Activity Report 2023

Team KERDATA

Scalable Storage for Clouds and Beyond

Joint team with Centre Inria de l'Université de Rennes

D1 – Large Scale Systems

Contents

Project-Team KERDATA

Creation of the Project-Team: 2012 July 01

Keywords

Computer sciences and digital sciences

- A1.1.1. – Multicore, Manycore
- A1.1.4. – High performance computing
- A1.1.5. – Exascale
- A1.1.9. – Fault tolerant systems
- A1.3. – Distributed Systems
- A1.3.5. – Cloud
- A1.3.6. – Fog, Edge
- A2.6.2. – Middleware
- A3.1.2. – Data management, quering and storage
- A3.1.3. – Distributed data
- A3.1.8. – Big data (production, storage, transfer)
- A6.2.7. – High performance computing
- A6.3. – Computation-data interaction
- A7.1.1. – Distributed algorithms
- A9.2. – Machine learning
- A9.7. – AI algorithmics

Other research topics and application domains

- B3.2. – Climate and meteorology
- B3.3.1. – Earth and subsoil
- B8.2. – Connected city
- B9.5.6. – Data science
- B9.8. – Reproducibility
- B9.11.1. – Environmental risks

1 Team members, visitors, external collaborators

Research Scientists

- Gabriel Antoniu [Team leader, INRIA, Senior Researcher, HDR]
- Silvina Caino Lores [INRIA, ISFP, from Oct 2023]
- Jakob Luettgau [INRIA, Starting Research Position, from Oct 2023]
- Guillaume Pallez [INRIA, Researcher, HDR]
- François Tessier [INRIA, ISFP]

Faculty Member

- Alexandru Costan [INSA RENNES, Associate Professor, HDR]

PhD Students

- Robin Boezennec [INRIA, from Jun 2023]
- Thomas Bouvier [INRIA]
- Julien Monniot [INRIA]
- Cedric Prigent [INRIA]
- Daniel Rosendo [INRIA, until Mar 2023]
- Mathis Valli [INRIA, from Jun 2023]

Technical Staff

- Joshua Bowden [INRIA, Engineer]
- Daniel Rosendo [INRIA, Engineer, from Apr 2023 until Nov 2023]

Interns and Apprentices

- Malvin Chevallier [INRIA, Intern, from Jun 2023 until Aug 2023]

Administrative Assistant

- Laurence Dinh [INRIA]

2 Overall objectives

Context: the need for scalable data management. For several years now we have been witnessing a rapidly increasing number of application areas generating and processing very large volumes of data on a regular basis. Such applications, called *data-intensive*, range from traditional large-scale simulation-based scientific domains such as climate modeling, cosmology, and bioinformatics to more recent industrial applications triggered by the Big Data phenomenon: governmental and commercial data analytics, financial transaction analytics, etc. Recently, the data-intensive application spectrum further broadened by the emergence of IoT applications that need to process data coming from large numbers of distributed sensors.

Our objective. The KerData project-team is focusing on designing innovative architectures and systems for *scalable data storage and processing*. We target three types of infrastructures: *pre-Exascale high-performance supercomputers*, *cloud-based* and *edge-based* infrastructures, according to the current needs and requirements of data-intensive applications. In addition, as emphasized by the latest Strategic Research Agenda of ETP4HPC [7], new complex applications have started to emerge: they combine simulation, analytics and learning and require hybrid execution infrastructures combining supercomputers, cloud-based and edge-based systems. Our most recent research aims to address the data-related requirements (storage, processing) for such complex workflows. They are structured in three research axes summarized below.

Challenges and goals related to the HPC-Big Data convergence. Traditionally, HPC and Big Data analytics have evolved separately, using different approaches for data storage and processing as well as for leveraging their respective underlying infrastructures. The KerData team has been tackling the convergence challenge from a data storage and processing perspective, trying to provide answers to questions like: what common storage abstractions and data management techniques could fuel storage convergence, to support seamless execution of hybrid simulation/analytics workflows on potentially hybrid supercomputer/cloud infrastructures? From a broader perspective, additional challenges are posed by the question: how does the emergence of the computing continuum impact the data storage and processing infrastructure on HPC systems? The team's activities in this area are grouped in Research Axis 1 (see 3.1).

Challenges and goals related to cloud-based and edge-based storage and processing. The growth of the Internet of Things is resulting in an explosion of data volumes at the edge of the Internet. To reduce costs incurred due to data movement and centralized cloud-based processing, cloud workflows have evolved from single-datacenter deployment towards multiple-datacenter deployments, and further from cloud deployments towards distributed, edge-based infrastructures.

This allows applications to distribute analytics while preserving low latency, high availability, and privacy. Jointly exploiting edge and cloud computing capabilities for stream-based processing leads however to multiple challenges.

In particular, understanding the dependencies between the application workflows to best leverage the underlying infrastructure is crucial for the end-to-end performance. We are currently missing models enabling this adequate mapping of distributed analytics pipelines on the Edge-to-Cloud Continuum. The community needs tools that can facilitate the modeling of this complexity and can integrate the various components involved. In particular, the need for such tools is increasing when considering AI-enabled data analytics pipelines (e.g., based on Federated Learning or Continual Learning). This is the challenge we address in Research Axis 2 (described in 3.2).

Challenges and goals related to storage and I/O for data-intensive HPC applications. Key research fields such as climate modeling, solid Earth sciences and astrophysics rely on very large-scale simulations running on post-Petascale supercomputers. Such applications exhibit requirements clearly identified by international panels of experts like IESP [43], EESI [41], ETP4HPC [42]. A jump of one order of magnitude in the size of numerical simulations is required to address some of the fundamental questions in several communities in this context. In particular, the lack of data-intensive infrastructures and methodologies to analyze the huge results of such simulations is a major limiting factor. The high-level challenge we have been addressing in Research Axis 3 (see 3.3) is to find scalable ways to store, visualize and analyze massive outputs of data during and after the simulations through asynchronous I/O and *in-situ processing*.

Approach, methodology, platforms. KerData's global approach consists in studying, designing, implementing and evaluating distributed algorithms and software architectures for scalable data storage and I/O management for efficient, large-scale data processing. We target three main execution infrastructures: edge and cloud platforms and pre-Exascale HPC supercomputers.

The highly experimental nature of our research validation methodology should be emphasized. To validate our proposed algorithms and architectures, we build software prototypes, then validate them at large scale on real testbeds and experimental platforms.

We strongly rely on the Grid'5000 platform. Moreover, thanks to our projects and partnerships, we have access to reference software and physical infrastructures.

In the cloud area, we use the Microsoft Azure and Amazon cloud platforms, as well as the Chameleon [38] experimental cloud testbed. In the post-Petascale HPC area, we are running our experiments on systems including some top-ranked supercomputers, such as Titan, Jaguar, Kraken, Theta, Pangea and Hawk. This provides us with excellent opportunities to validate our results on advanced realistic platforms.

Collaboration strategy. Our collaboration portfolio includes international teams that are active in the areas of data management for edge, clouds and HPC systems, both in Academia and Industry. Our academic collaborating partners include Argonne National Laboratory, University of Illinois at Urbana-Champaign, Universidad Politécnica de Madrid, Barcelona Supercomputing Center. In industry, through bilateral or multilateral projects, we have been collaborating with Microsoft, IBM, Total, Huawei, ATOS/Eviden.

Moreover, the consortia of our collaborative projects include application partners in multiple application domains from the areas of climate modeling, precision agriculture, earth sciences, smart cities or botanical science. This multidisciplinary approach is an additional asset, which enables us to take into account application requirements in the early design phase of our proposed approaches to data storage and processing, and to validate those solutions with real applications and real users.

Alignment with Inria's scientific strategy. Data-intensive applications exhibit several common requirements with respect to the need for data storage and I/O processing. We focus on some core challenges related to data management, resulting from these requirements. Our choice is perfectly in line with Inria's strategic objectives [44], which acknowledges in particular HPC-Big Data convergence as one of the Top 3 priorities of our institute.

In addition, we have engaged in collaborative projects with some of Inria's main strategic partners: DFKI (the main German research center in artificial intelligence) though the ENGAGE Inria-DFKI project started in 2022; and with ATOS, through the ACROSS and EUPEX H2020 EuroHPC projects, started in Merch 2021 and January 2022, respectively. Gabriel Antoniu, Head of the KerData team, serves as a scientific lead for Inria in these three projects. The ENGAGE project is carried out in collaboration with the DataMove and HiePACS teams, while the EUPEX project also involves the TADaaM and HiePACS teams.

3 Research program

The scientific landscape in the areas of High-Performance Computing and Cloud Computing has changed significantly over the last few years. Two evolutions strongly impacted this landscape.

First, while High-Performance Computing and Big Data analytics had already started their convergence movement before 2015, this phenomenon was further enforced by the increased usage of machine learning for data analytics. This led to a triple convergence in the end: HPC, Big Data ***and AI*** (where the term "AI" is actually mainly referring in practice to machine learning). This convergence was driven by the emergence of new, complex application workflows. Modern use cases such as autonomous vehicles, digital twins, smart buildings and precision agriculture, are contexts where such application workflows are useful. They typically combine physics-based simulations, analysis of large data volumes and machine learning.

Second, the execution of such workflows requires a hybrid infrastructure: edge devices create streams of input data, which are processed by data analytics and machine learning applications in the Cloud, and simulations on large, specialized HPC systems provide insights into and prediction of future system state. From these results, additional steps create and communicate output data across the infrastructure levels, and, for some use cases, devices or cyber-physical systems in the real world are controlled (as in the case of smart factories). Thus, such workflows need to create different requirements for every step of their execution and require a hybrid combination of interconnected underlying infrastructure subsystems: supercomputers, cloud data centers and edge-processing systems connected to sensors (emergence of the *computing continuum*).

To leverage the computing continuum, cooperation between multiple areas (HPC, Big Data analytics, AI, cyber-security, etc.) is necessary; in Europe, this motivated the creation of the TransContinuum Initiative (TCI), whose vision is summarized in [37]. We are proud to play a leading role in TCI, where Gabriel Antoniu co-leads the use case analysis working group, in charge of "Big Data" aspects. In addition, in the framework of ETP4HPC, we have contributed to the definition of a European vision on how the HPC area is being reshaped due to the emergence of the computing continuum by co-authoring the ETP4HPC agenda in 2020 [34] and in 2022 [33]. Very recently, we have also contributed to a community white paper [28] describing the challenges of creating an integrated software/hardware ecosystem for the computing continuum.

These two evolutions are the major factors that are directly impacting the definition of our scientific program for the upcoming years. In short, we maintain our three major research axes defined five years ago, while adapting them to cope with these important evolutions.

3.1 Research Axis 1: Convergence of Extreme-Scale Computing and Big Data infrastructures

This axis keeps HPC-Big Data convergence at storage infrastructure level as a major investigation area for the team, while shifting focus from storage abstractions to the convergence of the underlying storage resources (namely, HPC storage systems and cloud storage systems). In addition, we plan to focus on I/O orchestration on hybrid HPC/cloud infrastructures as part of the computing continuum.

Dynamic provisioning of hybrid storage resources. While for years high-performance computing (HPC) systems were the predominant means of meeting the requirements expressed by large-scale scientific workflows, today some components have moved away from supercomputers to cloud-type infrastructures. This migration has been mainly motivated by the cloud's ability to perform data analysis tasks efficiently.

From an I/O and storage perspective, this means having to deal with two very different worlds: the world of cloud computing, where direct access to resources is extremely limited due to a very high level of abstraction, and the world of on-premise supercomputers offering a low level approach requiring tight user control. The abstraction layer of clouds also allows storage, network and computing resources to have a certain elasticity and to be exclusively allocated.

In this context, we propose to converge these two worlds by exploring ways to provide storage resources distributed across hybrid HPC/cloud infrastructures to complex scientific workflows combining simulation and data analysis.

To do so, we continue our recently started work on scheduling algorithms dedicated to storage resources and implemented in a storage-aware scheduler developed in the team (simulator and scheduler). We also start a new research line focused on the abstraction of storage resources in order to provide a unified interface allowing to query any type of storage on an hybrid infrastructure.

I/O Orchestration over hybrid infrastructures. On hybrid infrastructures, in the same way as the amount of generated data increases, the need for persistence has stepped up. A broad variety of large-scale scientific applications and workflows in scientific domains such as materials, high energy physics or engineering have massive I/O needs.

On a HPC system, for instance, it is typically estimated that around 10 % to 20 % of the wall time of this class of applications is spent in I/O. In addition, in the case of workflows running on hybrid infrastructures, these I/O are extremely varied and are no longer restricted to a single system but are spread across complex architectures.

To take advantage of the capability of current systems and hope to leverage future ones, improving the I/O performance is decisive. The complexity of both the federation and the different underlying systems implies having a strong knowledge of the workloads' I/O behavior and adopting a topology-aware approach for data movement orchestration.

