

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC BÁCH KHOA
KHOA ĐIỆN – ĐIỆN TỬ
NĂM HỌC 2019 – 2020

-----*-----



KỸ THUẬT SỐ NÂNG CAO
BÁO CÁO BÀI TẬP LỚN

GVHD: TRẦN HOÀNG LINH

Tên	MSSV
Nguyễn Duy Tân	1713068
Nguyễn Mỹ Hằng	1711215

TP.HCM, ngày 15 tháng 12 năm 2019

LỜI NÓI ĐẦU

Chúng em gửi lời cảm ơn đến thầy Trần Hoàng Linh, người đã hướng dẫn cho chúng em một đề tài rất hay về mảng thiết kế vi mạch.

Mặc dù chúng em đều không theo hướng vi mạch nhưng mà đây cũng được xem như là trải nghiệm đáng nhớ, cũng như đem lại cho chúng em nhiều hiểu biết hơn về mảng thiết kế vi mạch.

Qua sự tìm hiểu hai tháng chúng em đã thiết kế thành công bộ FPAU, cùng với đó là các bộ lũy thừa, bộ căn bậc n.

Ngoài ra chúng em có tìm hiểu về Matlab, làm cách nào để chuyển từ số thực sang số floatting point 32bits cũng như chuyển đổi ngược lại.

MỤC LỤC

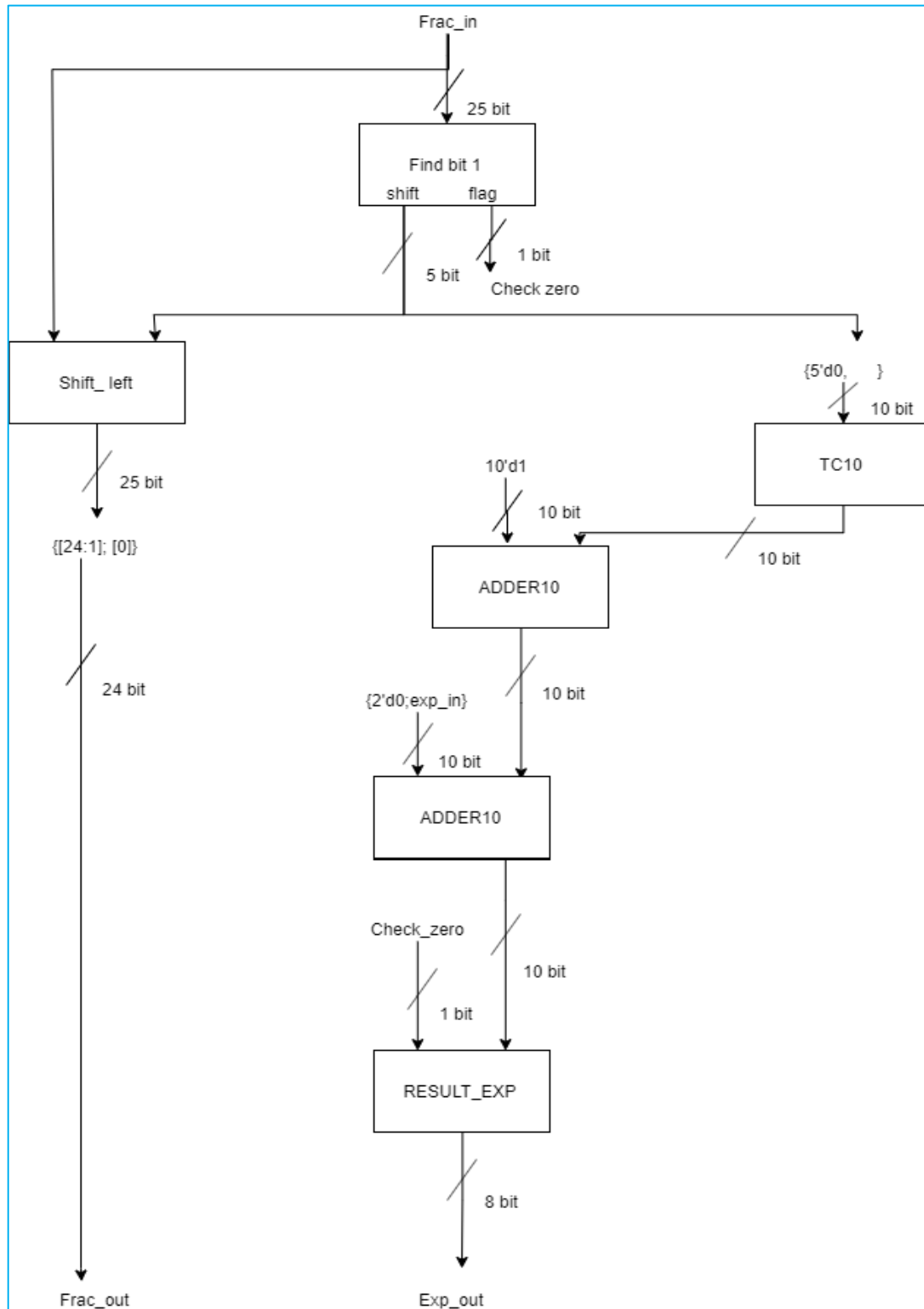
NỘI DUNG	5
Chương 1: Các Bộ Dùng Chung	5
1. Bộ Normalize	5
2. Bộ Mux.....	6
3. Bộ Shift Right	7
4. Bộ Shift Left.....	8
5. Bộ Check_input	9
6. Bộ Find Bit1	10
Chương 2:Bộ Cộng Trừ.....	11
1. Giải thuật CLA.....	11
2. Giải thuật RCA.....	12
3. Bộ tổng quát.....	13
4. Các module nhỏ	14
5. Kết quả mô phỏng	17
Chương 3: Bộ Nhân	18
1. Giải thuật Double	18
2. Giải thuật vedic	20
3. Bộ tổng quát.....	23
4. Các module nhỏ	24
5. Kết quả mô phỏng	25
Chương 4:Bộ Chia.....	26
1. Lưu đồ giải thuật Restoring-Division	26
2. Khối chính.....	27
3. Khối module thành phần.....	28
4. Kết quả mô phỏng	32
Chương 5: Bộ FPAU	33
1. Mô hình	33
2. Kết quả chạy mô phỏng	34
Chương 6 : Bộ Lũy Thừa.....	35
1. Module tổng quát	35
2. Module Element_N.....	36

3. Module Element_4	37
4. Result_Ex	38
5. Kết quả mô phỏng	39
Chương 7: Bộ Căn bậc 2	40
1. Module tổng quát	40
2. Module Rs_Element	41
3. Module Unit	42
4. Result_Rs	43
5. Kết quả mô phỏng	44
Chương 8: Bộ Căn bậc n	45
1. Mô hình chung	45
2. Module Rs_Element	46
3. Module Unit	47
4. Result_Rs	48
5. Kết quả mô phỏng	49
Chương 9: Chuyển Đổi Floating Point 32bits Trong Matlab	50
1. Chuyển từ số thực sang float 32bit	50
2. Chuyển từ số float 32bit sang số thực	51
3. Kết quả chạy trên matlab	52
KẾT LUẬN	53
TÀI LIỆU THAM KHẢO	54

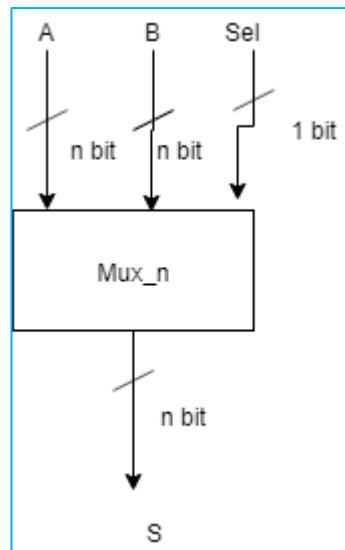
NỘI DUNG

Chương 1: Các Bộ Dùng Chung

1. Bộ Normalize

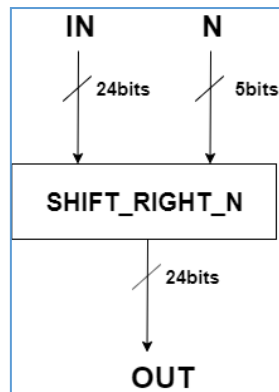


2. Bộ Mux

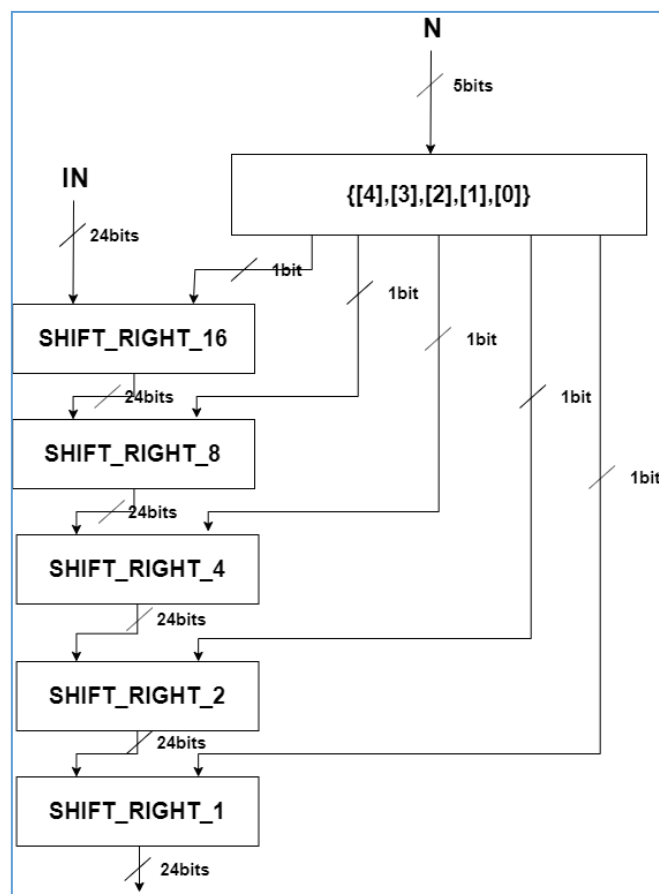


3. Bộ Shift Right

a. Khối chung:

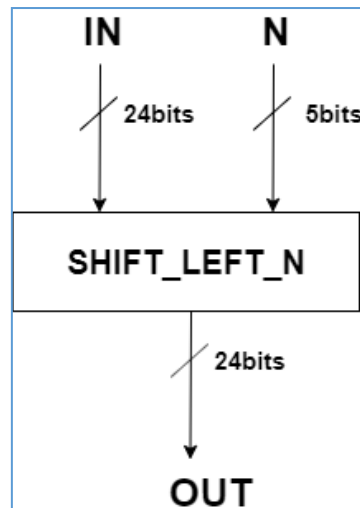


b. Khối Chi tiết:

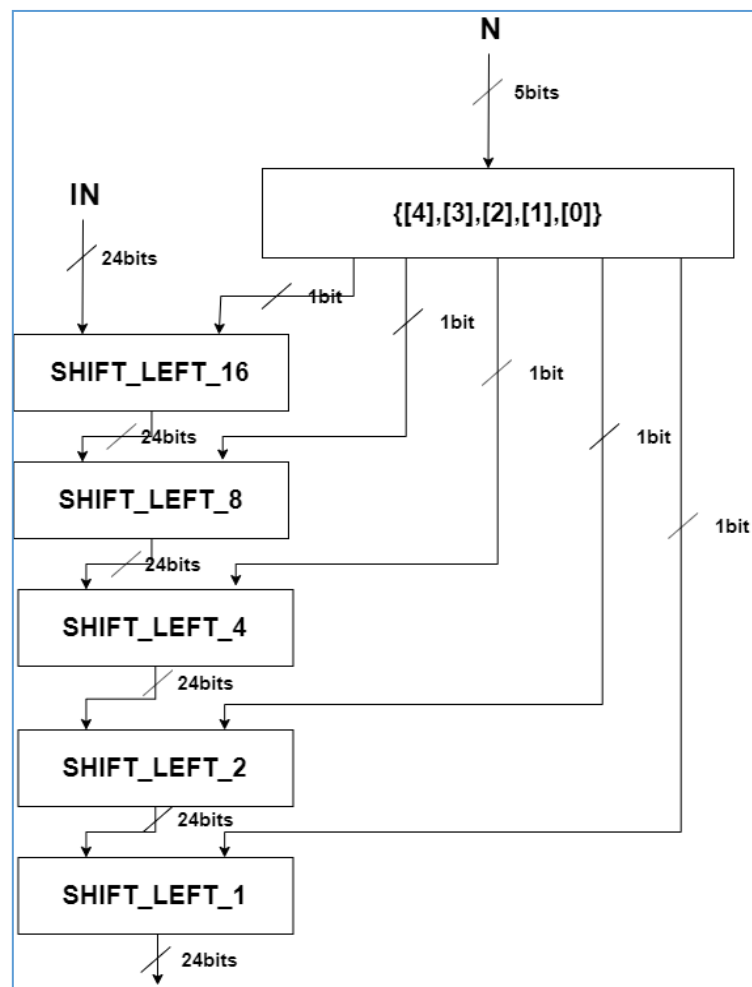


4. Bộ Shift Left

a. Khối chung:



b. Khối Chi tiết:



5. Bộ Check_input

Exponent	Fraction	Flag
00000000	000000000000000000000000	FlagZero
11111111	111111111111111111111111	FlagNaN
11111111	000000000000000000000000	FlagInf

6. Bộ Find Bit1

Sử dụng bộ mã hóa ưu tiên:

A24	A23	A22	A21	A20	A19	A18		S4	S3	S2	S1	S0
X	1	x	x	x	x	x		0	0	0	0	1
X	0	1	x	x	x	x		0	0	0	1	0
X	0	0	1	x	x	x		0	0	0	1	1
X	0	0	0	1	x	x		0	0	1	0	0
X	0	0	0	0	1	x		0	0	1	0	1
X	0	0	0	0	0	1		0	0	1	1	0
...

