Automating application deployment, scaling, and management in geo-distributed infrastructures: The Kubernetes case.

Start : ASAP
Funding : Inria
Director: Adrien Lebre, Prof. IMT Atlantique,
Stack research group (IMT Atlantique, Inria, LS2N)
http://stack.inria.fr

Keywords: resource management, distributed systems, edge computing, network disconnections, Kubernetes

Context

While it is clear that edge infrastructures are required for emerging use-cases related to smart-* applications, Industry 4.0, Virtual Reality, and others latency critical applications, there is currently no resource management system able to deliver all the features that make the success of cloud computing for a set of smaller data centers deployed at the edge of the internet, and with intermittent connectivity.

Although new initiatives have been proposed since our alert two year agos [1] (StarlingX [2], KubeEdge [3], and Kubefed [4] to name a few), the situation has not really evolved: proposals are still based either on a centralized approach or a federation of independent managers. The former lies on operating and exposing an edge infrastructure as a traditional single data center environment, the key difference being the wide-area network found between the control and compute nodes [6] . The latter consists in deploying one manager per site of the edge infrastructure and federating them through a brokering approach to give the illusion of a single coherent system as promoted by ETSI [7]. Due to frequent isolation risks of an edge site from the rest of the infrastructure, the federated approach presents a significant advantage: each site can continue to operate locally. However, the vanilla code of the manager does not provide any mechanism to deliver such a federation. In other words, delivering the expected global vision requires important development efforts.
To mitigate the overhead related to the geo-distribution aspects (i.e. the federation mechanisms), STACK members proposed in 2019 to investigate a new composition model aiming at delivering on demand collaboration between multiple instances of the same manger. By leveraging dynamic composition of services and a Domain Specific Language (DSL), Admins/DevOps can specify, on a per-request basis, which services from which instances should be used to execute the request. This information is then interpreted to dynamically recompose services between the different instances thanks to a reverse proxy. A first proof-of-concept has been implemented on the OpenStack framework, the defacto open source solution to operate and use data centers [8, 12].

Ultimately, the model we have in mind would enable developpers to deal with geo-distributed concerns after the implementation of the business logic. In other words, developpers would be able to program a complex application such as a resource management system without taking care whether the application will be deployed through one or mutliple sites : one instance of an application will be deployed on each location and the collaboration between the instances will be performed on demand and according to each situation. This a major change with respect to current proposals that investigate how a global view of the same application can be maintained permanently through mutlitple locations.

Following the encouraging results we obtained, we extended our initial study through a first PhD (granted by Inria 2019/2022). More precisely, we are extending the initial model with location operators (e.g., &, |) to enable the execution of multiple-site requests.

In this new PhD, we propose to investigate how two instances of the same service can collaborate. Indeed, this kind of collaboration cannot be addressed by dynamic composition as it consists generally in extending a service deployed on one site to one of its twin deployed on another site.

To move forward to our ultimate target, we took the assumption that such a collaboration between the same service was not critical and could be addressed later. Unfortunately, our current study on Kubernetes, the defacto open source management system for containers, has revealed the importance of such a capability. Although the architecture of Kubernetes follows the services-programming principle, all services are hidden behind a single one. In such as situation, the collaboration between two instances of Kubernetes can only be achieved by offering collaboration means between instances of the same service.

Objectives

The PhD position will focus on the design and implementation of new buidling blocks that enable the exposition of one resource created on one instance of an application to another one. Such an operation lies in the collaboration of multiple instances of the same service of an application.

As opposed to the state-of-the-art approaches that have been focusing on designing “glue” components in the form of a hierarchical broker, the innovative part lies in delivering generic building blocks that can be used with minimal code changes in the business logic of the application. While the collaboration between distinct services of different managers can be done without changes, our initial investigations [9] demonstrated that it would be mandatory to follow some principles in the way services of an application are coded to successfully separate geo-distributed concerns from the business logic.

