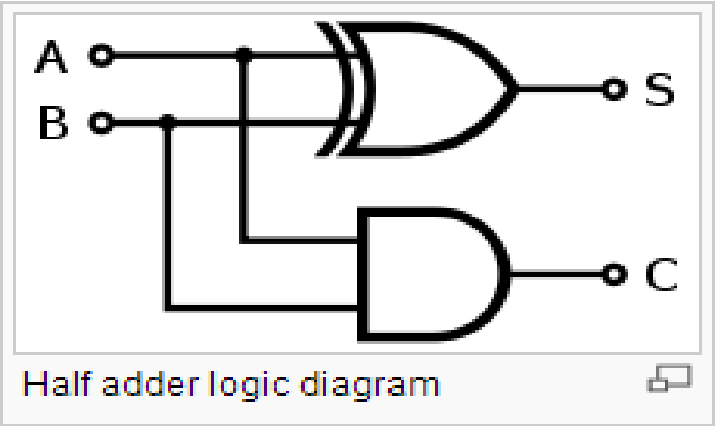
計算機結構
(算術單元 ALU 的設計)

陳鍾誠 於金門大學

半加器

S = A XOR B

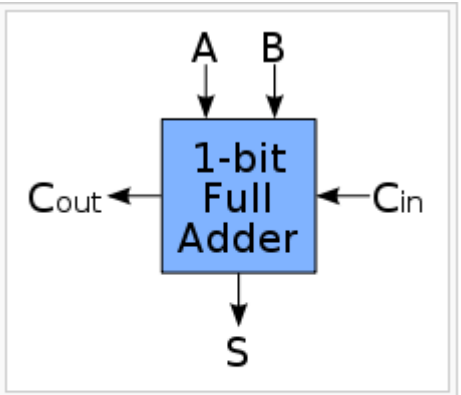
C = A AND B



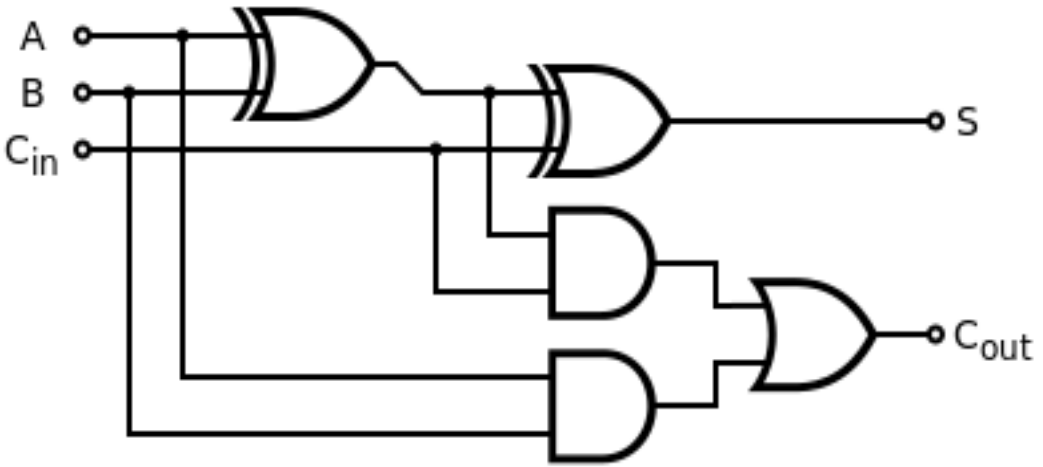
Half adder logic diagram

全加器

Inputs			Outputs	
A	B	C_{in}	C_{out}	S
0	0	0	0	0
1	0	0	0	1
0	1	0	0	1
1	1	0	1	0
0	0	1	0	1
1	0	1	1	0
0	1	1	1	0
1	1	1	1	1