Chương 2: Bộ Cộng Trừ

1. Giải thuật CLA

a. Bộ CLB

$$g_i = a_i b_i, P_i = a_i (xor) b_i$$

$$C_0 = C_{in}$$

$$C_1 = g_0 + p_0 C_0$$

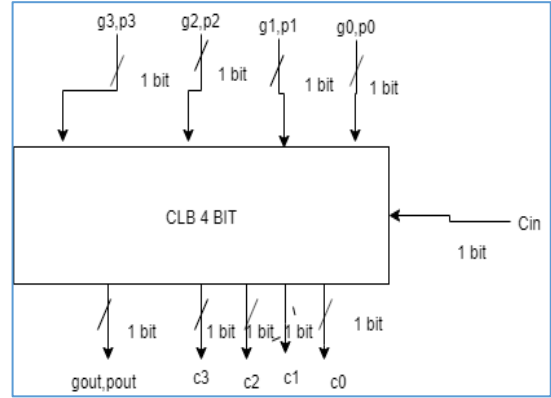
$$C_2 = g_1 + p_1 C_1$$

$$C_3 = g_2 + p_2 C_2$$

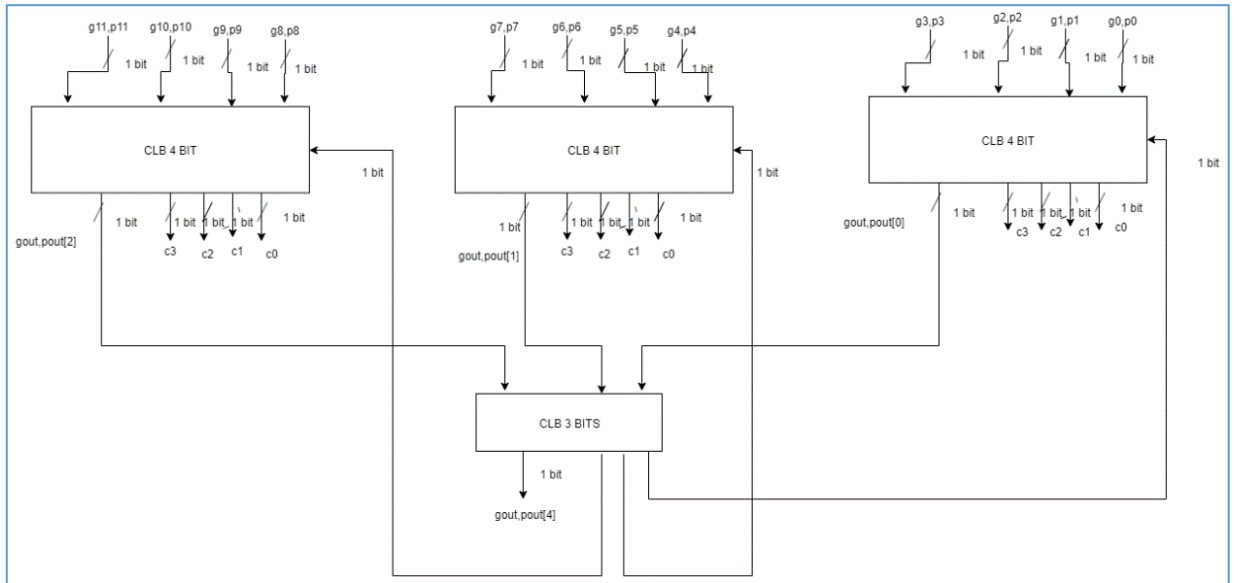
$$g_{out} = g_3 + p_2 g_2 + p_3 p_2 g_1 + p_3 p_2 p_1 g_0$$