We focus our effort on two research lines here. We first model the I/O behavior from the application and workflow's point of view. The parameters influencing I/O performance may be as diverse as the data size, the data model (multidimensional arrays, meshes, etc.), the data layout (array of structures, structure of arrays, etc.) or the access frequency.

The impact of each characteristic on I/O performance will be evaluated with benchmarks and real applications on the different systems of an hybrid infrastructure (HPC, cloud and edge later on) and an I/O workload model will be proposed.

Then, while using this I/O characterization, we focus our effort on data aggregation, taking into account the underlying topology, which consists of selecting a subset of intermediate resources to collect data before moving it from/to the destination (a storage system or a data processing system in case of in-transit workflows, for instance). This technique has several advantages: it increases the I/O bandwidth by reading or writing larger chunks of data, it highly reduces the number of concurrent streams to the destination and it minimizes the network contention.

3.2 Research Axis 2: Advanced data processing, analytics and AI in a reproducible way on the Edge-to-Cloud Continuum

This second axis explores challenges posed by the Computing Continuum to data processing. For the short term we will continue our current work investigating the best ways to leverage the Edge-to-Cloud continuum (using **E2C***lab* as an experimental platform), and we plan to extend the infrastructure scope to also include HPC subsystems (i.e., to cover the full computing continuum), in support to application workloads where machine learning will play an increasing role.

Supporting repeatable, replicable and reproducible automatic deployments across the continuum. As communities from an increasing number of scientific domains are leveraging the Computing Continuum, a desired feature of any experimental research is that its scientific claims are verifiable by others in order to build upon them. This can be achieved through repeatability, replicability, and reproducibility (3 Rs).

E2C*lab* is a first step towards enabling these goals and, to the best of our knowledge, it is the first platform to support the complete analysis cycle of an application on the Computing Continuum. We plan to further consolidate **E2C***lab* in order to make it a promising platform for future performance optimization of applications on the Edge-to-Cloud Continuum through reproducible experiments.

Specifically, we plan to focus on three main directions: (1) develop new, finer grained abstractions to model the components of the entire data processing pipeline across the continuum (from data production to permanent storage) and allow researchers to trade between different costs with increased accuracy; (2) enable built-in support for other large-scale experimental testbeds, besides Grid'5000, such as Vagrant and Chameleon and ultimately provide a community driven tool for large scale experimentation; and (3) develop a benchmark for processing frameworks within the Computing Continuum atop **E2C***lab*.

Many exciting research questions could then be explored leveraging such an enhanced deployment and optimization tool, especially in domains like machine and deep learning: how to improve the convergence speed of distributed algorithms (i.e., gradient descent) to reach good accuracy quickly? how to appropriately partition a model based on the capability of different cloud or edge devices?

Continual learning and inference in parallel across the Computing Continuum. As neural network architectures and their training data are getting more and more complex, so are the infrastructures that are needed to execute them sufficiently fast. Hyperparameter setting and tuning, training, inference, dataset handling are operations that are all putting a growing pressure on the underlying compute infrastructure and call for novel approaches at all levels of the workflow, including the algorithmic level, the middleware and deployment level, and the resource optimization level.

Our goal is to address the following specific research questions: how can the various possible deployment options of complex AI workflows on the available underlying infrastructure impact performance metrics? how can this infrastructure be best leveraged in practice, potentially through seamless integration of supercomputers, clouds, and fog/edge systems?

We will focus on the middleware and the deployment level. Our objective is to investigate various deployment strategies for complex AI workflows (e.g., potentially combining continual training, simulations and inference, all in parallel and in real-time) on hybrid execution infrastructures (e.g., combining supercomputers and cloud/fog/edge systems).

Efficient federated learning in heterogeneous and volatile environments. The latest technological advances in hardware accelerators like the GPUs enable the execution of machine and deep learning tasks on large volumes of data in a time that has become reasonable. Embedded systems make it possible to deploy some inference tasks as close as possible to the operational context. One of the major challenges of these heterogeneous distributed systems lies in the ability to have relevant data in a given place and at a given time.

One approach is to rely on the recent privacy-preserving Federated Learning paradigm that leverages the edge devices for training. However, such solutions raise some major challenges related to system and statistical heterogeneity, energy footprint and security.

Our goal is to identify and adapt such emerging approaches resulting from the Computing Continuum in order to respond to the problems of distribution of computations and processing, particularly in the case of workflows involving AI. This exploratory topic has concrete application use-cases such as with the smart autonomous vehicles or military and civilian warning systems.

3.3 Research Axis 3: I/O management, in situ visualization and analysis on HPC systems at extreme scales

Our third research axis (mainly dedicated to our HPC-centered activity during the past years) will now be redefined to address challenges posed by the increasing HPC/Big Data/AI convergence at the application level and the evolutions of the HPC infrastructures that are becoming hybrid as well, as CPU/GPU architectures become the norm for pre-Exascale/Exascale machines.

Towards unified data processing techniques for hybrid simulation/analytics workflows executed across potentially hybrid CPU/GPU infrastructures. In the high-performance computing area (HPC), the need to get fast and relevant insights from massive amounts of data generated by extreme-scale computations led to the emergence of *in situ/in transit processing*. In the Big Data area, the search for real-time, fast analysis was materialized through a different approach: *stream-based* processing. A major challenge is the joint use of these techniques in a unified data processing architecture.

Preliminary work already started within the "frameworks" work package of the HPC-Big Data Inria Challenge. It is also a core direction of our team's involvement in the ACROSS H2020 EuroHPC project. A typical scenario considered in ACROSS consists in executing hybrid workflows combining simulations and (potentially learning-based) analytics running concurrently.

The challenge is to integrate both stream and in-situ/in-transit processing tasks in the targeted workflows, leading to a decrease in execution times for data-intensive/deep learning like HPC simulations and modeling workloads. In particular, we will introduce programmatic support for on-demand data analytics on platforms that were traditionally used only for simulations. This new type of workflow (combining simulations with data analytics) could help anticipate the future behavior of the simulated systems.

Analyzing and exploiting stored data jointly with simulated data can provide a richer tool for much deeper interpretation of the targeted systems, enabling more reliable, transparent and innovative decision making. To this purpose, Damaris will be extended to support asynchronously Big Data analytics plugins, to enable in-situ and in transit analysis of simulation data, then to support hybrid (stream-based and batch-based) in transit data analysis. These new, hybrid workflows will allow on one hand to reduce the simulation time (by pre-analyzing some parts of the results locally, in-situ) and on the other hand to use simulations to train proxy models for optimization.

In the EUPEX EuroHPC Project, one goal is to introduce cross-application optimizations for data-driven stream parallel applications. This will rely on Damaris to orchestrate transfers, by leveraging various storage capabilities to provide scalable asynchronous I/O and non-intrusive in situ and in transit data processing on the data nodes. This provides another motivation to adapt Damaris to support workflows and Big Data analytics plugins by enabling in-situ and in-transit analysis of stream data.

Finally, as a piece of software considered for the Inria Exascale Software Task, in collaboration with the CEA, we plan to investigate new types of scenarios for hybrid CPU/GPU machines, where simulations could trigger on-demand analytics potentially run on GPU hardware.

4 Application domains

The KerData team investigates the design and implementation of architectures for data storage and processing across clouds, HPC and edge-based systems, which address the needs of a large spectrum of applications. The use cases we target to validate our research results come from the following domains.

4.1 Climate and meteorology

The European Centre for Medium-Range Weather Forecasts (ECMWF) [40] is one of the largest weather forecasting centers in the world that provides data to national institutions and private clients. ECMWF's production workflow collects data at the edge through a large set of sensors (satellite devices, ground and ocean sensors, smart sensors). This data, approximately 80 million observations per day, is then moved to be assimilated, i.e. analyzed and sorted, before being sent to a supercomputer to feed the prediction models.

The compute and I/O intensive large-scale simulations built upon these models use ensemble forecasting methods for the refinement. To date, these simulations generate approximately 60 TB per hour, while the center predicts an annual increase of 40 % of this volume. Structured datasets called "products" are then generated from this output data and are disseminated to different clients, such as public institutions or private companies, at a rate of 1PB per month transmitted.

In the framework of the ACROSS EuroHPC Project started in 2020, our goal is to participate in the design of a hybrid software stack for the HPC, Big Data and AI domains. This software stack must be compatible with a wide range of heterogeneous hardware technologies and must meet the needs of the trans-continuum ECMWF workflow.

4.2 Earth science

Earthquakes cause substantial loss of life and damage to the built environment across areas spanning hundreds of kilometers from their origins. These large ground motions often lead to hazards such as tsunamis, fires and landslides. To mitigate the disastrous effects, a number of Earthquake Early Warning (EEW) systems have been built around the world. Those critical systems, operating 24/7, are expected to automatically detect and characterize earthquakes as they happen, and to deliver alerts before the ground motion actually reaches sensitive areas so that protective measures can be taken.

One goal of our research is to improve the accuracy of Earthquake Early Warning (EEW) systems. These systems are designed to detect and characterize medium and large earthquakes before their damaging effects reach a certain location. Traditional EEW methods based on seismometers fail to accurately identify large earthquakes due to their low sensitivity to ground motion velocity. The recently introduced high-precision GPS stations, on the other hand, are ineffective to identify medium earthquakes due to their propensity to produce noisy data. In addition, GPS stations and seismometers may be deployed in large numbers across different locations and may produce a significant volume of data consequently, affecting the response time and the robustness of EEW systems.

Integrating and processing in a timely manner high-frequency data streams from multiple sensors scattered over a large territory requires high-performance computing techniques and equipments. We therefore design distributed machine learning-based approaches [6] to earthquake detection, jointly with experts in machine learning and Earth data. Our expertise in swift processing of data on edge and cloud infrastructures allows us to learn from the data from the large number of sensors arriving at high sampling rate, without transferring all data to a single point and thus enables real-time alerts.

4.3 Sustainable development through precision agriculture

Feeding the growing world's population is a on-going challenge, especially in view of climate change, which adds a certain level of uncertainty in food production. Sustainable and precision agriculture is one of the answers that can be implemented to partly overcome this issue. Precision agriculture consists in using new technologies to improve crop management by considering environmental parameters such as temperature, soil moisture or weather conditions, for example. These techniques now need to scale up to improve their accuracy. Over recent years, we have seen the emergence of precision agriculture workflows running across the digital continuum, that is to say all the computing resources from the edge to High-Performance Computing (HPC) and Cloud-type infrastructures. This move to scale is accompanied by new problems, particularly with regard to data movements.

CybeleTech [39] is a French company that aims at developing the use of numerical technologies in agriculture. The core products of CybeleTech are based on numerical simulation of plant growth through dedicated biophysical models and machine learning methods extracting knowledge through large databases. To develop its models, CybeleTech collects data from sensors installed on open agricultural plots or in crop greenhouses. Plant growth models take weather variables as input and the accuracy of agronomic indices estimation heavily rely on the accuracy of these variables.

To this purpose, CybeleTech wishes to collect precise meteorological information from large forecasting centers such as the European Center for Medium-Range Weather Forecasting (ECMWF) [40]. This data gathering is not trivial since it involves large data movements between two distant sites under severe time constraints. In the context of the EUPEX EuroHPC project, our team is exploring innovative data management techniques and data movement algorithms to accelerate the execution of these hybrid geo-distributed workflows running on large-scale systems in the area of precision agriculture.

4.4 Smart cities

The proliferation of small sensors and devices that are capable of generating valuable information in the context of the Internet of Things (IoT) has exacerbated the amount of data flowing from all connected objects to cloud infrastructures. In particular, this is true for Smart City applications. These applications raise specific challenges, as they typically have to handle small data (in the order of bytes and kilobytes), arriving at high rates, from many geographical distributed sources (sensors, citizens, public open data sources, etc.) and in heterogeneous formats, that need to be processed and acted upon with high reactivity in near real-time.

Our vision is that, by smartly and efficiently combining the data-driven analytics at the edge and in the cloud, it becomes possible to make a substantial step beyond state-of-the-art prescriptive analytics through a new, high-potential, faster approach to react to the sensed data of the smart cities. The goal is to build a data management platform that will enable comprehensive joint analytics of past (historical) and present (real-time) data, in the cloud and at the edge, respectively, allowing to quickly detect and react to special conditions and to predict how the targeted system would behave in critical situations. This vision is the driving objective of our SmartFastData associate team with Instituto Politécnico Nacional, Mexico.

In a similar context, smart homes by leveraging numerous sensors and connected devices aim at improving the quality of life, security and making better use of the energy. This is one target of the ENGAGE project.

4.5 Botanical Science

Pl@ntNet [32] is a large-scale participatory platform dedicated to the production of botanical data through AI-based plant identification. Pl@ntNet's main feature is a mobile app allowing smartphone owners to identify plants from photos and share their observations. It is used by around 10 million users all around the world (more than 180 countries) and it processes about 400K plant images per day. One of the challenges faced by Pl@ntNet engineers is to anticipate what should be the appropriate evolution of the infrastructure to pass the next spring peak without problems and also to know what should be done the following years.

