Scaling on the Edge: A Benchmarking Suite For Human-in-the-Loop Applications

M. Olguín, J. Wang, M. Satyanarayanan, J. Gross

ROYAL INSTITUTE OF TECHNOLOGY

Carnegie Mellon University Computer Science Department

Abstract

Benchmarking human-in-the-loop application is complex given their nature, which heavily depends on the actions taken by the *human* user. This limits reproducibility as well as feasibility of performance evaluations. We propose a methodology and present a benchmarking suite we call EdgeDroid that can address these challenges. Our core idea rests on recording traces of these applications which are played out in a controlled fashion based on an underlying model of human behavior. The traces are then exposed to the original backend compute process of the respective human-in-the-loop application, generating realistic feedback. This allows for an automated system which greatly simplifies benchmarking large scale scenarios.

Design & Implementation

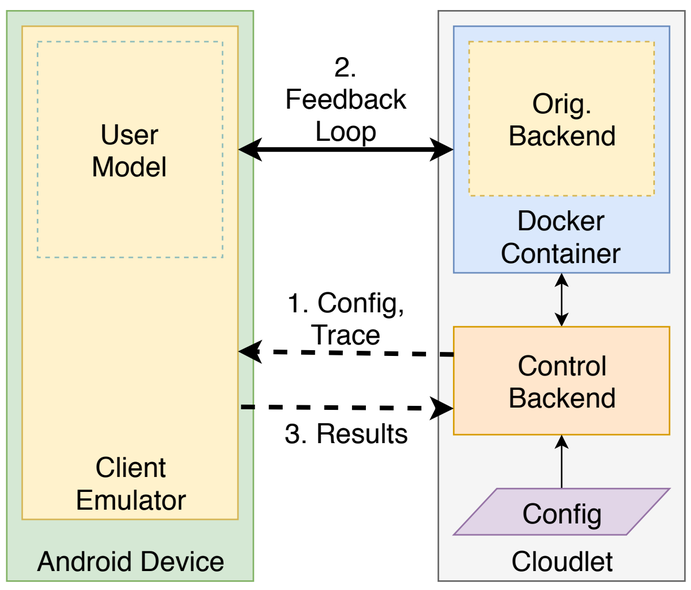


Figure 1: EdgeDroid Architecture

The architecture of our benchmarking suite consists of two main components:

- The *control backend*, which controls the experiments and collects measurements from the application and the client emulators.
- The *client emulators*, which play out a prerecorded sensory input trace over the network in a controlled fashion.

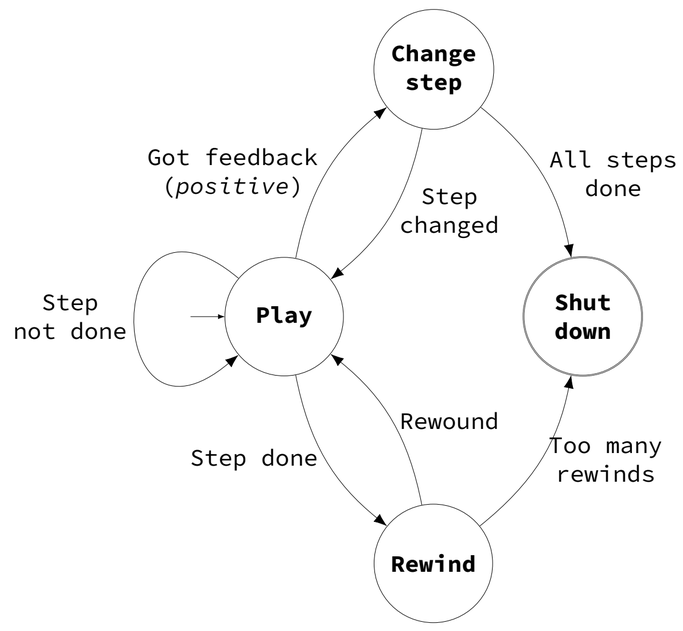
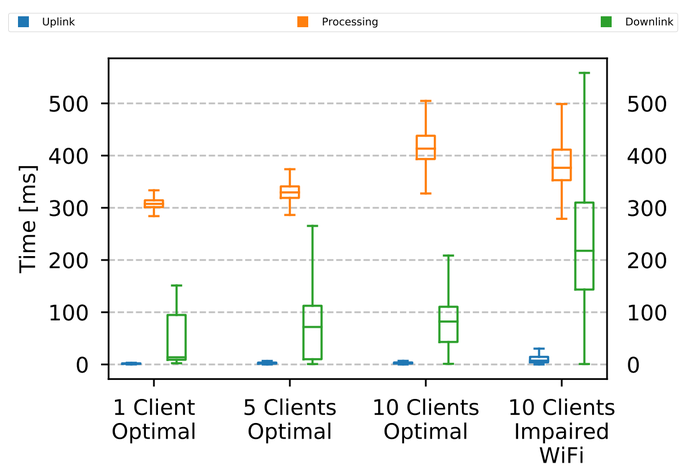
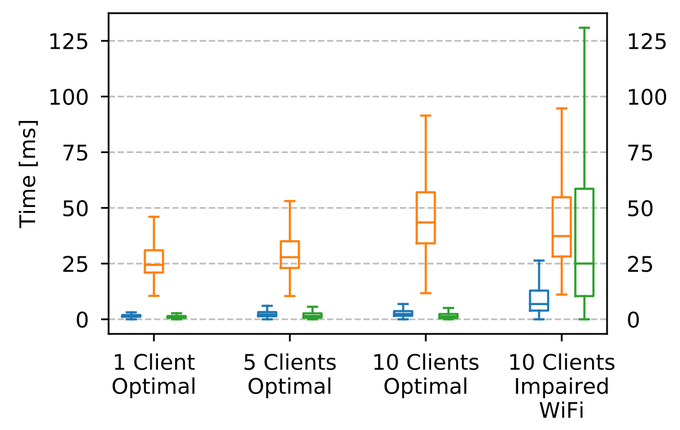


Figure 2: Simple user model used for the initial iteration of the suite.

Some Example Results



a: Inputs that triggered feedback.



b: Inputs that did not trigger feedback.

Figure 3: Comparison of latencies for a series of scenarios, differentiated by feedback-/lack of feedback.