Computer Science Tripos - Part II Project - Progress Report

Achieving Distributed Consensus with Paxos

Christopher Jones, Trinity Hall

`<clj40@cam.ac.uk>`

31st January 2018

Project Supervisor: Dr Richard Mortier

Director of Studies: Prof Simon Moore

Project Overseers: Dr Markus Kuhn & Prof Peter Sewell

Work accomplished

The implementation of Mutli Paxos has been completed such that a number of processes can participate with nominated roles. The messaging system, leveraging the Cap'n Proto RPC library, was completed first, and extended as necessary throughout development.

Both phases of the algorithm have been implemented. The first phase allows for clients to submit requests to replicas, which perform de-duplication and serialization before proposing requests to Leaders. Leaders then participate in the Synod Protocol with Acceptors. They ensure that a majority consnesus is achieved on how the requests are serialized, providing the fault-tolerant memory of Paxos.

Evalution has begun and involves the writing of scripts for simulation on Mininet, the network simulator used for testing and evaluation. Currently the scripts support the required setup, logging and teardown of simulated networks and a simple experimental setup for collecting latency data.

Difficulties encountered

Difficulties were encountered with Mininet. Getting an OCaml runtime with the necessary packages in place that would execute on the virtual machine proved difficult in the initial weeks of the project. A library dependency that can't be installed has prevented compilation of the project on the virtual machine running Mininet. As a workaround, I have recently managed to get a bytecode implementation to successfully run on the simulator.

These difficulties have meant not as much testing and evaluation have been performed in simulation as hoped at this stage of the project. During development, the processes were run on the development machine, communicating locally using the same RPC library.

Progress through work plan

The difficulties encountered with Mininet have delayed the work plan by 1-2 weeks. Although the evaluation has begun and an experimental setup is in place, this has not been finished and the comparative evaluation of Libpaxos still needs to be performed.

The extension of the project to modify the quorum system to implement Flexible Paxos may take less time than anticipated in the work plan, which will allow me to complete all of the necessary evaluation in time to begin writing the dissertation in week 8. The evaluation of the extension will be very similar to that of the Multi Paxos implementation so should take much less time to complete once the Multi Paxos evaluation has been performed.