes solver for compressible reactive multiphase flows specifically designed for leadership class High Performance Computing architectures. Based on domain decomposition (ParMETIS [6], PTSCOTCH [7], ZOLTAN [8]) and efficient utilisation of MPI, it has shown almost perfect scaling on most CPU based PRACE systems. It is the state-of-the-art application used in academia and industry for combustion simulations with over 300 users in Europe and the US. It is mainly developed by CERFACS, a full partner in the EXCELLERAT consortium, where we focus on two specific challenges in the combustion community, safety applications and accurate pollutant predictions.

The first use case **C3U1** focuses on using state of the art large eddy simulation for multi-scale combustion processes leading to explosions. The second use case **C3U2** singles out the prediction of pollutants in aerospace applications. These use-cases are the next challenges facing numerical combustion in the aerospace industry. They require improving user interactivity with the code by introducing efficient parallel mesh adaptation, efficient load balancing in an exascale framework as well as improving chemistry modelling.

Of course, core developments such as code porting and preparing the code for new architectures remain a high priority in EXCELLERAT. A specific effort is expected on current GPU port to devise efficient cost-effective GPU adaptation methods for legacy codes.

CoastED [9] is a branch of **Fluidity (C4)** [10] - an open source, general purpose, multiphase computational fluid dynamics code - optimised for coastal modelling. It solves the Navier-Stokes equation and accompanying field equations on arbitrary unstructured finite element meshes in one, two and three dimensions. It is parallelised using MPI and is capable of scaling to many thousands of processors. Although CoastED contains optimisations and extra functionality for coastal modelling, it is still capable of the full set of use cases of Fluidity. Thus, although the optimisations made during EXCELLERAT are targeted for coastal modelling, many of them will apply to general use cases of Fluidity.

The use-case for CoastED targets the multi-scale problem of modelling the tidal motion in and around a selected off-shore region, incorporating detailed models of tidal turbines to predict and analyse the impact of those turbines on sedimentation and the tidal flow. The challenge, which makes this case an exascale problem, is the ratio of the length scales that ranges from several tenth of centimetres in the region of the turbines to several hundreds of kilometres to cover the project area.

There is a large difference between turbine scale and large-scale tidal flow. Unstructured mesh models provide more flexibility to narrow that gap, but simplification and parameterisation of the effects of the turbines is still necessary at petascale computing. Existing studies have looked at parameterising the drag force due to turbines, and this can be modelled reasonably accurately. However, this requires a corrective factor to be applied as the resolution of the tidal flow is lowered in order to ensure the parameterisation remains realistic. The parameterisation of the turbulence still needs much work to achieve realistic wakes. Clearly a move to exascale will help reduce the reliance on such parametrisations (and any corrective factors) and thus more accurate models can be analysed. Unrealistic side effects introduced by the parameterisations can be minimised at the higher resolutions made possible at exascale.

To achieve these goals, we need to address three things. Firstly, we have identified optimisations to the matrix assembly data structures and routines which when implemented should obtain a significant speedup. Next, the main solver for the global momentum matrix needs to be considered, as this is where a large proportion of runtime is spent. Even though CoastED currently utilises the PETSc [11] library which is known to be designed for extreme scaling and high efficiency, the solver part of the code currently scales less well than other components. This prevents the given problem to be scaled to the needed problem size in order to make realistic predictions. We will either look at potential enhancements of the data structures and calling hierarchies to better make use of the extreme scale capabilities of PETSc or indeed use alternative solvers like e.g. geometric multigrid. Lastly, the parametrizations of the turbines need to be considered. The integration of the parametrized models into the larger tidal simulations needs to be examined in order to understand the best approach to take from the numerical as well as the technical point of view.

The use-case **C4U1** is a prototype for EXCELLERAT's service to deliver expertise and development support to a community code's user and developer in terms of multi scale modelling and the improvement of numerical methods, i.e. linear system solvers.