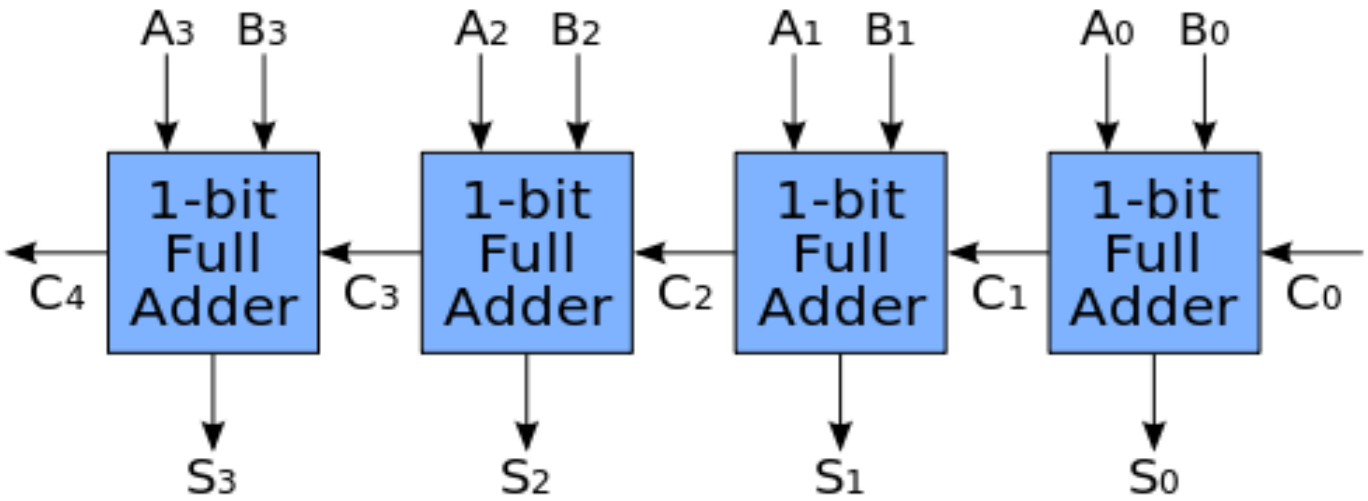
Schematic symbol for a 1-bit full adder with C_{in} and C_{out} drawn on sides of block to emphasize their use in a multi-bit adder



http://en.wikipedia.org/wiki/Adder_(electronics)

