## Micro-Architecture & RTL Design – NPU



- Microarchitecture development from scratch for a new Convolutional Neural Network algorithm
- RTL development of Neural Processing Unit sub-system using System Verilog, including distributed caching and computation, and data-recombination pathways
- Fine tuned optimizations for peak performance, increasing bandwidth, while drastically reducing area and power over previous architectures
- Implemented distributed RAMs topology for a multi-level caching mechanism to improve throughput further, while reducing peak power
- Developed modular Test-Benches for data path checks





7 months



Individual
Contribution at
client site