



Activity Report 2022

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	1
1 Team members, visitors, external collaborators	2
2 Overall objectives	2
3 Research program	4
3.1 Research Axis 1: Convergence of Extreme-Scale Computing and Big Data infrastructures	5
3.2 Research Axis 2: Advanced data processing, analytics and AI in a reproducible way on the Edge-to-Cloud Continuum	6
3.3 Research Axis 3: I/O management, in situ visualization and analysis on HPC systems at extreme scales	7
4 Application domains	8
4.1 Climate and meteorology	8
4.2 Earth science	9
4.3 Sustainable development through precision agriculture	9
4.4 Smart cities	10
4.5 Botanical Science	10
5 Social and environmental responsibility	10
5.1 Footprint of research activities	10
5.2 Impact of research results	11
6 Highlights of the year	11
6.1 Academic Award	11
6.2 Award Nominations	11
6.3 SC'24 Program Chair	11
6.4 PEPR projects: leading roles	11
6.5 New exploratory Action: Repas	11
7 New software and platforms	12
7.1 New software	12
7.1.1 Damaris	12
7.1.2 E2Clab	13
7.1.3 StorAlloc	13
8 New results	13
8.1 Convergence of HPC and Big Data Infrastructures as part of the Computing Continuum	13
8.1.1 Identifying challenges posed by the emergence of the Computing Continuum to the area of HPC	13
8.1.2 Provisioning storage resources for hybrid supercomputer/cloud infrastructures	14

8.1.3	Supporting seamless execution of HPC-enabled workflows across the Computing Continuum	15
8.2	Advanced data processing support for reproducibility and Artificial Intelligence across the Computing Continuum	15
8.2.1	Towards a collaborative environment for cost-effective reproducibility of Edge-to-Cloud experiments	15
8.2.2	Efficient workflow provenance capture on the Edge-to-Cloud Continuum	16
8.2.3	Supporting Efficient Workflow Deployment of Federated Learning Systems across the Computing Continuum	17
8.2.4	Towards data parallel rehearsal-based Continual Learning	17
8.2.5	Enabling better decision making for the selection of a configuration in distributed reinforcement learning	18
8.3	Scalable I/O, Communication, in-situ Visualization and Analysis on HPC Systems at Extreme Scales	19
8.3.1	Scalable asynchronous I/O and in-situ processing with Damaris for carbon sequestration	19
8.3.2	Using application grouping to improve I/O scheduling	19
9	Partnerships and cooperations	20
9.1	International initiatives	20
9.1.1	Associate Teams in the framework of an Inria International Lab or in the framework of an Inria International Program	20
9.1.2	Inria associate team not involved in an IIL or an international program	20
9.2	International research visitors	21
9.2.1	Visits of international scientists	21
9.2.2	Visits to international teams	21
9.3	European initiatives	21
9.3.1	H2020 projects	21
9.3.2	Other european programs/initiatives	24
9.3.3	Collaborations with Major European Organizations	25
9.4	National initiatives	25
10	Dissemination	27
10.1	Promoting scientific activities	27
10.1.1	Scientific events: organisation	27
10.1.2	Scientific events: selection	27
10.1.3	Journal	28
10.1.4	Invited talks	28
10.1.5	Leadership within the scientific community	29
10.2	Teaching - Supervision - Juries	29
10.2.1	Teaching	29
10.2.2	Supervision	30
10.2.3	Juries	31
10.3	Popularization	31
10.3.1	Internal or external Inria responsibilities	31

11 Scientific production	31
11.1 Major publications	31
11.2 Publications of the year	33
11.3 Cited publications	34

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]
- Guillaume Pallez [INRIA, Researcher, from Sep 2022]
- François Tessier [INRIA, Researcher]

Faculty Member

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

PhD Students

- Thomas Bouvier [INRIA]
- Julien Monniot [INRIA]
- Cédric Prigent [INRIA]
- Daniel Rosendo [INRIA]

Technical Staff

- Joshua Bowden [INRIA, Engineer]

Administrative Assistant

- Laurence Dinh [INRIA]

External Collaborators

- Loïc Cudennec [DGA, Bruz, PhD co-advisor for Cédric Prigent. ASC2 DGA Maîtrise de l'Information DGA/DT/MI/SDTI/ASC2/IA3D Intelligence Artificielle pour les Agents Autonomes & Datascience, HDR]
- Bogdan Nicolae [Argonne National Laboratory, Technical co-advisor for the PhD thesis of Thomas Bouvier. <https://www.anl.gov/profile/bogdan-nicolae>]

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 [33], EESI [31], ETP4HPC [32]. 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 [28] 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/BULL.

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 [34], 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 (main German research center in artificial intelligence) through the ENGAGE Inria-DFKI project started in 2022, with ATOS, through the ACROSS and EUPEX H2020 EuroHPC projects, started in March 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 [27]. 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 [26] and in 2022 [17]. Very recently, we have also contributed to a community white paper [22] 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 **E2Clab** 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).

E2Clab 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 **E2Clab** 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 **E2Clab**.

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) [30] 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.

Our research aims 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 [5] 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 [29] 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) [30]. This data gathering is not trivial since it involves large data movements between two distant sites

under severe time constraints. On 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 [25] 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 [23] by E2Clab [12]), 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 [3] 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 lives. 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 Academic Award

Luc Bougé, co-founder of the KerData team has been awarded the prestigious *Palme Académiques* of ENS Rennes. He retired in 2022.

6.2 Award Nominations

Our poster "Modeling Allocation of Heterogeneous Storage Resources on HPC Systems" [20], was one of the four nominated posters for the Best Research Poster Award, out of nearly hundred submissions to the SC'2022 conference held in November in Dallas, TX, USA.