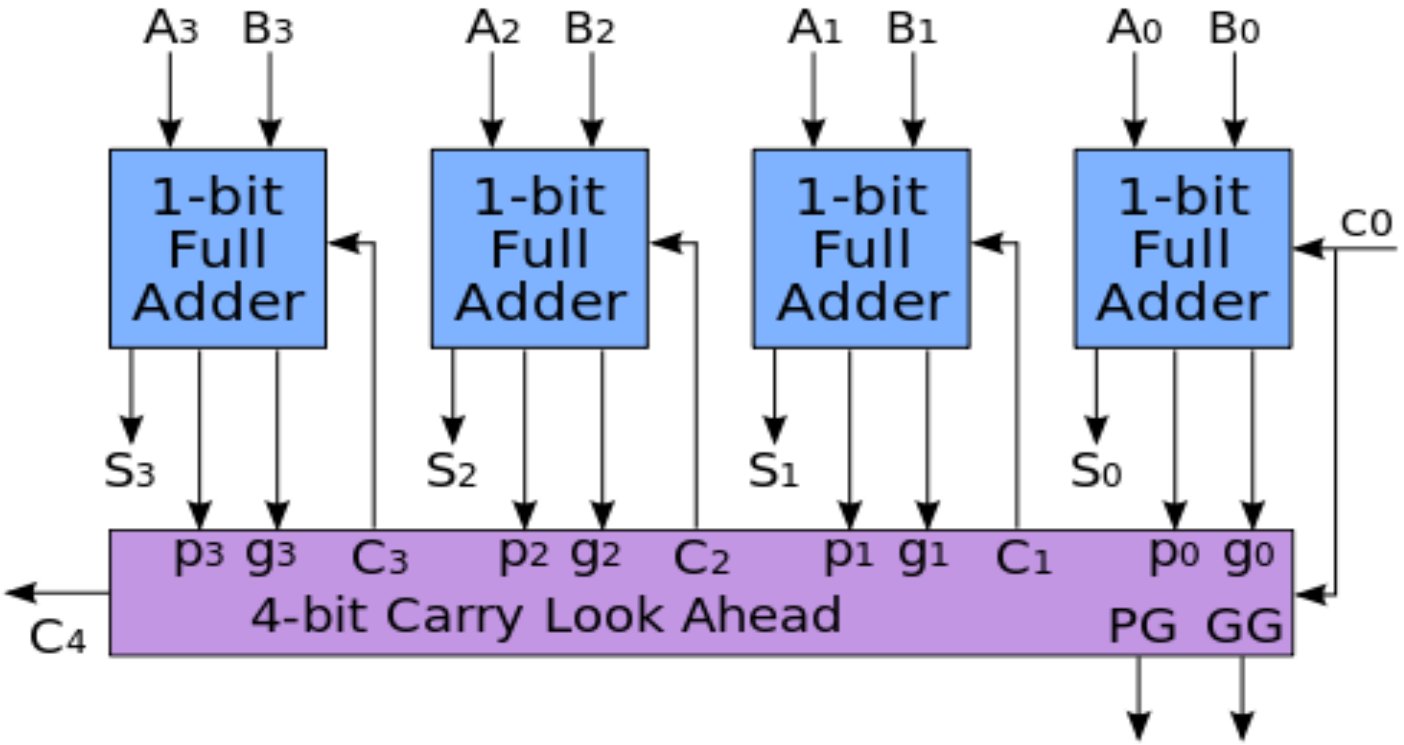
4 位元鏈波加法器

4-bit adder with logic gates shown



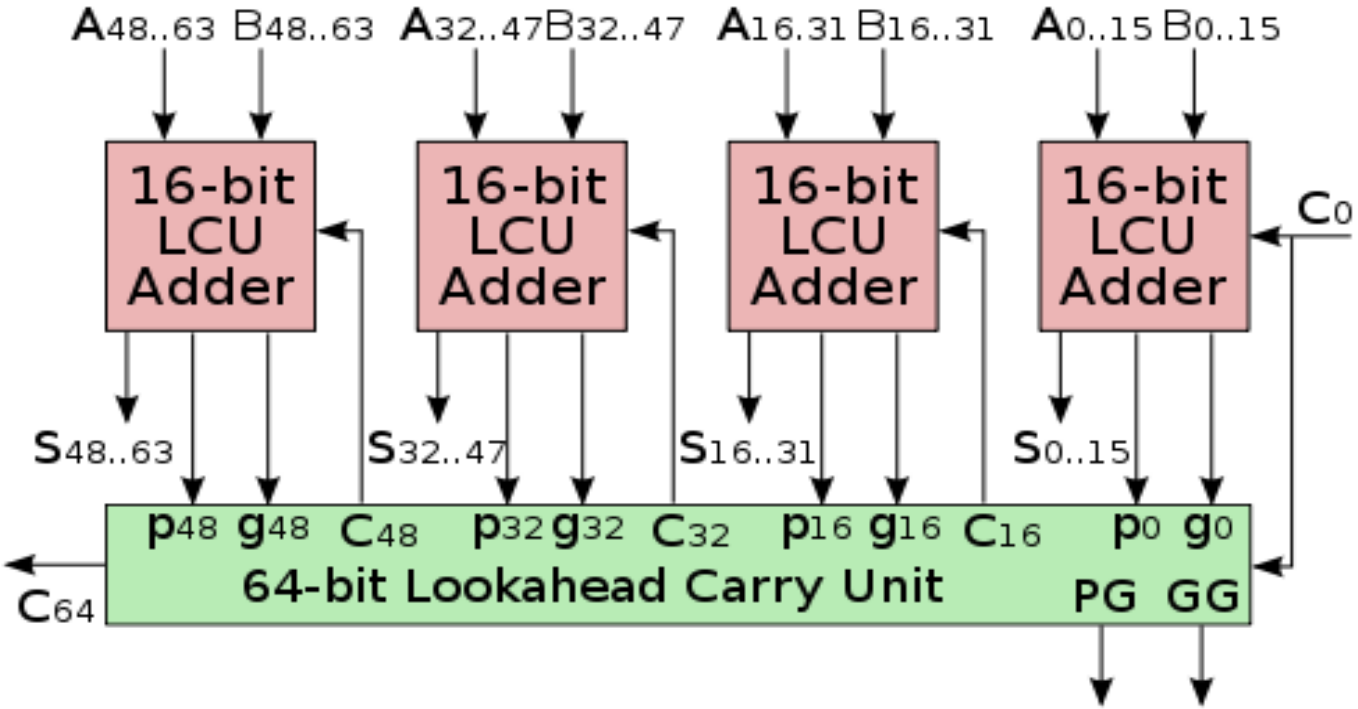
4 位元前瞻進位快速加法器

4-bit adder with carry lookahead



64 位元加法器

A 64-bit adder



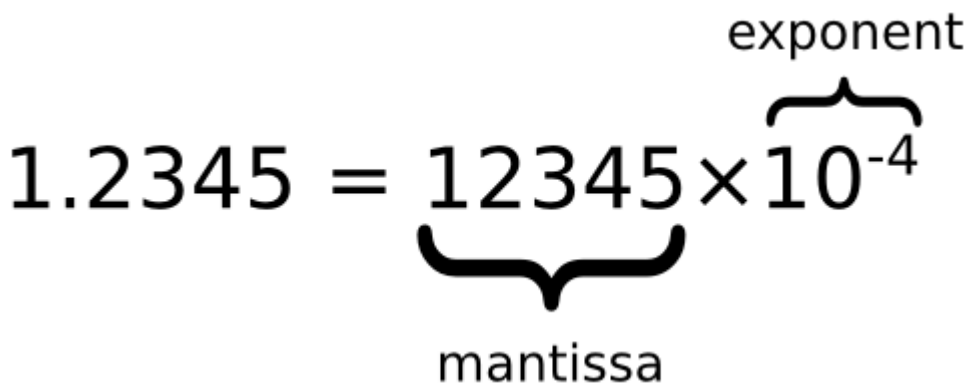
乘法的計算 (10 進位)

```
   123
 x 456
 =====
   738   (this is 123 x 6)
  615    (this is 123 x 5, shifted one position to the left)
+ 492    (this is 123 x 4, shifted two positions to the left)
 =====
 56088
```

乘法的計算 (2 進位)

```
     1011   (this is 11 in decimal)
   x 1110   (this is 14 in decimal)
   ======
     0000   (this is 1011 x 0)
    1011    (this is 1011 x 1, shifted one position to the left)
   1011     (this is 1011 x 1, shifted two positions to the left)
+ 1011      (this is 1011 x 1, shifted three positions to the left)
=========
 10011010   (this is 154 in decimal)
```

浮點數 (Floating Point)



http://en.wikipedia.org/wiki/Floating_point

浮點數加法（科學記號版）

```
123456.7 = 1.234567 x 10^5
101.7654 = 1.017654 x 10^2 = 0.001017654 x 10^5
```

```
Hence:
123456.7 + 101.7654 = (1.234567 x 10^5) + (1.017654 x 10^2)
                    = (1.234567 x 10^5) + (0.001017654 x 10^5)
                    = (1.234567 + 0.001017654) x 10^5
                    =  1.235584654 x 10^5
```

浮點數加法（定點運算版）

```