Our research aims to improve the performance of Pl@ntNet. Reproducible evaluations of Pl@ntNet on large-scale testbed (e.g., deployed on Grid'5000 [29] by **E2C***lab* [10]), aim to optimize its software configurations in order to minimize the user response time.

5 Social and environmental responsibility

5.1 Footprint of research activities

HPC facilities are expensive in capital outlay (both monetary and human) and in energy use. Our work on Damaris supports the efficient use of high performance computing resources. Damaris [4] can help minimize power needed in running computationally demanding engineering applications and can reduce the amount of storage used for results, thus supporting environmental goals and improving the cost effectiveness of running HPC systems.

5.2 Impact of research results

Social impact. One of our target applications is Early Earthquake Warning. We proposed a solution that enables earthquakes classification with an outstandingly perfect accuracy. By enabling accurate identification of strong earthquakes, it becomes possible to trigger adequate measures and save lifes. For this reason, our work was distinguished with an Outstanding Paper Award — Special Track for Social Impact at AAAI-20, an A* conference in the area of Artificial Intelligence. This result has been highlighted by the *Le Monde* journal in its edition of December 28, 2020, in a section entitled: *Ces découvertes scientifiques que le Covid-19 a masquées en 2020*. This collaborative work continued beyond 2020.

Environmental impact. As presented in Section 4, we are partners with CybeleTech in the framework of the EUPEX EuroHPC project. CybeleTech is a French company specialized in precision agriculture. Within the framework of our collaboration, we propose to focus our efforts on a scale-oriented data management mechanism targeting two CybeleTech use-cases. They address irrigation scheduling for orchards and optimal harvest date for corn, and their models require the acquisition of large volumes of remote data. The overall goal is to improve the accuracy of plant growth models and improve decision making for precision agriculture, which directly aims to contribute to sustainable development.

6 Highlights of the year

6.1 Team evolution

The year 2023 saw the KerData team grow with the physical arrival of Guillaume Pallez at Rennes. Guillaume joined the team since 2022 while working remotely from the Inria Bordeaux center until last July. The 2023 Inria recruitment campaign also enabled us to welcome Silvina Caino-Lores as a permanent researcher (ISFP) starting in October 2023. Finally, Jakob Luettgau joined us as a researcher on a Starting Research Position (SRP).

6.2 Academic Award

Daniel Rosendo won the 2nd PhD Thesis Award prize for his PhD work at the BDA conference.

6.3 SC'24 Program Chair

Guillaume Pallez has been selected to be the Program Chair of SC'24, the top conference in the area of HPC (around 13,000 participants every year).

6.4 PEPR projects

Two large-scale PEPR National programs started in 2023: NumPEx (dedicated to Exascale computing) and CLOUD (for cloud computing research). KerData is strongly involved in leading roles in two projects of these programs:

- Exa-DoST (Data-oriented Software and Tools for the Exascale) is a project of the NumPEx program started on 1/1/2023 for 82 months - overall budget: 6,125,000 €. Gabriel Antoniu is coordinating the project. François Tessier is involved as a Work Package leader.
- STEEL (Secure and efficient daTa storagE and procEssing on cLoud based infrastructures) - started on 1/6/2023 for 87 months - overall budget: 2,779,950 €. Gabriel Antoniu is coordinating the project. Alexandru Costan is involved a a Work Package leader.

7 New software, platforms, open data

7.1 New software

7.1.1 Damaris

Keywords: Visualization, I/O, HPC, Exascale, High performance computing

Scientific Description: Damaris is a middleware for I/O and data management targeting large-scale, MPI-based HPC simulations. It initially proposed to dedicate cores for asynchronous I/O in multicore nodes of recent HPC platforms, with an emphasis on ease of integration in existing simulations, efficient resource usage (with the use of shared memory) and simplicity of extension through plug-ins.

Over the years, Damaris has evolved into a more elaborate system, providing the possibility to use dedicated cores or dedicated nodes to in situ data processing and visualization. It proposes a seamless connection to the VisIt visualization framework to enable in situ visualization with minimum impact on run time. Damaris provides an extremely simple API and can be easily integrated into the existing large-scale simulations.

Damaris was at the core of the PhD thesis of Matthieu Dorier, who received an Accessit to the Gilles Kahn Ph.D. Thesis Award of the SIF and the Academy of Science in 2015. Developed in the framework of our collaboration with the JLESC – Joint Laboratory for Extreme-Scale Computing, Damaris was the first software resulted from this joint lab validated in 2011 for integration to the Blue Waters supercomputer project. It scaled up to 16,000 cores on Oak Ridge's leadership supercomputer Titan (first in the Top500 supercomputer list in 2013) before being validated on other top supercomputers. Active development is currently continuing within the KerData team at Inria, where it is at the center of several collaborations with industry as well as with national and international academic partners.

In 2023, in the context of the ACROSS EuroHPC project, we added an interface for Damaris to enable asynchronous analytics, in particular to support Dask (www.dask.org), a Python-based library for scalable analytics. Dask offers a suite of useful distributed analytic methods using familiar Python-like interfaces, similar to NumPy and Pandas. Our proposed Python interface has enabled access to the suite of Python based visualization libraries and Damaris has been successfully tested with new options for in situ visualization.

Damaris has been selected to be one of the key software pieces of software for the NumPEx PEPR project, which aims to provide the software infrastructure for the future Exascale machine to be hosted in France in 2025 (Jules Vernes project). The capabilities within Damaris will further studied in collaboration with CEA within the NumPEx exploratory PEPR project.

Functional Description: Damaris is a middleware for data management and in-situ visualization targeting large-scale HPC simulations. Damaris enables: - In-situ data analysis by using selected dedicated cores/nodes of the simulation platform. - Asynchronous and fast data transfer from HPC simulations to Damaris. - Semantic-aware dataset processing through Damaris plug-ins, - Writing aggregated data (by hdf5 format) or visualizing them either by VisIt or ParaView.

Release Contributions: v1.8 adds type introspection API and HDF5 select capability to reorder data as it is saved to HDF5 files. v1.9 adds extended support for unstructured mesh data types.

URL: https://project.inria.fr/damaris/

Contact: Gabriel Antoniu

Participants: Gabriel Antoniu, Lokman Rahmani, Luc Bouge, Matthieu Dorier, Orçun Yildiz, Hadi Salimi, Joshua Bowden

Partner: ENS Rennes

7.1.2 E2Clab

Name: Edge-to-Cloud lab

Keywords: Distributed systems, Cloud, Reproducibility, Experimentation, Computing Continuum, Evaluation, Large scale, Provenance

Scientific Description: E2Clab is a framework that implements a rigorous methodology that provides guidelines to move from real-life application workflows to representative settings of the physical infrastructure underlying this application in order to accurately reproduce its relevant behaviors and therefore understand and optimize end-to-end performance.

E2Clab allows a rigorous analysis of possible application configurations in a controlled testbed environment to understand their behavior and related performance trade-offs. E2Clab can be generalized to other applications in the Edge-to-Cloud Continuum. E2Clab is currently used by the Pl@ntNet team to understand and optimize the performance of the application. It is also used by our partners from Instituto Politécnico Nacional for automatic experiment deployments in the context of the SmartFastData associate team.

In 2023, in an effort to enhance the reproducibility capabilities of E2Clab, we extended it to enable efficient provenance date capture across the Edge-to-Cloud Continuum. Specifically, we leverage simplified data models, data compression and grouping, and lightweight transmission protocols to reduce overheads for collecting such data on the IoT/Edge. This integration makes E2Clab a promising platform for the performance optimization of applications through reproducible experiments.

Functional Description: E2Clab is a framework that implements a rigorous methodology that provides guidelines to move from real-life application workflows to representative settings of the physical infrastructure underlying this application in order to accurately reproduce its relevant behaviors and therefore understand end-to-end performance. Understanding end-to-end performance means rigorously mapping the scenario characteristics to the experimental environment, identifying and controlling the relevant configuration parameters of applications and system components, and defining the relevant performance metrics.

Release Contributions: Changelog https://gitlab.inria.fr/E2Clab/e2clab/-/blob/master/CHANGELOG.md?ref_type=heads

Features (release 1.0.0):

(i) the configuration of the experimental environment, libraries and frameworks, (ii) the mapping between the application parts and machines on the Edge, Fog and Cloud, (iii) the deployment of the application on the infrastructure, (iv) Edge-to-Cloud network emulation,

(v) the automated execution and monitoring, (vi) the application optimization, and (vii) the gathering of experiment metrics.

URL: https://gitlab.inria.fr/E2Clab/e2clab

Publications: hal-02916032, hal-03310540, hal-03269852, hal-03332524, hal-03270129, hal-03338520, hal-03324177, hal-03259975, hal-03409405, hal-03510012

Contact: Gabriel Antoniu

Participants: Daniel Rosendo, Gabriel Antoniu, Alexandru Costan, Mathieu Simonin

7.1.3 StorAlloc

Keywords: Simulation, HPC, Distributed Storage Systems

Functional Description: StorAlloc is a simulator of a job scheduler dedicated to heterogeneous storage resources. It allows to model storage infrastructures, to simulate their partitioning and allocation, and to evaluate various scheduling algorithms.

In practice, StorAlloc takes a storage request as input, which represents the presumed storage requirements of a job executed on a HPC system. It then offers to select some fitting storage resources, to be used by the client job. Storage resources are defined by the users, thanks to a YAML format with storage nodes and disks. Their selection happens by mean of an algorithm, also chosen by user (either from predefined algorithms, or user-developed). During simulation, various metrics are stored by StorAlloc all along the processing of storage requests, and eventually written to file when the simulation ends. Components of StorAlloc are independent and communicate through messages. They are easily extensible and new components may also be added.

URL: https://github.com/hephtaicie/storalloc

Contact: François Tessier

Participants: Julien Monniot, François Tessier, Gabriel Antoniu

7.1.4 Fives

Keywords: Simulation, HPC, Distributed Storage Systems

Scientific Description: Development of Fives began in 2023, given the limitations of our previous StorAlloc simulator. At the end of 2023, Fives is still in active development, while its design and initial results are being submitted to a conference in the field.

Functional Description: Fives is a storage resource scheduling simulator for supercomputers based on WRENCH and SimGrid, two state-of-the-art simulation frameworks. In particular, Fives can model a parallel file system such as Lustre, a computing partition, and simulate a set of jobs performing I/O on the resulting HPC system.

Fives is based on several components. Firstly, as part of the development of this simulator, an abstraction called "Compound Storage Service" was proposed to represent a distributed storage system, and integrated into WRENCH. Within Fives, a job model was designed to represent a history of jobs and submit them to the scheduler present in WRENCH. Finally, a model of an existing supercomputer, Theta at Argonne National Laboratory, and a reverse-engineered version of its Lustre file system were developed in our simulator.

Experiments are underway to calibrate and validate Fives.

Contact: François Tessier

8 New results

8.1 Convergence of HPC and Big Data Infrastructures for supporting Computing Continuum workflows

8.1.1 Provisioning storage resources for HPC and Cloud systems

Participants: François Tessier, Julien Monniot, Gabriel Antoniu.

Collaboration. *This work has been carried out in close co-operation with Henri Casanova (University of Manoa, HI, USA)*

One of the recent axis we are developing in the context of high-performance data access concerns the provisioning of storage resources. The way these resources are accessed on supercomputers and clouds opposes a complex low-level vision that requires tight user control (on supercomputers) and a very abstract vision that implies uncertainty for performance modeling (on clouds). Nevertheless, taking full advantage of all available resources is critical in a context where storage is central for coupling workflow components. Our goal is then to make heterogeneous storage resources distributed across HPC+Cloud infrastructures allocatable and elastic to meet the needs of I/O-intensive hybrid workloads.

This is the context of Julien Monniot's thesis (thesis started in October 2021). He explores the techniques for scheduling storage resources on large-scale systems though a simulator of a job scheduler developed in KerData. The modeling of storage infrastructures and the evaluation of storage-aware scheduling algorithms are the main contributions of this work.

A first paper published in 2022 introduced StorAlloc, a DES-based simulator, and demonstrated how it can help to size a burst-buffer partition for a top-tier supercomputer [36]. A research poster was also presented at SC'22 in Dallas [35] and was selected as a Best Research Poster finalist. An extended and significantly augmented version of this work were published in a journal of the field in 2023 [13].

In 2023, the StorAlloc project evolved into Fives (Simulator for Scheduling on Storage Systems at Scale), a new simulator implemented with WRENCH [30], a state-of-the-art simulation framework. Fives not only reproduces the results of StorAlloc, but goes far beyond it. Thanks to this simulator, we were able to model a Lustre parallel file system (both hardware and software, which we reverse-engineered), as well as a supercomputer mimicking Theta, a 11 PFlops HPC system at Argonne National Laboratory. Using a job abstraction layer we devised, Fives enabled us to simulate several weeks of job execution on Theta from an I/O point of view (Darshan traces). The aim of this simulation was to calibrate our simulator with a view to subsequently predicting the Lustre's I/O performance on any other dataset. This work is currently being submitted to an A-rank conference in the field.