FEniCS (C5) [12] is a high-level problem-solving environment for automated solution of partial differential equations (PDEs) by the finite element method. To manage the complexity of multiphysics problems FEniCS takes the weak form of a PDE as input in a near mathematical notation and automatically generates low-level source code, abstracting away implementation details and HPC concepts from domain scientists. FEniCS is supported by a large loosely coupled community, providing various versions of its core components. But while several implementations exist, its broad user base is connected by the common code generation capabilities, allowing users to easily move between various flavours within the community. FEniCS is further developed to be an exascale ready prototyping framework within the use-case **C5U1**, covering a complete engineering workflow for external aerodynamics simulation in the automotive sector, including shape optimization for a given quantity of interested e.g. drag. This propose several challenges from scalable numerical methods, post-processing and transfer and handling of potentially sensitive CAD data. Such a use-case is the prototype application of the EXCELLERAT service portfolio in which a code developer searches for expertise and co-developers within the engineering community to bring a promising approach, in this case the domain specific language (DSL) approach, which is different to classical direct problem specific discretization approaches, towards exascale scalability, and extend its features in several directions.

Flucs (C6) [13] is a CFD solver for aircraft design. It features innovative algorithms as well as advanced software technologies and concepts dedicated to HPC. DLR started the implementation from scratch in 2013 considering requirements for exascale like multi-level-parallelization and asynchronous execution among others already in the initial design phase. This development targets the replacement of the working horse TAU, a large legacy code designed for pure MPI use, which would have been hard to extend for efficient use on future hardware architectures with all its functionalities. The Flucs software is now the basis of a recently initiated strong partnership between Airbus, ONERA and DLR focusing on the development of a common next generation CFD code for aircraft flow predictions (with the new name CODA since January 2019). As a partner of this development team DLR has joined the EXCELLERAT consortium as a full member, taking the role as code developer and application user in order to further improve the CFD code. This will be demonstrated by two use cases, both targeting strong scaling scenarios demonstrating the parallel efficiency for simulations for a full aircraft. **C6U1** will demonstrate the CFD solver performance only, while **C6U2** deals with coupled multi-disciplinary simulations with fluid-structure interaction. Both use cases will allow for the development and demonstration of EXCELLERAT's capability to provide expertise and extended knowledge to engineering application developers (DLR) in terms of efficient performance engineering methodologies on node as well as system level. By this, the application developer team (DLR/ONERA/Airbus) will be supported to integrate and develop a highly efficient next generation CFD-code capable of executing simulation tasks on exascale level (Complete Aircraft Aerodynamics).

3 The service catalogue of EXCELLERAT

EXCELLERAT's service catalogue is based on a) the already existing competence of HPC and domain experts in the consortium and b) identified requirement from engineering applications. The envisioned service portfolio, evolving during the project runtime includes:

- Expertise consulting: Develop expertise and provide this information to possible stakeholders.
- Networking consulting: Support access to industry funded research opportunities
- Solution consulting: Take problems of stakeholders and solve them, communicate the results of simulations and its reliability including the visualization.
- Optimization of applications (codes) for the specific realm of engineering, (in synergy with the Centre of Excellence for Performance Optimization [14])
- Evolution of applications (codes), design of applications (codes), maintenance of applications (codes), testing and validation of applications (codes), porting of applications (codes), benchmarking of applications.

- Development of dedicated tools (I): Data analytics (also within simulations), visual analytics and data management for large data sets.
- Development of dedicated tools (II): Provisioning of visualization methods (e.g. plugins/ filters) for specific engineering work flows and supply of repositories.
- Promoting outstanding applications (runs): Exascale demonstration runs as well as high capacity runs.
- Training: Workshops and courses (for the user community but not only) about new methodologies and approaches
- Hosting of large data sets as well as data handling expertise.
- Co-Design SW / SW and HW / SW: Application optimization and consulting.
- Knowledge / conference hubs: Networking events and conferences, establishing knowledge hubs online with best practice guidelines.
- Unique and complete data base for all kinds of simulations, results and KPI of industrial impact
- Solution repository: Open and commercial solutions
- Career portal for HPC applications for engineering experts
- General support service (e.g. trouble ticket, etc.)