6.3 SC'24 Program Chair

Guillaume Pallez, who joined the KerData team on September 1, 2022, has been selected to be the Program Chair of SC'24, the top conference in the area of HPC.

6.4 PEPR projects: leading roles

KerData has been strongly involved in setting up two PEPR national projects: Cloud and NumPEX. In NumPEX, Gabriel Antoniu is coordinating the ExaDoST project dedicated to exascale data-oriented tools (Julien Bigot from CEA is co-coordinator), while François Tessier is serving as work package co-leader. In the PEPR Cloud, Gabriel Antoniu is coordinating the STEEL project, dedicated to secure, performant cloud storage. Both projects will start in 2023.

6.5 New exploratory Action: Repas

In 2022, Guillaume Pallez was awarded an exploratory action, REPAS to work on new representations of HPC applications in scheduling schemes. *Inria's exploratory actions programme* "was set

up to promote the emergence of new research themes, giving scientists the means with which to test original ideas. The aim of this scheme is to focus resources around highly innovative subjects, targeting approaches that are *off the beaten track*, risky and/or which represent a disruption in relation to traditional approaches."

7 New software and platforms

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.

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: This version adds support for Python as a method of in situ analysis for data passed to Damaris.

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 Applications, Distributed systems, Computing Continuum, Large scale, Experimentation, Evaluation, Reproducibility

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.

Furthermore, this methodology leverages research quality aspects such as the Repeatability, Replicability, and Reproducibility of experiments through a well-defined experimentation methodology and providing transparent access to the experiment artifacts and experiment results. This is an important aspect that allows that the scientific claims are verifiable by others in order to build upon them.

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

Contact: Gabriel Antoniu

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/heptaicie/storalloc>

Contact: François Tessier

Participants: Julien Monniot, François Tessier, Gabriel Antoniu

8 New results

8.1 Convergence of HPC and Big Data Infrastructures as part of the Computing Continuum

8.1.1 Identifying challenges posed by the emergence of the Computing Continuum to the area of HPC

Participants: Alexandru Costan, Gabriel Antoniu.

The emergence of the Edge-Cloud-HPC Continuum raises challenges at multiple levels: at the application level, innovative algorithms are needed to bridge simulations, machine learning and data-driven analytics; at the middleware level, adequate tools must enable efficient deployment, scheduling and orchestration of the workflow components across the whole distributed infrastructure; and, finally, a capable resource management system must allocate a suitable set of components of the infrastructure to run the application workflow, preferably in a dynamic and adaptive way, taking into account the specific capabilities of each component of the underlying heterogeneous infrastructure. To address the challenges, we foresee an increasing need for integrated software ecosystems which combine current “island” solutions and bridge the gaps between them. These ecosystems must facilitate the full lifecycle of Computing Continuum use cases, including initial modelling, programming, deployment, execution, optimisation, as well as monitoring and control. As part of collaborative work within the ETP4HPC association, we provided contributions to the ETP4HPC Strategic Research Agenda published in 2022 [17], consisting in describing and discussing the above challenges with respect to the process of building the corresponding software ecosystems (p. 42-49) and with respect to programming models and environments (p. 102-105). In addition, we published a detailed systematic review of state-of-the-art solutions for distributed ML-based analytics on edge devices, as step towards building integrated ML-based analytics ecosystems leveraging the Computing Continuum [14].

8.1.2 Provisioning storage resources for hybrid supercomputer/cloud infrastructures

Participants: François Tessier, Julien Monniot, Gabriel Antoniu.

Collaboration. *This work has been carried out in close co-operation with Rob Ross (Argonne National Laboratory, IL, USA) and Henri Casanova (University of Manoa, HI, USA)*

One of the recent axis we are developing in the context of HPC and Big Data convergence 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 through StorAlloc, 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 introduces StorAlloc and demonstrates how this simulator can help to size a burst-buffer partition for a top-tier supercomputer [15]. An extended and significantly augmented version of this work is currently being reviewed in a journal in the field[19]. A research poster was also presented at SC’22 in Dallas last November [20]. This poster was one of four nominees for the Best Research Poster (out of nearly 100 submissions).

Finally, a collaboration was started with Henri Casanova from the University of Manoa (HI, USA) in order to integrate the results obtained with StorAlloc into WRENCH [24], a simulator based on SimGrid.

8.1.3 Supporting seamless execution of HPC-enabled workflows across the Computing Continuum

Participants: Juliette Fournis d'Albiat, Alexandru Costan, François Tessier, Gabriel Antoniu.

Collaboration. *This work has been carried out in close co-operation with Rosa Badia, Barcelona Supercomputing Center*

Scientific applications have recently evolved to more complex workflows combining traditional HPC simulations with Data Analytics (DA) and ML algorithms. While these applications were traditionally executed on separate infrastructures (e.g., simulations on HPC supercomputers, DA/ML on cloud/edge), the new combined ones need to leverage myriads of resources from the edge to the HPC in order to promptly extract insights. Our goal is to enable seamless deployment, orchestration and optimization of HPC workflows across the Computing Continuum.

To this end, we have started to leverage the building blocks developed at BSC (Barcelona Supercomputing Center) and Inria. In particular, the **E2Clab** framework already enables the deployment and execution of DA/ML workflows on cloud and edge resources following a reproducibility-oriented methodology. We are exploring the possibility to extend this support to workflows including simulations and other compute intensive applications on HPC infrastructures. The layers and services abstractions introduced by **E2Clab** fit naturally to the approach based on containers for scientific codes pushed forward by BSC (i.e., the COMPSs task programming model for complex workflows). Beyond the ease of deployment, this approach aims to enable the reproducibility of HPC experiments, one of the core design principles of **E2Clab**.

