

## Comp arch #5

① taking the decimal number 12.0625 is showing it in 'Floating Point'

$$12.0625 = 2^3 + 2^2 + 2^{-4} + 2^{-5} (1 + 2^{-1} + 2^{-2})$$

Mantissa

because the number is pos, the sign bit will be 0.

The exponent is 3

$$3 = X - 127$$

it will be encoded as 130

he will write this out in 8 bits: 10000010

Write the Mantissa in 23 bits: 10000010000000000000000

⇓

0 10000010 10000010000000000000000

in HEX: 0x41410000

②  $0x40F30000 = (0100\ 0000\ 1111\ 0011\ 0000\ 0000\ 0000\ 0000)_2$   
 $= (0\ 10000001\ 11100110000000000000000)_2$

Sign bit = 0 - pos number

exponent = 129 - 127 = 2 = 4096

Mantissa = 11100110000000000000000

write it out in decimal

$(11100110000000000000000)_2 = 2^{-1} + 2^{-2} + 2^{-3} + 2^{-6} + 2^{-7} =$

using the formula

Decimal =  $(-1)^S \cdot [1 + \text{mantissa}] \cdot 2^E = (1 + 2^{-1} + 2^{-2} + 2^{-3} + 2^{-6} + 2^{-7}) \cdot 2^2 =$   
 $4 + 2 + 1 + 2^{-1} + 2^{-4} + 2^{-5} = 7.59375$

③  $0x40000000 = (0\ 10000000\ 00000000000000000001000)_2 = (1 + 2^{-20}) 2^1 = A$   
 $0x3E800000 = (0\ 01111001\ 0000000000000000001000)_2 = (1 + 2^{-10}) 2^{-2} = B$

$B = 0.00000000000000000001000$

$A = 1.00000000000000000001000$

add the exponents: 1.00100000000000000001010

HEX: 0x10000A

Exponent: 128

float: 0 10000000 00100000000000000001010

HEX: 0x40100000



$$\textcircled{4} \quad 5 = 2^2 + 2^0 = (1+2^{-2}) \cdot 2^2 = (1.01)_2 \cdot 2^2$$

$$-0.4375 = -(2^{-2} + 2^{-3} + 2^{-4}) \cdot 2^0 = -(1+2^{-1} + 2^{-2}) \cdot 2^{-2} = (-1.11)_2 \cdot 2^{-2}$$

$$(1.01_2 \cdot 2^2) \cdot (-1.11_2 \cdot 2^{-2}) = (1.01_2 \cdot -1.11_2) \cdot 2^{2-2} = -(10.0011)_2 \cdot 2^0$$

$$= -1.00011_2 \cdot 2^1$$

Float:  $S = 1$

Mantissa = 00011

exponent = 128

$$\textcircled{5} \quad \textcircled{a} \quad A = 10101010 \quad B = 11110000$$

PROD = 0000000000000000

A = 11111110101010

B = 11110000

i = 0

$b_1b_0 = 00$

A = 1111111010101000      move A a bit to the left

i = 1

$b_1b_0 = 00$

A = 1111111010101000      move A a bit to the left

i = 2

$b_1b_0 = 00$

A = 111111101010100000      move A a bit to the left

i = 3

$b_1b_0 = 00$

A = 111111101010100000      move A a bit to the left

i = 4

$b_1b_0 = 10$

PROD = PROD - A

add

PROD = 0000010101100000

i = 5       $b_1b_0 = 11$

A = 1110101010000000

move A a bit to the left

i = 7  $\Rightarrow$  end

PROD: 0000010101100000



⑤ all of which are either  $\infty$  or 11 except one which is 10  
time of multiplication is T

⑥ 27