$$p_{out} = p_3 p_2 p_1 p_0$$

$$S_i = p_i (xor) c_i$$

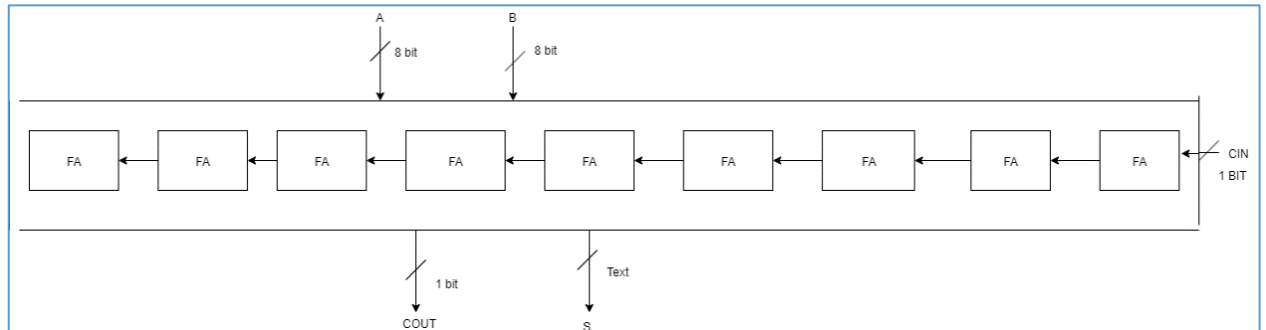


b. Bộ CLA 12bits

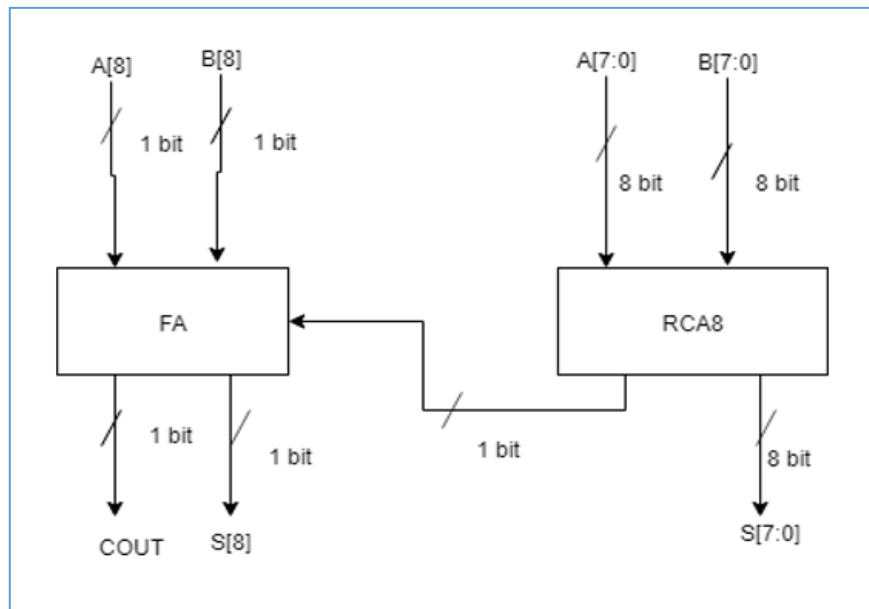


2. Giải thuật RCA

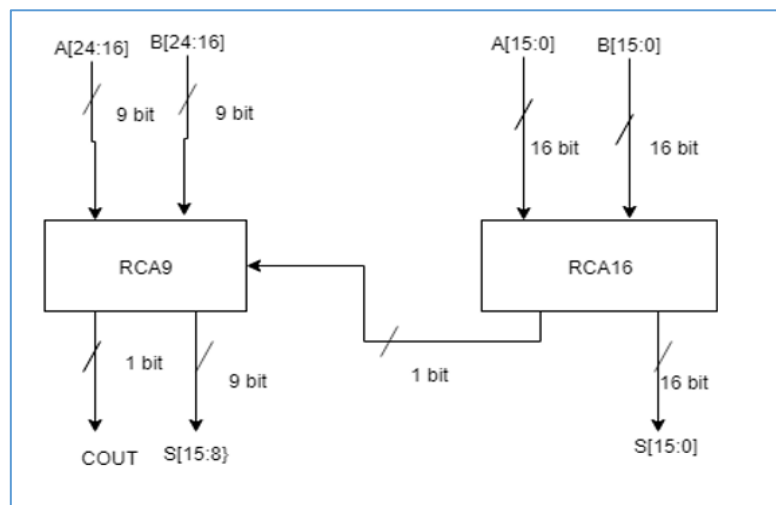
a. RCA 8bits



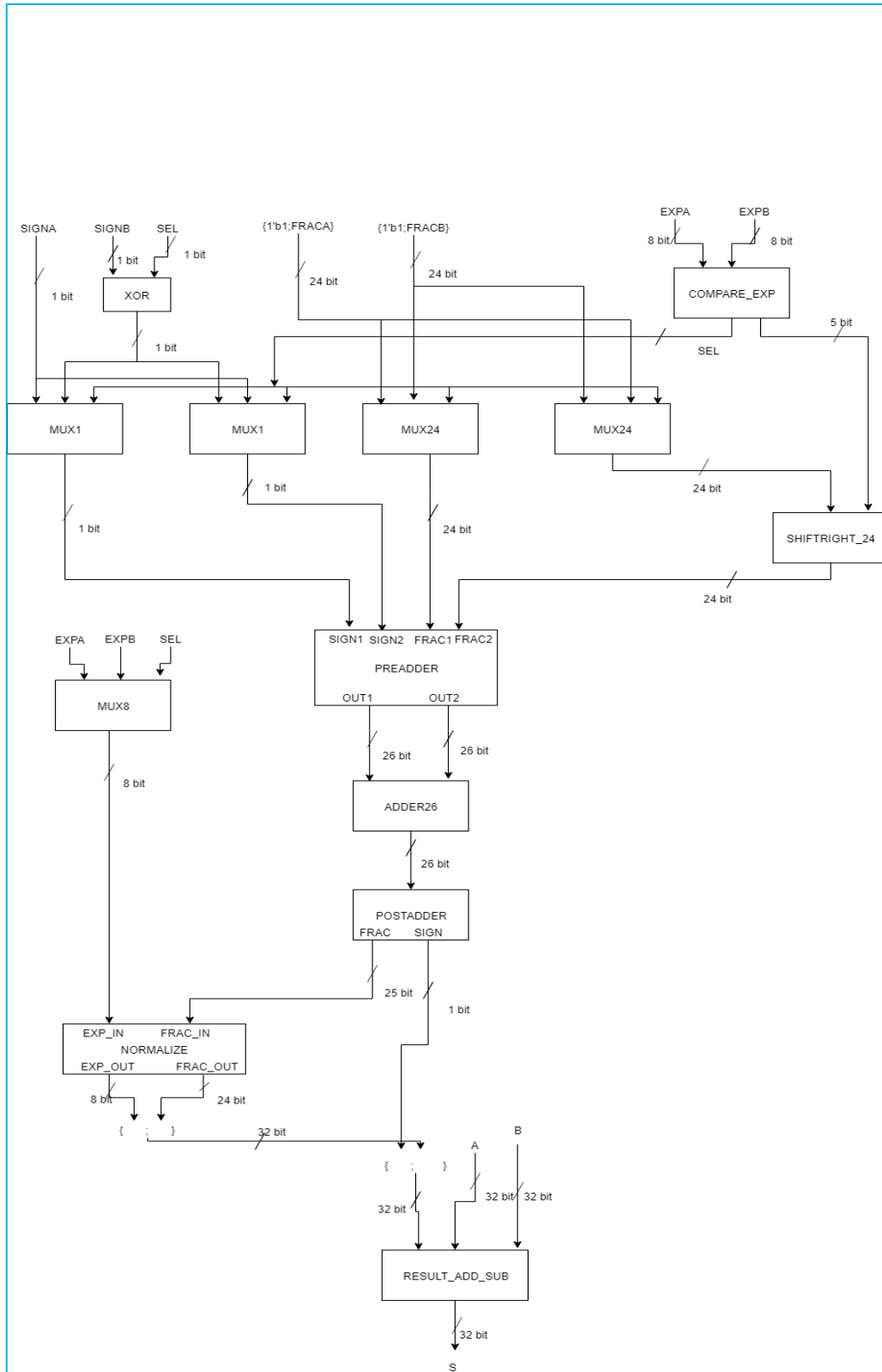
b. RCA 9bits



c. RCA 26bits



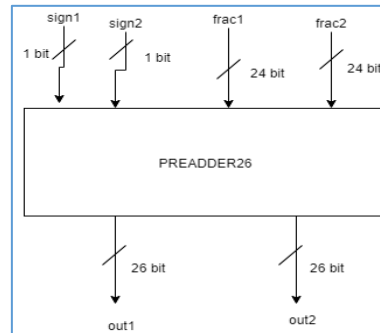
3. Bộ tổng quát



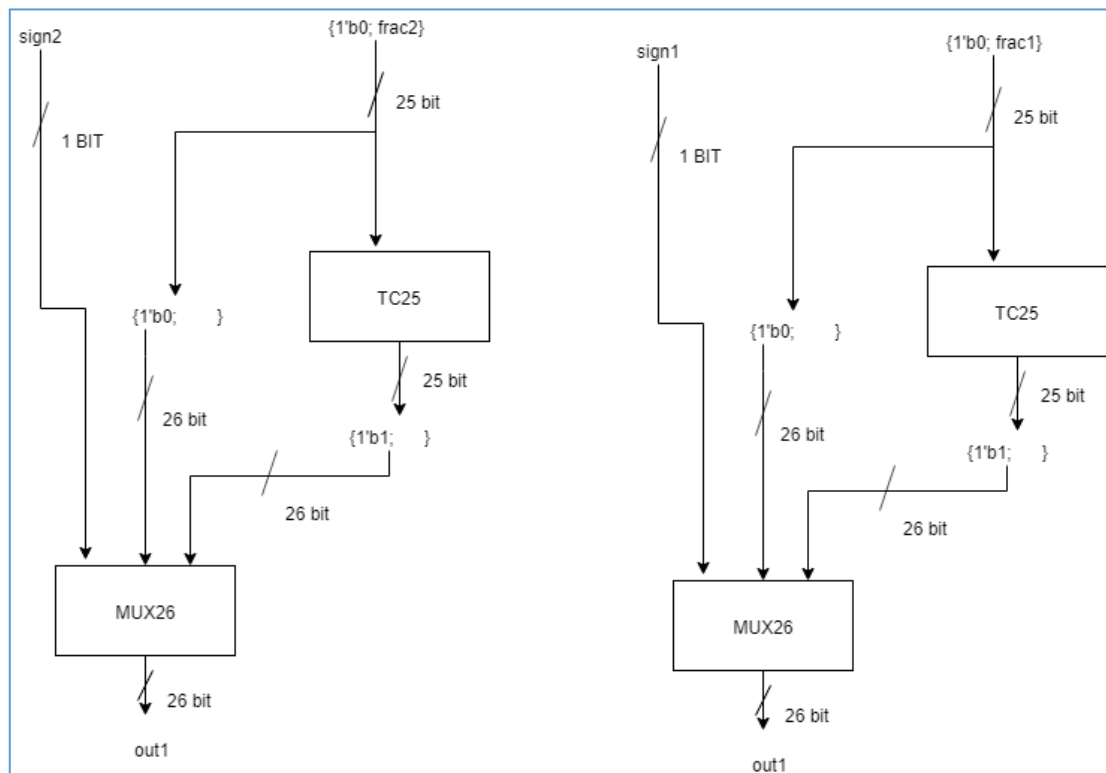
4. Các module nhỏ

a. Module Pre_Adder

- Mô hình chung

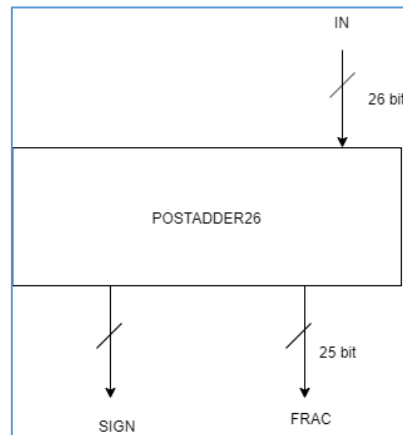


- Mô hình chi tiết

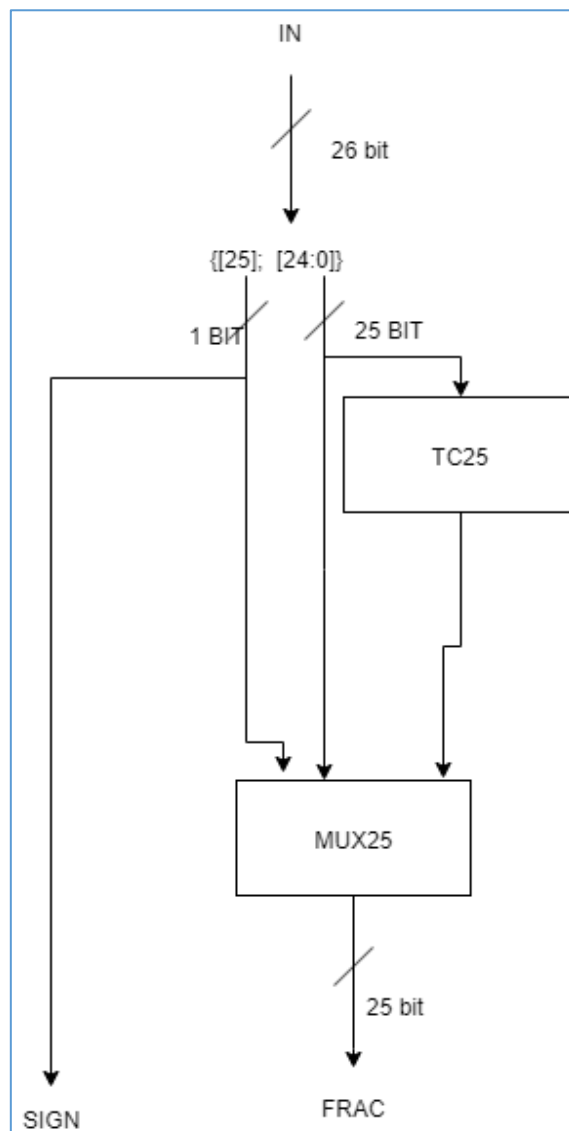


b. Module Post_Adder

- Mô hình chung



- Mô hình chi tiết



c. Module ResultS_addsub

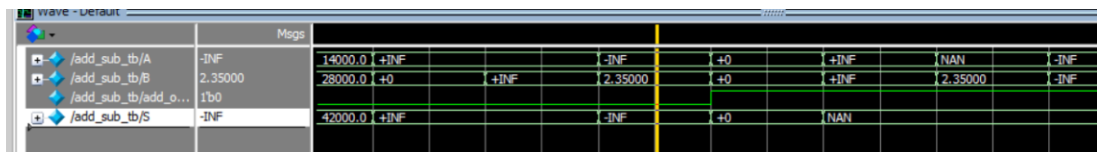
A	B	S
Inf	-Inf	NaN
-Inf	Inf	
NaN	NaN	
NaN	x	
x	NaN	
Inf	x	Inf
x	Inf	Inf
A	0	A
0	B	B
0	0	0
A	B	AnsS

5. Kết quả mô phỏng



Wave - Default		Msgs									
/add_sub_tb/A	-INF	25.3870	-30.1100	-15.2500	14000.0	+INF					
/add_sub_tb/B	2.35000	19.9375	62.2060	2.35000	28000.0	+0					+INF
/add_sub_tb/add_o...	1'b0										
/add_sub_tb/S	-INF	45.3245	32.0960	-17.6000	42000.0	+INF					

Hình 1. Mô phỏng với giải thuật CLA



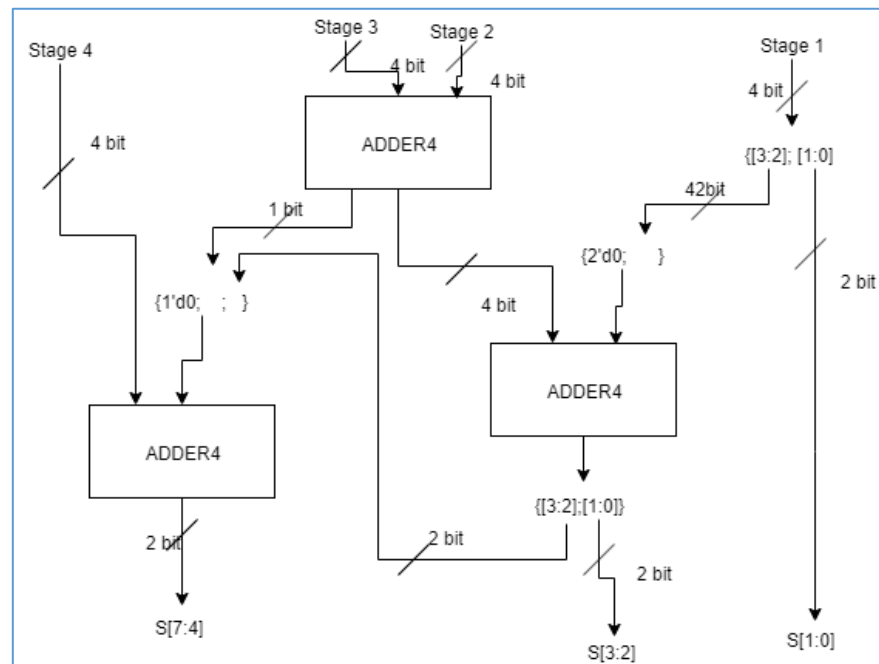
Wave - Default		Msgs									
/add_sub_tb/A	-INF	14000.0	+INF			-INF	+0		+INF	NAN	-INF
/add_sub_tb/B	2.35000	28000.0	+0		+INF	2.35000	+0		+INF	2.35000	-INF
/add_sub_tb/add_o...	1'b0										
/add_sub_tb/S	-INF	42000.0	+INF			-INF	+0		NAN		

Hình 2. Mô phỏng với giải thuật RCA

Chương 3: Bộ Nhân

1. Giải thuật Double

a. Double 2x2



b. Double 24x24

2. Giải thuật vedic

a. Mô tả Vedic 4bits

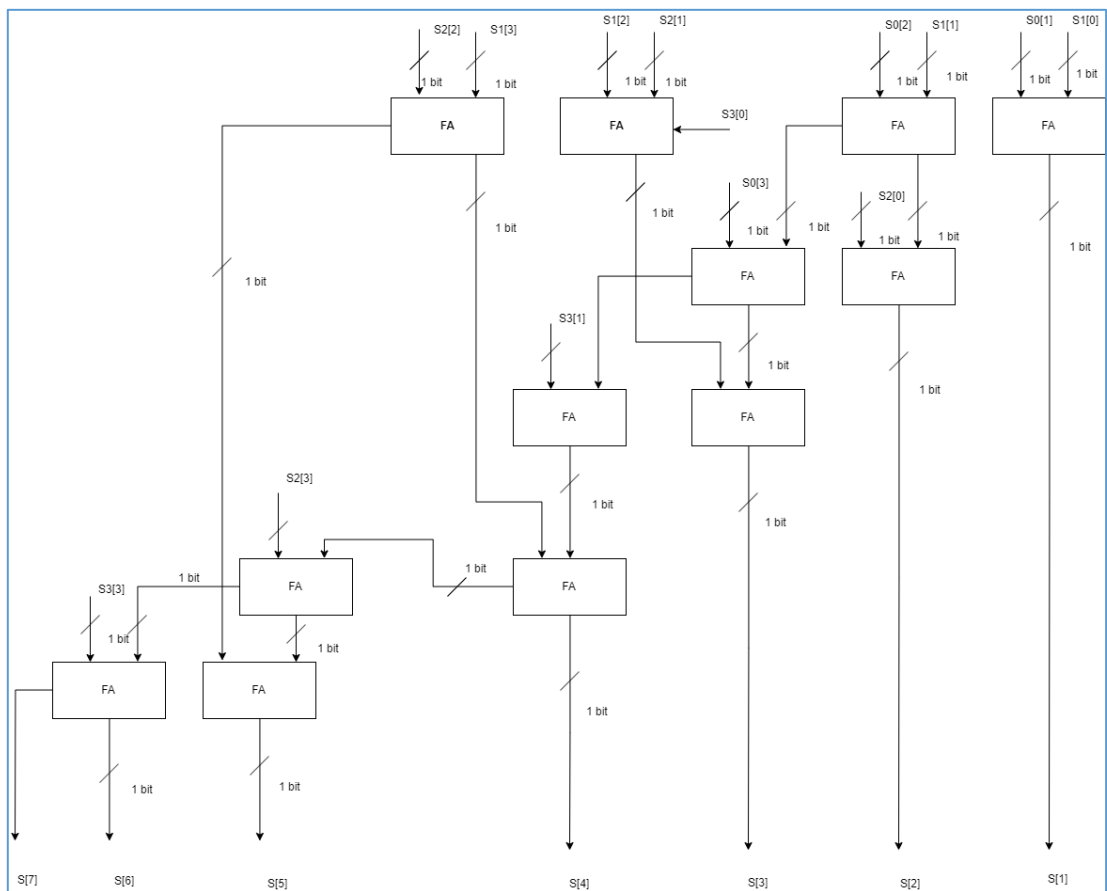
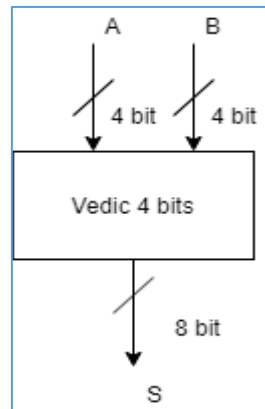
$$S = A * B$$