In the framework of the internship of Juliette Fournis d'Albiat, one direction explored during this year towards such integration is the support for rigorous deployment of experiments with HPC applications on distributed platforms. We leveraged the **E2Clab** methodology to deploy a COMPSs application and define the deployment workflow. We have further formulated an optimisation problem in order to identify the best configuration of the computing resources reducing the application execution time. Initial experiments using a K-means based application deployed on Grid'5000 show that this approach is able to reduce the overall execution time.

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

8.2.1 Towards a collaborative environment for cost-effective reproducibility 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 digital infrastructures for computation and analytics are now evolving towards an interconnected ecosystem allowing complex applications to be executed from IoT Edge devices to the HPC Cloud (aka the Computing Continuum). One important challenge is to accurately reproduce relevant behaviors of a given application workflow and representative settings of the physical infrastructure underlying this complex continuum.

This work aims to propose a collaborative environment for the practical reproducibility of experiments in a cost-effective way, that means: reproducing the exact same experiment environment, hardware/software versions, network topology, processing workflow; and experiment results. The ultimate goal is to lower the barrier to reproducing research by combining the reproducible artifacts and the experimental environments (e.g., Chameleon, CHI@Edge, Grid5000, and FIT IoT LAB testbeds).

We started by exploring the following research question: *What are the limitations of the existing collaborative environments?* We investigate the main state of the art environments, such as: Google Colab, Kaggle, and Code Ocean. We observed that they lack support to: (1) providing access to heterogeneous resources (e.g., IoT/Edge devices); (2) practical reproducibility of experiments (e.g., it is hard to reproduce experiments on the exact same hardware since the resources vary over time); and (3) executing experiments at large scale (e.g., users have to pay to access multiple machines).

Based on these limitations, we explored the following research question: *What would a good collaborative system look like?* In our vision, collaborative environments for enabling Computing Continuum research should provide mainly the following three features: (1) open access to research artifacts to allow other researchers to reproduce published experiments; (2) an interactive computing environment packaged with code, data, environment configurations, and experiment results; and (3) experiment methodologies exploring large-scale scientific testbeds.

Finally, we proposed and implemented our collaborative environment. It provides the following main features: (1) **Trovi sharing portal**: allows users to package code, data, environment configurations, and results and archive them in this portal, so artifacts can be easily shared and found by other users; (2) Grid'5000 Jupyter environment: guides users to systematically define the experiment workflow and execute experiments; (3) Our experiment methodology: abstracts all the complexities of deploying applications on multiple testbeds (e.g., Grid5000, Chameleon, FIT IoT lab, and CHI@Edge), as well as repeating experiments on the same infrastructure.

We illustrate our collaborative environment with an **Edge-to-Cloud experiment workflow deployed on multiple testbeds**, such as: Grid5000 + FIT IoT lab; and Chameleon + CHI@Edge. The use case refers to a monitoring system in the African savanna where various edge devices (Raspberry Pi available at FIT IoT lab / CHI@Edge) located in different regions take pictures of animals, perform some image preprocessing and then send images to the cloud server (available at Grid5000 / Chameleon). On the cloud there is a Deep Learning application that identifies the animals.

Evaluations show that our collaborative environment has proven useful for reproducing experiments on large-scale platforms from the IoT/Edge to the HPC/Cloud Continuum. It helps users to: (1) systematically configure the experimental environment; (2) easily deploy distributed applications on multiple testbeds; (3) repeat experiments on the same testbed configurations; and (4) make code, data, environment, and results shared easily.

8.2.2 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 *Computing Continuum*) to enable their optimized execution. Understanding and optimizing the performance of such complex Edge-to-Cloud workflows is challenging. Capturing provenance of key performance

indicators, with their related data and processes, may assist understanding and optimizing workflow executions. However, the provenance capture overhead can be prohibitive, particularly in resource-constrained devices, such as the ones on the IoT/Edge.

In this work, we explore provenance capture in workflows running on resource-constrained IoT/Edge devices. Based on a performance analysis of the existing systems, we propose ProvLight, a tool for the efficient provenance capture on the IoT/Edge. The integration of ProvLight into the E2Clab framework enables end-to-end workflow provenance capture of Edge-to-Cloud workflows. In addition, it makes E2Clab a promising platform for performance optimization of applications on the Computing Continuum through reproducible experiments.

We validate ProvLight with synthetic workloads on real-life IoT/Edge devices in the large-scale FIT IoT LAB testbed. Evaluations show that ProvLight outperforms ProvLake and DfAnalyzer provenance systems in resource-constrained devices. ProvLight is 20—28x 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.

8.2.3 Supporting Efficient Workflow Deployment of Federated Learning Systems across the Computing Continuum

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.

In the Computing Continuum (CC) context, FL raises several challenges such as supporting very heterogeneous devices and optimizing massively distributed applications.

In [21] we propose a workflow to better support and optimize FL systems across the Computing Continuum by relying on formal descriptions of the underlying infrastructure, hyperparameter optimization and model retraining in case of performance degradation. This approach is motivated by preliminary experiments using a human activity recognition dataset showing the importance of hyperparameter optimization and model retraining in the FL scenario.

Moreover, personalized FL, which aims at providing specialized models for each FL client, is considered to better fit the very heterogeneous demands in the CC context. While training a single model for a population of clients with diverging data distributions will likely not converge to the optimum for each client, adopting this personalization approach is a promising direction to build models that fit each client's needs.