As a starting point, an initial service catalogue is defined and mapped to the five ad-hoc tasks of WP4:

- Co-design
- Visualisation
- Data management
- Data analytics
- Usability

Using the core competences of the Centre together with the identified challenges of the reference applications, these five tasks will form a set of core services from which the service catalogue will evolve.

3.1 Core services

3.1.1 Co-design

In this section a number of services are described which could be offered by the CoE related to co-design and which could be part of a service catalogue that will generate revenue to sustain the CoE. The co-design service catalogue lead by UEDIN, will develop over time, and not all services described below will be necessarily part of it. There is a dependency on the business model chosen by the project, based on the work of WP6.

3.1.1.1 Periodic publications summarising future plans of hardware developers

The CoE could publish regular technical reports (whitepapers) which would summarise current and future developments in both hardware and software. The information would be gained from the network of contacts that the CoE has with hardware companies and with software developers. Analysis of the state of the art could be included in these reports, with a view to facilitating a two-way flow of information between the hardware and software communities and enabling co-design across these communities. Any information presented would have to be already publicly available or could be made public with the agreement of the relevant parties.

This service could be offered free, so as to build early interest in the CoE. Interested parties from outside the consortium could sign up to receive whitepapers. This could be extended to a premium model in which paying subscribers receive added-value material. This depends on the business model chosen by the project.

The value proposition in this case is to provide access to information in one place and expert analysis of that information.

3.1.1.2 Workshops

The service of providing whitepapers could be complemented by workshops organised by the CoE. These could be organised in conjunction with hardware vendors.

The business model could be either to provide these free to generate interest, or to charge a nominal fee to cover the costs of hosting the workshop, or a premium fee generating a surplus. This decision depends on the overall business model and plan from WP6.

The value proposition is similar to that in the previous section, with the added value of offering an opportunity to have direct contact with experts.

3.1.1.3 Consultancy to develop versions of code exploiting new architectures

A service could be offered to work directly on applications or software libraries brought in by external parties to exploit new technologies, possibly with collaboration from hardware vendors. This could be trialled on the CoE application codes, although the full business model cannot be trialled as the financial model of the project does not allow it. The project could be used as an incubator for this service by WP4 working in a service-oriented way with WP2 using selected codes (all of which drive this need) although as stated there can be no flow of money between the partners. However, it may be possible to estimate the costs of providing a service which could help formulate a business plan for offering the service externally. Ownership of IP would have to be in accordance with the business model selected by the project – it is anticipated that code owners would not wish to share IP that they have paid for with consultancy.

The value proposition here is to enhance existing assets (i.e. application codes) for the customer through expertise not available to them in-house.

3.1.1.4 Optimised computational kernels for different hardware

It is a potential revenue raising activity and therefore could be considered by the CoE as part of the business model and is included in the document.

In this scenario the CoE would generate code, for which it would own the IP, specifically targeted at new hardware architectures. The definition of ownership will depend on the legal status of the CoE. In addition, there would be an expectation for maintenance and support. A possible option here would be to develop IP jointly with a vendor. The vendor might then carry out the support function, with the CoE carrying out fixes and enhancements and both sharing the income from licences. Again, the precise model depends on the business model chosen by the project.

The value proposition here is the availability of high-performance software through a licensing model (for example sale of licenses or SaaS).

3.1.2 Visualization

A crucial part in understanding simulation results lies in their visualization. The traditional way these visualizations are done is to apply filters to the simulation data and map the results on a (3D) geometry. In the context of exascale simulation data reaches a new level in size, but whereas computing power has risen tremendously through parallelization, memory size has not. So, for exascale simulations post-processing simulation data will exceed the capabilities of single workstations or even small clusters.

