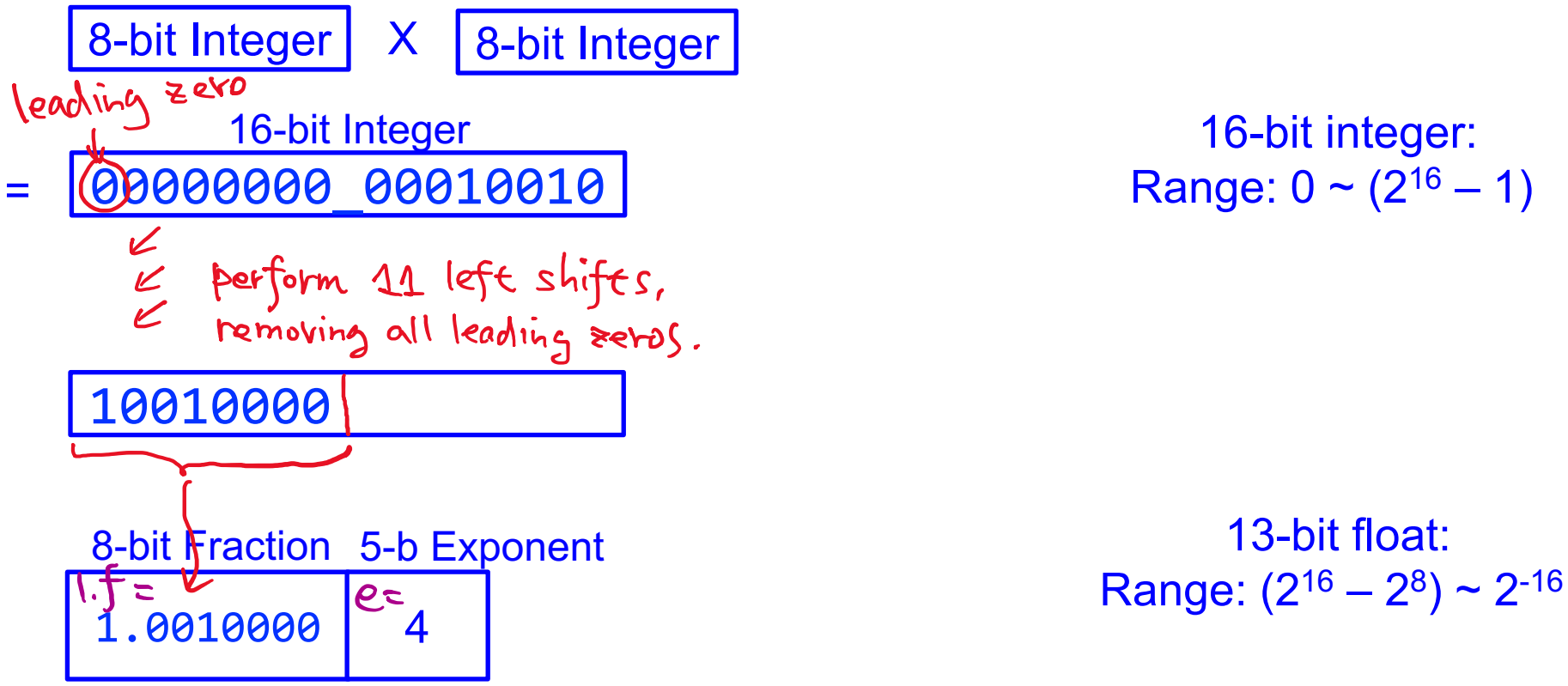
CS2102 02
Discussion of
The Floating Point Conversion

