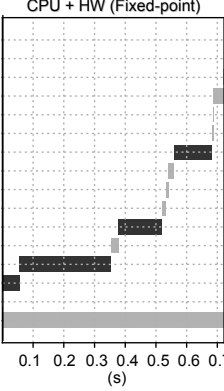a) Model A (8-bit fixed-point quantization)
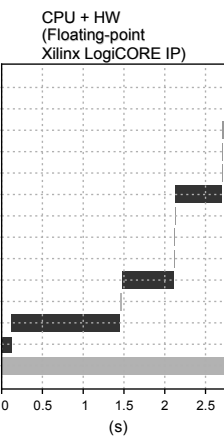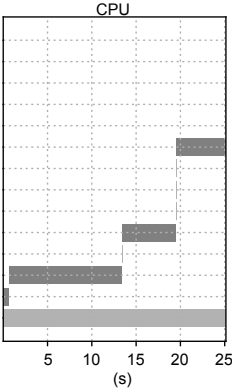
**Tensor operation**

CPU — operations listed top to bottom: DEQUANTIZE, SOFTMAX, FULLY_CONNECTED, FULLY_CONNECTED, RESHAPE, MAX_POOL_2D, **CONV_2D**, ADD, MUL, MAX_POOL_2D, **CONV_2D**, MAX_POOL_2D, **CONV_2D**, **CONV_2D**, QUANTIZE, Interpreter. x-axis: (s), ticks 5, 10, 15, 20, 25.
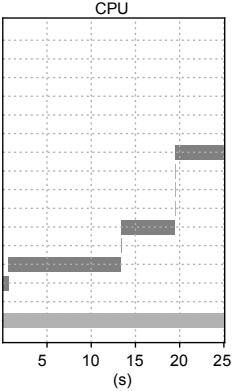
CPU + HW (Fixed-point). x-axis: (s), ticks 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7.

b) Model A (Floating-point)

CPU — operations: SOFTMAX, FULLY_CONNECTED, FULLY_CONNECTED, RESHAPE, MAX_POOL_2D, **CONV_2D**, ADD, MUL, MAX_POOL_2D, **CONV_2D**, MAX_POOL_2D, **CONV_2D**, **CONV_2D**, Interpreter. x-axis: (s), ticks 5, 10, 15, 20, 25.

CPU + HW (Floating-point Xilinx LogiCORE IP). x-axis: (s), ticks 0, 0.5, 1, 1.5, 2, 2.5.

c) Model A (Floating-point)

CPU + HW (Hybrid custom floating-point) — operations: SOFTMAX, FULLY_CONNECTED, FULLY_CONNECTED, RESHAPE, MAX_POOL_2D, **CONV_2D**, ADD, MUL, MAX_POOL_2D, **CONV_2D**, MAX_POOL_2D, **CONV_2D**, **CONV_2D**, Interpreter. x-axis: (s), ticks 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7.

CPU + HW (Hybrid logarithmic). x-axis: (s), ticks 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7.
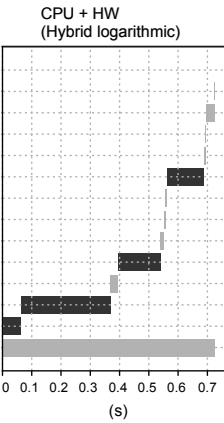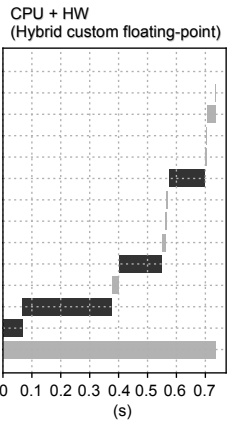
d) Model B (Floating-point)

CPU — operations: SOFTMAX, FULLY_CONNECTED, RESHAPE, MAX_POOL_2D, **CONV_2D**, ADD, MUL, MAX_POOL_2D, **CONV_2D**, **DEPTHWISE_CONV_2D**, MAX_POOL_2D, **CONV_2D**, **DEPTHWISE_CONV_2D**, Interpreter. x-axis: (s), ticks 0.2, 0.4, 0.6, 0.8, 1.0, 1.2.

CPU + HW (Hybrid custom floating-point). x-axis: (s), ticks 0, 0.2, 0.4, 0.6, 0.8, 0.1.