Title: Quality of Service in the Cloud

Summary
A Cloud can be seen as an execution environment of resources involving multiple stakeholders and providing tailored services at multiple levels of quality. In such environment, different roles for stakeholders can be identified: Cloud Providers that offer clouds to the customer and Cloud Services Providers that are hosts of cloud enhanced applications delivered as services over the Internet called Cloud Services or Software as a Service (SaaS) [1].

Cloud computing is a new research field in Internet computing that can deploy, allocate or reallocate computing resources dynamically and monitor their usage at execution time [2]. One of the new aspects in Cloud computing is the ability to pay-per-use of computing resources on a short-term basis as needed (e.g. processors by the hour, storage by the day,...) and release them as needed [3,4]. This pay-per-use model enables Cloud Services Providers to offer their services to customers in different Quality-of-Service (QoS) levels. Service Level Agreement (SLA) is a negotiated agreement between a Cloud Services Provider and a customer where QoS parameters specify the quality level of service that the Cloud Services Provider has to guarantee. However, due to the unreliable and dynamic nature of Cloud Services Providers and their environment (i.e. the Cloud), the quality values of Cloud Services could change frequently over time resulting in some SLA violations and penalties for the Cloud Services Provider. For instance, the Cloud Service response time at any particular moment could be significantly affected by the Cloud Service Provider load and network traffic at that moment. Such dynamism and uncertainty can lead to highly undesirable situations during Cloud Service execution (e.g. unfulfilled quality promises), and may request costly corrective actions.

The aim of this thesis is to develop solutions for QoS guarantee, which are able to dynamically adapt the resources of the Cloud in order to satisfy SLAs, by minimizing quality deviations of Cloud Services at execution time.

Advisory: Joyce El Haddad and Marta Rukoz

References
[7] Danilo Ardagna, Giuliano Casale, Michele Ciavotta, Juan F Pé