Locality-Aware Placement in Geographically Distributed Clouds

September 2017

Information

- Location: IRISA / Inria (Rennes, France)
- Team: Myriads (`https://team.inria.fr/myriads/`)
- Advising:
  - Matthieu Simonin (Research Engineer, Inria)
  - Cédric Tedeschi (Associate Professor, U. Rennes 1)
- Contact:
  - `matthieu.simonin@inria.fr`
  - `cedric.tedeschi@univ-rennes1.fr`
- Keywords:
  - Tactile Internet
  - Fog / Edge Computing
  - Distributed algorithms
  - Placement

Context

To satisfy requirements of the Industrial and Tactile Internet applications regarding very low latencies (for applications using augmented reality or haptic interactions in domains such as culture, healthcare and robotics), academics and industry experts advocate for moving from large, centralized infrastructures to federations of smaller ones massively distributed at the edge of the network. In other words, these infrastructures envision to put resources as close as possible to the clients consuming them. To favor the adoption of this decentralized model (referred to as "Fog/Edge Computing"), the development of a system in charge of turning a complex and diverse network of resources into a global unified infrastructure is mandatory. The architecture of such an infrastructure is currently an open research topic.

Towards this goal, the Discovery initiative [2], led by Inria, gathers both academic, industrial partners (from Orange) and network infrastructure providers (from Renater). More broadly speaking, the Discovery project is active in the OpenStack [1] Community which gathers a larger audience from Cloud operators to large telecommunication companies.

The applications targeted by the next generation of utility computing platforms advocated by the Discovery project need to cope with several dimensions brought by the users: they need to experience very low-latency interactions with the service hosted on the platform, while being mobile and geographically dispersed. As a consequence, a static placement of the applications and its service components is no longer suitable. This placement must now be dynamic and self-adapting while keeping a certain level of quality of services for the users.

Detailed Description

Few close problems have been addressed recently by the Discovery project around locality-awareness in VM placement (without dynamic adaptation) [5] or distributing OpenStack's database [4].

For the sake of illustration, let us consider the placement provided in Figure 1. A service located in Rennes accepts queries. At some point in time, the set of clients of this service tend to be geographically far from Rennes, calling for migration of the service closer to the client. Note that the Rennes-based service rely on some other service in Lille.

A new possible configuration that may improve delays and network utilization is displayed in Figure 2. In these new settings, the service has been migrated to some place minimizing the average delay between it and its clients (say in some cluster in Clermont-Ferrand).

Internship activity

This internship mixes a conceptual work about modeling the targeted platforms, specifying constraints to be tackled and designing distributed algorithms with development and (large-scale) experimentation to validate the conceptual contributions of the internship.

The envisioned phases of the internship are the following:

1. Devise a model of the targeted platform taking mobility and geographic dispersion into account

2. Design dynamic placement strategies minimizing average delay over the platform while being resilient to mobility

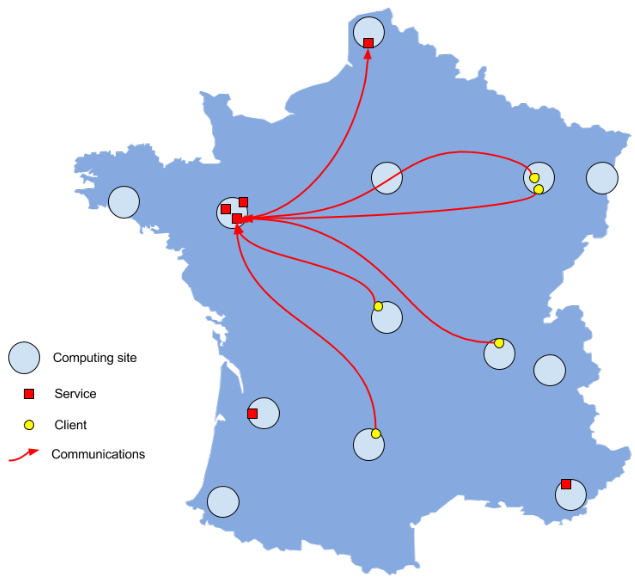

Figure 1: Initial placement of services and clients over the resources.

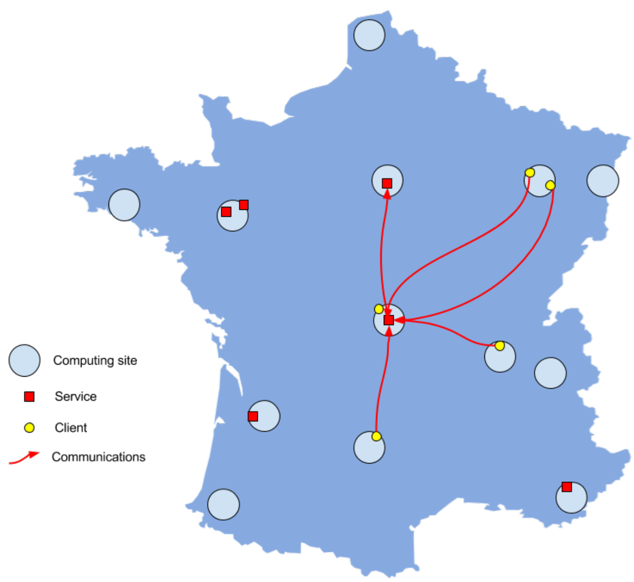

Figure 2: Placement of services and clients after reconfiguration.

3. Validate the algorithms defined over a real testbed such as the nation-wide Grid'5000 experimental platform [3]

References

[1] Openstack. `https://openstack.org`.

[2] Marin Bertier, Frédéric Desprez, Gilles Fedak, Adrien Lebre, Anne-Cécile Orgerie, Jonathan Pastor, Flavien Quesnel, Jonathan Rouzaud-Cornabas, and Cédric Tedeschi. *Beyond the Clouds: How Should Next Generation Utility Computing Infrastructures Be Designed?*, pages 325–345. Springer International Publishing, Cham, 2014.

[3] Raphaël Bolze, Franck Cappello, Eddy Caron, Michel Dayde, Frédéric Desprez, Emmanuel Jeannot, Yvon Jégou, Stephane Lanteri, Julien Leduc, Nouredine Melab, Guillaume Mornet, Raymond Namyst, Pascale Primet, Benjamin Quétier, Olivier Richard, Talbi El-Ghazali, and Iréa Touche. Grid'5000: A Large Scale And Highly Reconfigurable Experimental Grid Testbed. *International Journal of High Performance Computing Applications*, 20(4):481–494, 2006.

[4] Adrien Lèbre, Jonathan Pastor, Anthony Simonet, and Frédéric Desprez. Revising OpenStack to Operate Fog/Edge Computing Infrastructures. In *IEEE International Conference on Cloud Engineering (IC2E)*, pages 138–148, Vancouver, Canada, April 2017.

[5] Jonathan Pastor, Marin Bertier, Frédéric Desprez, Adrien Lebre, Flavien Quesnel, and Cédric Tedeschi. Locality-Aware Cooperation for VM Scheduling in Distributed Clouds. In *20th International Conference on Parallel Processing (EuroPar 2014)*, volume 8632, pages 330–341, Porto, Portugal, August 25-29 2014.