  e=5;   s=1.234567      (123456.7)
+ e=2;   s=1.017654      (101.7654)
```

```
  e=5;   s=1.234567
+ e=5;   s=0.001017654   (after shifting)
--------------------
  e=5;   s=1.235584654   (true sum: 123558.4654)
```

注意：會有捨去誤差 (rounded error)

如果數字太小，有時完全被捨去

```
  e=5;  s=1.234567
+ e=-3; s=9.876543
```

```
  e=5;  s=1.234567
+ e=5;  s=0.00000009876543 (after shifting)
----------------------
  e=5;  s=1.23456709876543 (true sum)
  e=5;  s=1.234567         (after rounding/normalization)
```

浮點數乘法

```
  e=3;   s=4.734612
x e=5;   s=5.417242
-----------------------
  e=8;   s=25.648538980104 (true product)
  e=8;   s=25.64854        (after rounding)
  e=9;   s=2.564854        (after normalization)
```

IEEE 754 浮點數標準 (2008)

Name	Common name	Base	Digits	E min	E max	Notes	Decimal digits	Decimal E max
binary16	Half precision	2	10+1	−14	+15	storage, not basic	3.31	4.51
binary32	Single precision	2	23+1	−126	+127		7.22	38.23
binary64	Double precision	2	52+1	−1022	+1023		15.95	307.95
binary128	Quadruple precision	2	112+1	−16382	+16383		34.02	4931.77
decimal32		10	7	−95	+96	storage, not basic	7	96
decimal64		10	16	−383	+384		16	384
decimal128		10	34	−6143	+6144		34	6144

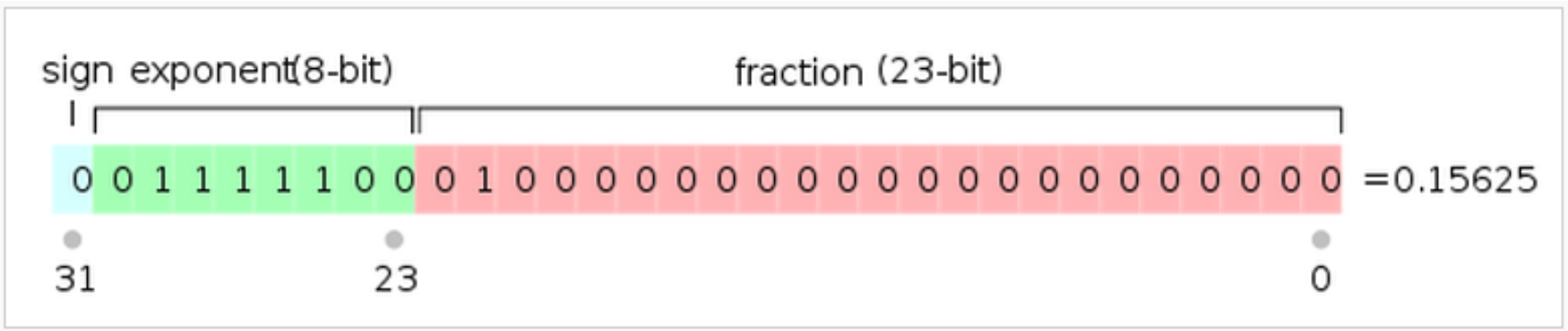
http://en.wikipedia.org/wiki/IEEE_754-2008