8.1.2 Efficient workflow provenance capture on the Edge-to-Cloud Continuum

Participants: Daniel Rosendo, Alexandru Costan, Gabriel Antoniu.

Collaboration. *This work has been carried out in close co-operation with Marta Mattoso, Federal University of Rio de Janeiro.*

Modern scientific workflows require hybrid infrastructures combining numerous decentralized resources on the IoT/Edge interconnected to Cloud/HPC systems (aka the Computing Continuum) to enable their optimized execution. Understanding and optimizing the performance of such complex Edge-to-Cloud workflows is challenging. Capturing the provenance of key performance indicators, with their related data and processes, may assist in understanding and optimizing workflow executions. However, the capture overhead can be prohibitive, particularly in resource-constrained devices, such as the ones on the IoT/Edge.

To address this challenge, based on a performance analysis of existing systems, we propose ProvLight [19], a tool to enable efficient provenance capture on the IoT/Edge. We leverage simplified data models, data compression and grouping, and lightweight transmission protocols to reduce overheads. We further integrate ProvLight into the E2Clab framework to enable workflow provenance capture across the Edge-to-Cloud Continuum. This integration makes E2Clab a promising platform for the performance optimization of applications through reproducible experiments.

We validate ProvLight at a large scale with synthetic workloads on 64 real-life IoT/Edge devices in the FIT IoT LAB testbed. Evaluations show that ProvLight outperforms state-of-the-art systems like ProvLake and DfAnalyzer in resource-constrained devices. ProvLight is 26—37x faster to capture and transmit provenance data; uses 5—7x less CPU; 2x less memory; transmits 2x less data; and consumes 2—2.5x less energy. ProvLight and E2Clab are available as open-source tools.

8.1.3 Data-centric workflow composition in the Computing Continuum

Participants: Silvina Caino-Lores, Alexandru Costan, Gabriel Antoniu.

Scientific workflows have become integral tools in broad scientific computing use cases. Science discovery is increasingly dependent on workflows to orchestrate large and complex scientific experiments that range from the execution of a cloud-based data preprocessing pipeline to multi-facility instrument-to-edge-to-HPC computational workflows. Given the changing landscape of scientific computing (often referred to as a computing continuum) and the evolving needs of emerging scientific applications, it is paramount that the development of novel scientific workflows and system functionalities seek to increase the efficiency, resilience, and pervasiveness of existing systems and applications.

As part of our leadership roles in the Workflows Community Initiative and workflow-specific venues like the WORKS workshop, we have identified the need to support data streams from the edge-to-cloud-to-HPC; orchestrate distributed services (workflows, instruments, data movement, provenance, publication, etc.) across computing and user facilities [31]; scale workflows to exascale computing systems and clouds; and modernize existing workflows to modern workflow management systems, amongst others [14]. We are investigating how to model workflow data in a way that allows the interoperability of existing programming models across the computing continuum, with the goal of defining a common data exchange layer serving workflow management building blocks (e.g., task scheduling, task control flow, data staging, provenance services). Specifically, we are exploring how the E2Clab framework can be expanded to support this vision due to its service-oriented nature. We are currently studying relevant application scenarios provided by the SKAO in the context of the ExaDoST project as a specific example of a large scale application in the computing continuum with mixed data and processing requirements.

8.2 Advanced data processing support for Artificial Intelligence across the Computing Continuum

8.2.1 Cost-effective Repeatability, Reproducibility, and Replicability of Edge-to-Cloud Experiments

Participants: Daniel Rosendo, Alexandru Costan, Gabriel Antoniu.

Collaboration. *This work has been carried out in close co-operation with Kate Keahey, University of Chicago / Argonne National Laboratory, who supervised the internship of Daniel Rosendo.*

Distributed infrastructures for computation and analytics are now evolving towards an interconnected ecosystem allowing complex scientific workflows to be executed across hybrid systems spanning from IoT Edge devices to Clouds, and sometimes to supercomputers (the Computing Continuum). Understanding the performance trade-offs of large-scale workflows deployed on such complex Edge-to-Cloud Continuum is challenging. To achieve this, one needs to systematically perform experiments, to enable their reproducibility and allow other researchers to replicate the study and the obtained conclusions on different infrastructures. This breaks down to the tedious process of reconciling the numerous experimental requirements and constraints with low-level infrastructure design choices.

To address the limitations of the main state-of-the-art approaches for distributed, collaborative experimentation, such as Google Colab, Kaggle, and Code Ocean, we propose KheOps[18], a collaborative environment specifically designed to enable cost-effective reproducibility and replicability of Edge-to-Cloud experiments. KheOps is composed of three core elements: (1) an experiment repository; (2) a notebook environment; and (3) a multi-platform experiment methodology.

We illustrate KheOps with a real-life Edge-to-Cloud application. The evaluations explore the point of view of the authors of an experiment described in an article (who aim to make their experiments reproducible) and the perspective of their readers (who aim to replicate the experiment). The results show how KheOps helps authors to systematically perform repeatable and reproducible experiments on the Grid5000 + FIT IoT LAB testbeds. Furthermore, KheOps helps readers to cost-effectively replicate authors experiments in different infrastructures such as Chameleon Cloud + CHI@Edge testbeds, and obtain the same conclusions with high accuracies (> 88% for all performance metrics).

8.2.2 Formalisation of workflow provenance for trustworthy and explainable AI

Participants: Silvina Caino-Lores, Alexandru Costan.

Collaboration. *This work is part of an ongoing collaboration with Renan Souza (Oak Ridge National Laboratory, USA), Rafael Ferreira da Silva (Oak Ridge National Laboratory, USA), Ana Trisovic (Massachusetts Institute of Technology, USA) and Neil Thompson (Massachusetts Institute of Technology, USA).*

Artificial Intelligence (AI) is driving scientific discovery and economic growth in all kinds of application domains while impacting from routine daily tasks to societal-level challenges. However, research communities, industry players and social actors are expressing increasing concern about the potential ethical and practical implications of the pervasive presence of AI. Of particular concern are the explainability of AI, or making AI's decision-making process understandable, and transparency of AI, ensuring clarity in AI's design, data and operation. Working towards advancing explainability and transparency of AI is currently a priority, essential for responsible and trustworthy AI applications.

Existing approaches to attain a degree of explainability of AI outputs mainly focus on analyzing the features in the data that lead to a given output or studying the layers and changes in neurons of the network throughout training. We are currently exploring a different approach by investigating the associations between neural network (NN) provenance metadata, model behavior, and outcome. Specifically, we seek to understand the correlation and causation between various architectural and algorithmic characteristics of the NN and their performance. Additionally, we aim to identify the most relevant architectural features serving as the highest predictors for NN behavior and accuracy. Our goal is to investigate to what extent we can leverage these characteristics and relationships as a proxy to improve transparency and explainability of NN models.

To achieve this, we rely heavily on the understanding of provenance metadata captured at runtime during the execution of AI workflows. Provenance refers to the origin or history of data and models, capturing the lineage of their creation, modification, and usage. It provides a detailed record of data sources, processing steps, and model configurations, ensuring transparency and traceability throughout the AI lifecycle. A challenging aspect of working with AI workflows is that today there are no comprehensive formalisms able to capture the complexity and relationships in workflow and model provenance data. Our ongoing work towards the definition of ontologies and taxonomies for AI workflow provenance data aims to fill this gap and serve as a theoretical foundation for developing tailored provenance data management systems tailored for the different stakeholders involved in AI applications.

8.2.3 Selective parameter aggregation for poisoning attack mitigation in federated learning

Participants: Cédric Prigent, Alexandru Costan, Gabriel Antoniu.

Collaboration. *This work has been carried out in close co-operation with Loïc Cudennec (DGA), who is co-advising the PhD thesis of Cédric Prigent, and with DFKI in the context of the ENGAGE Inria-DFKI project.*

Federated Learning (FL) is a distributed Machine Learning paradigm aiming to collaboratively learn a shared model while considering privacy preservation. Clients do the training process locally with their private data while a central server updates the global model by aggregating local models.

Minimizing the attack surface of FL systems is a field of active research. FL turns out to be highly vulnerable to various threats coming from the edge of the network, especially in highly distributed and volatile environment. Current approaches rely on robust aggregation, anomaly detection and generative models for defending against poisoning attacks. Yet, they either have limited defensive capabilities due to their underlying design or are impractical to use as they rely on constraining building blocks.

In [16], we introduce FedGuard, a novel FL framework that utilizes the generative capabilities of Conditional Variational AutoEncoders (CVAE) to effectively defend against poisoning attacks with tuneable overhead in communication and computation. Whilst the idea of hardening a FL system using generative models is not entirely new, FedGuard's original contribution is in its selective parameter aggregation operator with parameter selection being driven by synthetic validation data sampled from the CVAEs trained locally by each participating party. Experimental evaluations in a 100-client setup demonstrates FedGuard to be more effective than previous approaches against several types of attacks (label and sign flipping, additive noise, same value attacks). FedGuard successfully defends in scenarios with up to 50% malicious peers where other strategies fail. In addition, FedGuard does not require auxiliary datasets or centralized (pre-) training. It provides resilience against poisoning attacks from the very first round of federated training.

8.2.4 Efficient Data-Parallel Continual Learning with Asynchronous Distributed Rehearsal Buffers

Participants: Thomas Bouvier, Alexandru Costan, Gabriel Antoniu.

Collaboration. *This work has been carried out in close co-operation with Bogdan Nicolae (Argonne National Laboratory- ANL, USA), who co-advised the internship of Thomas Bouvier at ANL and serves as a technical advisor for his PhD work.*

Deep learning has emerged as a powerful method for extracting valuable information from large volumes of data. However, when new training data arrives continuously (i.e., is not fully available from the beginning), incremental training suffers from catastrophic forgetting (i.e., new patterns are reinforced at the expense of previously acquired knowledge). Training from scratch each time new training data becomes available would result in extremely long training times and massive data accumulation. Rehearsal-based continual learning has shown promise for addressing the catastrophic forgetting challenge, but research to date has not addressed performance and scalability. To fill this gap, we propose an approach based on a distributed rehearsal buffer that efficiently complements data-parallel training on multiple GPUs to achieve high accuracy, short runtime, and scalability. It leverages a set of buffers (local to each GPU) and uses several asynchronous techniques for updating these local buffers in an embarrassingly parallel fashion, all while handling the communication overheads necessary to augment input mini-batches (groups of training samples fed to the model) using unbiased, global sampling. We further propose a generalization of rehearsal buffers to support both classification and generative learning tasks, as well as more advanced rehearsal strategies (notably dark experience replay, leveraging knowledge distillation). We illustrate this approach with a real-life HPC streaming application from the domain of ptychographic image reconstruction. We run extensive experiments on up to 128 GPUs of the ThetaGPU supercomputer to compare our approach with baselines representative of training-from-scratch (the upper bound in terms of accuracy) and incremental training (the lower bound). Results show that rehearsal-based continual learning achieves a top-5 validation accuracy close to the upper bound, while simultaneously exhibiting a runtime close to the lower bound.

8.2.5 Local data compression for resource-constrained clustered federated learning in the Edge-to-Cloud Continuum

Participants: Cédric Prigent, Alexandru Costan, Gabriel Antoniu.

Collaboration. *This work has been carried out in close co-operation with Loïc Cudennec (DGA), and with DFKI in the context of the ENGAGE Inria-DFKI project.*

Although Federated Learning (FL) can be a very effective tool for large scale collaborative training of Machine Learning (ML) models, there exist some fundamental problems which reduce the efficacy of this framework, primarily the problem of clients having vastly varying data distributions. As a result, training a single global model that satisfies data distributions of the entire federation is usually unfeasible.

Clustered FL has been proposed to tackle this problem by grouping clients with similar data distributions and train personalized models in each identified cluster. While providing an efficient framework for training ML models in heterogeneous networks, existing approaches mainly focus on the accuracy of the clustering mechanism, overlooking system constraints. This leads to sustainability problems for resource-constrained devices as encountered in a Edge-to-Cloud Continuum.

We introduce a new method for clustered FL involving privacy-secured low dimensional data projections of user data to derive optimal clusters for personalized FL. We conduct a preliminary study to demonstrate the robustness of our approach against data reconstruction attacks. Afterwards, we conduct extensive experiments on Grid'5000 to evaluate the performance of the proposed approach in various scenarios introducing concept drift and label shift. Our approach achieves the best clustering performance in 3 scenarios, and second best clustering performance in a last scenario while competing with state-of-the-art clustering strategies. Computation and communication overhead analysis shows that with the help of a centralized AutoEncoder pre-training policy, our clustering approach introduces negligible system overhead compared to standard FL. In comparison, IFCA the second best performing clustering approach induces up to 100% communication overhead and 87% training time overhead. This work is currently in submission.

8.2.6 Comparative Analysis of Federated Learning: Simulations Versus Real-World Testbeds

Participants: Mathis Valli, Alexandru Costan, Gabriel Antoniu.