Therefore, HLRS is developing the VISualization Testing Laboratory for Exascale computing Vistle [15] since 2012 within the European project CRESTA [16] and bwVisu [17]. Vistle is an extensible software environment that integrates simulations on supercomputers, post-processing and parallel interactive visualization. It is under active development at HLRS since 2012 within the European project CRESTA and bwVisu. This modular data-parallel visualization system provides highly scalable algorithms for exploiting data, task and pipeline parallelism in hybrid shared and distributed memory environments with acceleration hardware. Domain decompositions used during simulation can be reused for visualization. For rendering, Vistle uses OpenCOVER [18], a rendering engine also developed at HLRS, that is able to process user interactions and can handle immersive and non-immersive presentation.

Even though this system allows post-processing of huge datasets, handling this amount of data can be very expensive, especially in regard to Exascale simulations. Therefore, the system will be expanded by a user-friendly interface for in-Situ data-filtering and visualization. This will not only reduce necessary disk space but also enables specialists to detect simulation errors early and react accordingly.

Since the data-filtering methods will be integrated in the simulation loop, they must be as efficient as possible. Furthermore, on-line interaction should make it possible to restrict the filtering on features and fields the viewer is truly interested in. In order to support decision making processes, a VR interface to data analytics processes developed in Task 4.3 will be implemented in OpenCOVER. Due to the insights that can be gathered during simulations the next step will be to integrate methods for steering the running simulation into the in-situ interface.

These measures will be applied to the simulation codes Nek5000 (C1U1, C1U3) and Alya (C2U1, C2U2) to get practically oriented use-cases for testing and evaluating the in-situ visualization system.

3.1.3 Data analytics

Big Data Analysis is a general name for large scale data acquisition and its exploitation, where in engineering the focus is on sensor and simulation data from technical applications. It is generally believed that today's simulation capabilities provide unique opportunities for product development, thereby allowing a paradigm shift in the way products are designed and developed in engineering applications.

The activities within EXCELLERAT will focus on the specific problems connected to the analysis of exascale simulation data. The value of such data lies only to a small extent in accurate integral values such as forces, or heat loads, it will be of much higher importance for the engineer to understand complex physical mechanisms, sensitivities of the performance with respect to design parameters and uncertainties connected to the simulation. In EXCELLERAT, in-situ post-processing and visualization techniques will be established as a standard monitoring mechanism such that results are already available during the runtime of a simulation avoiding large amounts of disk I/O.

In most cases an engineer, who wants to evaluate more than standard quantities from the simulation data, will employ the ansatz to store the large-scale data to disk with a subsequent data processing by application specific software. Such a procedure will not be feasible any more for real exascale applications due to the huge data volumes. Only reduced data with possibly a considerable loss of information, e.g., due to coarsening or lossy compression can be stored for post-processing.

Of special importance will be the development of new and adaptation of existing advanced analysis tools, which are indispensable for an efficient work flow delivering the relevant information from a simulation to the engineer, as well as machine learning (ML) approaches, as an alternative to long existing response surface methods, which are nowadays based on surrogate models for numerical simulation results.

The work in EXCELLERAT will generate several services that aim at both industrial and academic applications. The requirements as stated above will be addressed in the following five services.

3.1.3.1 Plugin for in-situ data analysis

A common requirement for several core codes of EXCELLERAT is the reduction of massive data I/O and post-processing efforts during large-scale simulations. An efficient approach for doing so is the analysis of the data while it is still in memory. This, so called, in-situ analysis has two major advantages compared to classical post-processing. First, it provides useful insights into the simulation and can indicate important trends already during run time. A comparison with previous simulations can then identify sensitivities based on influential parameters such as numerical settings, mesh resolution or boundary conditions. The second advantage of an in-situ analysis is the availability of the data in a higher spatial and temporal resolution than is usually stored for post-processing. This allows the computation of sensitive quantities that can only be derived from the high-resolution flow field based on physical definitions or purely data-driven approaches. While the calculation of physical quantities requires distinct expert knowledge about the physical phenomena before the simulation is carried out, this is not the case for advanced data-driven methods. The latter have the potential to capture the most import features in the solution, even if they were not expected in advance. Thus, data-driven approaches are more appropriate for the analysis of simulation data when no a priori knowledge about the expected results is available. By storing the derived features and results of the analysis instead of entire snapshots, data I/O can be reduced considerably.

