1 Conventions

1.1 Variables

s : Valeur du bit de signe
e : Valeur binaire de l'exposant
m : Valeur binaire de la mantisse
w : Nombre de bits de l'exposant
t : Nombre de bits de la mantisse exculant le LSB
p : Nombre de bits de la mantisse incluant le LSB ommit
E : Valeur de l'exposant
M : Valeur de la mantisse
b : biais de l'exposant

1.2 Équations

$$b = 2^{w-1} - 1$$

$$E = e–b$$

$$Emin = 2 - 2^{w-1}$$

$$Emax = 2^{w-1} - 1$$

$$Inf : E = 2^{w-1} \& m = 0$$

$$qNaN : E = 2^{w-1} \& m/ = 0 \& m[-1] = 0$$

$$sNaN : E = 2^{w-1} \& m/ = 0 \& m[-1] = 1$$

$$Subnormal : E = 0$$

$$Normal : (-1)^s \cdot 2^E \cdot (1 + m \cdot 2^{-t})$$

$$Subnormal : (-1)^s \cdot 2^{Emin} \cdot m \cdot 2^{-t}$$

2 FMUL

Le multiplicateur exécute la fonction $Y = A * B$

2.1 Operation standard

1. normal x normal = normal

$$Y = A \times B$$

$$(-1)^{s_y} \cdot 2^{E_y} \cdot (1+m_y \cdot 2^{-t}) = (-1)^{s_a} \cdot 2^{E_a} \cdot (1+m_a \cdot 2^{-t}) \times (-1)^{s_b} \cdot 2^{E_b} \cdot (1+m_b \cdot 2^{-t})$$

$$(-1)^{s_y} \cdot 2^{E_y} \cdot (1+m_y \cdot 2^{-t}) = (-1)^{s_a+s_b} \cdot 2^{E_a+E_b} \cdot (1+m_a \cdot 2^{-t}) \times (1+m_b \cdot 2^{-t})$$

$$s_y = s_a + s_b$$

$$E_y = E_a + E_b$$

$$(1+m_y \cdot 2^{-t}) = (1+m_a \cdot 2^{-t}) \times (1+m_b \cdot 2^{-t})$$

Agorithme exposant: lorsque les 2 exp sont additionné, un bias en extra se trouve dans le resulat, on peut soustraire ce biais en effectuant -2**

2. normal x normal = subnormal
3. normal x subnormal = subnormal
4. normal x subnormal = normal
5. subnormal x subnormal = subnormal
6. subnormal x subnormal = normal impossible

2.2 Operation spéciales

1. NaN x NaN La valeur de A est propagée
2. NaN x Valeur/0/InF La valeur NaN est propagée
3. 0 x InF NaN est généré
4. 0 x normal/subnormal 0 est généré
5. InF x InF InF est généré
6. InF x valeur InF est généré