黃稚存

國立清華大學
資訊工程學系

HW 06

A Simple Floating Point Converter

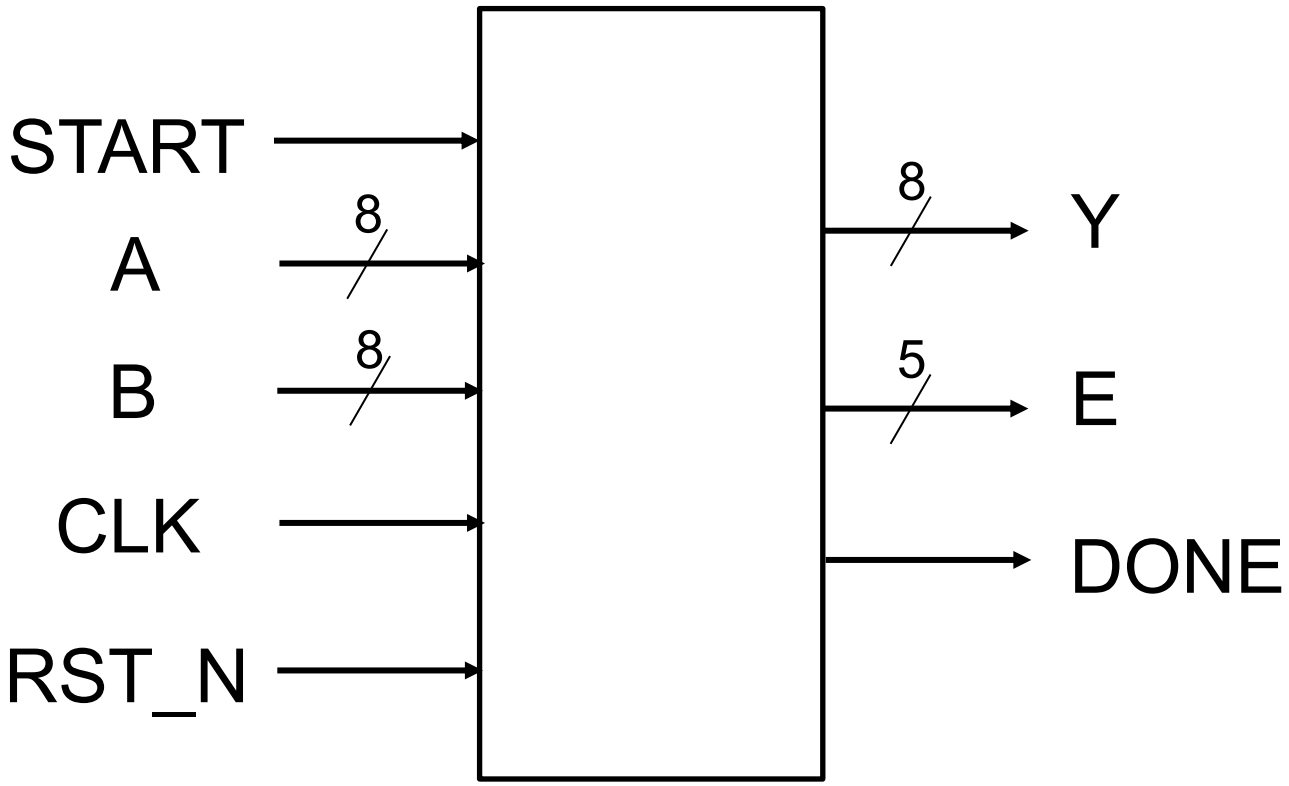


16-bit integer:
Range: 0 ~ $(2^{16} - 1)$

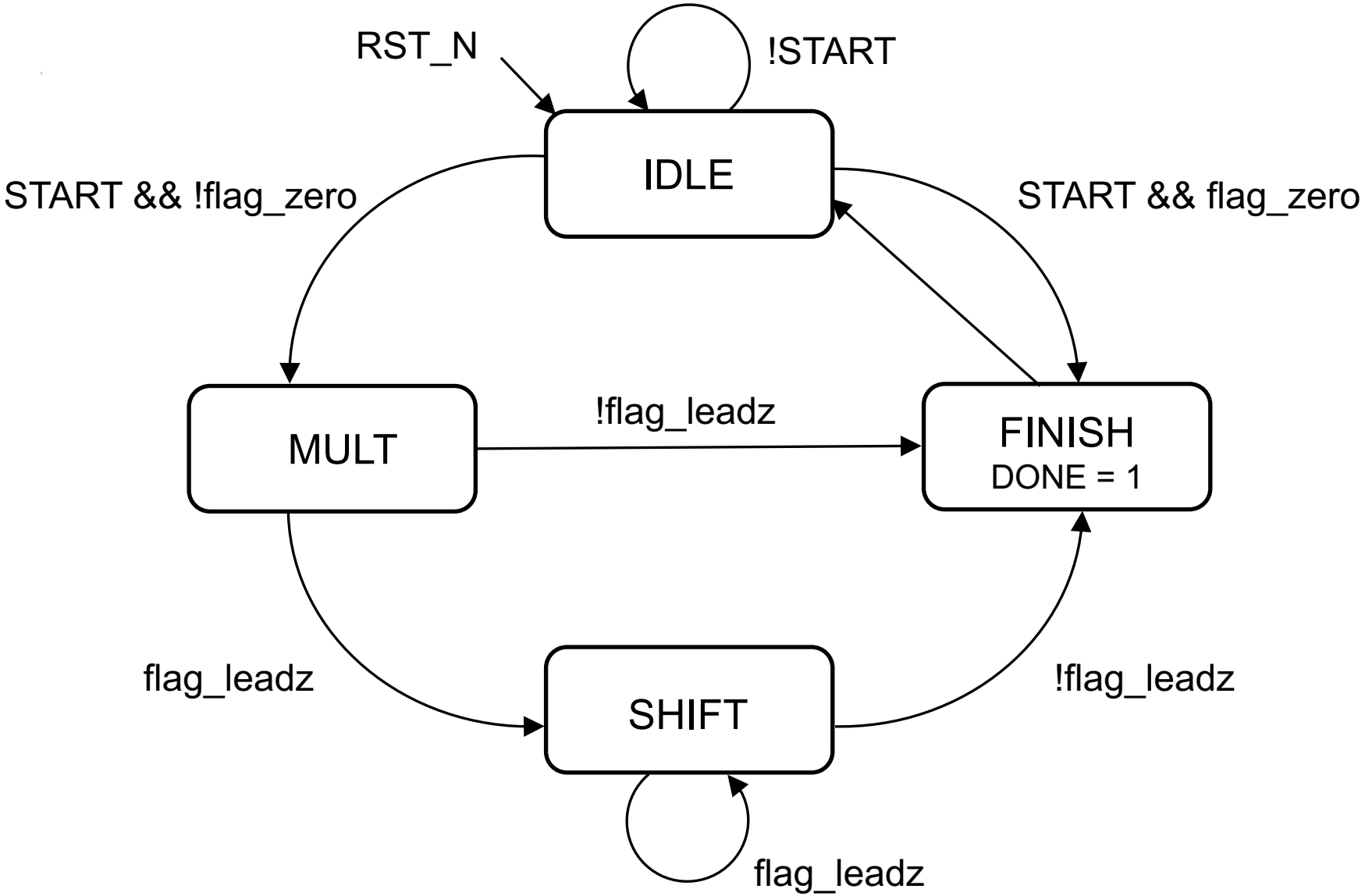
13-bit float:
Range: $(2^{16} - 2^{8}) \sim 2^{-16}$

$$1.f \times 2^{e} = 1.0010000 \times 2^{4}$$

Block Diagram



Finite State Machine (Control Unit)