8.2.4 Towards data parallel rehearsal-based Continual Learning

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.*

During the past decade, Deep learning (DL) supported the shift from rule-based systems towards statistical models. Deep Neural Networks (DNNs) revolutionized how we address problems in a wide range of applications by extracting patterns from complex datasets. Larger models and centralized datasets demand for distributed strategies to leverage multiple compute nodes.

Most existing supervised learning algorithms operate under the assumption that all the input data is available at the beginning of the training process. However, this constraint goes against many real-life scenarios where the aforementioned datasets are replaced by high volume, high velocity data streams generated over time by distributed (sometimes geographically) devices. It is unfeasible to keep training the models in an offline fashion from scratch every time new data arrives, as this would lead to prohibitive time and/or resource constraints. Also, typical DNNs suffer from catastrophic forgetting in this context, a phenomenon causing them to reinforce new patterns at the expense of previously acquired knowledge (i.e., a bias towards new samples). Some authors have shown that memory replay methods are effective in mitigating accuracy degradation in such settings. However, their performance is still far from that of oracles with full access to the static dataset. The problem of Continual Learning (CL) remains an open research question.

This is the context of Thomas Bouvier's thesis. We are interested in how CL methods can take advantage of data parallelization across nodes, which is one of the main techniques to achieve training scalability on HPC systems. The aggregated memory could benefit the accuracy achieved by such algorithms by instantiating distributed replay buffers. The main research goals of this project are the (i) design and implementation of a distributed replay buffer leveraging distributed systems effectively and the (ii) study of trade-offs introduced by large-scale CL in terms of training time, accuracy and memory usage.

8.2.5 Enabling better decision making for the selection of a configuration in distributed reinforcement learning

Participants: Cédric Prigent, Alexandru Costan, Gabriel Antoniu.

Collaboration. *This work has been carried out in close co-operation with Loïc Cudennec (DGA).*

As Artificial Intelligence-based systems are getting more complex with time, they are also increasing in computing demand and become time consuming. Therefore speeding up the learning phase by distributing the training process using multiple nodes appears adequate to address this problem. Unfortunately, ML also brings new development frameworks, methodologies and high-level programming languages that do not fit to the regular high-performance computing design flow.

In [16], a methodology to build a decision making tool that allows ML experts to arbitrate between different frameworks and deployment configurations is proposed to fulfill project objectives such as the accuracy of the resulting model, the computing speed or the energy consumption of the learning computation. The proposed methodology is applied to an industrial-grade case study in which reinforcement learning is used to train an autonomous steering model for a cargo airdrop system.

Results are presented within a Pareto front that lets ML experts choose an appropriate solution, a framework and a deployment configuration, based on the current operational situation. While the proposed approach can effortlessly be applied to other ML problems, as for many decision making systems, the selected solutions involve a trade-off between several antagonist evaluation criteria and require experts from different domains to pick the most efficient solution from the short list. Nevertheless, this methodology speeds up the development process by clearly discarding, or, on the contrary, including combinations of frameworks and configurations, which has a significant impact for time and budget-constrained projects.

8.3 Scalable I/O, Communication, in-situ Visualization and Analysis on HPC Systems at Extreme Scales

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 such 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 process for next generation HPC systems. The 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, 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 are being developed and include a demonstration of multi-simulation sensitivity analysis of full simulation field data, and the in-planning development of on-line training of deep learning based auto-encoders with possibilities for clustering of simulation data. 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.

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

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

Duration: 2019 – 2022

Coordinator: Tom PETERKA (tpeterka@mcs.anl.gov)

Partners:

- Argonne National Laboratory Argonne (États-Unis)

Inria contact: Gabriel Antoniu

Summary: The landscape of scientific computing is being radically reshaped by the explosive growth in the number and power of digital data generators, ranging from major scientific instruments to the Internet of Things (IoT) and the unprecedented volume and diversity of the data they generate. This requires a rich, extended ecosystem including simulation, data analytics, and learning applications, each with distinct data management and analysis needs. Science activities are beginning to combine these techniques in new, large-scale workflows, in which scientific data is produced, consumed, and analyzed across multiple distinct steps that span computing resources, software frameworks, and time. This paradigm introduces new data-related challenges at several levels. The UNIFY Associate Team aims to address three such challenges. First, to allow scientists to obtain fast, real-time insight from complex workflows combining extreme-scale computations with data analytics, we will explore how recently emerged Big Data processing techniques (e.g., based on stream processing) can be leveraged with modern in situ/in transit processing approaches used in HPC environments. Second, we will investigate how to use transient storage systems to enable efficient, dynamic data management for hybrid workflows combining simulations and analytics. Finally, the explosion of learning and AI provides new tools that can enable much more adaptable resource management and data services than available today, which can further optimize such data processing workflows.

9.1.2 Inria associate team not involved in an IIL or an international program

SmartFastData

Title: Efficient Data Management in Support of Hybrid Edge/Cloud Analytics for Smart Cities

Duration: 2019 – 2022

Coordinator: Rolando Menchaca-Mendez (rolando.menchaca@gmail.com)

Partners:

- Instituto Politécnico Nacional (Mexico)

Inria contact: Alexandru Costan

Summary: 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 private and public cloud infrastructures. In particular, this is true for Smart City applications, which cover a large spectrum of needs in public safety, water and energy management. Unfortunately, the lack of a scalable data management subsystem is becoming an important bottleneck for such applications, as it increases the gap between their I/O requirements and the storage performance.

The vision underlying the SmartFastData associated team 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. 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.

