

HPC-Cloud-Big Data Convergent Architectures and Research Data Management: The LEXIS Approach

Thursday, 25 March 2021 16:00 (30 minutes)

The LEXIS project (Large-scale EXecution for Industry & Society, H2020 GA825532) provides a platform for optimized execution of Cloud-HPC workflows, reducing computation time and increasing energy efficiency. The system will rely on advanced, distributed orchestration solutions (Bull Ystia Orchestrator, based on TOSCA and Alien4Cloud technologies), the High-End Application Execution Middleware HEAppE, and new hardware capabilities for maximizing efficiency in data processing, analysis and transfer (e.g. Burst Buffers with GPU- and FPGA-based data reprocessing).

LEXIS handles computation tasks and data from three Pilots, based on representative and demanding HPC/Cloud-Computing use cases in Industry (SMEs) and Science: i) Simulations of complex turbomachinery and gearbox systems in Aeronautics, ii) Tsunami simulations and earthquake loss assessments which are time-constrained to enable immediate warnings and to support well-informed decisions, and iii) Weather and Climate simulations where massive amounts of in-situ data are assimilated to improve forecasts. A user-friendly LEXIS web portal, as a unique entry point, will provide access to data as well as workflow-handling and remote visualization functionality.

As part of its back-end, LEXIS builds an elaborate system for the handling of input, intermediate and result data. At its core, a Distributed Data Infrastructure (DDI) ensures the availability of LEXIS data at all participating HPC sites, which will be federated with a common LEXIS AAI (with unified security model, user database and authorization policy). The DDI leverages best of breed data-management solutions from EUDAT, such as B2SAFE (based on iRODS) and B2HANDLE. REST APIs on top of it will ensure a smooth interaction with LEXIS workflows and the orchestration layer. Last, but not least, the DDI will provide functionalities for Research Data Management following the FAIR principles ("Findable, Interoperable, Accessible, Reusable"), e.g. DOI acquisition, which helps to publish and disseminate open data products.

Summary

This talk presents the LEXIS ("Large-Scale Execution for Industry and Society") platform for running orchestrated compute-intensive workflows in Science, Engineering, Industry, Governmental Applications and beyond. These workflows handle Big Data with a hybrid, geographically distributed Cloud/HPC computing environment and an EUDAT/iRODS-based Distributed Data Infrastructure (DDI). Particular emphasis in the presentation is put on the DDI architecture and its integration with European (EUDAT/EOSC) research data management. Currently, the DDI connects two large Supercomputing Centres (IT4I/CZ and LRZ/DE), with additional significant European computing and data sites being added to the data federation.

Primary author: Dr HACHINGER, Stephan (Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities, Garching, Germany)

Co-authors: Dr SCIONTI, Alberto (Advanced Computing and Applications, LINKS Foundation, Torino, Italy); Dr PARODI, Antonio (CIMA Research Foundation, Savona, Italy); Mr D'AMICO, Carmine (Advanced Computing and Applications, LINKS Foundation, Torino, Italy); Mr KOCH-HOFER, Cédric (ATOS, Paris, France); Prof. SCHORLEMMER, Danijel (GFZ German Research Centre for Geosciences, Potsdam, Germany); Dr MAGARIELLI, Donato (AvioAero, Torino, Italy); Mr APOPEI, Florin Ionut (TESEO, Torino, Italy); Dr MARTINOVIČ, Jan (IT4Innovations, VŠB - Technical University of Ostrava, Ostrava-Poruba, Czech Republic); Dr LEVRIER, Marc (ATOS, Paris, France); Mr GOLASOWSKI, Martin (IT4Innovations, VŠB - Technical University of Ostrava, Ostrava-Poruba, Czech Republic); Mr HAYEK, Mohamad (Leibniz Supercomputing Centre (LRZ), Bavarian Academy of Sciences and Humanities, Garching, Germany); Dr TERZO, Olivier (Advanced Computing and Applications, LINKS Foundation, Torino, Italy); Dr HARSH, Piyush (Cyclops Labs, Zurich, Switzerland); Dr MURPHY, Seán (School of Engineering, Zürcher Hochschule für Angewandte Wissenschaften (ZHAW), Zurich, Switzerland); Dr CICCIA, Simone (Advanced Computing and Applications, LINKS Foundation, Torino, Italy); Dr GOUBIER, Thierry (CEA, LIST, Paris, France)

Presenter: Dr HACHINGER, Stephan (Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities, Garching, Germany)

Session Classification: Converging High Performance infrastructures: Supercomputers, clouds, accelerators Session

Track Classification: Converging High Performance infrastructures: Supercomputers, clouds, accelerators