Besides, the proposed building blocks should follow on initial assumption which is to perform collaboration only on demand and between the relevant instances. The creation of a global knowledge base is to be prescribed. The Phd student will investigate to which extent conflict-free replicated data type (CRDT) [10] can satisfy this need.

The pilot use-case would be the namespace resources of Kubernetes [13] with the ultimate goal of being able to reattach a resource created on one instance to another one.

It is noteworthy that this work is tighly coupled with two ongoing PhDs granted by Inria and Orange Labs as well as the activities of STACK members. Our ultimate goal is (i) to make the implementation of tomorrow’s applications easier and (ii) to) deliver a framework capable of managing the life cycle of such applications across infrastructures where the location of computations/data is a key element and where the disconnection from the Internet would be the norm [11].

References

[1] R. A. Cherrueau, A. Lebre, D. Pertin, F. Wuhib, and J. Monteiro Soares. Edge computing resource management system: a critical building block! Initiating the debate via openstack. In on Proceedings of Hot Topics in Edge Computing, (HotEdge Series), Boston, MA, July 10, 2018. USENIX Association, 2018.
[2] StarlingX, a complete cloud infrastructure software stack for the edge. https://www.starlingx.io. Accessed: 03/2020.
[3] KubeEdge, a Kubernetes Native Edge Computing Framework. https://kubernetes.io/blog/2019/03/19/kubeedge-k8s-based-edge-intro. Accessed: 03/2020.
[4] KubeFed, Kubernetes Cluster Federation. https:// github.com/kubernetes-sigs/kubefed. Accessed: 03/2020.
[6] R. A. Cherrueau, A. Lebre,and P. Riteau. Toward Fog, Edge, and NFV Deployments: Evaluating OpenStack WANwide. OpenStack Summit, Boston, USA, https://www.youtube.com/watch?v=xwT08H02Nok, May 2017. Accessed: 03/2020.
[7] D. Sabella, A. Vaillant, P. Kuure, U. Rauschenbach, and F. Giust. Mobile-Edge Computing architecture: The role of MEC in the Internet of Things. *IEEE Consumer Electronics Magazine*, 5(4):84–91, Oct 2016
[8] R. A. Cherrueau ,J. Rojas Balderrama, and A. Lebre.OpenStackOïd: Collaborative OpenStack Clouds On-Demand. Open Infrastructure Summit, Denver USA, https://www.youtube.com/watch?v=kDcwToguKK0, May 2019. Accessed: 02/2020.
[9] D. Espinel Sarmiento, A. Chari, L. Nussbaum, and A. Lebre. Multi-site Connectivity for Edge Infrastructures - DIMINET:DIstributed Module for Inter-site NETworking. To appear in the proceedings of the 20th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID), Nov 2020, Sydney, Australia.
[10] M. Shapiro,N. Preguiça, C. Baquero, and M. Zawirski. Conflict-free replicated data types. In *Symposium on Self-Stabilizing Systems* Oct. 2011, Grenoble, France.
[11] G. Tato, M. Bertier, E. Rivière, C. Tedeschi. Split and migrate: Resource-driven placement and discovery of microservices at the edge. OPODIS 2019 : 23rd International Conference On Principles Of Distributed Systems, Dec 2019, Neuchâtel, Switzerland.
[12] RA Cherrueau, M Delavergne, A Lebre, JR Balderrama, M Simonin, Edge Computing Resource Management System: Two Years Later! Inria RR 9336, https://hal.inria.fr/hal-02527366
[13] K. Manaouil, A. Lebre, Kubernetes and the Edge? Inria RR 9370 https://hal.inria.fr/hal-02972686/

Skills and profiles

Knowledge/experience on programming languages and software engineering
Knowledge/experience on distributed systems (OpenStack/Kubernetes would be a plus)
Experimentation skills
Autonomy / Curiosity
English mandatory

Additional information: The candidates are invited to contact adrien.lebre@inria.fr.

Duration: 36 months Location: Nantes, France

Monthly gross salary of around 2000 euros (~1550 after taxes)