2018-00680 - A Data-driven microservices architecture for Deep Learning applications (Campagne Doctorant Grenoble Rhône-Alpes PhD Campaign)

Contract type : Public service fixed-term contract
Level of qualifications required : Graduate degree or equivalent
Function : PhD Position

About the research centre or Inria department
Grenoble Rhône-Alpes Research Center groups together a few less than 800 people in 35 research teams and 9 research support departments.

Staff is localized on 5 campuses in Grenoble and Lyon, in close collaboration with labs, research and higher education institutions in Grenoble and Lyon, but also with the economic players in these areas.

Present in the fields of software, high-performance computing, Internet of things, image and data, but also simulation in oceanography and biology, it participates at the best level of international scientific achievements and collaborations in both Europe and the rest of the world.

Context
Localization : Ecole Normale Supérieure de Lyon

Encadrants :
Eddy Caron (Ens de Lyon, Inria équipe Avalon)
Daniel Balouek-Thomert (Rutgers University. USA New Jersey. RD2I)
Eddy.Caron@inria.fr <mailto:Eddy.Caron@inria.fr>
http://graal.ens-lyon.fr/~ecaron

International reputation: We plan to improve the visibility of this CORDIS Phd across the collaboration built-in Inria associate team SUSTAM. Moreover the shared knowledges of the both team will be very fruitful.

Assignment
While our capacity for collecting data is expanding dramatically, our ability to manage, manipulate, and analyze this data to transform it into knowledge and understanding has not kept pace [AZZ+17].

This research work aims at realizing a fluid ecosystem where distributed resources and microservices are aggregated on-demand using deep learning technologies to support emerging data-driven applications.

We have been finding better ways to build systems, helped by the mature aspects of Cloud technologies and the emergence of Big Data. The concept of microservices [DGL+17] is a trend pattern that emerged from real world use, at the intersection of distributed systems and Service-Oriented Architecture [Erl05]. It promotes the use of finely grained services with their own life cycles, which collaborate together. These services need to be able to change independently of each other, and be deployed by themselves. Moreover microservices is a good candidate for the needs of Edge computing platform [SCQ+16].

Large distributed systems are a complex interaction of heterogeneous resources, changing demands and data products. The large number of possible operating configurations and nonlinear interdependencies make it difficult to understand and optimize data-driven decisions. In recent years, machine learning technologies have demonstrate to be an effective way to leverage performance in pattern recognition, feature extraction and classification [Sch15]. We aim to use such techniques to continuously generate infrastructure configurations that matches application performance and quality of experience.

General Information
- Theme/Domain : Distributed and High Performance Computing
- Town/city : Lyon
- Inria Center : CRI Grenoble - Rhône-Alpes
- Starting date : 2018-10-01
- Duration of contract : 3 years
- Deadline to apply : 2018-04-30

Contacts
- Inria Team : AVALON
- Recruiter : Caron Eddy / eddy.caron@inria.fr

The keys to success
Master degree.

Knowledge on Distributed Systems and/or Cloud and Edge platform and/or Data Management and/or Deep learning will be appreciate.

Conditions for application
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.

Warning : you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.
This research work calls for novel solutions for programming applications and services, which are capable of reacting in real-time to unpredictable data products. This becomes particularly challenging when providing scalable applications with changing demands and bounded response time while maintaining available services according to resource and users events. In this context, we aim at formalizing and evaluating the end-to-end performance and data management of microservices architectures.

Such architecture could benefit from the use of machine learning techniques by expressing complex relations between users requests and data products. This would enable the establishment of predictive models and failure patterns to automate the process of continuous composition of services.

We believe that such hybrid approach of microservices architectures and deep learning could benefit infrastructure providers (1) and service users (2):

(1) Independent, yet collaborative services will ease the development and adoption of future Cloud-natives and IoT applications. Some large, successful organizations like Amazon and Google espoused the view of small teams owning the full lifecycle of their services. And, more recently, Netflix has shared with us ways of building antifragile systems at a scale that would have been hard to comprehend just 10 years ago [BMM16].

(2) The low granularity of the microservice give a modular and an efficient solution. With this kind of architecture we propose a solution where the resource management are fully hidden for the final user. Thus service catalogs could be based entirely on data information.

Bibliography:


Keywords : Microservices, Modelization, Data streaming, Deep Learning, Edge computing

Main activities

Our goal is to formalize principles of a microservices infrastructure that integrates deep learning features and build a prototype able to support industrial (preliminary discussions are in progress) and scientific use cases.

Skills

Technical skills and level required :

Languages :

Relational skills :

Other valued appreciated :

Benefits package

- Subsidised catering service
- Partially-reimbursed public transport
- Social security
- Paid leave
- Flexible working hours
- Sports facilities

**Remuneration**

Gross salary: 1982 € per month (1st year), 2085 € 2nd & 3rd years.