9.2 International research visitors

9.2.1 Visits of international scientists

Other international visits to the team

Pr. Osamu Tatebe

Status: Professor

Institution of origin: University of Tsukuba

Country: Japan

Dates: From Thursday June 2nd to Friday June 3rd

Context of the visit: Seminar and open discussions with the team

Mobility program/type of mobility: Seminar

9.2.2 Visits to international teams

In September 2022, Gabriel Antoniu, Alexandru Costan and François Tessier visited the Argonne National Laboratory (in conjunction with the mission to the 14th JLESC workshop) to give presentations and discuss potential collaborations.

9.3 European initiatives

9.3.1 H2020 projects

EUPEX [EUPEX project on cordis.europa.eu](https://cordis.europa.eu/project/eupez)

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 (UNIVERSITY OF ZAGREB FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING), Croatia
- UNIVERSITA DEGLI STUDI DI TORINO (UNITO), Italy
- CYBELETECH (Cybeletech), France
- UNIVERSITA DI PISA (UNIFI), 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
- 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:

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.

The EUPEX WP5 goals are to support the tools and libraries that focus on data storage and transfer on the upcoming European designed Modular System Architecture (MSA) supercomputing architecture. In 2022, partners of the consortium have been presented details of the Damaris software library and various project startup tasks have been completed including requirements specification and software details for project dissemination.

ACROSS [ACROSS project on cordis.europa.eu](https://cordis.europa.eu/project/ACROSS)

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, Netherlands
- EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS (ECMWF), United Kingdom
- 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:

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 Other european programs/initiatives

ENGAGE: NExt GeNeration ComputinG Environments for Artificial intelliGence

Participants: Cedric Prigent, Gabriel Antoniu, Alexandru Costan.

Collaboration. *This is a collaborative project with the German Research Centre for Artificial Intelligence — DFKI (Germany).*

Project Acronym: ENGAGE

Project Title: NExt GeNeration ComputinG Environments for Artificial intelliGence

Coordinator: Gabriel Antoniu

Duration: 2021–2024

The goal of this project is to leverage efficient collaboration of experts in the AI and HPC areas 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?

In 2022 we proposed a workflow to support efficient deployment of Federated Learning (FL) systems across the Computing Continuum through formal descriptions of the underlying infrastructure, hyperparameter optimization and model retraining in case of performance degradation. This approach is taking advantage of E2Clab [12] to automate the deployment of FL experiments. Moreover personalization approaches are targeted to better fit the needs of each federated device.

9.3.3 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 [17]. 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

HPC-Big Data Inria Inria Challenge (ex-IPL)

Participants: Daniel Rosendo, Gabriel Antoniu, Alexandru Costan.

Collaboration. *This work has been carried out in close co-operation with Pedro De Souza Bento Da Silva, formerly a post-doc student in the team, and now at Hasso Plattner Institute, Berlin, Germany.*

Project Acronym: HPC-BigData

Project Title: The HPC-BigData INRIA Challenge
Coordinator: Bruno Raffin
Duration: 2018–2022
Web site: project.inria.fr/hpcbigdata

The goal of this HPC-BigData IPL is to gather teams from the HPC, Big Data and Machine Learning (ML) areas to work at the intersection between these domains. Research is organized along three main axes: high performance analytics for scientific computing applications, high performance analytics for big data applications, infrastructure and resource management. Gabriel Antoniu is a member of the Advisory Board and leader of the Frameworks work package.

In 2022 we extended the E2Clab framework for reproducible experimentation on the edge-to-cloud continuum with support for multiple testbeds and metadata management (see Section 8.2).

ADT Damaris 2

Participants: Joshua Charles Bowden, Gabriel Antoniu.

Project Acronym: ADT Damaris 2
Project Title: Technology development action for the Damaris environment
Coordinator: Gabriel Antoniu
Duration: 2019–2022
Web site: project.inria.fr/damaris

This action aims to support the development of the Damaris software. Inria's *Technological Development Office* (D2T, *Direction du Développement Technologique*) provided 3 years of funding support for a senior engineer.

In April 2020, Joshua Bowden was hired on this position. He introduced a support for un-structured mesh model types to Damaris. This capability opens up the use of Damaris for a large number of simulation types that depend on this data structure, in the areas of Computational Fluid dynamics, which has applications in energy production and combustion modeling, electric modeling and atmospheric and flow.

The capability has been developed and tested using Code_Saturne, a finite volume computational fluid dynamics (CFD) simulation environment. Code_Saturne is an open source CFD modeling environment which supports both single phase and multi-phase flow and includes modules for atmospheric flow, combustion modeling, electric modeling and particle tracking. This work is being validated on PRACE Tier-0 computing infrastructure in the framework of the PRACE-6IP project.

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

10.1.1 Scientific events: organisation

General chair, scientific chair

François Tessier: Co-Chair of [SuperCompCloud](#), the 6th Workshop on Interoperability of Supercomputing and Cloud Technologies, held in conjunction with SC'22, Dallas (TX, USA).

Alexandru Costan: Co-Chair of [FlexScience](#), 12th Workshop on AI and Scientific Computing at Scale using Flexible Computing Infrastructures, held in conjunction with ACM HPDC'22, Minneapolis (MS,USA).

Gabriel Antoniu: Steering Committee Chair of the [ESSA](#) Workshop series on High-Performance Storage, held in conjunction with the IEEE IPDPS conference since 2020. General Co-Chair of ESSA'22.

Guillaume Pallez: [SC'24](#) Technical program chair.

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.
- Poster and Student Program Co-Chair of [ICPP 2022](#), the International Conference on Parallel Processing.