Collaboration. *This work has been carried out in close co-operation with Loïc Cudennec (DGA). Mathis Valli PhD is co-supervised by Cédric Tedeschi (Myriads team).*

Federated Learning (FL), while a breakthrough in decentralized machine learning, is predominantly validated through simulations. These simulations often fail to capture the complexities and unpredictability of real-world environments. They usually overlook factors like network variability, device heterogeneity, and real-time operational challenges, leading to a gap between theoretical efficiency and practical applicability.

Addressing these limitations, our research shifts focus from simulations to real testbed evaluations. We aim to deploy and assess FL models in environments that mimic more realalistic conditions, considering network instability, varied data distributions, and device capabilities. By comparing these real-testbed outcomes with simulation results, our study intends to highlight the disparities and refine the practical applicability of FL models, bridging the gap between theory and practice.

We deploy FL models across a heterogeneous mix of nodes within Grid'5000, managing and monitoring their performance using e2clab's orchestration tools. Key performance metrics such as model accuracy, training time, network usage, and energy efficiency are measured to evaluate the effectiveness of FL in a real-world environment. We aim to conduct tests to assess the gap between simulated environnements and real deployment.

While our current research emphasizes bridging the theoretical and practical aspects of Federated Learning through real-world testbed evaluations, a key future direction is to investigate the resilience and adaptability of FL models in dynamic environments. This involves understanding how these models react and adjust to changing conditions such as fluctuating network stability, varied data distributions, and shifting device capabilities.

8.3 Scalable I/O, in-situ Visualization and Resource Management at Large Scale

8.3.1 Scalable asynchronous I/O and in-situ processing with Damaris for carbon sequestration

Participants: Joshua Bowden, Alexandru Costan, François Tessier, Gabriel Antoniu.

Collaboration. *This work has been carried out in close co-operation with Atgeirr Rasmussen (SINTEF) and his team within the framework of the EuroHPC H2020 ACROSS project.*

Carbon capture and storage (CCS) is one of the technologies that have a large potential for mitigating CO2 emissions, and can also enable carbon-negative processes. Before one can commit to large-scale carbon storage operations, it is essential to do simulation studies to assess the storage potential and safety of the operation and to optimize the placement and operation of injection wells. Such a simulation is done by computer programs that solve the equations that describe the motion and state of the fluids within the porous rocks. In the ACROSS project, we use OPM Flow, an open-source reservoir simulator program suitable for both industrial uses as well as research.

As large-scale simulations can take a long time to run, and require significant high-performance computing resources, we investigate how asynchronous I/O and in situ processing can help improve the performance, scaling, and efficiency of OPM Flow and of workflows using the program. The ACROSS project is using a co-development method, where software requirements inform the hardware design choice for next generation HPC systems. The OPM Flow software is typical of MPI based simulation software where I/O inhibits the scaling of the simulation to larger machine sizes due to its serialized nature. We started to investigate how the Damaris approach could be leveraged by OPM Flow to provide asynchronous I/O.

We proposed an interface for Damaris to enable asynchronous analytics, in particular to support Dask (www.dask.org), a Python-based library for scalable analytics. Dask offers a suite of useful distributed analytic methods using familiar Python-like interfaces, similar to NumPy and Pandas. Our proposed Python interface has enabled access to the suite of Python based visualization libraries and Damaris has been successfully tested with new options for in situ visualization.

The EuroHPC ACROSS project has supported this work and the results are benefiting the OPM Flow simulation software which integrates Damaris in a public release. Workflows that use the Python and Dask analytics capabilities have been developed and include a demonstration of multi-simulation sensitivity analysis of full simulation field data, and an on-line training of neural network based auto-encoders for reduced order model development. These results have been described in ACROSS deliverable D7.3 and D7.6. The capabilities within Damaris are to be further studied in collaboration with CEA within the future NumPEx exploratory PEPR project.

8.3.2 Using application grouping to improve I/O scheduling

Participants: Guillaume Pallez.

Collaboration. *This work has been carried out in the framework of the EuroHPC H2020 Admire project. It started while Guillaume was in the Tadaam team and has been carried out since Guillaume has joined KerData (but is still located in Tadaam).*

Previous work has shown that, when multiple applications perform I/O phases at the same time, it is best to grant exclusive access to one of them at a time, which limits interference. That strategy is especially well adapted for a situation where applications have similar periods (they perform I/O phases with a similar frequency). However, when that is not the case, applications with shorter I/O phases present a higher stretch. We have been investigating a strategy where applications are grouped according to their I/O frequency. The idea is that applications from the same group should be executed one at a time, while different groups should share the available bandwidth. We are also working to determine a good priority-assigning policy.

In this topic we have made several contribution:

- In [11], we present the potential of grouping strategies through a scheduling heuristic called SET-10, which is simple and requires only minimal information. Our extensive experimental campaign shows the importance of IO-Sets and the robustness of SET-10 under various workloads. We also provide insights on using our proposal in practice.

- Then we have provided FTIO [27], a tool for characterizing the temporal I/O behavior of an application using frequency techniques such as DFT and autocorrelation. FTIO imposes generate only a modest amount of information and hence imposes minimal overhead. We also proposed metrics that quantify the confidence in the obtained results and further characterize the I/O behavior based on the identified period. The results can be directly included in the algorithmic strategy proposed in [11].

8.3.3 Scheduling distributed I/O resources in HPC systems

Participants: Guillaume Pallez.

Parallel file systems cut files into fixed-size stripes and distribute them across a number of storage targets (OSTs) for parallel access. Moreover, a layer of I/O nodes is often placed between compute nodes and the PFS. In this context, it is important to notice both OST and I/O nodes are potentially shared by running applications, which may lead to contention and low I/O performance.

Contention-mitigation approaches usually see the shared I/O infrastructure as a single resource capable of a certain bandwidth, whereas in practice it is a distributed set of resources from which each application can use a subset. In addition, using X% of the OSTs, for example, does not grant a job X% of the PFS' peak performance. Indeed, depending on their characteristics, each application will be impacted differently by the number of used I/O resources.

We conducted a comprehensive study of the problem of scheduling shared I/O resources — I/O nodes, OSTs, etc — to HPC applications. We tackled this problem by proposing heuristics to answer two questions: 1) how many resources should we give each application (allocation heuristics), and 2) which resources should be given to each application (placement heuristics). These questions are not independent, as using more resources often means sharing them. Nonetheless, our two-step approach allows for simpler heuristics that would be usable in practice.

In addition to overhead, an important aspect that impacts how "implementable" algorithms are is their input regarding applications' characteristics, since this information is often not available or at least imprecise. Therefore, we proposed heuristics that use different input and studied their robustness to inaccurate information.

8.3.4 Enabling Agile Analysis of I/O Performance Data with PyDarshan

Participants: Jakob Luettgau.

Modern scientific applications utilize numerous software and hardware layers to efficiently access data. This approach poses a challenge for I/O optimization because of the need to instrument and correlate information across those layers. The Darshan I/O characterization tool seeks to address this challenge by providing efficient, transparent, and compact runtime instrumentation of many common I/O interfaces. It also includes command-line tools to generate actionable insights and summary reports. However, the extreme diversity of today's scientific applications means that not all applications are well served by one-size-fits-all analysis tools. We have designed and implemented PyDarshan, a Python-based library that enables agile analysis of I/O performance data. PyDarshan caters to both novice and advanced users by offering ready-to-use HTML reports as well as a rich collection of APIs to facilitate custom analyses. In [17] we present the design of PyDarshan and demonstrate its effectiveness in four diverse real-world analysis use cases.

8.3.5 Analyzing Qualitatively Optimization Objectives in the Design of HPC Resource Manager

Participants: Guillaume Pallez, Robin Boezennec.

A correct evaluation of scheduling algorithms and a good understanding of their optimization criterias are key components of resource management in HPC. In [15, 25] we discuss bias and limitations of the most frequent optimization metrics from the literature. We provide elements on how to evaluate performance when studying HPC batch scheduling. We experimentally demonstrate these limitations by focusing on two use-cases: a study on the impact of runtime estimates on scheduling performance, and the reproduction of a recent highimpact work that designed an HPC batch scheduler based on a network trained with reinforcement learning. We demonstrate that focusing on quantitative optimization criterion ("our work improve the literature by X%") may hide extremely important caveat, to the point that the results obtained are opposed to the actual goals of the authors. Key findings show that mean bounded slowdown and mean response time are irrelevant objectives in the context of HPC. Despite some limitations, mean utilization appears to be a good objective. We propose to complement it with its standard deviation in some pathologic cases. Finally, we argue for a larger use of area-weighted response time, that we find to be a very relevant objective.

9 Partnerships and cooperations

9.1 International initiatives

9.1.1 Associate Teams in the framework of an Inria International Lab or in the framework of an Inria International Program

UNIFY 2

Participants: Gabriel Antoniu, Thomas Bouvier, Alexandru Costan, Julien Monniot, François Tessier.

Title: Intelligent Unified Data Services for Hybrid Workflows Combining Compute-Intensive Simulations and Data-Intensive Analytics at Extreme Scales - 2

Duration: 2023 ->

Coordinator: Gabriel Antoniu

Partners:

- Argonne National Laboratory Argonne (États-Unis)

Inria contact: Gabriel Antoniu

Summary: Since several years we have been witnessing the emergence of complex workflows combining simulations with data analysis, potentially leveraging machine-learning techniques. Such complex workflows seem to naturally need to jointly use supercomputers interconnected with clouds and potentially Edge-based systems. This assembly is called the Computing Continuum. In a general scheme, Edge devices create streams of input data, which are processed by data analytics and machine learning applications in the Cloud, whereas simulations on large, specialised HPC systems provide insights into and prediction of future system state. The emergence of such workflows is reshaping the traditional vision on the areas involved, as described in the ETP4HPC Research Agenda published in 2020. Building software ecosystems addressing the needs of such workflows poses multiple challenges at several levels. In this context, this Associate Team will focus on three related challenges: 1) How to adequately handle the heterogeneity of storage resources within the Computing Continuum to support complex science workflows? 2) How to efficiently support deep-learning workloads across the Computing Continuum? 3) How to provide reproducibility support for experimentation across the Computing Continuum?

9.2 International research visitors

9.2.1 Visits of international scientists

Ewa Deelman

Status: Research Professor and Research Director

Institution of origin: University of South California

Country: USA

Dates: 1 June 2023

Context of the visit: PhD defense of Daniel Rosendo, meetings with the KerData team

Manish Parashar

Status: Director of the Scientific Computing and Imaging Institute and Presidential Professor

Institution of origin: University of Utah

Country: USA

Dates: 24 November 2024

Context of the visit: Visit of several Inria teams in Nantes and Rennes in the area of distributecd computing and AI

9.2.2 Visits to international teams

In September 2023, Julien Monniot, François Tessier and Gabriel Antoniu visited the University of Manoa, HI, USA, to collaborate with Henri Casanova on simulation of distributed storage systems. The outcome of this visit and the following work is going to be submitted in a rank-A conference.

9.3 European initiatives

9.3.1 H2020 projects

EUPEX

Participants: Joshua Bowden, Gabriel Antoniu, François Tessier.

EUPEX project on cordis.europa.eu

Title: EUROPEAN PILOT FOR EXASCALE

Duration: From January 1, 2022 to December 31, 2025

Partners:

- INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET AUTOMATIQUE (INRIA), France
- GRAND EQUIPEMENT NATIONAL DE CALCUL INTENSIF (GENCI), France
- VSB - TECHNICAL UNIVERSITY OF OSTRAVA (VSB - TU Ostrava), Czechia
- FORSCHUNGSZENTRUM JULICH GMBH (FZJ), Germany
- COMMISSARIAT A L ENERGIE ATOMIQUE ET AUX ENERGIES ALTERNATIVES (CEA), France
- IDRYMA TECHNOLOGIAS KAI EREVNAS (FOUNDATION FOR RESEARCH AND TECHNOLOGYHELLAS), Greece
- SVEUCILISTE U ZAGREBU FAKULTET ELEKTROTEHNIKE I RACUNARSTVA (UNIVERSITYOF ZAGREB FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING), Croatia
- UNIVERSITA DEGLI STUDI DI TORINO (UNITO), Italy
- Consortium Ubiquitous Technologies S.c.a.r.l. (CUBIT), Italy
- CYBELETECH, France
- UNIVERSITA DI PISA (UNIPI), Italy
- GRAN SASSO SCIENCE INSTITUTE (GSSI), Italy
- ISTITUTO NAZIONALE DI ASTROFISICA (INAF), Italy
- UNIVERSITA DEGLI STUDI DEL MOLISE, Italy
- E 4 COMPUTER ENGINEERING SPA (E4), Italy
- UNIVERSITA DEGLI STUDI DELL'AQUILA (UNIVAQ), Italy
- CONSIGLIO NAZIONALE DELLE RICERCHE (CNR), Italy
- JOHANN WOLFGANG GOETHE-UNIVERSITAET FRANKFURT AM MAIN (GUF), Germany
- EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS (ECMWF), United Kingdom
- Eviden, formerly ATOS/BULL SAS (BULL), France
- POLITECNICO DI MILANO (POLIMI), Italy
- EXASCALE PERFORMANCE SYSTEMS - EXAPSYS IKE, Greece
- ALMA MATER STUDIORUM - UNIVERSITA DI BOLOGNA (UNIBO), Italy
- PARTEC AG (PARTEC), Germany
- ISTITUTO NAZIONALE DI GEOFISICA E VULCANOLOGIA, Italy
- CINECA CONSORZIO INTERUNIVERSITARIO (CINECA), Italy
- SECO SPA (SECO SRL), Italy
- CONSORZIO INTERUNIVERSITARIO NAZIONALE PER L'INFORMATICA (CINI), Italy

Inria contact: Olivier Beaumont

Coordinator: Etienne Walter, Eviden

Summary: The EUPEX consortium aims to design, build, and validate the first EU platform for HPC, covering end-to-end the spectrum of required technologies with European assets: from the architecture, processor, system software, development tools to the applications. The EUPEX prototype will be designed to be open, scalable and flexible, including the modular OpenSequana-compliant platform and the corresponding HPC software ecosystem for the Modular Supercomputing Architecture. Scientifically, EUPEX is a vehicle to prepare HPC, AI, and Big Data processing communities for upcoming European Exascale systems and technologies. The hardware platform is sized to be large enough for relevant application preparation and scalability forecast, and a proof of concept for a modular architecture relying on European technologies in general and on European Processor Technology (EPI) in particular. In this context, a strong emphasis is put on the system software stack and the applications.