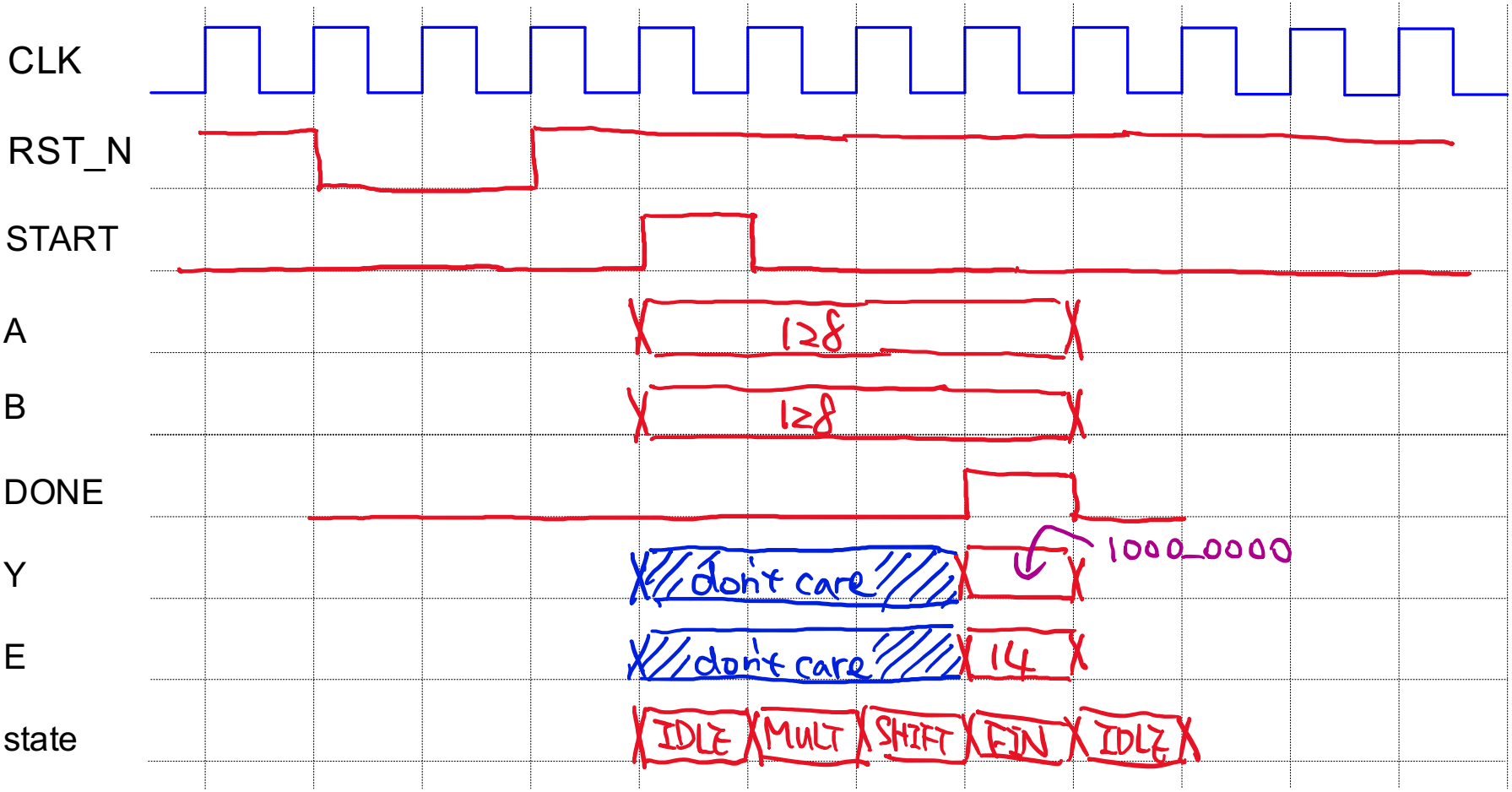


Further Discussion

- How do you define the control flags?
  - flag_allzero
  - flag_leadz
- Any additional flags to facilitate the control flow?
- How do you define the datapath computation?
  - How to sample A and B?
  - How to compute the multiplication?
  - How to adjust the result Y, and calculate E?

