Chair of Communication Networks
TUM Department of Electrical and Computer Engineering
Technical University of Munich

Master's Thesis

Implementing and Evaluating a Neural Network-based Routing Protocol

The goal of traffic engineering in communication networks is steering traffic such that a objective is optimized. The objective depends on the application and can be among others minimizing the maximum link utilization, and minimizing the flow completion time. To optimize any objective, a TE solution maps the current state of the network into a forwarding decision. That is, given the network state, source, and destination, forward this traffic along that path. Currently, the decision making in TE system is based on simple, hand-crafted algorithms. The reason lies in the strict computational requirements towards any TE algorithm (decisions at (sub)millisecond scale), and the necessity to realize the TE system as a distributed protocol.

Recent work shows that Neural Networks (NNs) can learn a distributed protocol from examples. The NN uses decisions of a TE system and synthesizes a distributed protocol out of those examples. In the process, the NN learns how information on a node should be encoded, which nodes need to exchange information to make decisions, and how to map the exchanged network state into a forwarding decision. For fast inference, the NN that makes the decisions is a fully binarized NN, i.e., input, weights, and activations are binary. While a Binary Neural Network (BNN) accelerates the evaluation, a practical implementation in existing hardware is still missing.

The goal of this thesis is to fill this gap and realize the BNN on a physical system. Your task is to develop a host based implementation (Appendix C.2 in [1]) using the extended Berkeley Packet Filter (eBPF). Concretely, the eBPF implementation must process update messages, and use the NN to make forwarding decisions for outbound traffic, and signal the forwarding decisions to the Network. The creation and the sending of update messages on switches is not part of the thesis and will be provided. The final deliverable is a small VM-based Clos-Topology (e.g., two pods of a k=4 fat-tree topology) in which traffic is routed based on the path determined by the NN in the end-hosts.

Advisors

Patrick Krämer