A: binary 4 bits

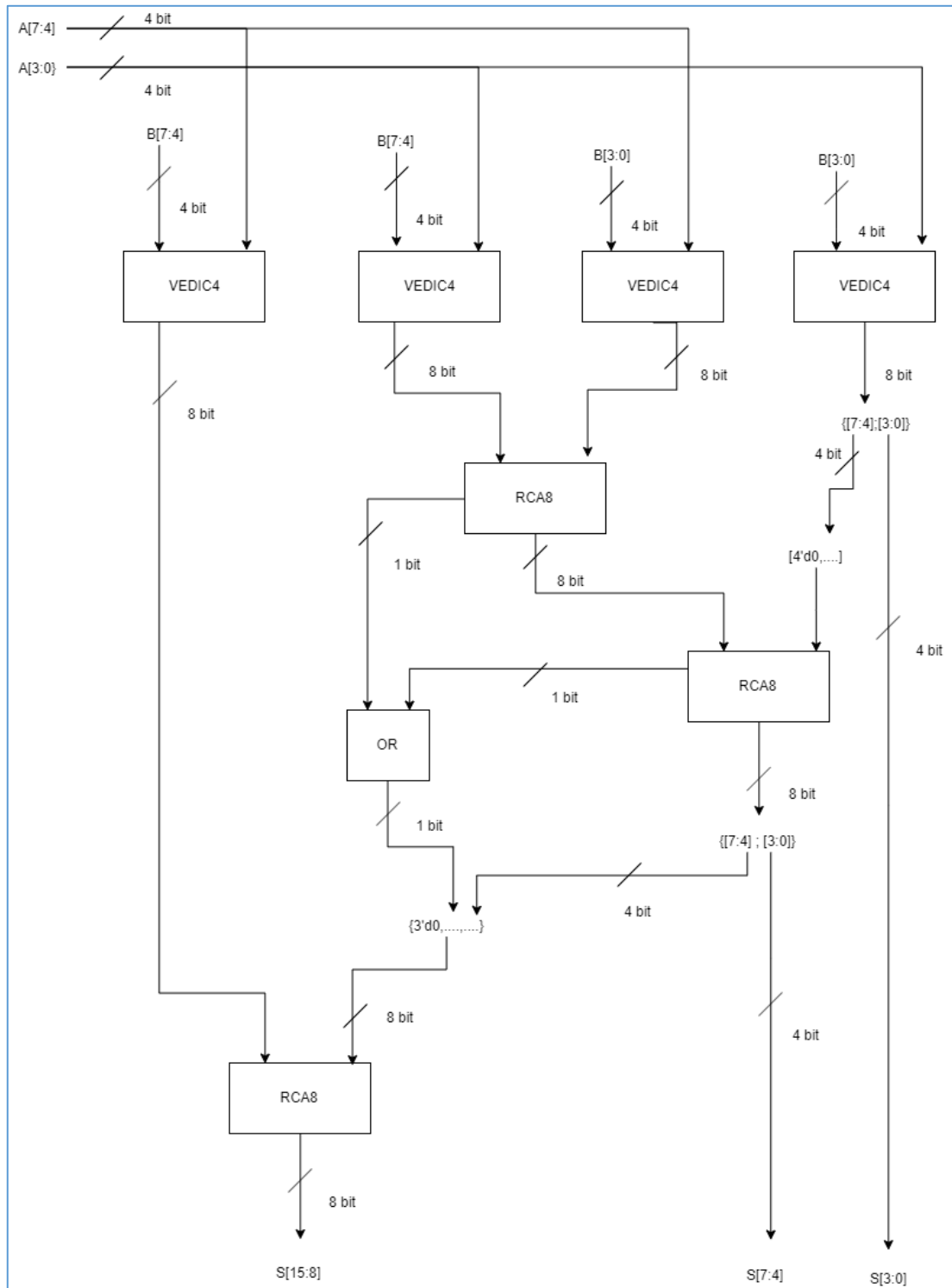
B: binary 4 bits

S: binary 8 bits

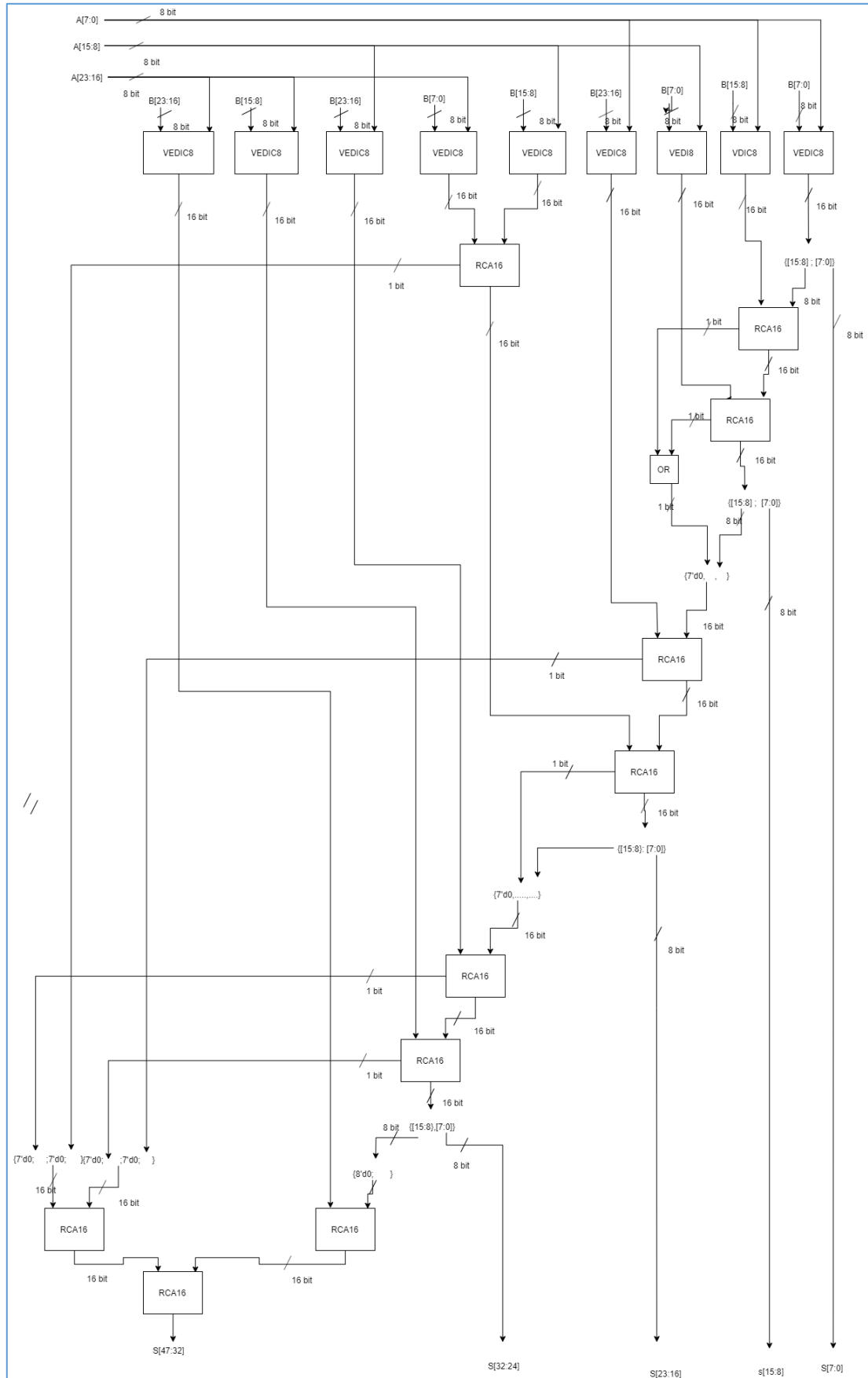
b. Chi tiết



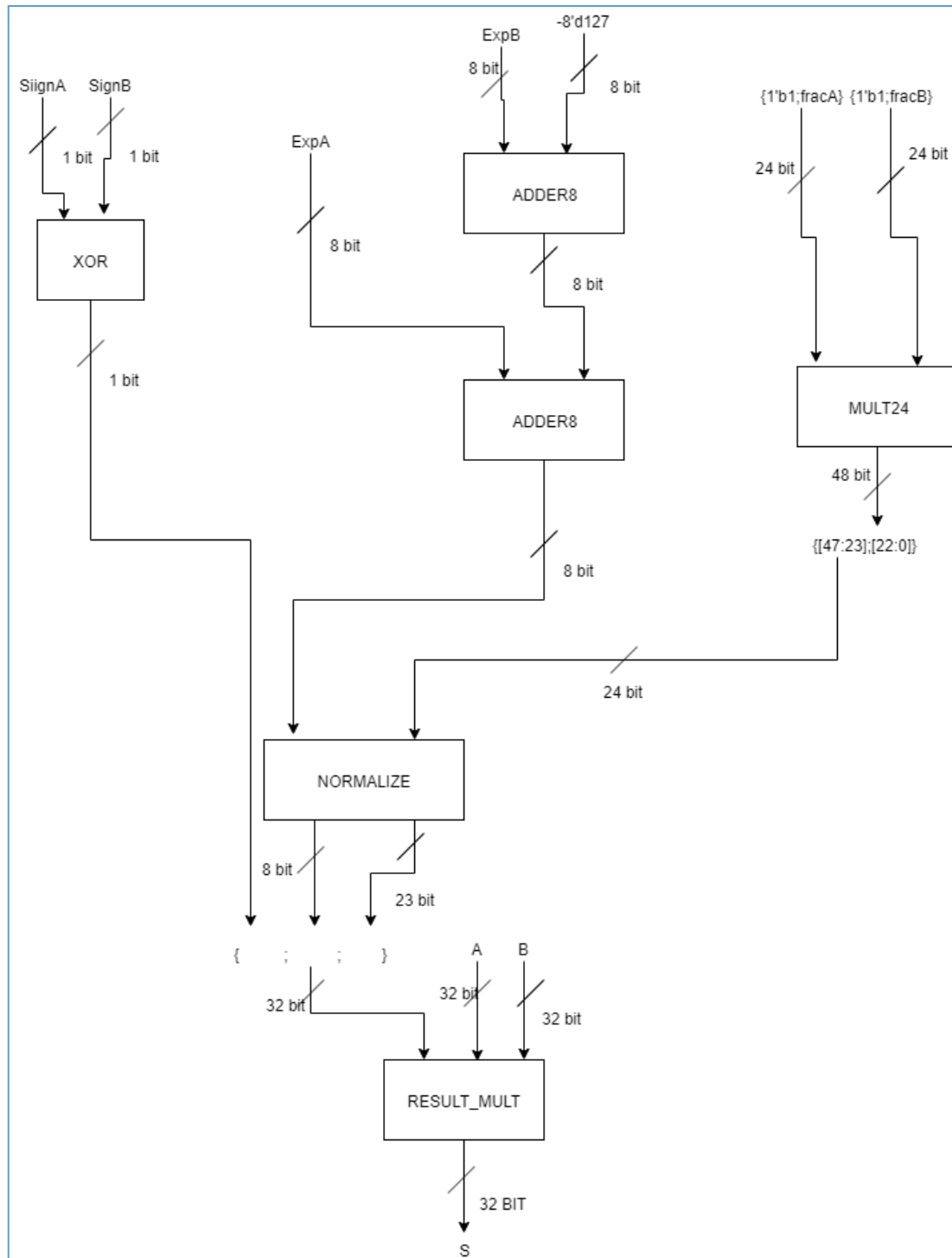
c. Vedic 8bits



d. Vedic 24bits



3. Bộ tổng quát

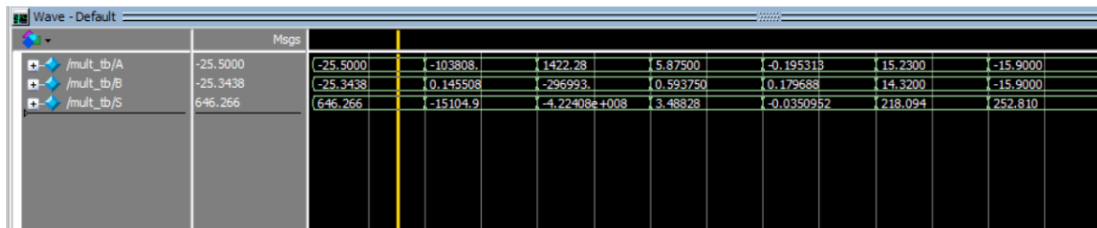


4. Các module nhỏ

a. Result_mult

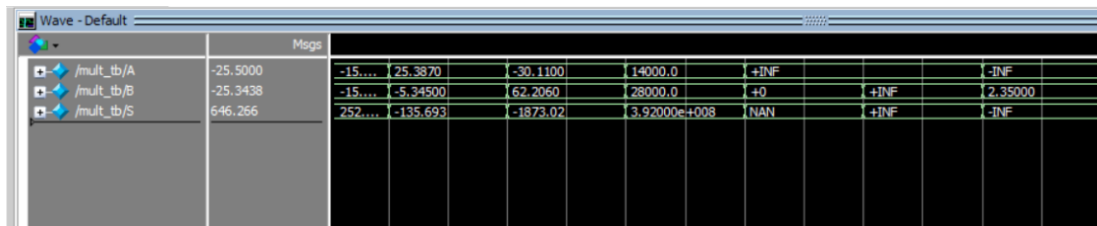
A	B	S
inf	inf	Inf
0	inf	NaN
inf	0	
NaN	X	
X	NaN	
NaN	NaN	
0	X	0
X	0	
0	0	
A	B	AnsS