Being the first of its kind, EUPEX sets the ambitious challenge of gathering, distilling and integrating European technologies that the scientific and industrial partners use to build a production-grade prototype. EUPEX will lay the foundations for Europe's future digital sovereignty. It has the potential for the creation of a sustainable European scientific and industrial HPC ecosystem and should stimulate science and technology more than any national strategy (for numerical simulation, machine learning and AI, Big Data processing).

The EUPEX consortium – constituted of key actors on the European HPC scene – has the capacity and the will to provide a fundamental contribution to the consolidation of European supercomputing ecosystem. EUPEX aims to directly support an emerging and vibrant European entrepreneurial ecosystem in AI and Big Data processing that will leverage HPC as a main enabling technology.

ACROSS

Participants: Gabriel Antoniu, Alexandru Costan, François Tessier, Joshua Bowden, Thomas Bouvier, Daniel Rosendo.

ACROSS project on cordis.europa.eu

Title: HPC BIG DATA ARTIFICIAL INTELLIGENCE CROSS STACK PLATFORM TOWARDS EXASCALE

Duration: From March 1, 2021 to February 29, 2024

Partners:

- INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET AUTOMATIQUE (INRIA), France
- VSB - TECHNICAL UNIVERSITY OF OSTRAVA (VSB - TU Ostrava), Czechia
- MORFO DESIGN SRL, Italy
- NEUROPUBLIC AE PLIROFORIKIS & EPIKOINONION (NEUROPUBLIC SA), Greece
- UNIVERSITA DEGLI STUDI DI FIRENZE (UNIFI), Italy
- UNIVERSITA DEGLI STUDI DI TORINO (UNITO), Italy
- SINTEF AS (SINTEF), Norway
- INSTITUT NATIONAL DES SCIENCES APPLIQUEES DE RENNES (INSA RENNES), France
- STICHTING DELTARES (Deltares), Netherlands
- EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS (ECMWF), United Kingdom

- Eviden, formerly ATOS/BULL SAS (BULL), France
- GE AVIO SRL (GE AVIO SRL), Italy
- FONDAZIONE LINKS - LEADING INNOVATION & KNOWLEDGE FOR SOCIETY (FONDAZIONE LINKS), Italy
- UNIVERSITA DEGLI STUDI DI GENOVA (UNIGE), Italy
- MAX-PLANCK-GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV (MPG), Germany
- CINECA CONSORZIO INTERUNIVERSITARIO (CINECA), Italy
- CONSORZIO INTERUNIVERSITARIO NAZIONALE PER L'INFORMATICA (CINI), Italy

Inria contact: Gabriel Antoniu

Coordinator: Olivier Terzo, LINKS

Summary: Supercomputers have been extensively used to solve complex scientific and engineering problems, boosting the capability to design more efficient systems. The pace at which data are generated by scientific experiments and large simulations (e.g., multiphysics, climate, weather forecast, etc.) poses new challenges in terms of capability of efficiently and effectively analysing massive data sets. Artificial Intelligence, and more specifically Machine Learning (ML) and Deep Learning (DL) recently gained momentum for boosting simulations' speed. ML/DL techniques are part of simulation processes, used to early detect patterns of interests from less accurate simulation results. To address these challenges, the ACROSS project will co-design and develop an HPC, BD, and Artificial Intelligence (AI) convergent platform, supporting applications in the Aeronautics, Climate and Weather, and Energy domains. To this end, ACROSS will leverage on next generation of pre-exascale infrastructures, still being ready for exascale systems, and on effective mechanisms to easily describe and manage complex workflows in these three domains. Energy efficiency will be achieved by massive use of specialized hardware accelerators, monitoring running systems and applying smart mechanisms of scheduling jobs. ACROSS will combine traditional HPC techniques with AI (specifically ML/DL) and BD analytic techniques to enhance the application test case outcomes (e.g., improve the existing operational system for global numerical weather prediction, climate simulations, develop an environment for user-defined in-situ data processing, improve and innovate the existing turbine aero design system, speed up the design process, etc.). The performance of ML/DL will be accelerated by using dedicated hardware devices. ACROSS will promote cooperation with other EU initiatives (e.g., BDVA, EPI) and future EuroHPC projects to foster the adoption of exascale-level computing among test case domain stakeholders.

9.3.2 Collaborations with Major European Organizations

Participants: Gabriel Antoniu, Alexandru Costan.

Appointments by Inria in relation to European bodies

Big Data Value Association (BDVA) and ETP4HPC: Since 2018, Gabriel Antoniu and Alexandru Costan have been serving as Inria representatives in the working groups dedicated to HPC-Big Data convergence.

Community service at European level in response to external invitations

ETP4HPC: Since 2019: Gabriel Antoniu has served as a co-leader of the working group on Programming Environments and co-leader of two research clusters, contributing to two successive versions of the Strategic Research Agenda of ETP4HPC, the latest one being published in 2022 14. Alexandru Costan served as a member of these working groups.

Transcontinuum Initiative (TCI): As a follow-up action to the publication of its Strategic Research Agenda, ETP4HPC initiated a collaborative initiative called TCI (Transcontinuum Initiative). It gathers major European associations in the areas of HPC, Big Data, AI, 5G, Cybersecurity, including ETP4HPC, BDVA, CLAIRE, HIPEAC, 5G IA, ECO). It aims to strengthen research and industry in Europe to support the Digital Continuum - infrastructure (including HPC systems, clouds, edge infrastructures) by helping to define a set of research focus areas/topics requiring interdisciplinary action. The expected outcome of this effort is the co-editing of multidisciplinary calls for projects to be funded by the European Commission. Gabriel Antoniu is in charge of ensuring the BDVA-ETP4HPC coordination and of co-animating the working group dedicated to the definition of representative application use cases.

Big Data Value Association: Gabriel Antoniu was asked by BDVA to coordinate Diva's contribution to the TCI initiative recently started (see above).

9.4 National initiatives

Exa-DoST

Participants: Gabriel Antoniu, François Tessier, Julien Monniot, Joshua Bowden, Silvina Caino Lores, Guilaume Pallez.

Exa-DoST project of the NumPEx PEPR program

Title: Data-oriented Software and Tools for the Exascale

Duration: From January 1, 2023 to April 1, 2030

Partners:

- Inria
- CEA
- CNRS
- University of Bordeaux
- Observatoire de Paris
- Observatoire de la Côte d'Azure
- Data Direct Networks France (DDN)

Coordinator: Gabriel Antoniu (KerData Team, Inria)

Summary: The advent of future Exascale supercomputers raises multiple data-related challenges. To enable applications to fully leverage the upcoming infrastructures, a major challenge concerns the scalability of techniques used for data storage, transfer, processing and analytics. Additional key challenges emerge from the need to adequately exploit emerging technologies for storage and processing, leading to new, more complex storage hierarchies. Finally, it now becomes necessary to support more and more complex hybrid workflows involving at the same time simulation, analytics and learning, running at extreme scales across supercomputers interconnected to clouds and edgebased systems. The Exa-DoST project will address most of these challenges, organized in 3 areas:

- Scalable storage and I/O;

- Scalable in situ processing;
- Scalable smart analytics.

As part of the NumPEx program, Exa-DoST targets a much higher technology readiness level than previous national projects concerning the HPC software stack. It will address the major data challenges by proposing operational solutions co-designed and validated in French and European applications. This will allow filling the gap left by previous international projects to ensure that French and European needs are taken into account in the roadmaps for building the data-oriented Exascale software stack.

STEEL

Participants: Gabriel Antoniu, Alexandru Costan, Jakob Luettgau, François Tessier, Mathis Valli.

Title: Secure and efficient daTa storagE and procEssing on cLoud-based infrastructures

Duration: From June 1, 2023 to 31 August 2030

Partners:

- Inria
- CNRS
- Institut Mines Télécom (IMT)
- University of Bordeaux
- University of Rennes
- INSA Rennes
- INSA Lyon

Coordinator: Gabriel Antoniu (KerData Team, Inria)

Summary: The strong development of cloud computing since its emergence in 2007 and its massive adoption for the storage of unprecedented volumes of data in a growing number of domains has brought to light major technological challenges. In this project we will address several of these challenges, organized in three research directions. The first direction concerns the exploitation of emerging technologies for efficient storage on cloud infrastructures. We will address this challenge through NVRAM-based distributed performance storage solutions, as close as possible to data production and consumption locations (disaggregation principle) and develop strategies to optimize the trade-off between data consistency and access performance. The second direction concerns the efficient storage and processing of data on hybrid, heterogeneous infrastructures within the digital edge-cloud-supercomputer continuum. In many domains (autonomous cars, predictive maintenance, intelligent buildings, etc.) we are witnessing the emergence of hybrid workflows combining simulations, analysis of sensor data flows and machine learning. Their execution requires storage resources ranging from the edge to cloud infrastructures, and even to supercomputers, which poses challenges for unified data storage and processing. The third research direction is dedicated to confidential storage, in connection with the need to store and analyze large volumes of data of strategic interest or of a personal nature. For all of these directions, the project will take into account the need to propose and validate interoperable approaches with a potential for transfer to major French or European industrial players in cloud computing.

Grid'5000 We are members of Grid'5000 community and run experiments on the Grid'5000 platform on a daily basis.

Inria Exploratory program: Repas

Participants: Guillaume Pallez.

Project Acronym: REPAS
Title: New Portrayal of HPC Applications
Coordinator: Guillaume Pallez
Collaboration: This is done in collaboration with the team DATAMOVE (Inria Grenoble)
Duration: 2022-2025

What is the right way to represent an application in order to run it on a highly parallel (typically exascale) machine? The idea of project is to completely review the models used in the development scheduling algorithms and software solutions to take into account the real needs of new users of HPC platforms.

10 Dissemination

10.1 Promoting scientific activities

Participants: Gabriel Antoniu, Silvina Caino-Lores, Alexandru Costan, Guillaume Pallez, François Tessier, Julien Monniot, Cédric Prigent.

10.1.1 Scientific events: organisation

General chair, scientific chair

François Tessier: General Chair of ESSA 2024, the 5th Workshop on Extreme-Scale Storage and Analysis, to be held in 2024 in conjunction with IPDPS 2024.

Gabriel Antoniu : Steering Committee Chair of the ESSA Workshop series on High-Performance Storage, held in conjunction with the IEEE IPDPS conference since 2020.

Silvina Caino-Lores: General Co-Chair of WORKS 2024, the 18th Workshop on Workflows in Support of Large-Scale Science, held in conjunction with SC'23, Denver (CO, USA).

Guillaume Pallez: Technical Program Chair of SC 2024

Alexandru Costan:

- Co-chair of the SC 2023 ACM Student Research Competition Graduate Posters
- General Co-Chair of FlexScience 2024, the 13th Workshop on AI and Scientific Computing at Scale using Flexible Computing Infrastructures, held in conjunction with ACM HPDC'23, Orlando (FL, USA).

Member of the organizing committees

François Tessier:

- Web and Publicity Chair of ESSA 2023, the 4th Workshop on Extreme-Scale Storage and Analysis, to be held in 2023 in conjunction with IPDPS 2023.
- Co-organizer of the SuperCompCloud BoF on Interoperability of Supercomputing and Cloud Technologies, held in conjunction with SC'23, Denver (CO, USA).

