

①  $0 \mid 10011 \mid 00000001010$   
 sign exponent significand

~~01001100000001011~~

Exponent:  $2^0 \times 1 + 2^1 \times 1 + 2^2 \times 0 + 2^3 \times 0 + 2^4 \times 1$   
 $= 1 + 2 + 16$   
 $= 19.$

subtract bias

Power of 2:  $19 - 15 = 4$

Significand

$1.00000001010 \times 2^{19-15}$   
 $\downarrow \quad \downarrow \quad \downarrow \quad \downarrow$   
 $2^{-1} \times 0 \quad 2^{-2} \times 0 \quad 2^{-7} \times 1 \quad 2^{-9} \times 1$

Fraction  $\left(1 + \frac{1}{2^7} + \frac{1}{2^9}\right) \times 2^4$   
 $= \left(1 + \frac{1}{128} + \frac{1}{512}\right) \times 2^4$   
 $= \left(\frac{512}{512} + \frac{4}{512} + \frac{1}{512}\right) \times 2^4$   
 $= \frac{517}{512} \times 16 = \boxed{\frac{517}{32}} \approx \boxed{16.16}$

- ② The next biggest number should increment the significand, leaving the same exponent.

a)  $0|10011|0000000\underline{1011}$

Here, we used the fact that 1011 is the next biggest binary number after 1010

(b) and (c) using a similar procedure:

$$1.\overset{19-15}{000000}1011 \times 2$$

$2^{-1} \times 0$     $2^{-2} \times 0$     $2^{-7} \times 1$     $2^{-9} \times 1$     $2^{-10} \times 1$

$$= \left( 1 + \frac{1}{2^7} + \frac{1}{2^9} + \frac{1}{2^{10}} \right) \times 2^4$$

$$= \left( \frac{1024}{1024} + \frac{8}{1024} + \frac{2}{1024} + \frac{1}{1024} \right) \times 2^4$$

$$= \frac{1035}{1024} \times 2^4 = \boxed{\frac{1035}{64}} \approx \boxed{16.17}$$

Machine Epsilon : difference between 1 and next bigger number.

$$\begin{array}{l} 001111100000000000 \\ 001111100000000001 \end{array} \longrightarrow (1 + 2^{-10}) - 1 = \frac{1}{2^{10}} \approx 9.7 \times 10^{-4}$$

$$\begin{array}{l} 001111111100\dots 0 \\ 001111111100\dots 1 \end{array} \longrightarrow (1 + 2^{-23}) - 1 = \frac{1}{2^{23}} \approx 1.2 \times 10^{-7}$$

23 bits

$$\begin{array}{l} 001111111111100\dots 0 \\ 001111111111100\dots 1 \end{array} \longrightarrow (1 + 2^{-52}) - 1 = \frac{1}{2^{52}} \approx 2.2 \times 10^{-16}$$

52 bits

## Largest number

We set both exponent and significand to be the highest  
They can be without showing up as "infinity" or "NaN".

$$\begin{array}{l} 011111011111111111 \\ 011111011111111111 \end{array}$$

Exponent = 30  
minus bias = 15

$$= 1.111\dots 1 \times 2^{15}$$

$$= \left(1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots + \frac{1}{1024}\right) \times 2^{15} = \frac{2047}{1024} \times 2^{15}$$

$$= 65,504$$

$$\begin{array}{l} 01111111011\dots 1 \\ 01111111011\dots 1 \end{array}$$

Exponent = 254  
minus bias = 127

$$= 1.111\dots 1 \times 2^{127}$$

$$= \left(1 + \frac{1}{2} + \frac{1}{4} + \dots + \frac{1}{2^{23}}\right) \times 2^{127}$$

$$\rightarrow = 1.999\dots \times 2^{127}$$

$$\approx 2 \times 2^{127} = 2^{128}$$

$$= 3.4 \times 10^{38}$$

$$\begin{array}{l} 011111111111011\dots 1 \\ 011111111111011\dots 1 \end{array}$$

Exponent = 2046  
minus bias = 1023

$$= 1.111\dots 1 \times 2^{1023}$$

$$= \left(1 + \frac{1}{2} + \frac{1}{4} + \dots + \frac{1}{2^{52}}\right) \times 2^{1023}$$

$$\rightarrow = 1.99\dots \times 2^{1023}$$

$$\approx 2 \times 2^{1023} = 2^{1024}$$

$$= 1.8 \times 10^{308}$$

# Smallest Number (greater than zero)

000...1      significand: 00...1  
                 exponent: 00...0

General form

$$0.00...1 \times 2^{0-\text{bias}+1}$$

16-bit  $0.\overbrace{00...1}^{10} \times 2^{-15+1}$

$$= \left(0 + \frac{1}{2^{10}}\right) \times 2^{-14} = 2^{-24} \approx 5.97 \times 10^{-8}$$

32-bit  $0.\overbrace{00...1}^{23} \times 2^{-127+1}$

$$= \left(0 + \frac{1}{2^{23}}\right) \times 2^{-126} = 2^{-149} \approx 1.4 \times 10^{-45}$$

64-bit  $0.\overbrace{00...1}^{52} \times 2^{-1023+1}$

$$= \left(0 + \frac{1}{2^{52}}\right) \times 2^{-1022} = 2^{-1074} \approx 5 \times 10^{-324}$$

Number of Numbers including zero but excluding  $\infty$  and NaN.

16-bit : possible numbers:  $2^{16}$

When exponent is all 1's, we don't have numbers

0 11111 0000000000  
1 bit                      10 bits

Whether these are zero or one, no admissible numbers.

$2^2$  possible numbers here.

$$\Rightarrow 2^{16} - 2^2 = 63,488$$

32-bit possible numbers:  $2^{32}$

When exponent is all 1's, we don't have numbers

0 11111 000000...000  
1 bit                      28 bits

Whether these are zero or one, no admissible numbers.

$2^{24}$  possible numbers here.

$$\Rightarrow 2^{32} - 2^{24} \approx 4.27 \times 10^9$$

64-bit possible numbers:  $2^{64}$

When exponent is all 1's, we don't have numbers

0 11111 000000...000  
1 bit                      52 bits

Whether these are zero or one, no admissible numbers.

$2^{53}$  possible numbers here.

$$\Rightarrow 2^{64} - 2^{53} \approx 1.84 \times 10^{19}$$