In EXCELLERAT, we will concentrate on the in-situ calculation of global and data–driven features. With the help of advanced methods for non-linear dimensional reduction, we can derive compact representations of the data that contain essential information and physical effects. Besides, no a priori knowledge of the solution is necessary, because new occurring effects can be tracked automatically if appropriate thresholds are available. By investigating the low dimensional representation of the data, one can identify trends more easily and find correlations with input parameters. Finally, the low dimensional representation can be used to compress the simulation data and store it for further investigations.

EXCELLERAT will rely on existing in-situ libraries that are developed in the context of visualization and extend them with additional data analytics - functionality, e.g. LibSim [19], Catalyst [20] or Vistle. Providing plugins to widely-used software solutions will increase the acceptance and usability of new analysis methods. With USTUTT, the consortium has a partner with many years of experience in the development of visualization tools. Therefore, a seamless integration of in-situ data analysis with the visualization workflow can be offered.

3.1.3.2 Dedicated tools for meta modelling and comparative analysis of simulation bundles

In the engineering design process, there are numerous needs and applications for accurate and cheap-to-evaluate meta models. Not only are they useful for an efficient design optimization and sensitivity analysis, but also applicable to active flow control, the integration of experimental and simulation data or sparse sampling to identity the states of complex dynamic systems.

The generation of meta models is based on the assumption that the intrinsic dimension of the underlying high dimensional simulation (or experimental) data is much lower than the formal degree of freedom, i.e. the number of grid points. Recent advances in non-linear machine learning methods have proven very successful in identifying low dimensional manifolds in simulation data. This low dimensional representation of the dynamic system can be interpolated at many combinations of input parameters and provides a cheap but accurate prediction of the solution. As an example, in the case of design optimization, a few initial, high resolution simulations are performed that are used to generate a meta model. Afterwards, the model can be used for an efficient optimization of a given cost function and provide a set of optimized input parameters. The results can be confirmed by only a few expensive high-resolution simulations.

In the field of structural mechanics, FRAUNHOFER has developed two dedicated tools, Dataviewer and DesParO [21], that provide methods for non-linear dimensional reduction, meta modelling and sensitivity analysis. They have been applied successfully to finite element data, such as generated in car crash simulations. The software tools will be used as a starting point in EXCELLERAT and extended for the specific needs that are imposed by the use-case data, namely in the field of fluid dynamics.

A demonstration of the derived methods within EXCELLERAT will be carried out on the use-case data from FEniCS. The optimization of a geometrical shape to minimize aerodynamic drag can be regarded as a representative engineering problem. It imposes several challenges, namely a complex multi-dimensional optimization and multi-scale turbulent flow data. Another potential use-case to benchmark results is the active flow control for fully aircraft aerodynamics driven by EXCELLERAT's core code Alya.

3.1.3.3 Dedicated tools for the calculation of advanced flow features

With the increasing availability of computational resources, high-resolution numerical simulations have become an indispensable tool in fundamental academic research as well as engineering product design. Based on large eddy or direct numerical simulations, turbulent flows including several orders of magnitude of length and time scales can be predicted with high fidelity, producing extremely large data sets, when results have to be stored on disk. The value of such data lies only to a small extent in accurate integral values such as aerodynamic forces or heat loads, it will be of much higher importance for the researcher to understand complex physical mechanisms hidden in the simulation results. In research, highly specialized and advanced analysis tools are developed, but their application is often non-trivial and may require complicated set up procedures and specialized data formats such that they are only accessible to a few experts and consequently not widely employed. Additionally, the calculation of these features becomes computational expensive for large data sets, which imposes the need for an efficient parallelization of the algorithms.