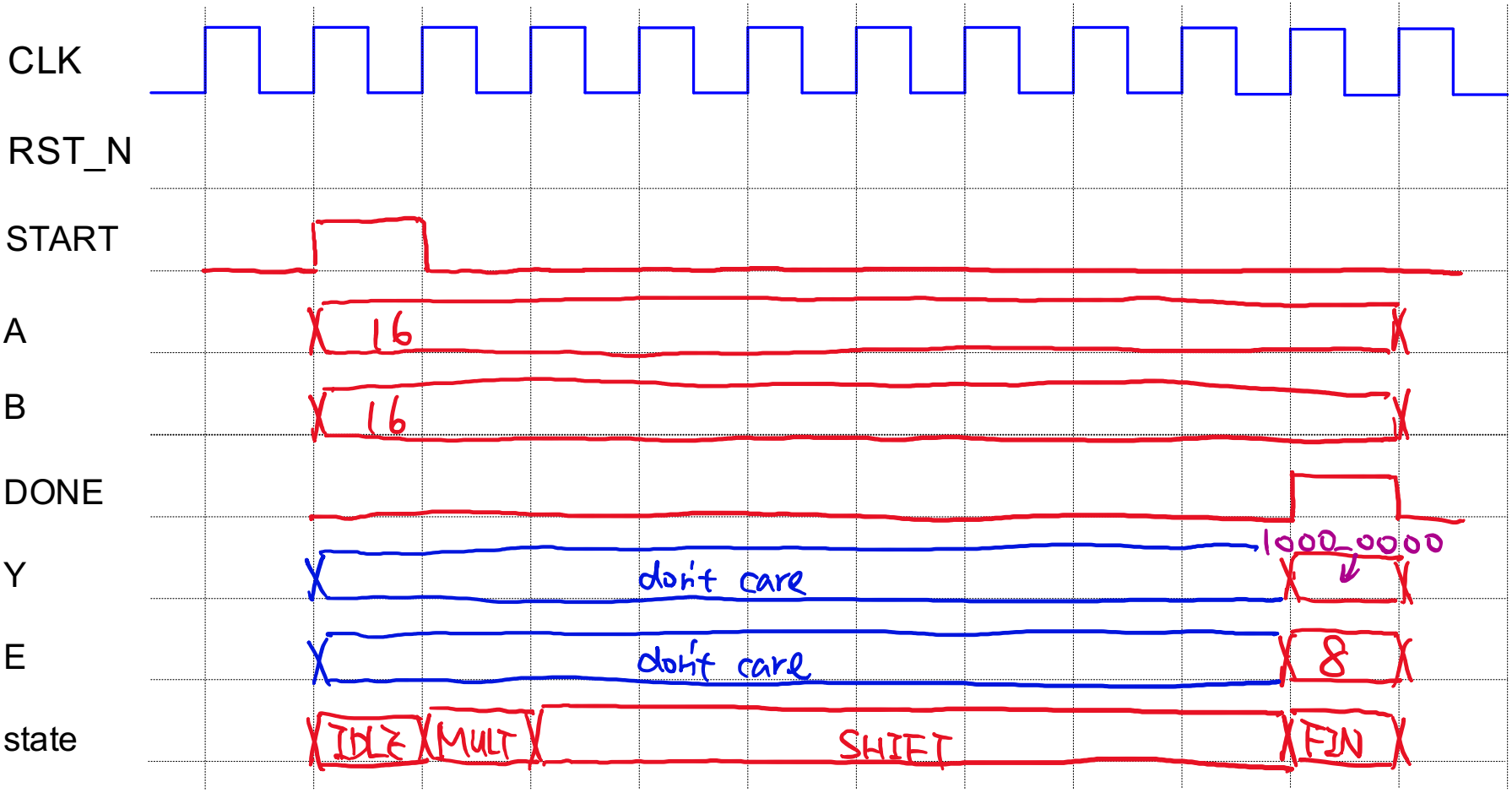
Some Examples and Timing Diagrams

- A = 128, B = 128, A x B = 16384
- Y = 1000_0000, E = 14 ($1.0 \times 2^{14} = 16384$)