5. Kết quả mô phỏng



Wave - Default	Msgs									
/mult_tb/A	-25.5000	-25.5000	-103808.	1422.28	5.87500	-0.195313	15.2300	-15.9000		
/mult_tb/B	-25.3438	-25.3438	0.145508	-296993.	0.593750	0.179688	14.3200	-15.9000		
/mult_tb/S	646.266	646.266	-15104.9	-4.22408e+008	3.48828	-0.0350952	218.094	252.810		

Hình 1. Mô phỏng với giải thuật Vedic

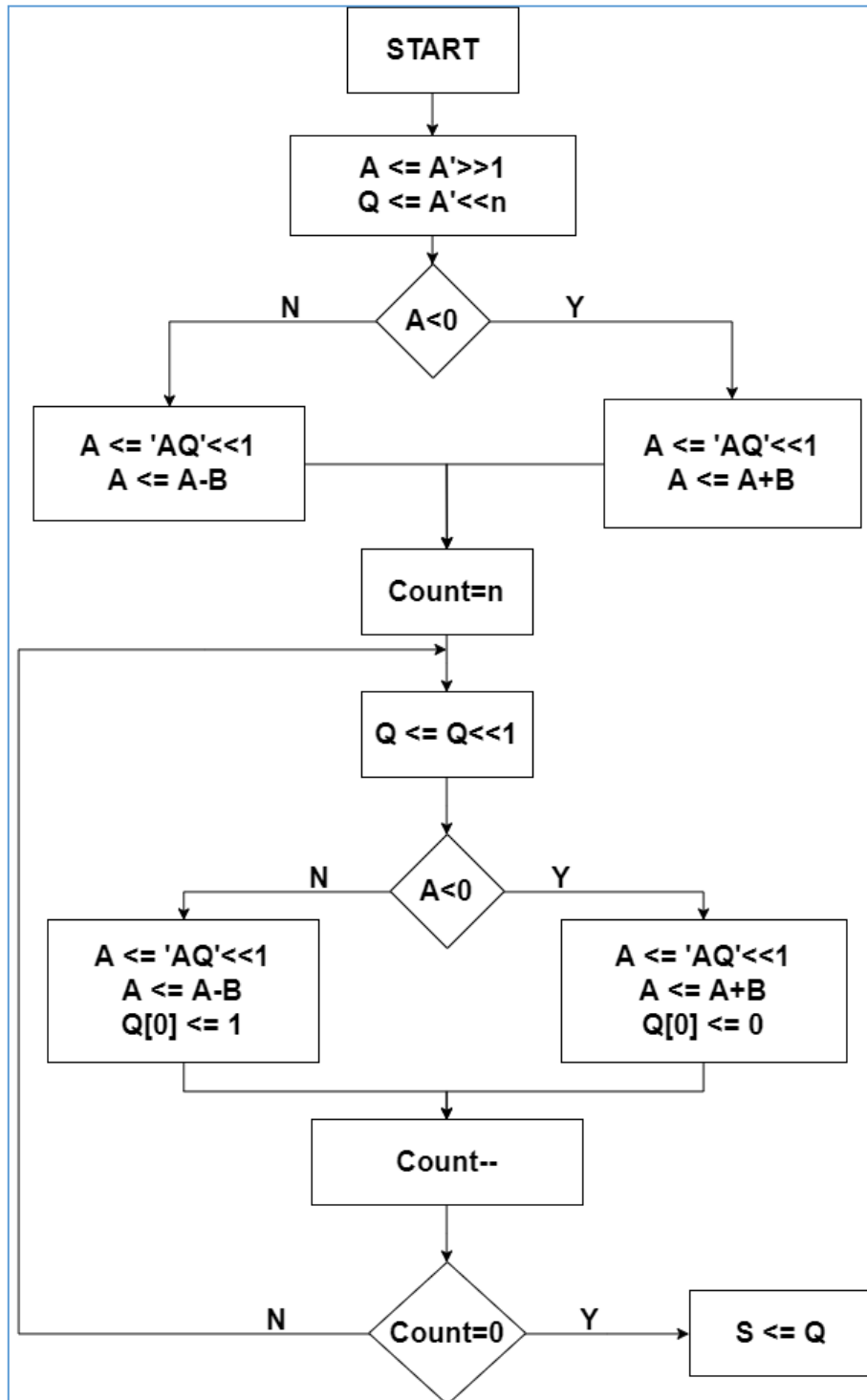


Wave - Default	Msgs									
/mult_tb/A	-25.5000	-15....	25.3870	-30.1100	14000.0	+INF		-INF		
/mult_tb/B	-25.3438	-15....	-5.34500	62.2060	28000.0	+0	+INF	2.35000		
/mult_tb/S	646.266	252....	-135.693	-1873.02	3.92000e+008	NAN	+INF	-INF		