In turbulent research, dynamic mode decompositions (DMD) and proper orthogonal decompositions (POD) provide sophisticated approaches to investigate complex flow behaviour. Based on the competence of RWTH Aachen University in this field, EXCELLERAT will further develop existing software tools for the application of DMD and POD tools with a focus on parallel execution and a tight integration into the visualization tools developed by HLRS. Furthermore, they will be applied to selected use-case data from EXCELLERAT to improve their general applicability. By this means, the access to advanced algorithms for DMD and POD will be simplified for external partners and industry.

3.1.3.4 Development of ML-methods for the modelling of advanced physics (potential service)

An additional requirement for data analysis was identified by the core codes Alya and AVBP. In an emerging research field, machine learning methods were applied successfully to the modelling of physical quantities such as turbulence, chemical reactions and the evolution of boundary layers. In classical solvers, these quantities are either approximated by simplified physical-based models or explicitly resolved on highly refined meshes. However, physical models are derived from simplified assumptions that are only valid in a limited range and thus they can become quite inaccurate. On the other hand, the direct and accurate calculation of the physics on all length scales requires very large meshes which results in high computational costs.

To overcome these shortcomings, the application of machine learning models is a promising approach. Thereby, the training of the models can be performed on highly resolved simulations for a larger range of boundary conditions. If successfully trained, the derived model will capture the intrinsic behaviour of the underlying physics and can provide accurate prediction even for conditions that are not contained in the training data. This way, one can expect to derive models that are quite accurate for more general conditions then the simplified physical based models – eventually reducing computational costs.

As research in this field is still at an early stage, the outcome of this service will concentrate on the communication of ongoing research results. As the machine learning models have to be integrated directly with the numerical solver, the development will be carried out within Alya and AVBP. Nevertheless, results can potentially be transferred to other applications at a later point in time.

3.1.3.5 Developing training modules for data analytics for simulations (Output of WP5)

Besides the development of technologies towards exascale readiness, a fundamental goal of EXCELLERAT is the setup of a central access point for technologies and related knowledge. Therefore, typical workflows for analysis of engineering data will be summarized in notebooks that can build the basis for educational services. Trainings will be designed in a modular concept covering basic and advanced data analytics topics in order to satisfy the specific needs of industrial end-users. These training modules will be provided as open or in-house training courses by the CoE.

3.1.4 Data management

SSC-Services GmbH is mainly working in "WP3.6 Data dispatching through data transfer" and "WP4.4 Data Management". Therefore, we want to combine data transfer and data management.

The overriding goal (*vision*) is to provide an all-in-one platform (*portal solution*) on which the data, that needs to be calculated, is uploaded and the rest is done automatically in the background. The great added value (*service*) is that e. g. manual command line entries can be completely dispensed with; result data could be visualized, and data could be transmitted in encrypted form. It should be a platform with which all HPC systems can be addressed. At any time, there should be traceability of what happens to the data or where the data is located. As soon as it is clear what kind of data will be transmitted, by whom the data will be transmitted and which data will be calculated where, it is possible to advance the degree of automation rapidly. The figure below gives an initial overview of how the infrastructure could look like.