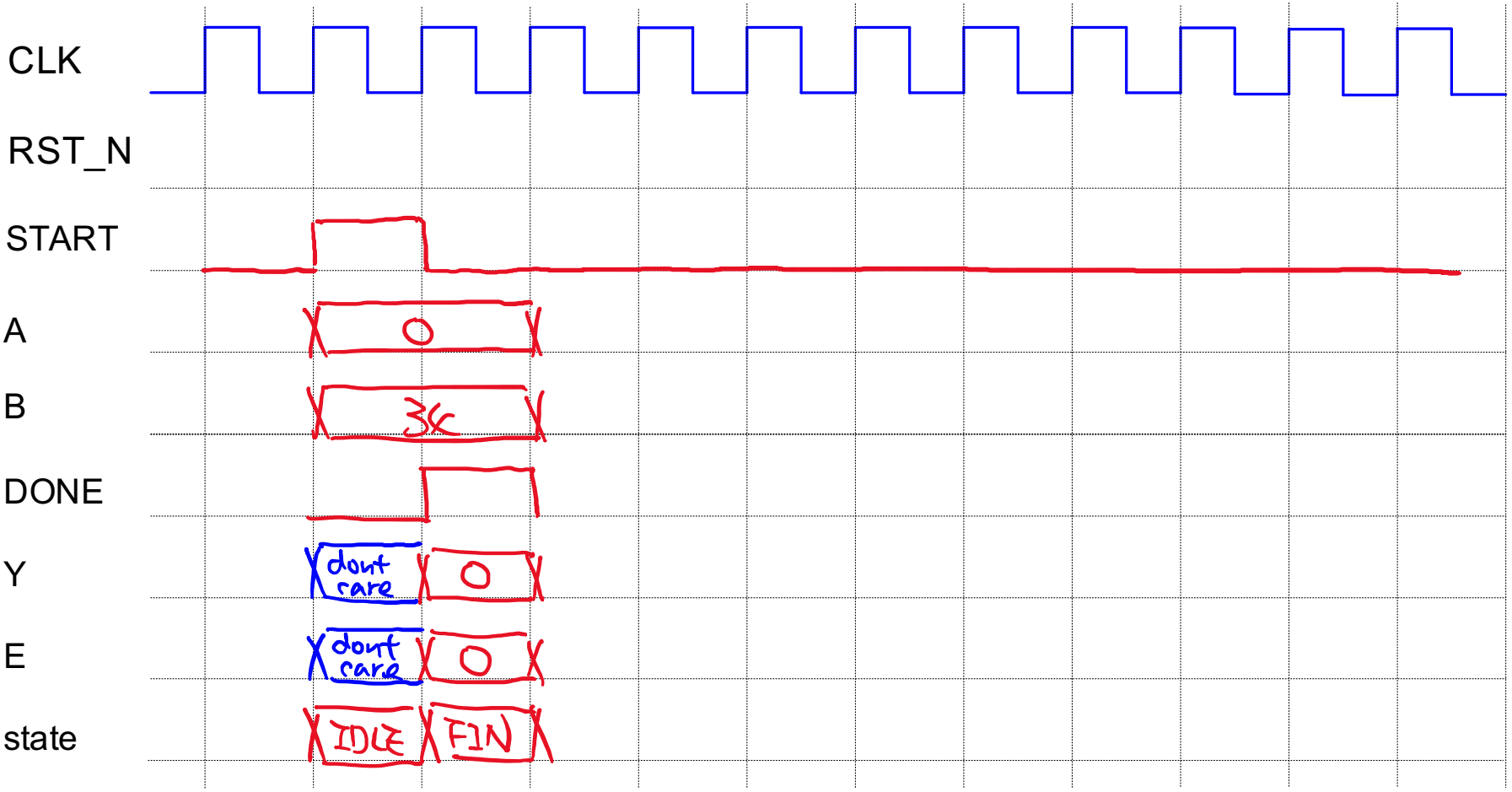
Some Examples and Timing Diagrams

- A = 16, B = 16, A x B = 256
- Y = 1000_0000, E = 8 (1.0 x 2^8 = 256)



