

SLA Based Profit Optimization in Autonomic Computing Systems

Li Zhang
IBM

T.J. Watson Research Center
P.O. Box 704, Yorktown Heights, NY 10598
zhangli@us.ibm.com

Daniilo Ardagna
Politecnico di Milano

Dipartimento di Elettronica e Informazione
Via Ponzio 34/5, 20133 Milano, Italy
ardagna@elet.polimi.it

ABSTRACT

With the development of the Service Oriented Architecture (SOA), organizations are able to compose complex applications from distributed services supported by third party providers. Under this scenario, large data centers provide services to many customers by sharing available IT resources. This leads to the efficient use of resources and the reduction of operating costs. Service providers and their customers often negotiate utility based Service Level Agreements (SLAs) to determine costs and penalties based on the achieved performance levels. Data centers often employ an autonomic computing infrastructure and use a centralized dispatch and control component (a dispatcher) to distribute the user requests to backend servers, and to set the scheduling policies at each server. This dispatcher can also decide to turn ON or OFF servers depending on the system load. This paper designs a set of dispatching and control policies for the dispatcher in such service oriented environments. The objective is to maximize the provider's profits associated with multiple class of SLAs. We show that the overall problem is NP-hard, and develop meta-heuristic solutions based on the tabu-search algorithm. Experimental results are presented to show the benefits of our approach.