Deferred-Update Database Replication: Theory and Algorithms

Rodrigo Schmidt
EPFL, Switzerland

Abstract:

The deferred update technique is a widely used model to implement efficient database replication in which transactions are initially executed in a single replica. Passive transactions, those that do not change the state of the database, can commit locally to the replica they execute. Active transactions, which change the database state, must be synchronized with the transactions running on other replicas. In this talk, I will first introduce an abstract specification that generalizes the deferred update technique. This specification provides a strong model to prove lower bounds on replication algorithms, design new correct-by-construction protocols tailor-made for specific settings, and prove existing protocols correct more easily, in a standard way. Using this model, we show that the problem of termination of active transactions in deferred-update protocols is highly related to the problem of sequence agreement among a set of processes. In the second part of the talk, I will present a novel, highly-dynamic, algorithm to solve sequence agreement that can quickly adapt to system changes in order to preserve its optimal latency. Our algorithm is based on a new agreement problem we introduce that seems to be more suitable to solve problems like sequence agreement than previously used abstractions. In the last part of the talk, I will present two new fault-tolerant protocols derived from our general abstraction. The first algorithm uses no extra assumptions about database replicas. Yet, it has very little overhead associated with the termination of active transactions, propagating only strictly necessary information to replicas. Our second protocol uses strong assumptions about the concurrency control mechanism used by database replicas to reduce even more the latency and the burden associated with transaction termination. These algorithms are good examples of how our general abstraction can be extended to create new protocols and prove them correct.