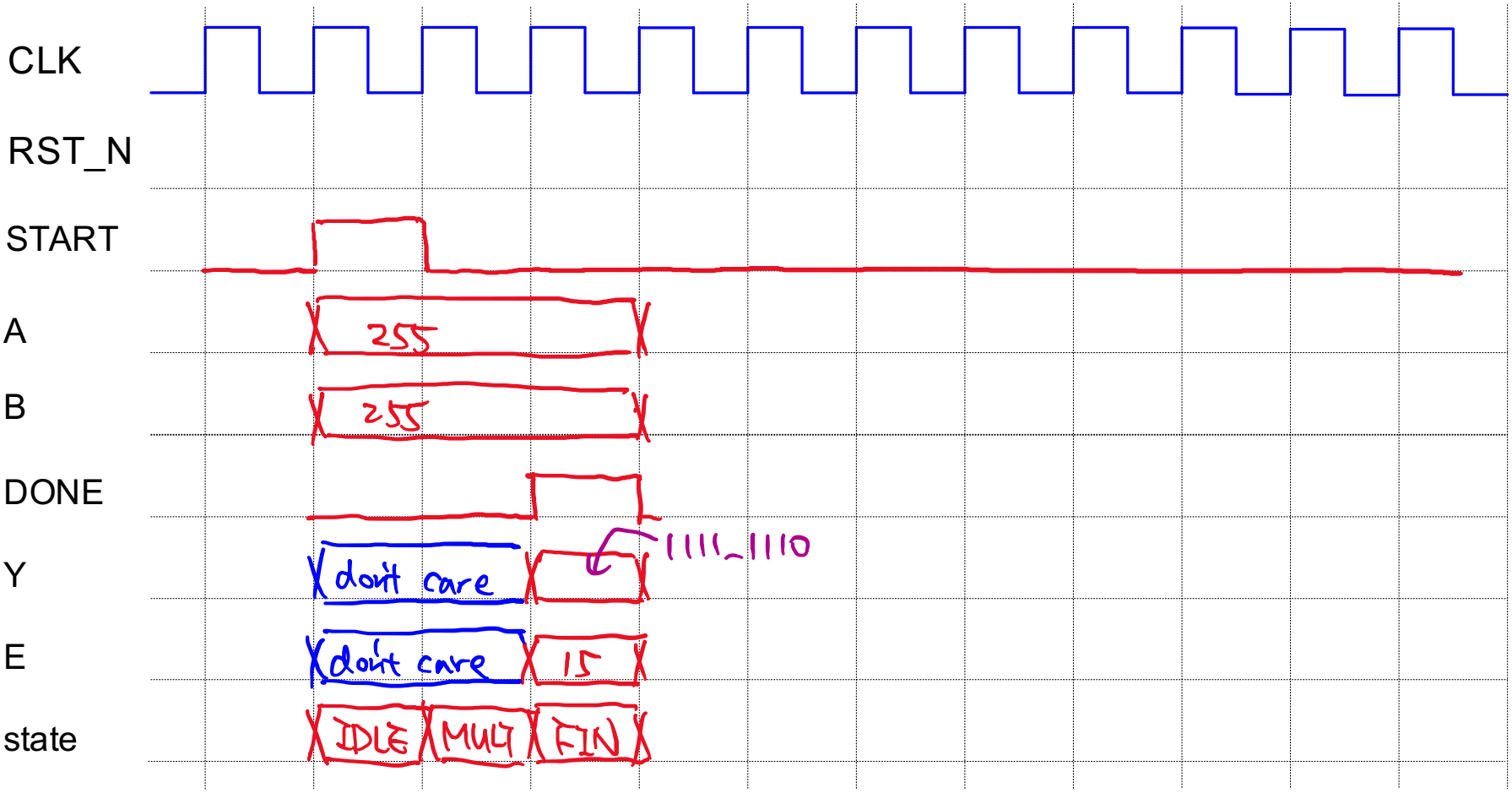
Some Examples and Timing Diagrams

- A = 0, B = 34, A x B = 0
- Y = 0000_0000, E = 0 (0 x 2^0 = 0)



Some Examples and Timing Diagrams

- A = 255, B = 255, A x B = 65025
- Y = 1111_1110, E = 15 (1.1111110 x 2^{15} = 65024)



Some Examples and Timing Diagrams

- A = 76, B = 95, A x B = 7220
- Y = 1110_0001, E = 12 (1.1100001 x 2^{12} = 7200)