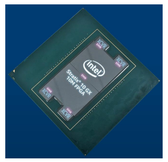
Machine Learning-Assisted Early Timing Prediction for Rapid FPGA Design Closure

Ecenur Ustun
eu49@cornell.edu

Rise of FPGA-Based Computing

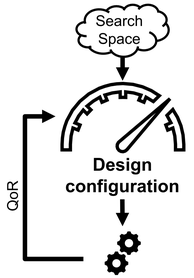
High performance and energy-efficiency

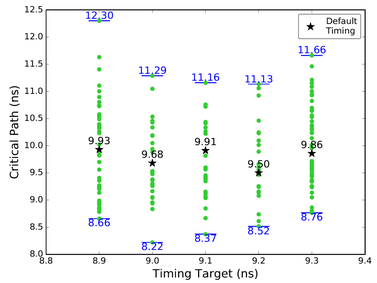
35 billion transistors
8.9 million logic elements
4.1 million LUTs



43.3 billion transistors
10.2 million logic elements
3.5 million LUTs

Configuring FPGA Design



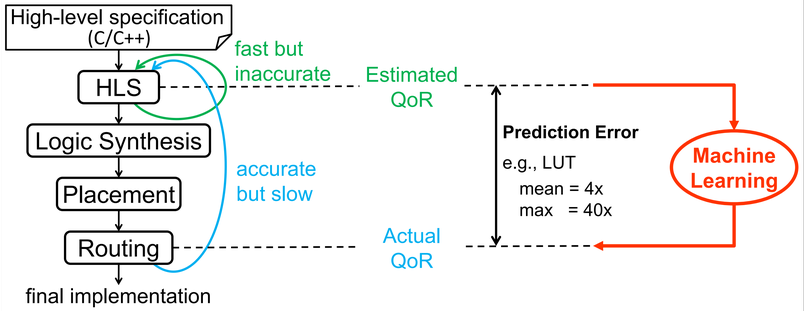


Timing distribution for various configurations

FPGA Design Flow

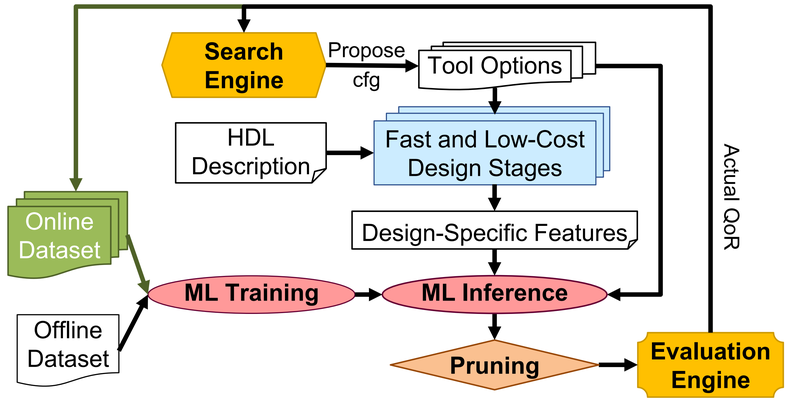
Main challenge behind rapid design closure is the runtime of end-to-end design flow.

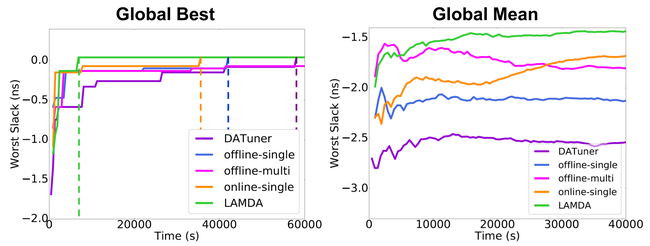
- Hours to synthesize a medium-scale design
- Weeks to explore 100 design configurations



LAMDA: Learning-Assisted Multi-Stage Design Autotuning

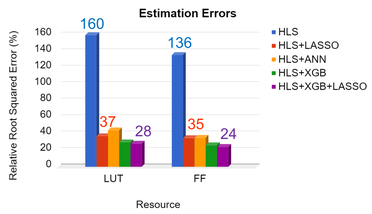
Balance the trade-off between **computing effort** and **prediction accuracy**





Fast and Accurate Estimation of QoR in HLS with Machine Learning

1300+ samples collected from 65 designs, 4 Xilinx FPGA device families, 5 clock periods



Received the Best Short Paper Award at FCCM 2018

Dataset and flow is open-sourced:

Accurate Operation Delay Prediction for HLS Using Graph Neural Networks

- Learn mapping of HLS operations onto FPGA device resources using graph neural networks
- Characterize delay in HLS based on learned mapping patterns
- 72% improvement in HLS operation delay prediction