Laboratory : LSIIT, UMR 7005 – ICPS team
Doctoral school : MSII (no 269)
City : Strasbourg
Supervisors : Stéphane Genaud (HDR) and Julien Gossa
Contact : genaud@unistra.fr, gossa@unistra.fr
Subject : Cloud Computing for data driven applications

UNIVERSITÉ DE STRASBOURG

PhD Thesis Proposal
April, 20th 2010

Description :

Cloud computing [1, 2] is a remarkable evolution of distributed computing using sets of remote and abstract resources taken from clouds, which provide on-demand capacity as an utility. Most often, clouds are virtualized so as to provide more flexibility to exploit physical servers. Examples of such clouds are Amazon EC2 or the Science Clouds3. An interesting property of virtualization is also to allow embedding complete system images in virtual machines, allowing to easily execute legacy codes out of their original software environment. Thus, clouds provide a platform mainly characterized by on-the-fly extensibility.

On the other hand, mobile devices like sensors can provide real-time field data like temperatures, pressure, and humidity. Such data can be processed for a wide range of valuable applications like weather forecasting or emergency management. The amount of produced data is characterized by an important dynamicity, while the applications are characterized by different execution constraints like response time or resources reservation fees.

In this context, one particularly important issue concerns on-the-fly cloud resources provisioning. The provisioning should meet the application requirements and ensuire the expected performance while making a tradeoff with respect to the financial cost of using resources. Moreover, this problem implies to be able to transparently use different cloud platforms to meet their characteristics with the target application profile and requirements. This is made especially difficult by the heterogeneity of both cloud platforms and target applications, but also by the high dynamicity implied by data provided by sensors, under potential strong execution time constraints.

The main objectives of this PhD is to study the different available cloud solutions and to propose an environment able to transparently manage resources for dynamic-data-driven applications. The work will have to tackle specific issues such as handling on-the-fly provisioning of resources suitable to one given application, as well as their reservation, deployment and execution of processes.

Requirements :

- Knowledge and interests in distributed computing
- Good skills in C, Java and Unix

[1] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia. Above the clouds : A berkeley view of cloud computing. Technical Report Tech. Rep. UCB/EECS-2009-28, EECS Department, University of California, Berkeley, February 2009.

[2] Luis M. Vaquero, Luis Rodero-Merino, Juan Caceres, and Maik Lindner. A break in the clouds : towards a cloud definition. *SIGCOMM Comput. Commun. Rev.*, 39(1) :50–55, 2009.