Self-Optimized Communication of Replicated Services

Johannes Köstler, Hans P. Reiser
{jk,hr}@sec.uni-passau.de

To achieve high availability, todays applications and services require measures to ensure a fault-tolerant, or even intrusion-tolerant, execution. State Machine Replication (SMR) is a common approach that is able to provide both properties. Therefore, SMR systems replicate existing services which follow the classical client-server architecture over multiple servers. Those replicas are modeled as deterministic state machines, so that each of them has its own state, operates on a set of inputs and produces a set of outputs. If the replicas start in the same state and receive the same client requests in the same order all replicas will eventually reach the same state having produced the same set of responses. This procedure provides fault tolerance since, even if particular replicas fail, the remaining ones will still able to process the client requests – without any interruption. And intrusion tolerance as the replicas do a majority voting on the results and are therefore able to detect corrupted or tampered values. However, the request ordering and distribution introduces a huge communication overhead and limits concurrency in the execution of requests.

Our current research project, OptSCORE, mitigates those issues by adapting the communication and scheduling of replicated services in a way that they become self-configuring and self-optimizing. Therefore, we constantly monitor the health state of our system as well as the state of the execution environment to be able to detect changing conditions and react accordingly. To find optimal configurations we analyzed existing SMR systems. We extracted the particular configuration parameters and measured their effects on the system's throughput and latency in different situations and environmental conditions. Those isolated heuristics will be combined in a future step into an autonomously operating global optimization component. Apart from that, we invented several optimizations for the applied algorithms and procedures that help increasing the overall performance of such context-aware systems.

This talk will present the intermediate results of our work on the communication system. These consists of several contributions. The first one is a novel architecture for a group communication system that is able to monitor its own health state as well as the state of the environment and provides means for reconfiguring various aspects of the system at runtime. This is complemented by a classification of the several measurement and configuration parameters as well as the relationships among each other that we identified during our analysis. Besides that, we give further insights into several applied optimization strategies including a dynamically adapted weighted voting scheme that allows the formation of faster operating quorums.