

Scaling Up the Evaluation of Distributed Systems

Simon Brulhart

Master thesis in Computer Science

Large-scale distributed systems are difficult to test in real-world conditions. Testing systems and frameworks tend to break in surprising ways when increasing the number of concurrent nodes. Access to enough computing power may be unavailable or unaffordable. We motivate this problem by running EpTO, a large-scale dissemination protocol that previously could not run with more than 100 nodes. We start by trying out docker network drivers with lower overhead. Our measurements show great performance for these drivers. However the complexity and inflexibility they bring is a deal-breaker. Furthermore, we notice that the previously used overlay driver performs well enough after switching to a testbed based on EC2. Another angle is to increase the resources available for testing. To grow this testbed while keeping it affordable, we use automation to rent spare virtual machines for the duration of the tests. We finally harness this new testbed by running the push/pull implementation of EpTO against JGroups SEQUENCER with up to 500 nodes. The results reveal a higher network efficiency for EpTO at larger scales, with some caveats.

Prof. Pascal Felber