

ED STIC - Proposition de Sujets de Thèse

pour la campagne d'Allocation de thèses 2011

Titre du sujet :

Mention de thèse :

HDR Directeur de thèse inscrit à l'ED STIC :

Co-encadrant de thèse éventuel :

Nom :

Prénom :

Email :

Téléphone :

Email de contact pour ce sujet :

Laboratoire d'accueil :

Description du sujet :

Cloud computing covers a wide range of applications, from online services for the end user (e.g., Google doc or dropbox) to full IT outsourcing (data and applications) for companies. A high level description of a cloud computing architecture consists of three key elements: (i) a computational/storage facility that generally consists of a bunch of high end servers that are heavily virtualized, (ii) an interconnection network within the datacenter to connect the physical servers to the Internet gateways (of the datacenter) with whom clients interfere, (iii) the paths between the gateway and the end users, that is not under the control of the entity which runs the cloud computing service, but is the major source of delay and can constitute the bottleneck of the overall system.

In the context of this thesis, we aim at investigating (some of) the performance problems raised by cloud computing architectures. Those problems relate to the three elements listed above. First, virtualization, which consists in packing several servers on the same physical servers, leads

to a number of problems. Several competing virtualization technologies exist (para-virtualization, full virtualization, container-based solutions) and a key point for those solutions in a datacenter context, is the technique used to share the physical network interfaces among the virtual machines (VMs). Also, with advanced virtualization techniques, some network segments are further virtualized, inside the virtual machine (inter VM communication) and between physical servers or even between datacenters in scenarios where several datacenters are involved. Within this thesis, we aim at investigating the scaling capacity of virtualization technologies when the number of VMs increases [1].

A second set of problems relates to the interconnection network of the datacenter. Several recent studies, e.g. [2], have suggested that the burstiness of traffic within datacenters put the transport layer (TCP) under pressure and needs special care. One intends to study the means by which such a limited but complex network can be monitored and how the performance of new TCP flavors can be tuned in such environments.

A third key component for cloud computing architecture is the network path from the datacenter to the end user premise. This component is a key one in the end-to-end performance as it imposes the end-to-end latency and restricts the end-to-end bandwidth. Several techniques have been proposed to mitigate the impact of the path characteristics, e.g., proxy data-center [3], which are small-scale datacenters located close to the client, and as such, offer a lower latency when transferring data. Whatever the solution is (single data-center or mix of a central and proxy datacenters), a precise assessment of the last miles characteristics is a key component. Using so-called active non-cooperative measurements techniques, e.g. [4], where measurements are carried out from the datacenter gateway without the explicit cooperation of the end user can be a key component in sustaining the service level agreement (SLA) of cloud computing architectures.

This thesis will be executed in cooperation with Orange Lab Sophia-Antipolis. The candidate should have a solid background in networking and programming.

References :

- [1] Omesh Tickoo, Ravi Iyer, Ramesh Illikkal, Don Newell, Modeling virtual machine performance: challenges and approaches SIGMETRICS Perform. Eval. Rev., Vol. 37, No. 3. (2009), pp. 55-60.
- [2] Theophilus Benson, Aditya Akella, and David A. Maltz, Network Traffic Characteristics of Data Centers in the Wild, in Internet Measurement Conference, Association for Computing Machinery, Inc., November 2010
- [3] A. Pathak, Y. A. Wang, C. Huang, A. Greenberg, Y. C. Hu, J. Li, K. W. Ross, Measuring and Evaluating TCP Splitting for Cloud Services, Passive and Active Measurement Conf (PAM), Zurich, 2010
- [4] D. Croce, T. En-Najjary, G. Urvoy-Keller and E. Biersack, Capacity Estimation of ADSL links, 4th ACM CoNEXT conference, Madrid, Spain, December 2008.

English version:

--