Hình 2. Mô phỏng với giải thuật Double

Chương 4: Bộ Chia

1. Lưu đồ giải thuật Restoring-Division



2. Khối chính

a. Khối chung

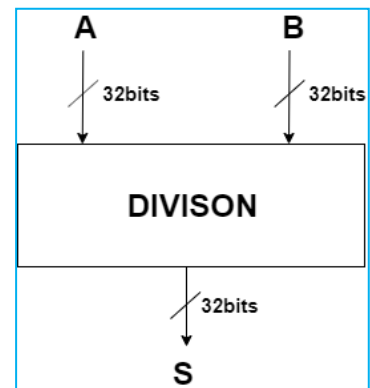
$$S = \frac{A}{B}$$

A: floating point 32bits

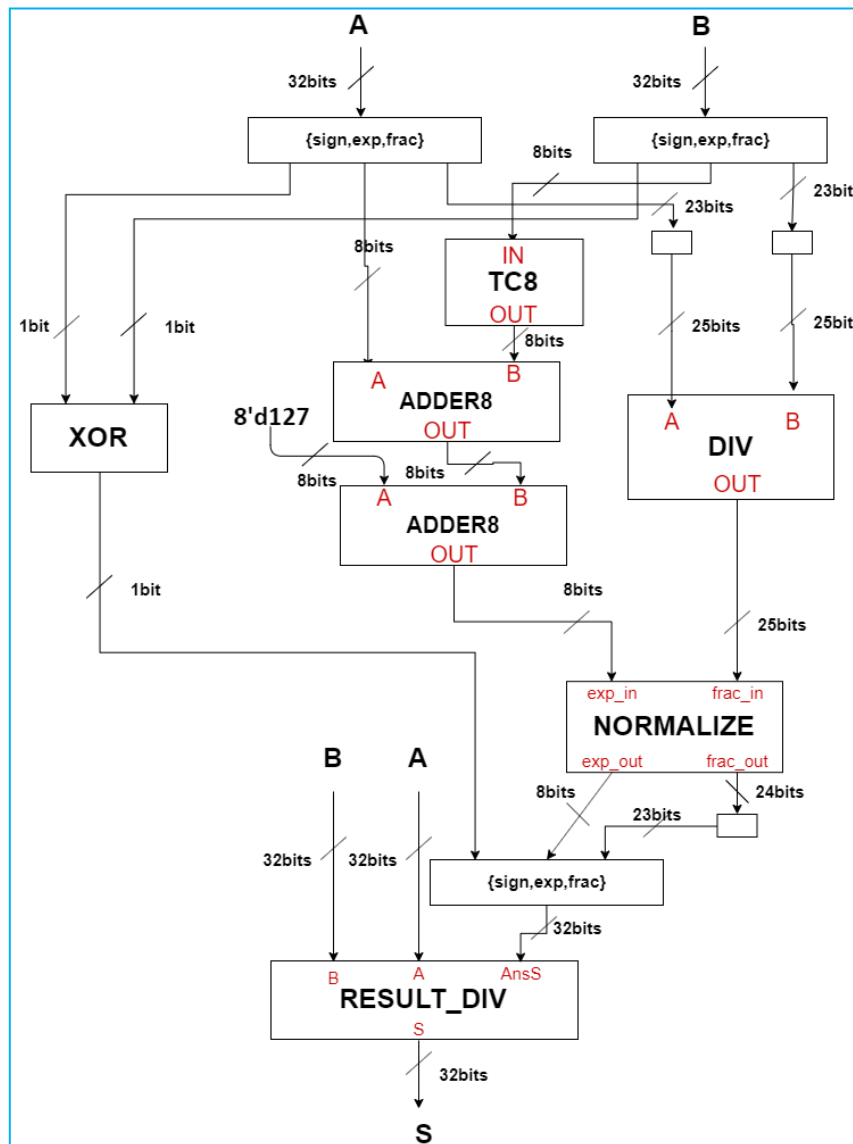
S: floating point 32bits

B: floating point 32bits

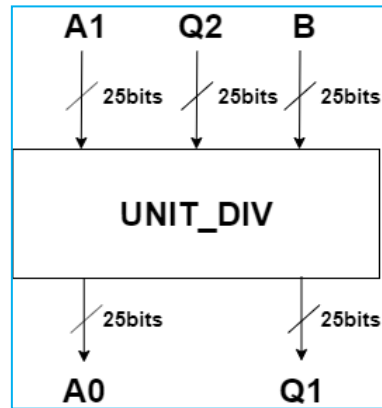
Giải thuật: Restoring_Division



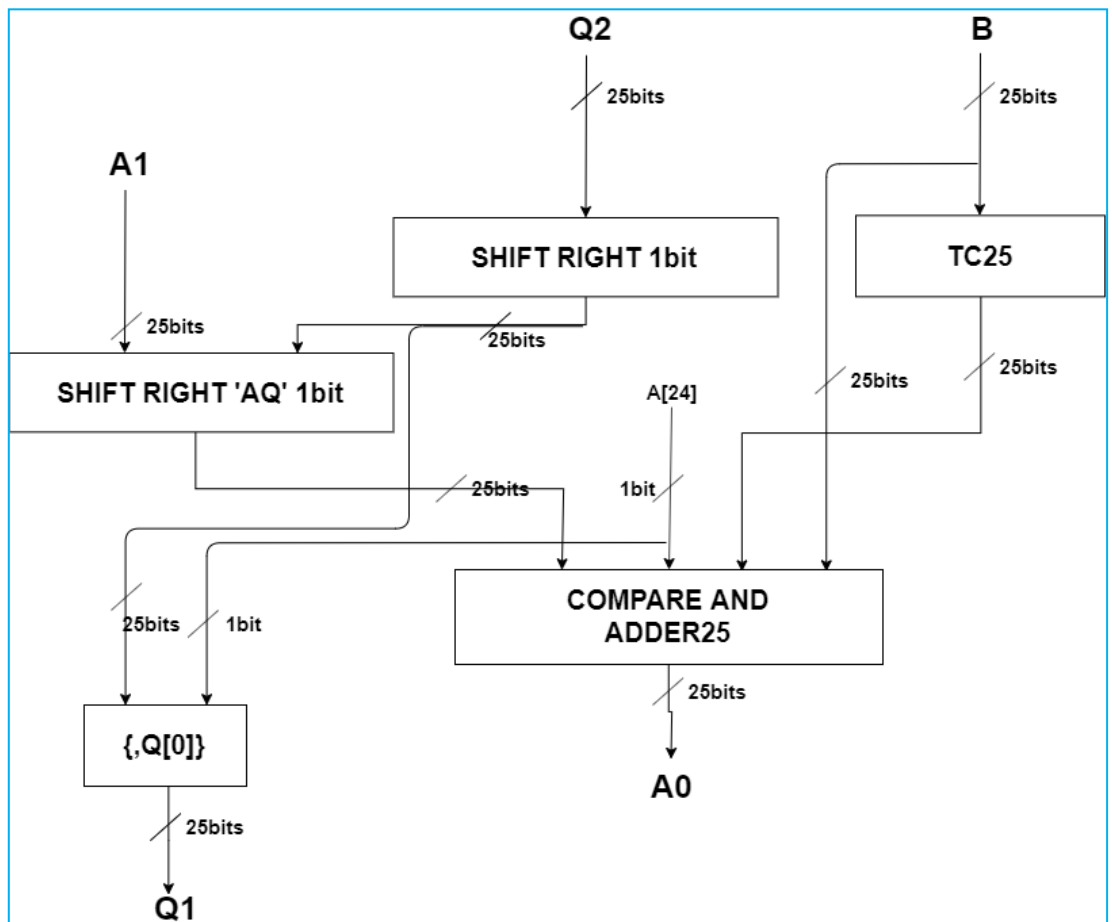
b. Khối chi tiết



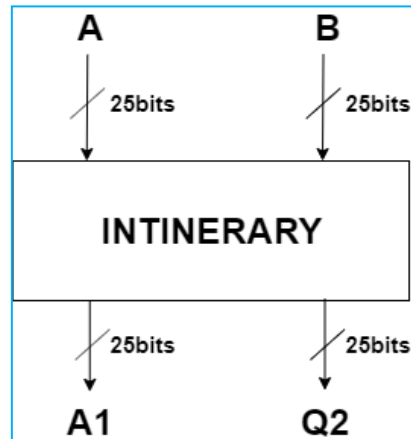
b. Module Unit_Div



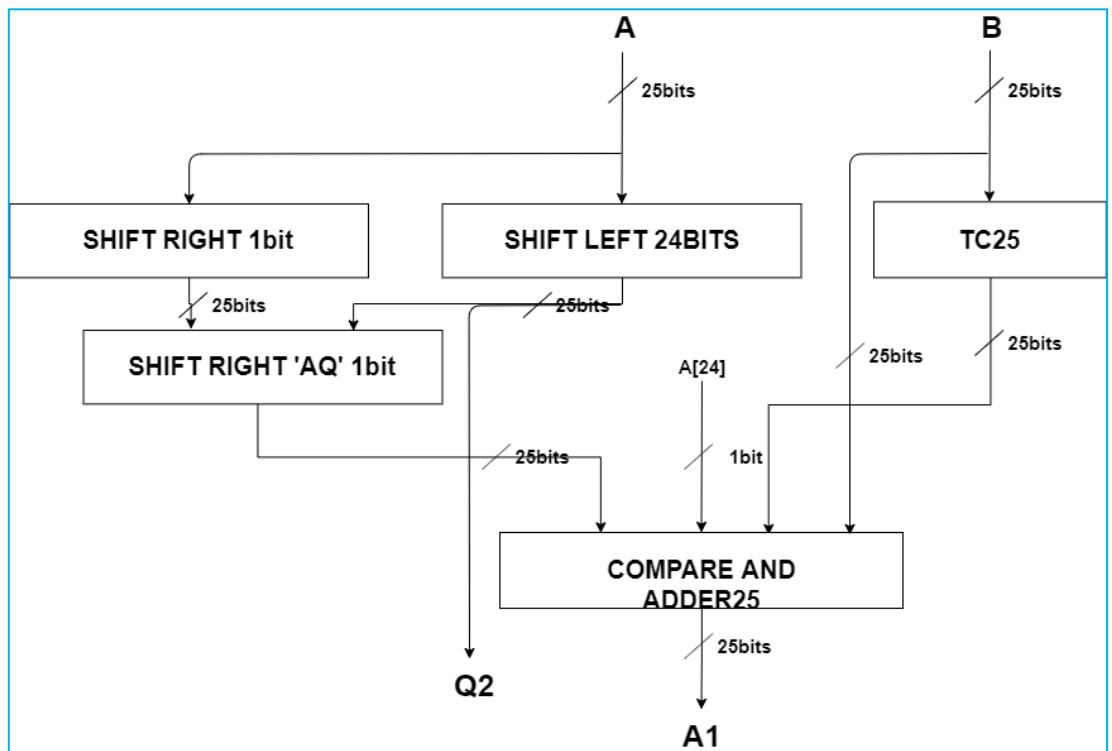
Chi tiết :



c. Module Initinray



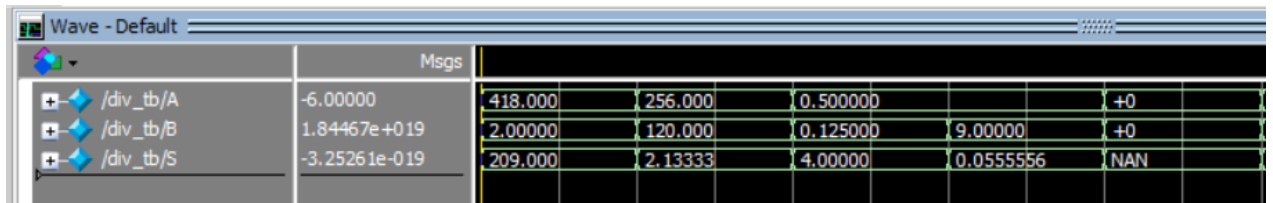
Chi tiết:



d. Module Results_Div

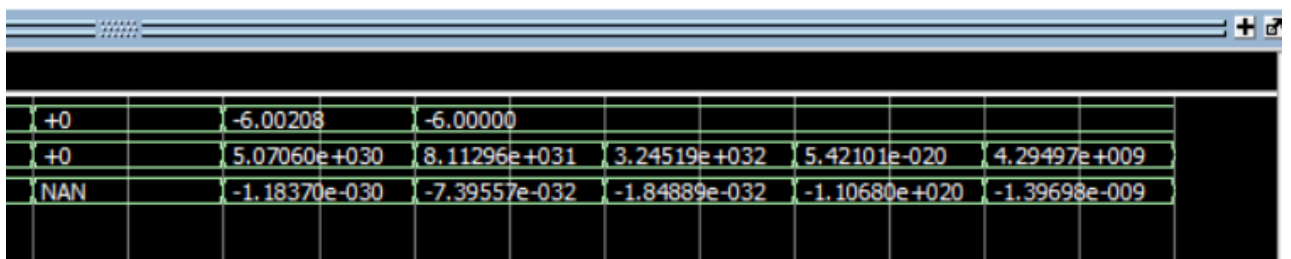
A	B	S
Inf	x	Inf
X	0	Inf
0	X	0
NaN	NaN	NaN
NaN	X	
X	NaN	
Inf	Inf	± 1
0	0	NaN
A	B	AnsS

4. Kết quả mô phỏng



Variable	Value
/div_tb/A	-6.00000
/div_tb/B	1.84467e+019
/div_tb/S	-3.25261e-019

Hình 1. Kết quả chạy mô phỏng lần 1



Variable	Value
/div_tb/A	-6.00208
/div_tb/B	5.07060e+030
/div_tb/S	-1.18370e-030

Hình 2. Kết quả chạy mô phỏng lần 2

Chương 5: Bộ FPAU

1. Mô hình

A: floating point 32bits
 S: floating point 32bits
 B: floating point 32bits
 SEL: unsigned 2bits

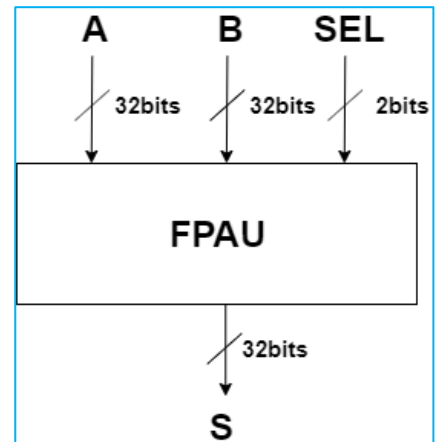
SEL:

+ 0 : $S = A + B$

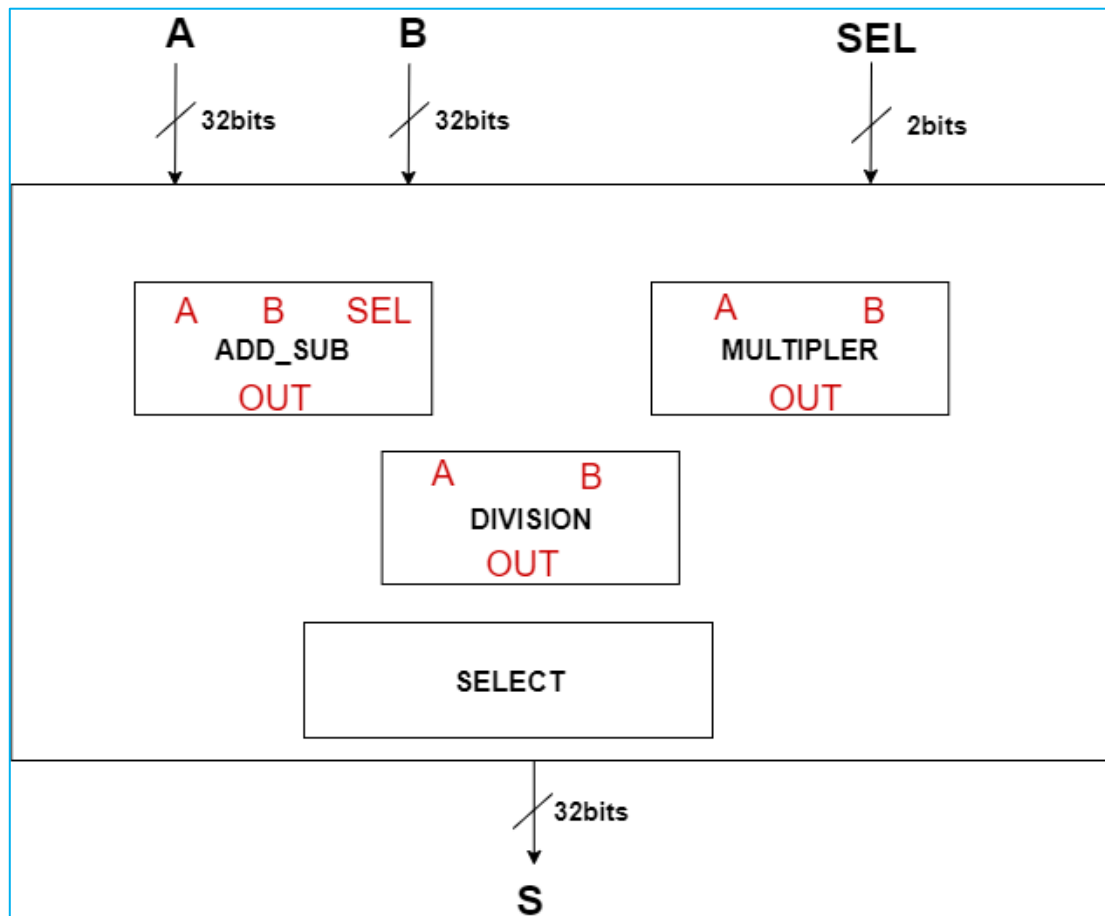
+ 1 : $S = A - B$

+ 2 : $S = A \times B$

+ 3 : $S = \frac{A}{B}$



- Chi tiết :



2. Kết quả chạy mô phỏng

	Msgs								
/FPAU_tb/A	NAN	-25.5000	-103808.	1422.28	5.87500	-0.195313			
/FPAU_tb/B	-15....	-25.3438	0.145508	-296993.	0.593750	0.179688			
/FPAU_tb/sel	Zd0	Zd2	Zd1	Zd3	Zd0	Zd2			
/FPAU_tb/S	NAN	646.266	-103808.	-0.00478893	6.46875	-0.0350952			

Hình 1. Kết quả chạy mô phỏng lần 1

15.2300	-15.9000	15.3200	-15.9000	25.3870					
14.3200	-15.9000	15.3200	-15.9000	-5.34500					
	Zd1	Zd2	Zd3	Zd1					
218.094	+0	234.702	1.00000	30.7320					

Hình 2. Kết quả chạy mô phỏng lần 2

15.3200	-15.9000	25.3870	-30.1100	14000.0	+INF	-INF			
15.3200	-15.9000	-5.34500	62.2060	28000.0	+0	+INF	2.35000		
Zd2	Zd3	Zd1	Zd3	Zd2	Zd1	Zd0	Zd3		
234.702	1.00000	30.7320	-0.484037	3.92000e+008	+INF	-INF			