Alexandru Costan:

- Co-chair of the [SC'23](#) SRC Graduate Posters
- Program Committee Co-chair of [ISPDC 2023](#), 22nd IEEE International Symposium on Parallel and Distributed Computing
- Publicity Chair of IEEE [ISPDC 2022](#), 21st IEEE International Symposium on Parallel and Distributed Computing.

Guillaume Pallez: [SC'22](#) awards-vice chair, [Cluster'23](#) finance chair

Gabriel Antoniu: Panel organizer and moderator at the [14th JLESC workshop](#) in Urbana, IL, USA, on HPC/Cloud convergence.

10.1.2 Scientific events: selection

Chair of conference program committees

François Tessier: Organizer of a mini-symposium about "Storage Systems at Extreme Scale" at the [PASC'22 conference](#) — Platform for Advanced Scientific Computing, digital event.

Member of the conference program committees

François Tessier: ComPAS 2022, HPCAsia 2022, ESSA 2022, SC'22 (Invited Speakers Committee)

Alexandru Costan: IEEE/ACM SC'22 (Clouds and Distributed Computing track, Posters and ACM Student Research Competition), IEEE/ACM UCC 2022, IEEE Big Data 2022, IEEE CloudCom 2022, IEEE ISPDC 2022, BDA 2022.

Gabriel Antoniu: ACM/IEEE SC22, IEEE CCGrid 2022, HPC Asia 2022

Guillaume Pallez: IPDPS'23.

Reviewer

François Tessier: IEEE/ACM CCGrid 2022, IEEE IPDPS 2022

Alexandru Costan: IEEE/ACM CCGrid 2022, IEEE IPDPS 2022, ACM HPDC 2022

Thomas Bouvier: IEEE/ACM CCGrid 2022, IEEE/ACM SC22, IEEE Big Data 2022

Julien Monniot: IEEE Cluster22, IEEE/ACM SC22

Cédric Prigent: IEEE/ACM SC22, IEEE/ACM UCC 2022

10.1.3 Journal

Member of the editorial boards

Alexandru Costan: Frontiers in HPC

Reviewer - reviewing activities

François Tessier: IEEE Transactions on Parallel and Distributed Systems - SC'22 Reproducibility Journal Special Issue

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

François Tessier:

- Panelist at the **14th JLESC workshop** in Urbana, IL, USA, to discuss HPC/Cloud convergence
- Short talk entitled "Investigating allocation of heterogeneous storage resources on HPC systems" given during the **14th JLESC workshop** in Urbana, IL, USA

Alexandru Costan:

- Short talk entitled "Supporting Efficient Workflow Deployment of Federated Learning Systems on the Computing Continuum" given during the given during the **14th JLESC workshop** in Urbana, IL, USA

Guillaume Pallez was invited to give the talk: "Model Accuracy in HPC System Software Algorithmic" at **CCDSC'22**

Gabriel Antoniu:

- Short talk entitled "Towards Integrated Hardware/Software Ecosystems for the Edge-Cloud-HPC Continuum: the Transcontinuum Initiative" given during the given during the **14th JLESC workshop** in Urbana, IL, USA

Julien Monniot:

- Lightning talk on StorAlloc at **Per3S Workshop**, IMT Palaiseau, France

Daniel Rosendo:

- Project talk entitled "Towards a Collaborative Environment for the Cost-effective Reproducibility of Edge-to-Cloud Experiments on Multi-platforms" given during the [14th JLESC workshop](#) in Urbana, IL, USA
- Short talk entitled "Reproducible Performance Optimization of Complex Applications on the Edge-to-Cloud Continuum" given during the [38th BDA Conference](#) in Clermont-Ferrand, France

Thomas Bouvier:

- Project talk entitled "Data Parallel Rehearsal-based Continual Learning" given during the [14th JLESC workshop](#) in Urbana, IL, USA.

10.1.5 Leadership within the scientific community**Gabriel Antoniu:**

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 [17]).

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).

Technology development project management: Coordinator of the Damaris 2 ADT project (2019–2022).

National project management: *Coordinator* of ExaDoST, one of the 5 targeted projects of the NumPEX PEPR project (accepted in 2022, to start in 2023, budget). Coordinator of STEEL, one of the 7 high-priority projects of the CLOUD PEPR project (accepted in 2022, to start in 2023).

François Tessier:

- Work package co-leader with Francieli Zanon-Boito (Associate Professor, University of Bordeaux) within the [NumPEX ExaDoST](#) project.
- Coordinator for KerData of the team's contributions to the Inria/ ATOS Grand Challenge

Alexandru Costan:

- Work package leader within the ENGAGE project.
- International Associate Team management: Leader of the [SmartFastData Associate Team](#) with Instituto Politécnico Nacional, Mexico City (2019–2022).

10.2 Teaching - Supervision - Juries**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): Big Data, 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.
- Master (Engineering Degree, 5th year): Big Data, 16 hours (lectures), M2 level, IMT Atlantique, Nantes.

François Tessier:

- Bachelor: Computer science discovery, 18 hours (lab sessions), L1 level, DIE Module, ISTIC, University of Rennes.
- Master: Cloud Computing and Big Data, 14 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.

Daniel Rosendo:

- Master: Miage BDDA, 24 hours (lab sessions), M2, ISTIC Rennes.
- Bachelor: Algorithms for Big Data, 10 hours (lectures, lab sessions), INSA Rennes.

Thomas Bouvier:

- Master: Stream Processing, 12 hours (lectures and lab sessions), M2, INSA Rennes.
- Bachelor: Introduction to Algorithms, 42 hours (lectures and lab sessions), L1, INSA Rennes.