Silvina Caino-Lores:

- Workshops Co-Chair of Euro-Par 2024, the 30th International European Conference on Parallel and Distributed Computing, to be held in 2024.
- Reproducibility Challenge Co-Chair of SC 2024, the International Conference for High Performance Computing, Networking, Storage, and Analysis, to be held in 2024.

Guillaume Pallez:

- Finance Chair of Cluster 2023

10.1.2 Scientific events: selection

Chair of conference program committees

Alexandru Costan: Program Committee Co-chair of IEEE ISPDC 2023, 22nd IEEE International Symposium on Parallel and Distributed Computing, Bucharest (Romania).

François Tessier: Organizer of a mini-symposium about "Data Management across the Computing Continuum" at the PASC'23 conference — Platform for Advanced Scientific Computing, digital event.

Member of the conference program committees

François Tessier: CLUSTER 2023, HiPC 2023, ICPP 2023, ISC2023, REX-IO 2023, WORKS 2023, SC'23 (Reproducibility Initiative)

Guillaume Pallez: IPDPS 2024

Alexandru Costan: SC'23 (Clouds and Distributed Computing track), IPDPS 2024, UCC 2023, HPC Asia 2023, ISC 2023 (Posters), IEEE BigData 2023, CloudCom 2023

Gabriel Antoniu : SC'23 (Data analytics, visualization and Storage track), CLOUD 2023, Cluster 2023, IPDPS 2024, CCGRID 2024, HPDC 2024.

Reviewer

Julien Monniot: IEEE/ACM SC23, ISC23, Cloud 23

Cédric Prigent: IEEE BigData 2023, IEEE/ACM SC23, IEEE IPDPS 2024

10.1.3 Journal

Reviewer - reviewing activities

Alexandru Costan: IEEE Transactions on Parallel and Distributed Systems, Future Generation Computer Systems, Concurrency and Computation Practice and Experience, IEEE Transactions on Cloud Computing, Journal of Parallel and Distributed Computing.

10.1.4 Invited talks

Gabriel Antoniu :

- Keynote Speaker at the ISPDC 2023 conference in Bucharest, Romania.

François Tessier:

- Talk at the Per3S workshop on High-Performance Storage in Paris, France.

Silvina Caino-Lores:

- Panel "Runtimes and Workflow Systems for Extreme Heterogeneity: Challenges and Opportunities" at SC 23 (Denver, CO, USA).

10.1.5 Leadership within the scientific community

Gabriel Antoniu:

Large-wingspan National project management: Coordinator of ExaDoST, one of the 5 targeted projects of the NumPEx PEPR project (started in 2023, budget: 6.2 M€). Coordinator of STEEL, one of the 7 high-priority projects of the CLOUD PEPR project (started in 2023, budget: 2.8 M€).

ETP4HPC: Since 2019, co-leader of the working group on Programming Environments, lead co-author of the corresponding chapter of the Strategic Research Agenda of ETP4HPC (latest edition published in 2022, next edition to be published in 2024).

TCI: Since 2020, co-leader of the Use-Case Analysis Working Group. TCI (The Transcontinuum Initiative) emerged as a collaborative initiative of ETP4HPC, BDVA, CLAIRE and other peer organizations, aiming to identify joint research challenges for leveraging the HPC-Cloud-Edge computing continuum and make recommendations to the European Commission about topics to be funded in upcoming calls for projects.

International lab management: Vice Executive Director of JLESC for Inria. JLESC is the Joint Inria-Illinois-ANL-BSC-JSC-RIKEN/AICS Laboratory for Extreme-Scale Computing. Within JLESC, he also serves as a Topic Leader for Data storage, I/O and in situ processing for Inria.

Bilateral Inria-DFKI project management: French coordinator of the ENGAGE project (2022-2024).

Team management: Head of the KerData Project-Team (INRIA-ENS Rennes-INSA Rennes).

International Associate Team management: Leader of the UNIFY Associate Team with Argonne National Lab (2019–2022).

Alexandru Costan:

- Work package co-leader with René Schubotz(DFKI) within the ENGAGE Inria-DFKI project.
- Work package leader witin the PEPR CLOUD STEEL project.

François Tessier:

- Work package co-leader with Francieli Zanon-Boito (Associate Professor, University of Bordeaux) within the NumPEX ExaDoST project.

10.2 Teaching - Supervision - Juries

Participants: Gabriel Antoniu, Thomas Bouvier, Silvina Caino-Lores, Alexandru Costan, François Tessier, Guillaume Pallez, Cédric Prigent.

10.2.1 Teaching

Alexandru Costan:

- Bachelor: Software Engineering and Java Programming, 28 hours (lab sessions), L3, INSA Rennes.
- Bachelor: Databases, 68 hours (lectures and lab sessions), L2, INSA Rennes.
- Bachelor: Practical case studies, 24 hours (project), L3, INSA Rennes.
- Master: Big Data Storage and Processing, 28h hours (lectures, lab sessions), M1, INSA Rennes.
- Master: Algorithms for Big Data, 28 hours (lectures, lab sessions), M2, INSA Rennes.
- Master: Big Data Project, 28 hours (project), M2, INSA Rennes.

Gabriel Antoniu:

- Master (Engineering Degree, 5th year): NoSQL and Cloud technologies, 20 hours (lectures), M2 level, ENSAI (*École nationale supérieure de la statistique et de l'analyse de l'information*), Bruz.
- Master: Scalable Distributed Systems, 10 hours (lectures), M1 level, SDS Module, EIT ICT Labs Master School.
  M2 level, IBD Module, SIF Master Program, University of Rennes.
- Master: Cloud Computing and Big Data, 14 hours (lectures), M2 level, Cloud Module, MIAGE Master Program, University of Rennes.

François Tessier:

- Bachelor: Computer science discovery, 15 hours (lab sessions), L1 level, DIE Module, ISTIC, University of Rennes.
- Master: Cloud Computing and Big Data, 15 hours (lectures), M2 level, Cloud Module, MIAGE Master Program, University of Rennes.
- Master (Engineering Degree, 4th year): Storage on Clouds, 5 hours (lecture and lab session), M2 level, IMT Atlantique, Rennes.

Silvina Caino-Lores:

- Master: Processing Artificial Intelligence and Machine Learning Workloads at Scale, 9 hours (lectures), M2 level, Big Data Storage and Processing Infrastructures Module, Cloud and Network Infrastructures Master Program, EIT Digital School, ISTIC, University of Rennes.

Thomas Bouvier:

- Master: Stream Processing, 12 hours (lectures and lab sessions), M2 level, INSA Rennes.
- Master: Database optimizations, 30 hours (lab sessions), M1 level, ISTIC, University of Rennes.

Cédric Prigent:

- Master: Cloud Computing and Big Data, 36 hours (lab sessions), M2 level, Cloud Module, MIAGE Master Program, University of Rennes.

10.2.2 Supervision

PhD:

- Daniel Rosendo Prigent, *"Methodologies for Reproducible Analysis of Workflows on the Edge-to-Cloud Continuum"*, thesis defended in June 2023, co-advised by Gabriel Antoniu, Alexandru Costan and Patrick Valduriez.

PhD in progress:

- Thomas Bouvier, *"Supporting Continual Learning across the Computing Continuum"*, thesis started in January 2021, co-advised by Alexandru Costan and Gabriel Antoniu.
- Julien Monniot, *"Dynamic provisioning of intermediate storage resources across hybrid HPC/Cloud infrastructures"*, thesis started in November 2021, co-advised by François Tessier and Gabriel Antoniu.
- Cédric Prigent, *"Supporting Online Learning and Inference in Parallel across the Digital Continuum"*, thesis started in November 2021, co-advised by Alexandru Costan, Gabriel Antoniu and Loïc Cudennec (DGA).
- Mathis Valli, *"Comparative Analysis of Federated Learning: Simulations Versus Real-World Testbeds in dynamic settings"*, thesis started in April 2023, co-advised by Alexandru Costan, Cédric Tedeschi (Myriads) and Loïc Cudennec (DGA).
- Robin Boezennec, *"Towards a New Representation of HPC Workloads"*, thesis started in 2022, co-advised by Guillaume Pallez and Fanny Dufossé (DataMove).

Internships:

- Malvin Chevallier, *"Study of regularization techniques applied to rehearsal-based continual learning"*, two-and-a-half-month internship started in June 2023, co-advised by Thomas Bouvier and Alexandru Costan.

10.2.3 Juries

Alexandru Costan:

- Reviewer for the PhD thesis of Quentin Guilloteau (Université Grenoble Alpes)
- GDR RSD - Member of the juries for the PhD award and Young Researcher award.

10.3 Popularization

Participants: Gabriel Antoniu, Alexandru Costan, François Tessier, Guillaume Pallez.

10.3.1 Internal or external Inria responsibilities

François Tessier:

- Member of the Commission on Health, Safety and Working Conditions (now called FSS) within the Inria center of Rennes
- Member of the COURSE committee in charge of gathering user needs for the future SLICES platform

Alexandru Costan:

- In charge of internships at the Computer Science Department of INSA Rennes.
- In charge of the management of the KerData team access to Grid'5000. for the future SLICES platform

Gabriel Antoniu: Inria representative in the BDVA working group dedicated to HPC-Big Data convergence.

Guillaume Pallez: Member of the Inria Evaluation Committee. As part of this role:

- Copywriter of the "Recommendations on "Grey-Zone Publishers from the Inria Evaluation Committee" [23, 24]
- Bilan de la mandature 2019-2023 de la Commission d'Évaluation Inria [26]

11 Scientific production

11.1 Major publications

[1] G. Antoniu, P. Valduriez, H.-C. Hoppe and J. Krüger. *Towards Integrated Hardware/Software Ecosystems for the Edge-Cloud-HPC Continuum*. 2021. DOI: 10.5281/zenodo.5534464. URL: https://hal.archives-ouvertes.fr/hal-03358930.

[2] F. Boito, G. Pallez, L. Teylo and N. Vidal. 'IO-SETS: Simple and efficient approaches for I/O bandwidth management'. In: *IEEE Transactions on Parallel and Distributed Systems* 34.10 (15th Aug. 2023), pp. 2783–2796. DOI: 10.1109/TPDS.2023.3305028. URL: https://inria.hal.science/hal-03648225.

[3] N. Cheriere, M. Dorier and G. Antoniu. 'How fast can one resize a distributed file system?' In: *Journal of Parallel and Distributed Computing* 140 (June 2020), pp. 80–98. DOI: 10.1016/j.jpdc.2020.02.001. URL: https://hal.archives-ouvertes.fr/hal-02961875.

[4] M. Dorier, G. Antoniu, F. Cappello, M. Snir, R. Sisneros, O. Yildiz, S. Ibrahim, T. Peterka and L. Orf. 'Damaris: Addressing Performance Variability in Data Management for Post-Petascale Simulations'. In: *ACM Transactions on Parallel Computing* 3.3 (2016), p. 15. DOI: 10.1145/2987371. URL: https://hal.inria.fr/hal-01353890.

[5] M. Dorier, S. Ibrahim, G. Antoniu and R. Ross. 'Using Formal Grammars to Predict I/O Behaviors in HPC: the Omnisc'IO Approach'. In: *IEEE Transactions on Parallel and Distributed Systems* (2016). DOI: 10.1109/TPDS.2015.2485980. URL: https://hal.inria.fr/hal-01238103.

[6] K. Fauvel, D. Balouek-Thomert, D. Melgar, P. Silva, A. Simonet, G. Antoniu, A. Costan, V. Masson, M. Parashar, I. Rodero and A. Termier. 'A Distributed Multi-Sensor Machine Learning Approach to Earthquake Early Warning'. In: In Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York, United States, 7th Feb. 2020, pp. 403–411. DOI: 10.1609/aaai.v34i01.5376. URL: https://hal.archives-ouvertes.fr/hal-02373429.

[7] M. Malms, L. Cargemel, E. Suarez, N. Mittenzwey, M. Duranton, S. Sezer, C. Prunty, P. Rossé-Laurent, M. Pérez-Harnandez, M. Marazakis, G. Lonsdale, P. Carpenter, G. Antoniu, S. Narasimharmurthy, A. Brinkman, D. Pleiter, U.-U. Haus, J. Krueger, H.-C. Hoppe, E. Laure, A. Wierse, V. Bartsch, K. Michielsen, C. Allouche, T. Becker and R. Haas. *ETP4HPC's SRA 5 - Strategic Research Agenda for High-Performance Computing in Europe - 2022*. Zenodo, 2022. DOI: 10.5281/zenodo.7347009. URL: https://hal.science/hal-03937825.

[8] O.-C. Marcu, A. Costan, G. Antoniu and M. S. Pérez-Hernández. 'Spark versus Flink: Understanding Performance in Big Data Analytics Frameworks'. In: Cluster 2016 - The IEEE 2016 International Conference on Cluster Computing. Taipei, Taiwan, 12th Sept. 2016. DOI: 10.1109/cluster.2016.22. URL: https://hal.inria.fr/hal-01347638.

