a-bit Act w-bit Weight Assume mul: amn accumulation: (m+n)/100/ N: # of accumulation of layer Conv: R'Cin Cout Hart Wout [mul] (k2 Cin-1) Cout Have Wout OCC N= k2Cin Cin Cout Imul 1 lace N= On (Cin-1) Cout Same layer (f) layer 1 fz layer 2 fi layer 2 fz layer N FI lays H fz