10.2.2 Supervision

PhD in progress:

- Daniel Rosendo, "*Enabling HPC-Big Data Convergence for Intelligent Extreme-Scale Analytics*", thesis started in October 2019, co-advised by Gabriel Antoniu, Alexandru Costan and Patrick Valduriez (Inria).
- Thomas Bouvier, "*Reproducible deployment and scheduling strategies for AI workloads on the Digital Continuum*", thesis started in January 2021, co-advised by Alexandru Costan and Gabriel Antoniu.
- Julien Monnot, "*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).
- Robin Boezennec, "*Vers de nouvelles représentations des applications HPC*". Started in September 2022, co-advised by Guillaume Pallez and Fanny Dufossé (Datamove, Inria).

Internships:

- Valentin Haudiquet: Student in 4th year of ENS Rennes. Research project (6h/week for 6 months) on the use of node-local SSD-based caches to accelerate I/O on HPC systems. Co-advised by François Tessier and Gabriel Antoniu.
- Juliette Fournis d'Albiat (ENS Rennes): *Supporting Seamless Execution of HPC-enabled Workflows Across the Computing Continuum*, 10-month internship started in October 2021 (6 months in France, 4 months in Spain), co-advised by Alexandru Costan, Gabriel Antoniu, François Tessier and Rosa Badia, Jorge Ejarque (Barcelona Supercomputing Center, Spain).

10.2.3 Juries

François Tessier: Member of the mid-term evaluation jury (CSI) of the thesis of Amal Geroudji, PhD student at the Maison de la Simulation

Alexandru Costan: GDR RSD - Member of the juries for the PhD award and Young Researcher award.

10.3 Popularization**10.3.1 Internal or external Inria responsibilities****François Tessier:**

- Member of the sustainable development committee within the Inria center
- Member of the jury for the allocation of incentive resources at the Inria center in Rennes (*Moyens incitatifs*)

Alexandru Costan:

- In charge of internships at the Computer Science Department of INSA Rennes.
- In charge of the organization of the IRISA D1 Department Seminars.
- In charge of the management of the KerData team access to Grid'5000.

Guillaume Pallez is an elected member of the Inria evaluation committee.

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

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](https://doi.org/10.5281/zenodo.5534464). URL: <https://hal.archives-ouvertes.fr/hal-03358930>.
- [2] N. Cherièr, 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](https://doi.org/10.1016/j.jpdc.2020.02.001). URL: <https://hal.archives-ouvertes.fr/hal-02961875>.
- [3] 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](https://doi.org/10.1145/2987371). URL: <https://hal.inria.fr/hal-01353890>.

- [4] M. Dorier, S. Ibrahim, G. Antoniu and R. Ross. ‘Using Formal Grammars to Predict I/O Behaviors in HPC: the OmniscIO Approach’. In: *IEEE Transactions on Parallel and Distributed Systems* (2016). DOI: [10.1109/TPDS.2015.2485980](https://doi.org/10.1109/TPDS.2015.2485980). URL: <https://hal.inria.fr/hal-01238103>.
- [5] 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](https://doi.org/10.1609/aaai.v34i01.5376). URL: <https://hal.archives-ouvertes.fr/hal-02373429>.
- [6] M. Malms, L. Cargemel, E. Suarez, N. Mittenzwey, M. Duranton, S. Sezer, C. Prunty, P. Rossé-Laurent, M. Pérez-Hernandez, M. Marazakis, G. Lonsdale, P. Carpenter, G. Antoniu, S. Narasimhamurthy, 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](https://doi.org/10.5281/zenodo.7347009). URL: <https://hal.science/hal-03937825>.
- [7] M. Malms, M. Ostasz, M. Gilliot, P. Bernier-Bruna, L. Cargemel, E. Suarez, H. Cornelius, M. Duranton, B. Koren, P. Rosse-Laurent, M. S. Pérez-Hernández, 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*. ETP4HPC White Papers. 3rd Mar. 2020. DOI: [10.5281/zenodo.4605343](https://doi.org/10.5281/zenodo.4605343). URL: <https://hal.inria.fr/hal-03354396>.
- [8] O.-C. Marcu, A. Costan, G. Antoniu, M. S. Pérez-Hernández, B. Nicolae, R. Tudoran and S. Bortoli. ‘KerA: Scalable Data Ingestion for Stream Processing’. In: ICDCS 2018 - 38th IEEE International Conference on Distributed Computing Systems. Vienna, Austria: IEEE, 2nd July 2018, pp. 1480–1485. DOI: [10.1109/ICDCS.2018.00152](https://doi.org/10.1109/ICDCS.2018.00152). URL: <https://hal.inria.fr/hal-01773799>.
- [9] 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](https://doi.org/10.1109/cluster.2016.22). URL: <https://hal.inria.fr/hal-01347638>.
- [10] 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](https://doi.org/10.1016/j.future.2018.07.035). URL: <https://hal.archives-ouvertes.fr/hal-01892682>.
- [11] P. Matri, A. Costan, G. Antoniu, J. Montes and M. S. Pérez. ‘Tyr: Blob Storage Meets Built-In Transactions’. In: IEEE ACM SC16 - The International Conference for High Performance Computing, Networking, Storage and Analysis 2016. Salt Lake City, United States, 13th Nov. 2016. DOI: [10.1109/sc.2016.48](https://doi.org/10.1109/sc.2016.48). URL: <https://hal.inria.fr/hal-01347652>.
- [12] 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](https://doi.org/10.1109/CLUSTER49012.2020.00028). URL: <https://hal.archives-ouvertes.fr/hal-02916032>.
- [13] Y. Taleb, R. Stutsman, G. Antoniu and T. Cortes. ‘Tailwind: Fast and Atomic RDMA-based Replication’. In: ATC ‘18 - USENIX Annual Technical Conference. Boston, United States, 13th July 2018, pp. 850–863. URL: <https://hal.inria.fr/hal-01676502>.