[9] P. Matri, Y. Alforov, A. Brandon, M. Pérez, A. Costan, G. Antoniu, M. Kuhn, P. Carns and T. Ludwig. 'Mission Possible: Unify HPC and Big Data Stacks Towards Application-Defined Blobs at the Storage Layer'. In: *Future Generation Computer Systems* 109 (Aug. 2020), pp. 668–677. DOI: 10.1016/j.future.2018.07.035. URL: https://hal.archives-ouvertes.fr/hal-01892682.

[10] D. Rosendo, P. Silva, M. Simonin, A. Costan and G. Antoniu. 'E2Clab: Exploring the Computing Continuum through Repeatable, Replicable and Reproducible Edge-to-Cloud Experiments'. In: Cluster 2020 - IEEE International Conference on Cluster Computing. Kobe, Japan, 14th Sept. 2020, pp. 1–11. DOI: 10.1109/CLUSTER49012.2020.00028. URL: https://hal.archives-ouvertes.fr/hal-02916032.

11.2 Publications of the year

International journals

[11] F. Boito, G. Pallez, L. Teylo and N. Vidal. 'IO-SETS: Simple and efficient approaches for I/O bandwidth management'. In: *IEEE Transactions on Parallel and Distributed Systems* 34.10 (15th Aug. 2023), pp. 2783–2796. DOI: 10.1109/TPDS.2023.3305028. URL: https://inria.hal.science/hal-03648225.

[12] E. Jeannot, G. Pallez and N. Vidal. 'IO-aware Job-Scheduling: Exploiting the Impacts of Workload Characterizations to select the Mapping Strategy'. In: *International Journal of High Performance Computing Applications* (2023), pp. 1–13. DOI: 10.1177/10943420231175854. URL: https://inria.hal.science/hal-04098706.

[13] J. Monniot, F. Tessier, M. Robert and G. Antoniu. 'Supporting Dynamic Allocation of Heterogeneous Storage Resources on HPC Systems'. In: *Concurrency and Computation: Practice and Experience*. Special Issue:S2 World 2020. CUG 2021 & 2022. PN_HCP. HeteroPar 2022 35.28 (16th Aug. 2023), pp. 1–16. DOI: 10.1002/cpe.7890. URL: https://hal.science/hal-03922866.

International peer-reviewed conferences

[14] J. Bader, J. Belak, M. Bement, M. Berry, R. Carson, D. Cassol, S. Chan, J. Coleman, K. Day, A. Duque, K. Fagnan, J. Froula, S. Jha, D. S. Katz, P. Kica, V. V. Kindratenko, E. Kirton, R. Kothadia, D. Laney, F. Lehmann, U. Leser, S. Lichołai, M. Malawski, M. Melara, E. Player, M. Rolchigo, S. Sarrafan, S.-J. Sul, A. Syed, L. Thamsen, M. Titov, M. Turilli, S. Caino-Lores and A. Mandal. 'Novel Approaches Toward Scalable Composable Workflows in Hyper-Heterogeneous Computing Environments'. In: *Proceedings of the SC'23 Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis*. SC-W 2023 - Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis. Denver (CO), United States: ACM, 12th Nov. 2023, pp. 2097–2108. DOI: 10.1145/3624062.3626283. URL: https://hal.science/hal-04385285.

[15] R. Boëzennec, F. Dufossé and G. Pallez. 'Optimization Metrics for the Evaluation of Batch Schedulers in HPC'. In: JSSPP 2023 - 26th edition of the workshop on Job Scheduling Strategies for Parallel Processing. St. Petersburg, Florida, United States, 23rd Mar. 2023, pp. 1–19. URL: https://inria.hal.science/hal-04042591.

[16] M. Chelli, C. Prigent, R. Schubotz, A. Costan, G. Antoniu, L. Cudennec and P. Slusallek. 'FedGuard: Selective Parameter Aggregation for Poisoning Attack Mitigation in Federated Learning'. In: Cluster 2023 - IEEE International Conference on Cluster Computing. Santa Fe, New Mexico, United States: IEEE, 2023, pp. 1–10. URL: https://hal.science/hal-04208787.

[17] J. Luettgau, S. Snyder, T. Reddy, N. Awtrey, K. Harms, J. L. Bez, R. Wang, R. Latham and P. Carns. 'Enabling Agile Analysis of I/O Performance Data with PyDarshan'. In: SC-W 2023: Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis. Denver CO USA, France: ACM, 12th Nov. 2023, pp. 1380–1391. DOI: 10.1145/3624062.3624207. URL: https://inria.hal.science/hal-04390702.

[18] D. Rosendo, K. Keahey, A. Costan, M. Simonin, P. Valduriez and G. Antoniu. 'KheOps: Cost-effective Repeatability, Reproducibility, and Replicability of Edge-to-Cloud Experiments'. In: *ACM REP '23: Proceedings of the 2023 ACM Conference on Reproducibility and Replicability*. REP 2023 - ACM Conference on Reproducibility and Replicability. Santa Cruz, CA, United States: ACM, 28th June 2023, pp. 62–73. DOI: 10.1145/3589806.3600032. URL: https://hal.science/hal-04157720.

[19] D. Rosendo, M. Mattoso, A. Costan, R. Souza, D. Pina, P. Valduriez and G. Antoniu. 'ProvLight: Efficient Workflow Provenance Capture on the Edge-to-Cloud Continuum'. In: IEEE Cluster 2023 - IEEE International Conference on Cluster Computing. Santa Fe, New Mexico, United States: IEEE, 2023, pp. 1–13. URL: https://hal.science/hal-04161546.

Doctoral dissertations and habilitation theses

[20] G. Pallez. 'Model Design and Accuracy for Resource Management in HPC'. Université de Bordeaux, 11th July 2023. URL: https://theses.hal.science/tel-04189199.

[21] D. Rosendo. 'Methodologies for Reproducible Analysis of Workflows on the Edge-to-Cloud Continuum'. INSA RENNES, 1st June 2023. URL: https://hal.science/tel-04167278.

Reports & preprints

[22] A. Bandet, F. Boito and G. Pallez. *Scheduling distributed I/O resources in HPC systems*. 15th Jan. 2024. URL: https://inria.hal.science/hal-04394004.

[23] F. Blanqui, A. Canteaut, H. de Jong, S. Imperiale, N. Mitton, G. Pallez, X. Pennec, X. Rival and B. Thirion. *Recommandations sur les « éditeurs de la zone grise »*. Inria, 25th Jan. 2023, pp. 1–3. URL: https://inria.hal.science/hal-04001505.

[24] F. Blanqui, A. Canteaut, H. D. Jong, S. Imperiale, N. Mitton, G. Pallez, X. Pennec, X. Rival and B. Thirion. *Recommendations on "Grey-Zone Publishers": Recommendations from the Inria Evaluation Committee, translated from https://hal.inria.fr/hal-04001505*. Inria, 25th Jan. 2023, pp. 1–3. URL: https://inria.hal.science/hal-04201298.

[25] R. Boëzennec, F. Dufossé and G. Pallez. *Analyzing Qualitatively Optimization Objectives in the Design of HPC Resource Manager*. 21st Aug. 2023. URL: https://hal.science/hal-04187517.

[26] A. Canteaut, M. Serrano, C. Grandmont, G. Pallez, V. Perrier, X. Rival and E. Thomé. *Bilan de la mandature 2019-2023 de la Commission d'Évaluation Inria*. Inria, 31st Aug. 2023. URL: https://inria.hal.science/hal-04193082.

[27] A. Tarraf, A. Bandet, F. Zanon Boito, G. Pallez and F. Wolf. *FTIO: Detecting I/O Periodicity Using Frequency Techniques*. 14th June 2023. URL: https://inria.hal.science/hal-04382142.

11.3 Cited publications

[28] G. Antoniu, P. Valduriez, H.-C. Hoppe and J. KrÃŒger. *Towards Integrated Hardware/Software Ecosystems for the Edge-Cloud-HPC Continuum*. ETP4HPC White Papers. ETP4HPC: European Technology Platform for High Performance Computing, 2021. DOI: 10.5281/zenodo.5534464. URL: https://hal.archives-ouvertes.fr/hal-03358930.

[29] R. Bolze, F. Cappello, E. Caron, M. Dayde, F. Desprez, E. Jeannot, Y. Jégou, S. Lanteri, J. Leduc, N. Melab, G. Mornet, R. Namyst, P. Primet, B. Quétier, O. Richard, E.-G. Talbi and I. Touche. 'Grid'5000: A Large Scale And Highly Reconfigurable Experimental Grid Testbed'. In: *International Journal of High Performance Computing Applications* 20.4 (2006), pp. 481–494. DOI: 10.1177/1094342006070078. URL: https://hal.inria.fr/hal-00684943.

[30] H. Casanova, R. Ferreira da Silva, R. Tanaka, S. Pandey, G. Jethwani, W. Koch, S. Albrecht, J. Oeth and F. Suter. 'Developing Accurate and Scalable Simulators of Production Workflow Management Systems with WRENCH'. In: *Future Generation Computer Systems* 112 (2020), pp. 162–175. DOI: 10.1016/j.future.2020.05.030.

[31] R. Ferreira Da Silva, R. Badia, V. Bala, D. Bard, P.-T. Bremer, I. Buckley, S. Caino-Lores, K. Chard, C. Goble, S. Jha et al. *Workflows Community Summit 2022: A Roadmap Revolution*. Tech. rep. Oak Ridge National Laboratory (ORNL), Oak Ridge, TN (United States), 2023.

[32] A. Joly, P. Bonnet, H. Goëau, J. Barbe, S. Selmi, J. Champ, S. Dufour-Kowalski, A. Affouard, J. Carré, J.-F. o. Molino, N. Boujemaa and D. Barthélémy. 'A look inside the Pl@ntNet experience'. In: *Multimedia Systems* 22.6 (2016), pp. 751–766. DOI: 10.1007/s00530-015-0462-9. URL: https://hal.inria.fr/hal-01182775.

[33] M. Malms, L. Cargemel, E. Suarez, N. Mittenzwey, M. Duranton, S. Sezer, C. Prunty, P. Rossé-Laurent, M. Pérez-Harnandez, M. Marazakis, G. Lonsdale, P. Carpenter, G. Antoniu, S. Narasimharmurthy, A. Brinkman, D. Pleiter, U.-U. Haus, J. Krueger, H.-C. Hoppe, E. Laure, A. Wierse, V. Bartsch, K. Michielsen, C. Allouche, T. Becker and R. Haas. *ETP4HPC's SRA 5 - Strategic Research Agenda for High-Performance Computing in Europe - 2022*. Zenodo, 2022. DOI: 10.5281/zenodo.7347009. URL: https://hal.science/hal-03937825.

[34] M. Malms, M. Ostasz, M. Gilliot, P. Bernier-Bruna, L. Cargemel, E. Suarez, H. Cornelius, M. Duranton, B. Koren, P. Rosse-Laurent, M. S. PÃ©rez-HernÃ¡ndez, M. Marazakis, G. Lonsdale, P. Carpenter, G. Antoniu, S. Narasimhamurthy, A. Brinkman, D. Pleiter, A. Tate, J. Krueger, H.-C. Hoppe, E. Laure and A. Wierse. *ETP4HPC's Strategic Research Agenda for High-Performance Computing in Europe 4*. Ed. by w. t. s. o. t. E.-2. p. ETP4HPC: European Technology Platform for High Performance Computing. ETP4HPC White Papers. 2020. DOI: 10.5281/zenodo.4605343. URL: https://hal.inria.fr/hal-03354396.

[35] J. Monniot, F. Tessier and G. Antoniu. *Modeling Allocation of Heterogeneous Storage Resources on HPC Systems*. SC 2022 - International Conference for High Performance Computing, Networking, Storage, and Analysis (Posters). Poster. Nov. 2022. URL: https://hal.science/hal-03878252.

[36] J. Monniot, F. Tessier, M. Robert and G. Antoniu. 'StorAlloc: A Simulator for Job Scheduling on Heterogeneous Storage Resources'. In: *HeteroPar 2022*. Glasgow, United Kingdom, Aug. 2022. URL: https://inria.hal.science/hal-03683568.

[37] *The TransContinuum Initiative vision paper*. 2020. URL: https://www.etp4hpc.eu/tci-vision.html.

[38] *Chameleon Cloud*. 2021. URL: https://www.chameleoncloud.org/.

[39] *Cybeletech - Digital technologies for the plant world*. 2021. URL: https://www.cybeletech.com/en/home/.

[40] *ECMWF - European Centre for Medium-Range Weather Forecasts*. 2021. URL: https://www.ecmwf.int/.

[41] *European Exascale Software Initiative*. 2013. URL: http://www.eesi-project.eu/.

[42] *The European Technology Platform for High-Performance Computing*. 2012. URL: http://www.etp4hpc.eu/.

[43] *International Exascale Software Program*. 2011. URL: http://www.exascale.org/iesp/.

[44] *Inria's strategic plan "Towards Inria 2020"*. 2016. URL: https://www.inria.fr/fr/recherche-innovation-numerique.