Floating-Point Multiply-Accumulate Unit: Assessing Non-Optimized Operations

IEEE 754 single-precision floating-point

Sign   Exponent   Mantissa
       8-bit   23-bit

S E E E E E E E E M M M M M M M M M M M M M M M M M M M M M M M

Sign   Exponent   Mantissa
       8-bit   23-bit

S E E E E E E E E M M M M M M M M M M M M M M M M M M M M M M M

Feature

Weight

MAC
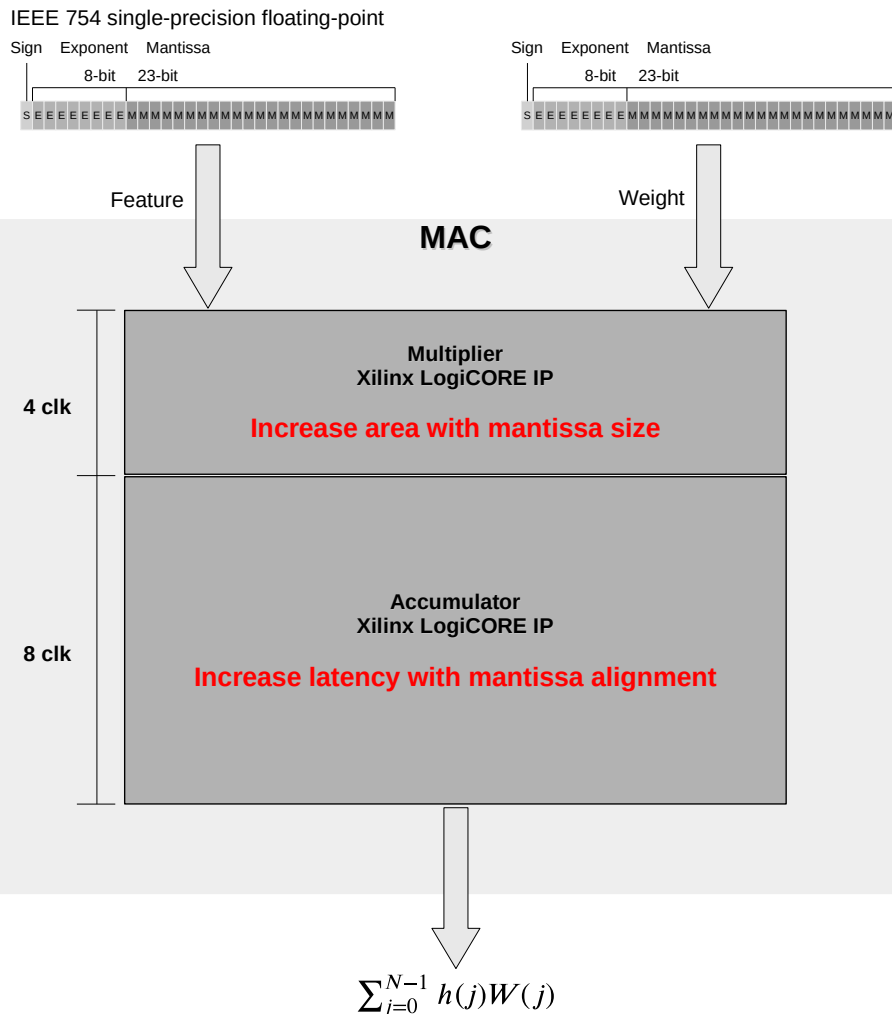
Loop pipelining:

Latency $= (II \times (N - 1)) + IL$

Latency $= 8N + 4$

N: Vector length

Infinity and **NaN:** Not expected

Handles subnormal cases

4 clk

Multiplier
Xilinx LogiCORE IP

Increase area with mantissa size

8 clk

Accumulator
Xilinx LogiCORE IP

Increase latency with mantissa alignment

$$\sum_{j=0}^{N-1} h(j)W(j)$$