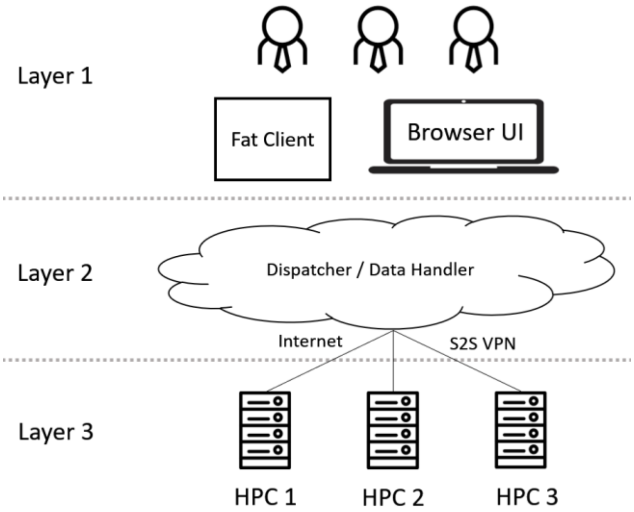

Figure 1: Overview of the data management infrastructure

The complete service is distributed over three layers: clients (*layer 1*), neutral data handling zone (*layer 2*) and the HPC systems (l*ayer 3*). The first layer concerns the users and the available interfaces they can use. There are two interfaces: a web interface (UI) and a fat client. The web interface can be used for handling smaller amounts of data and the fat client can perform more complex actions like delta comparison or data compression. With both interfaces it is possible to upload data, configure jobs and download or view the result data. In the second layer, a central distribution mechanism would be installed (*data handler* or *dispatcher*), which would also be responsible for data management. This layer can either run in one of the HPC systems or in the cloud (e.g. Google, Amazon), where a very high data throughput is possible. The data would not be stored and only passed to the appropriate HPC system. The transition between the second and the third layer could be controlled either directly via the Internet or via a Site-2-Site VPN. The lowest layer is formed by the individual HPC systems on which the code is executed. Each HPC system needs to run a small application, which can communicate with the second layer.

Furthermore, the formation of data deltas could contribute to a good data management system. For example, it would be conceivable to divide a file into several blocks, each with a unique checksum. If only a single block in the file has changed, only this block would be transferred, since the remaining blocks are unchanged. Before uploading the data, the decomposition mechanism and the checksum comparison must take place. This means that at the end there is only a huge collection of checksums left, which are reduced to a minimum. The prerequisite is that before a user uploads data, it must be known by the *data handler* which data is available everywhere.

Further services could be the encryption or the visualization of the data. With end-to-end encryption, each block would be encrypted during the upload and then being transferred. The user could use remote viewing to display the result data via the web interface, e.g. video, images or thumbnails. The vision of an *all-in-one* portal solution described above is intended to have maximum added value for all users.

3.1.5 Usability

The last of the core services evolving from WP4 is the usability service lead by KTH, offering support and best practices for an engineering simulation's entire life cycle, from pre-processing, including modelling and meshing, execution of simulations to post-processing with scalable visualisation and proper data management. Acting more like an umbrella service, it will gather the expertise and methods developed in the CoE to prototype and develop efficient and general simulation workflows.

3.1.5.1 Periodic publication of simulation best practices guides

To complement the regular planned deliverables (with updates on the services catalogue) the CoE could additionally publish best practices guides as whitepapers on a regular basis. These whitepapers will provide both information on tools (scripts, pipelines etc) and guidelines on how to efficiently perform a large-scale engineering simulation. These workflows should cover the entire life cycle of a simulation and should provide solutions built on established tools (possibly enhanced by EXCELLERAT) for an easier uptake in already established simulation workflows.

Similar to the co-design service, this service could be offered as a free complement to the public deliverables to build interest in the CoE, or as a subscription model, where paying subscribers receive more a detailed in-depth description of the workflows, or possibly access to developed tools, scripts and visualisation pipelines from the CoE's portal. This depends on the business model chosen by the project.

3.1.5.2 Workshop and training activities

The service of providing best practices guides could be complemented by organising workshops and training on how to apply formulated simulation workflows. Depending on the business model, these sessions could be free (or at a low fee to cover hosting costs), or provided for a premium fee, allowing for more in-depth custom sessions, possibly tailored towards a particular application.

