Offer #2024-07418

PhD Position F/M Decision models for the Edge-Cloud Computing Continuum

Contract type : Fixed-term contract

Level of qualifications required : Graduate degree or equivalent

Fonction : PhD Position

About the research centre or Inria department

The Inria Centre at Rennes University is one of Inria's eight centres and has more than thirty research teams. The Inria Centre is a major and recognized player in the field of digital sciences. It is at the heart of a rich R&D and innovation ecosystem: highly innovative PMEs, large industrial groups, competitiveness clusters, research and higher education players, laboratories of excellence, technological research institute, etc.

Context

Computing is shifting from the traditionally centralized cloud to a distributed set of heterogenous resources located between the edge, the cloud and in-between. As computing has moved to this *Computing Continuum*, the tradeoffs between performance, availability and cost have become increasingly complicated.

This thesis aims at contributing to the programming support and the resource management of emerging infrastructures and applications. The rationale that underlies the proposed research is the separation between the concerns of domain experts (familiar with the applications) and these of resource managers (familiar with the computing infrastructures).

PhD Advisor: Daniel Balouek (Inria)

This thesis will be hosted by the STACK team at the IMT Atlantique of Nantes. This thesis will also include collaborations with industrial and international partners, especially from the US.

Assignment

Introduction

With the advent of distributed infrastructures, the Cloud computing paradigm is progressively moving towards a full continuum from IoT devices and sensors to the centralized Cloud, with Edge (edge of the network) and Fog computing (core network) in between [1]. Simultaneously, distributed applications also evolve. Urgent computing tackles services that require time-critical decisions that improve quality of life, monitor civil infrastructures, respond to natural disasters and extreme events, and accelerate science (e.g., autonomous cars, disaster response, precision medicine, etc.). These services are typically sensitive to latency and response time and are among the best candidate for the IoT to Cloud computing continuum [2].

In this thesis, we consider a new breed of urgent intelligent services using the IoT-to-Cloud Continuum, combined with the recent advances in Artificial Intelligence and Big Data Analytics. First, these services and applications require a large computing capacity to perform well, while often being under the constraints to move data from the edge of the network to the Cloud [3]. Second, these services and applications require system support to program reactions that occur at runtime, especially when the target infrastructure capacities and capabilities are unknown during the design [4].

Thesis Proposal

There is a critical gap in the knowledge base that pertains to programming urgent analytics while managing important trade-offs of cost and quality. The successful candidate will be in charge of proposing and evaluating decision models integrating the content of data, cost of computations, and urgency of the results for the management of applications deployed across edge and cloud resources.

References

[1] Daniel Balouek-Thomert, Eduard Gibert Renart, Ali Reza Zamani, Anthony Simonet, and Manish Parashar. Towards a computing continuum : Enabling edge-to-cloud integration for data-driven workflows. The International Journal of High Performance Computing Applications, 33(6) :1159–1174, 2019.

[2] Daniel Balouek-Thomert, Ivan Rodero, and Manish Parashar. Harnessing the computing continuum for urgent science. SIGMETRICS Perform. Eval. Rev., 48(2) :41–46, November 2020.

[3] Kevin Fauvel, Daniel Balouek-Thomert, Diego Melgar, Pedro Silva, Anthony Simonet, Gabriel Antoniu, Alexandru Costan, Véronique Masson, Manish Parashar, Ivan Rodero, and Alexandre Termier. A distributed multi-sensor machine learning approach to earthquake early warning. Proceedings of the AAAI Conference on Artificial Intelligence, 34(01) :403–411, Apr. 2020.

[4] Eduard Gibert Renart, Daniel Balouek-Thomert, and Manish Parashar. An edge-based framework for enabling data-driven pipelines for iot systems. In 2019 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), pages 885–894, 2019.

Main activities

The main challenge consists in proposing novel models for facilitating decision-making models for edge-to-cloud applications, and deducing management policies based on their impact. The research work can be decomposed into two main objectives:

- The first objective is to build and validate a set of models describing the dynamics of the computing continuum.
- The second objective is to augment the common operations of data-driven analytics (e.g. collection, filtering, processing, delivery) with tunable parameters.

Through these two objectives, the successful candidate is expected to propose approaches for building intelligent services that manage the availability and efficiency of the infrastructure. Building on analytical models and tunable software abstractions, this project will inform infrastructure managers and application developers with insights on what data and services to run, where to run them, and how to run them across the Computing Continuum.

We expect the successful candidate to create repeatable processes and artifacts that will be used at scale to develop and evolve edge computing designs. Experiments and validation will occur on Grid'5000, the biggest share network in Europe dedicated to research in Computer Science.

Skills

- A Master degree in computer science or equivalent
- Solid background in Software Engineering and Distributed Systems
- Completion of a teaching unit in high-performance computing or distributed computing is an advantage
- Knowledge of edge computing, streaming systems or modeling approaches based on graphs or Petri Nets is an asset
- Programming skills in C/C++ and Python
- Good communication skills in oral and written English, with the aim to publish and present research results in high-level international journals and conferences
- Autonomous, curious and strongly motivated candidates are expected

Benefits package

- Subsidized meals
- Partial reimbursement of public transport costs
- Possibility of teleworking (90 days per year) and flexible organization of working hours
- Partial payment of insurance costs

Remuneration

Monthly gross salary amounting to 2100 euros for the first and second years and 2190 euros for the third year

General Information

- **Theme/Domain :** Distributed Systems and middleware
System & Networks (BAP E)
- **Town/city :** Nantes
- **Inria Center :** Centre Inria de l'Université de Rennes
- **Starting date :** 2024-10-01
- **Duration of contract :** 3 years
- **Deadline to apply :** 2024-05-24

Contacts

- **Inria Team :** STACK
- **PhD Supervisor :**
Balouek Daniel / daniel.balouek-thomert@inria.fr

About Inria

Inria is the French national research institute dedicated to digital science and technology. It employs 2,600 people. Its 200 agile project teams, generally run jointly with academic partners, include more than 3,500 scientists and engineers working to meet the challenges of digital technology, often at the interface with other disciplines. The Institute also employs numerous talents in over forty different professions. 900 research support staff contribute to the preparation and development of scientific and entrepreneurial projects that have a worldwide impact.

Instruction to apply

Please submit online : your resume, cover letter and letters of recommendation eventually

Defence Security :
This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical potential (PPST).Authorisation to enter an area is granted by the director of the unit, following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

Recruitment Policy :
As part of its diversity policy, all Inria positions are accessible to people with disabilities.