11.2 Publications of the year

International journals

- [14] D. Rosendo, A. Costan, P. Valduriez and G. Antoniu. ‘Distributed intelligence on the Edge-to-Cloud Continuum: A systematic literature review’. In: *Journal of Parallel and Distributed Computing* 166 (Aug. 2022), pp. 71–94. DOI: [10.1016/j.jpdc.2022.04.004](https://doi.org/10.1016/j.jpdc.2022.04.004). URL: <https://hal.archives-ouvertes.fr/hal-03654722>.

International peer-reviewed conferences

- [15] 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, 22nd Aug. 2022. URL: <https://hal.inria.fr/hal-03683568>.
- [16] C. Prigent, L. Cudennec, A. Costan and G. Antoniu. ‘A Methodology to Build Decision Analysis Tools Applied to Distributed Reinforcement Learning’. In: *ScaDL 2022 - Scalable Deep Learning over Parallel and Distributed Infrastructures - An IPDPS Workshop*. Lyon / Virtual, France, 3rd June 2022, pp. 1–10. URL: <https://hal.science/hal-03613558>.

Scientific books

- [17] M. Malms, L. Cargemel, E. Suarez, N. Mittenzwey, M. Durantou, S. Sezer, C. Prunty, P. Rossé-Laurent, M. Pérez-Hernandez, 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](https://doi.org/10.5281/zenodo.7347009). URL: <https://hal.science/hal-03937825>.

Edition (books, proceedings, special issue of a journal)

- [18] O. Tatebe, G. Antoniu, K. Sato and M. Emani, eds. *3rd IEEE International Workshop on Extreme-Scale Storage and Analysis (ESSA 2022)*. IEEE, 2022, pp. 1098–1099. DOI: [10.1109/IPDPSW55747.2022.00178](https://doi.org/10.1109/IPDPSW55747.2022.00178). URL: <https://hal.inria.fr/hal-03933407>.

Reports & preprints

- [19] J. Monniot, F. Tessier, M. Robert and G. Antoniu. *Supporting Dynamic Allocation of Heterogeneous Storage Resources on HPC Systems*. 1st Dec. 2022. URL: <https://hal.science/hal-03922866>.

Other scientific publications

- [20] J. Monniot, F. Tessier and G. Antoniu. ‘Modeling Allocation of Heterogeneous Storage Resources on HPC Systems’. In: *SC 2022 - International Conference for High Performance Computing, Networking, Storage, and Analysis (Posters)*. Dallas, United States, 13th Nov. 2022, pp. 1–1. URL: <https://hal.science/hal-03878252>.
- [21] C. Prigent, G. Antoniu, A. Costan and L. Cudennec. ‘Supporting Efficient Workflow Deployment of Federated Learning Systems across the Computing Continuum’. In: *SC 2022 - International Conference for High Performance Computing, Networking, Storage, and Analysis (Posters)*. Dallas, United States, 13th Nov. 2022. URL: <https://hal.inria.fr/hal-03878254>.

11.3 Cited publications

- [22] G. Antoniu, P. Valduriez, H.-C. Hoppe and J. KrÄEger. *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](https://doi.org/10.5281/zenodo.5534464). URL: <https://hal.archives-ouvertes.fr/hal-03358930>.
- [23] 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](https://doi.org/10.1177/1094342006070078). URL: <https://hal.inria.fr/hal-00684943>.
- [24] 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](https://doi.org/10.1016/j.future.2020.05.030).
- [25] 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 PI@ntNet experience’. In: *Multimedia Systems* 22.6 (2016), pp. 751–766. DOI: [10.1007/s00530-015-0462-9](https://doi.org/10.1007/s00530-015-0462-9). URL: <https://hal.inria.fr/hal-01182775>.
- [26] M. Malms, M. Ostasz, M. Gilliot, P. Bernier-Bruna, L. Cargemel, E. Suarez, H. Cornelius, M. Duranton, B. Koren, P. Rosse-Laurent, M. S. PÄ©rez-HernÄ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. Mar. 2020. DOI: [10.5281/zenodo.4605343](https://doi.org/10.5281/zenodo.4605343). URL: <https://hal.inria.fr/hal-03354396>.
- [27] *The TransContinuum Initiative vision paper*. 2020. URL: <https://www.etp4hpc.eu/tci-vision.html>.
- [28] *Chameleon Cloud*. 2021. URL: <https://www.chameleoncloud.org/>.
- [29] *Cybeletech - Digital technologies for the plant world*. 2021. URL: <https://www.cybeletech.com/en/home/>.
- [30] *ECMWF - European Centre for Medium-Range Weather Forecasts*. 2021. URL: <https://www.ecmwf.int/>.
- [31] *European Exascale Software Initiative*. 2013. URL: <http://www.eesi-project.eu/>.
- [32] *The European Technology Platform for High-Performance Computing*. 2012. URL: <http://www.etp4hpc.eu/>.
- [33] *International Exascale Software Program*. 2011. URL: <http://www.exascale.org/iesp/>.
- [34] *Inria’s strategic plan “Towards Inria 2020”*. 2016. URL: <https://www.inria.fr/fr/recherche-innovation-numerique>.