3.2 Evolution of the catalogue

In the initial phase of the project the catalogue will be very focused on the reference applications from WP2 and the core services related to WP4's tasks. The creation of the core services will be based upon the interaction between WP2 and WP4, in terms of the challenges/requirements identified in the reference application. This approach allows us to create new services from a more or less code agnostic viewpoint, allowing us to easy create a sustainable service portfolio for future applications.

Once an initial set of core services has been established, the work to develop more common and marketable services from the envisioned catalogue will commence. This evolution of the portfolio will be tracked through a set of periodical updates, released as a set of deliverables during the project's lifetime.

4 Conclusion

This deliverable puts into context all relevant application fields and competences covered by the Centre and its offered services. The key competences of the EXCELLERAT consortium are defined together with a presentation of the reference applications, their application fields and their respective challenges moving towards exascale. An initial service catalogue is defined for the Centre, with an outlined vision for each service together with links to one or several of the reference applications used to drive the evolution of the service portfolio. Finally, an outlook is given on the evolution of the catalogue.

5 References

[1] P. F. Fischer, J. W. Lottes and S. G. Kerkemeier, “Nek5000 Web page,” 2008. https://nek5000.mcs.anl.gov/.

[2] Alya. http://bsccase02.bsc.es/alya/.

[3] OmpSs. https://pm.bsc.es/ompss.

[4] T. Schønfeld and M. Rudgyard, “Steady and Unsteady Flow Simulations Using the Hybrid Flow Solver AVBP,” *AIAA Journal,* vol. 37, no. 11, 1999.

[5] AVBP. http://www.cerfacs.fr/avbp7x/ .

[6] K. Schloegel, K. George and V. Kumar, “Parallel Multilevel Algorithm for Multi-Constraint Graph Partitioning,” in *Euro-Par,* Munich, 2000.

[7] C. Chevalier and F.Pellegrini, “PT-SCOTCH: A tool for efficient parallel graph ordering,” in *PMAA*, Rennes, 2006.

[8] E. Boman, U. Catalyurek, C. Chevalier and K. Devine, “The Zoltan and Isorropia Parallel Toolkits for Combinatorial Scientific Computing: Partitioning, Ordering, and Coloring,” *Scientific Programming,* vol. 20, no. 2, 2012.

[9] CostED. https://github.com/CoastED/coasted.

[10] Fluidity. http://fluidityproject.github.io.

[11] S. Balay, S. Abhyankar, M. F. Adams, J. Brown, P. Brune, K. Buschelman, Lis, r. Dalcin, A. Dener, V. Eijkhout, W. D. Gropp, D. Karpeyev, D. Kaushik, M. G. Knepley, D. A. May, L. C. McInnes and R, “PETSc Web page,” 2019. [Online]. Available: http://www.mcs.anl.gov/petsc.

[12] M. S. Alnaes, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. E. Rognes and G. N. Wells, “The FEniCS Project Version 1.5,” *Archive of Numerical Software*, vol. 3, 2015.

[13] T. Leicht, D. Vollmer, J. Jägersküpper, A. Schwöppe, R. Hartmann, J. Fiedler and T. Schlauch, “DLR-Project DIGITAL-X - Next Generation CFD Solver FLUCS,” in *DGLR 2016*, Braunschweig, 2016.

[14] The Performance Optimisation and Productivity Centre of Excellence in Computing Applications, https://pop-coe.eu/.

[15] Vistle. https://www.hlrs.de/vistle.

[16] CRESTA. https://www.hlrs.de/about-us/research/past-projects/cresta.

[17] bwVisu. https://urz.uni-heidelberg.de/de/bwvisu.

[18] OpenCOVER. https://www.hlrs.de/solutions-services/service-portfolio/visualization/covise/opencover.

[19] Lawrence Livermore National Laboratory, “VisIt,” https://visit.llnl.gov/.

[20] Kitware, Inc., “ParaView,” https://www.paraview.org/.

[21] SCAI, Fraunhofer, “DesParO,” https://www.scai.fraunhofer.de/en/business-research-areas/numerical-data-driven-prediction/products/desparo.html.