IEEE 754 浮點數標準 (1985)

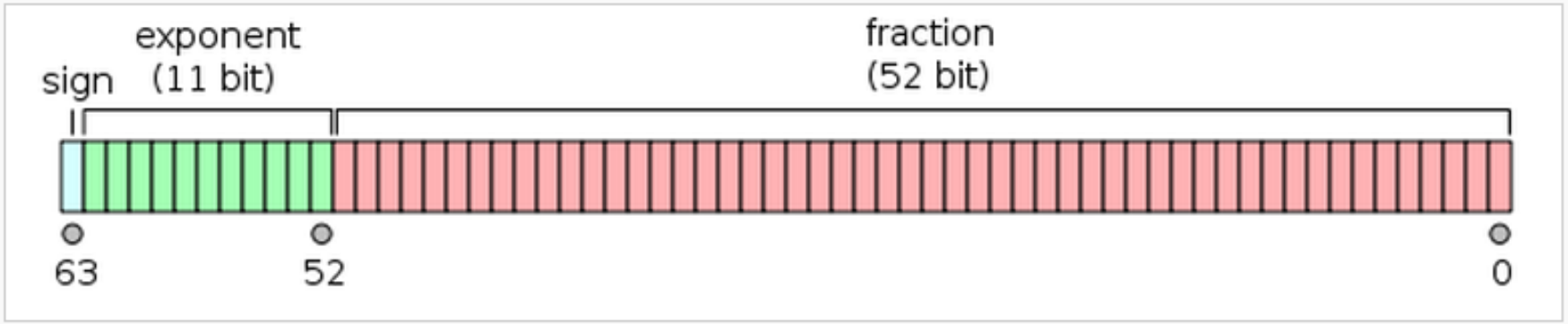
level	width	range	precision*
single precision	32 bits	$\pm 1.18 \times 10^{-38}$ to $\pm 3.4 \times 10^{38}$	approx. 7 decimal digits
double precision	64 bits	$\pm 2.23 \times 10^{-308}$ to $\pm 1.80 \times 10^{308}$	approx. 15 decimal digits

http://en.wikipedia.org/wiki/IEEE_754-1985

IEEE 754 單精度浮點數

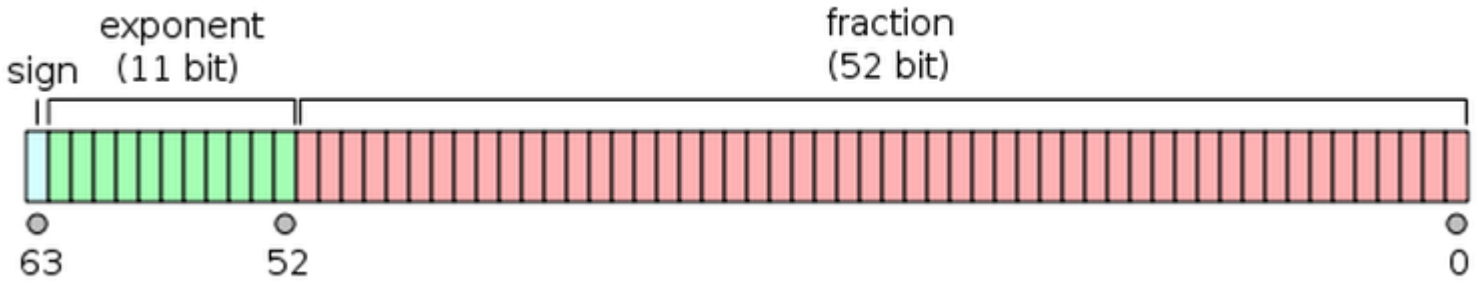


The number 0.15625 represented as a single-precision IEEE 754-1985 floating-point number. See text for explanation.



The three fields in a 64bit IEEE 754 float

IEEE 754 雙精度浮點數



$$\text{value} = (-1)^{\text{sign}}(1 + \sum_{i=1}^{52} b_{52-i} 2^{-i}) \times 2^{e-1023}$$