Autonomic mechanisms for transactional replication in elastic cloud environments⋆

Paolo Romano

INESC-ID, Lisbon, Portugal

Abstract

Over the last years, we have witnessed the blistering proliferation of in-memory, transactional data platforms, often referred to as NoSQL data grids. Thanks to their ability to achieve higher performance, scalability and elasticity levels than classical relational database management systems, these platforms are emerging as the reference solution for data management in Cloud computing environments.

In these in-memory platforms, replication plays a role of paramount importance for fault-tolerance purposes, given that it represents the key means to ensure data durability in face of unavoidable nodes' failures. Unsurprisingly, the replication schemes employed by these platforms take inspiration from the vast literature on replication of transactional systems (both traditional database systems and transactional memory systems). Unfortunately, despite the abundance of literature in the area, existing replication architectures still fall short in addressing one of the key requirements of Cloud computing environments: ensuring optimal efficiency when deployed over elastic infrastructures that dynamically expand/reduce the number of (physical or virtualized) resources in response to fluctuations of the workload characteristics.

The point is that current in-memory transactional data platforms rely on single-mode replication strategies, each of which designed to ensure optimal performance only in presence of specific types of workloads and scales of the data platform, but can suffer from significant performance degradations when operating in unfavorable scenarios. At the light of these considerations, we advocate that, in order to match the elasticity requirements proper of Cloud computing infrastructures, and maximize efficiency (i.e. reduce operational costs in the Cloud's pay-for-what-you-use model), in-memory transactional data grid should entail multi-modal replication strategies able to self-tune their operating mode.

In this talk, I will first discuss the implications associated with the main choices in the design space of transactional replication protocols. Next, I will present ongoing research work aimed at developing autonomic mechanisms to self-tune the replication scheme employed in elastic transactional data grids by exploiting both analytical performance models and machine learning techniques.