Hình 2. Kết quả chạy mô phỏng lần 3

Chương 6 : Bộ Lũy Thừa

1. Module tổng quát

a. Mô tả:

$$S = A^n$$

A: floating point 32bits

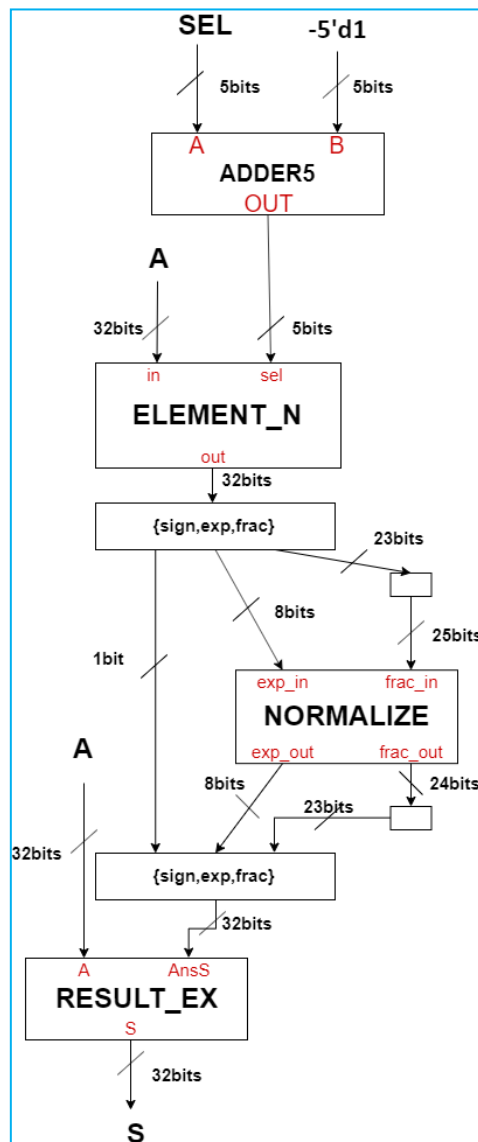
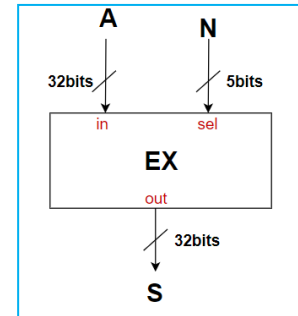
S: floating point 32bits

N: unsigned 5bits

b. Giải thuật :

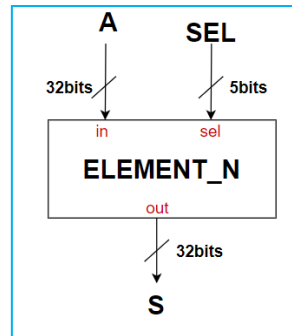
Cổ điển: $S_{n+1} = S_n * A; S_0 = A$

c. Chi tiết

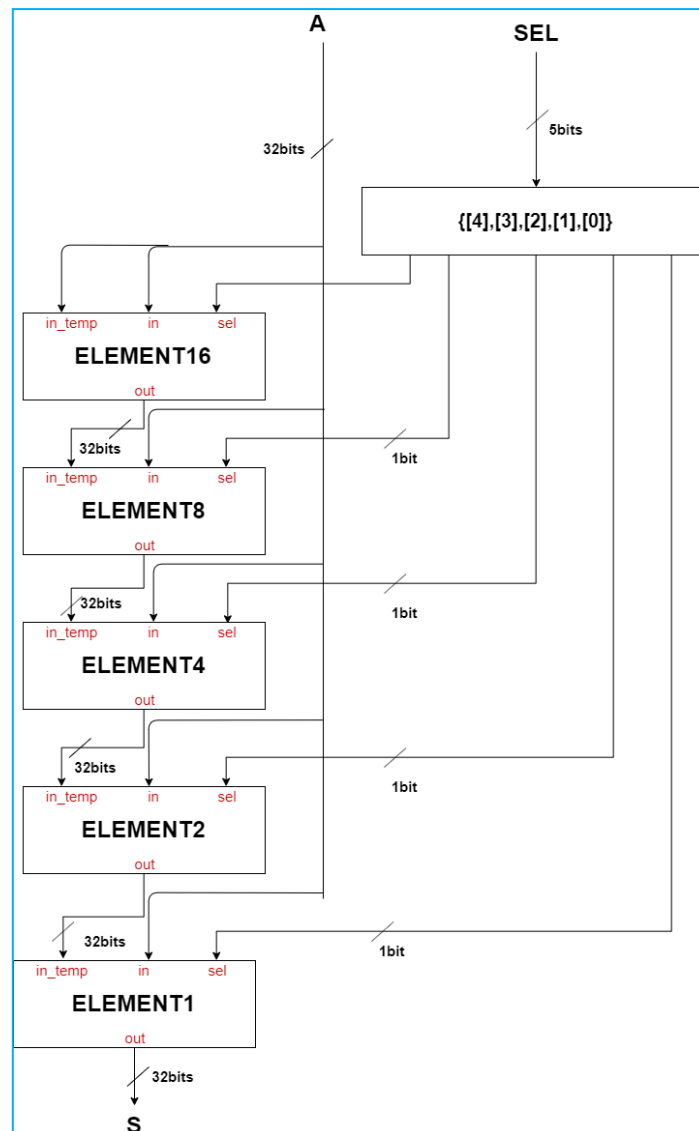


2. Module Element_N

a. Mô hình chung



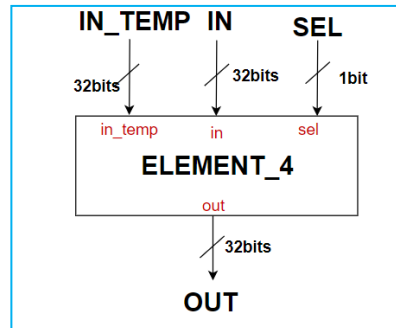
b. Mô hình chi tiết



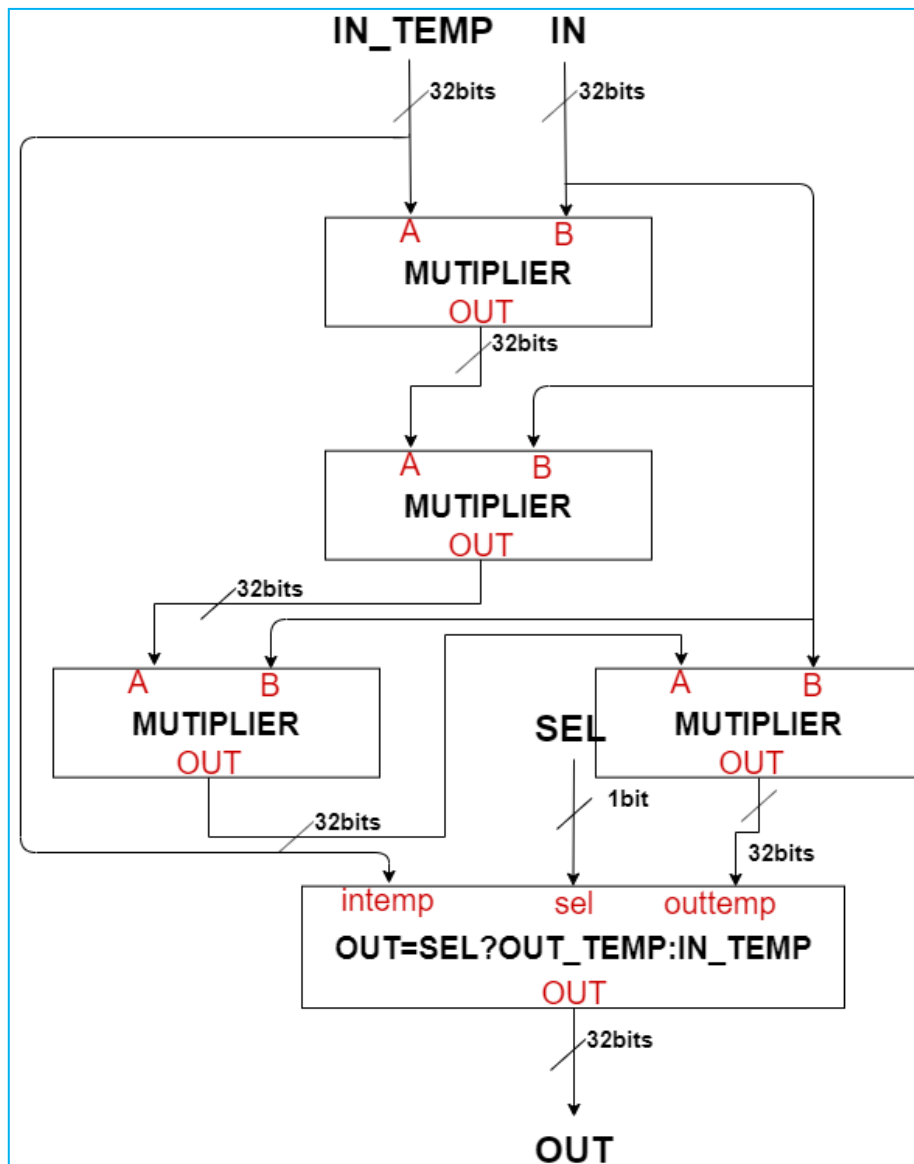
3. Module Element_4

a. Mô tả:

$$Out = In_temp.In$$



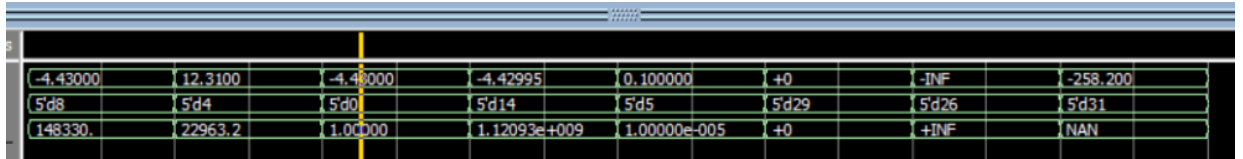
b. Chi tiết



4. Result_Ex

n	A	S
Odd	inf	signA,inf
Event		inf
X	NaN	NaN
0	0	NaN
0	A	1
X	A	Ans

5. Kết quả mô phỏng



-4.43000	12.3100	-4.43000	-4.42995	0.100000	+0	-INF	-258.200
5d8	5d4	5d0	5d14	5d5	5d29	5d26	5d31
148330.	22963.2	1.00000	1.12093e+009	1.00000e-005	+0	+INF	NAN

Hình1. Kết quả mô phỏng bộ lũy thừa

Chương 7: Bộ Căn bậc 2

1. Module tổng quát

a. Mô tả:

$$S = \sqrt{A}$$

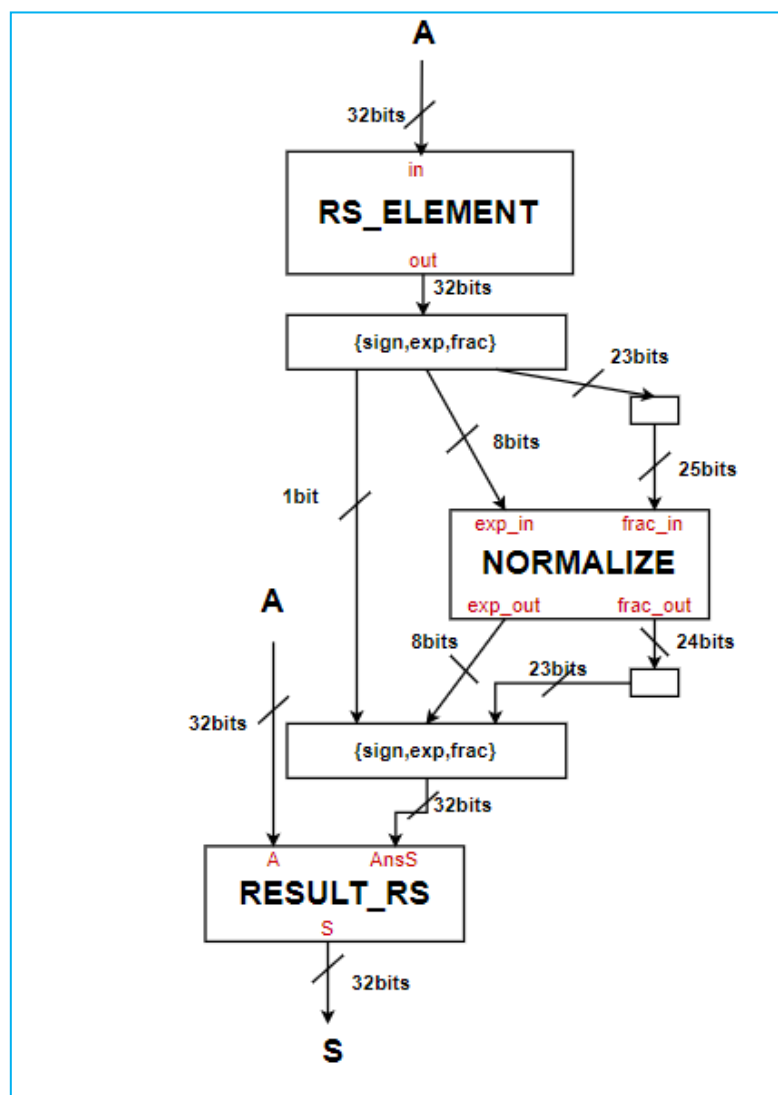
A: floating point 32bits

S: floating point 32bits

b. Giải thuật:

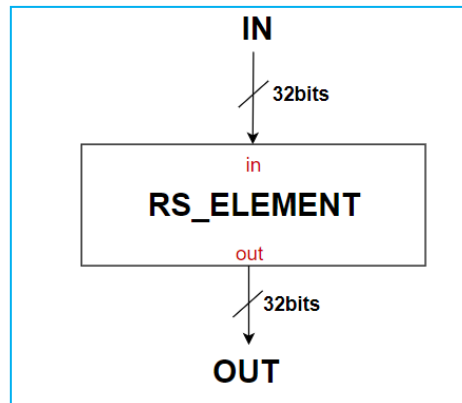
$$\text{Approximately: } S_{n+1} = S_n + \left(\frac{A}{S_n} - S_n \right) / 2$$

c. Chi tiết:

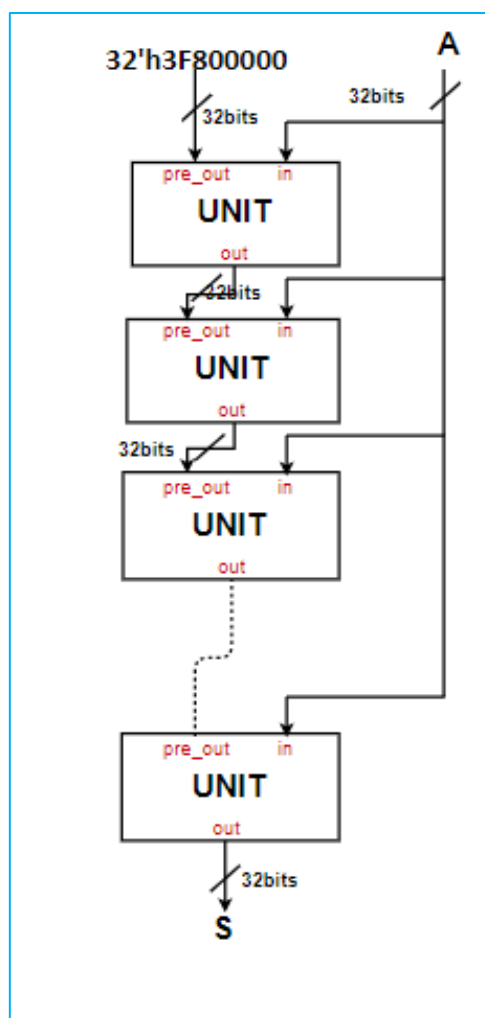


2. Module Rs_Element

a. Mô hình chung



b. Mô hình chi tiết

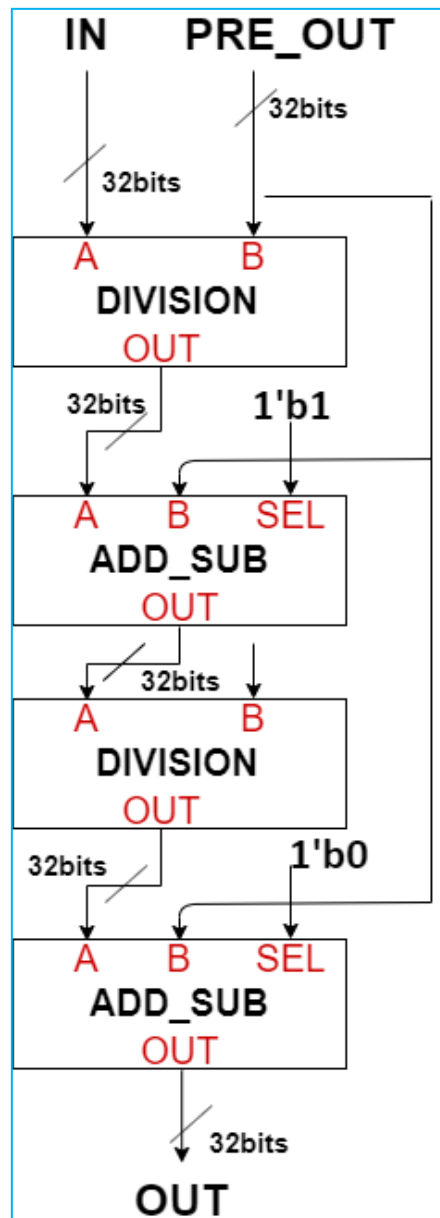


3. Module Unit

a. Mô tả

$$OUT = PRE_OUT + \left(\frac{A}{PRE_OUT} - PRE_OUT \right) / 2$$

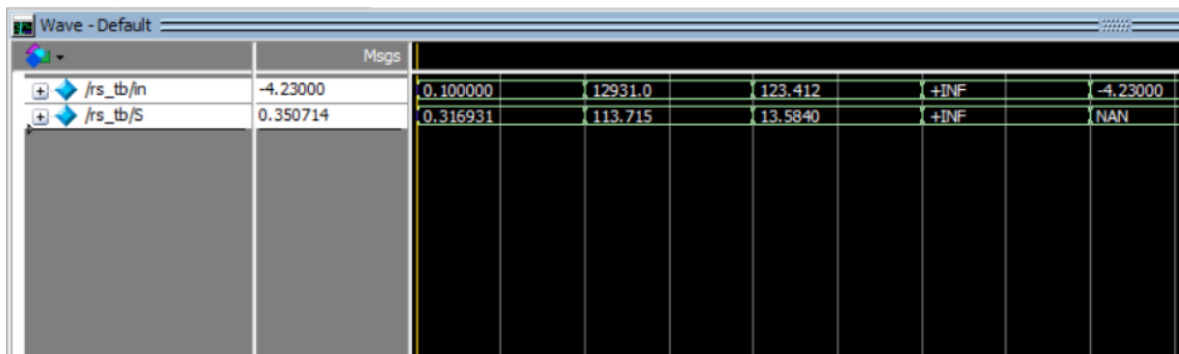
b. Chi tiết



4. Result_Rs

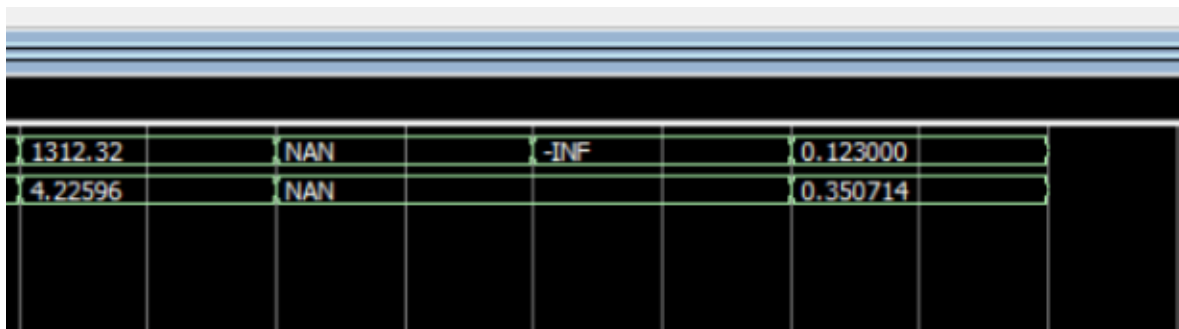
A	S
Inf	Inf
-Inf	NaN
0	0
NaN	NaN
$A > 0$	AnsS
$A < 0$	NaN

5. Kết quả mô phỏng



Wave - Default		Msgs							
+	/rs_tb/in	-4.23000	0.100000	12931.0	123.412	+INF	-4.23000		
+	/rs_tb/S	0.350714	0.316931	113.715	13.5840	+INF	NAN		

Hình 1. Kết quả mô phỏng lần 1



1312.32	NAN	-INF	0.123000		
4.22596	NAN		0.350714		

Hình 2. Kết quả mô phỏng lần 2

Chương 8: Bộ Căn bậc n

1. Mô hình chung

a. Mô tả:

$$S = \sqrt[n]{A}$$

A: floating point 32bits

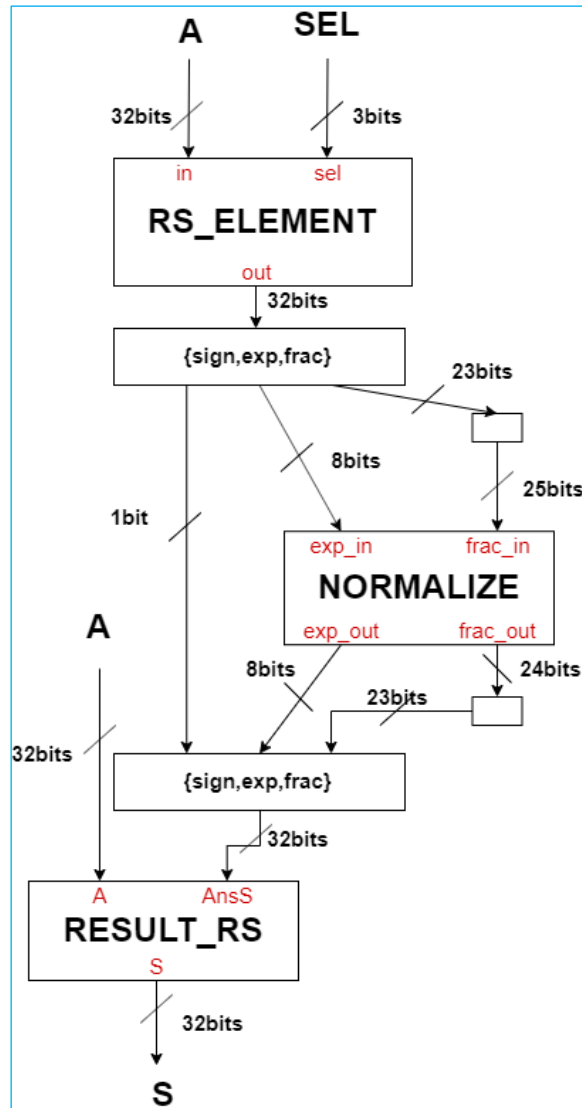
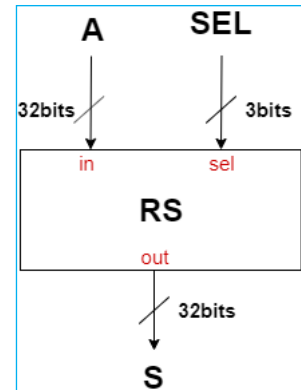
S: floating point 32bits

N: unsigned 3bits

b. giải thuật :

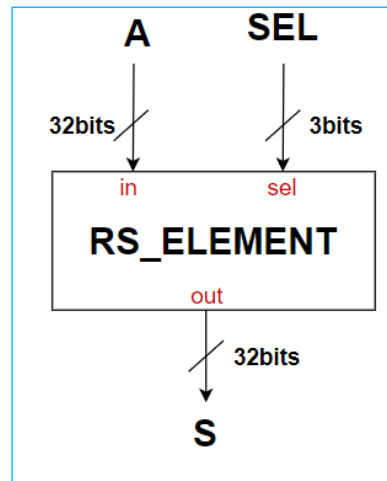
$$\text{Approximately: } S_{n+1} = S_n - \frac{S_n}{n} + \frac{A}{S_n^{n-1} \cdot n}$$

c. Chi tiết:

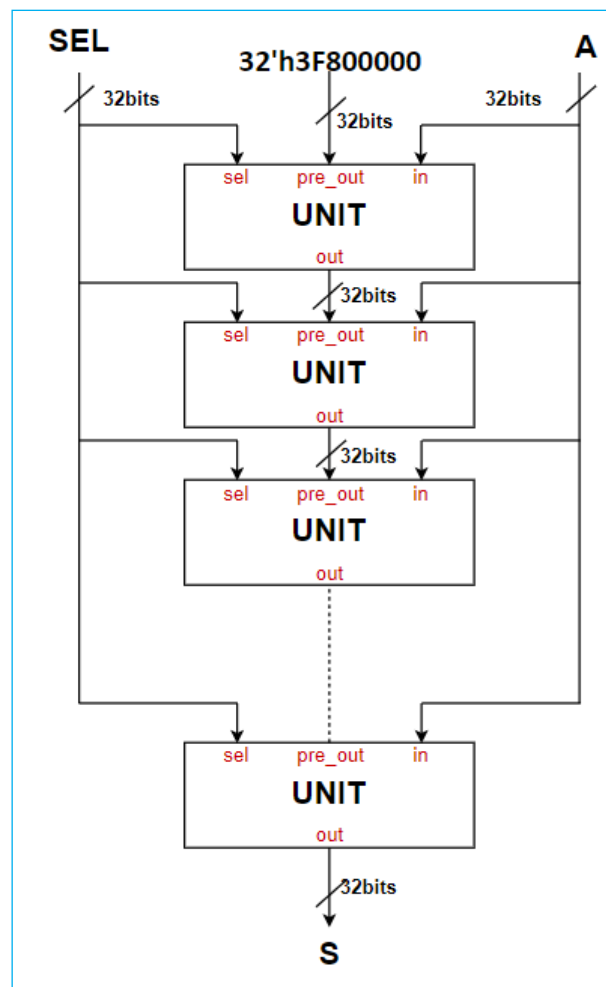


2. Module Rs_Element

a. Mô hình chung



b. Mô hình chi tiết

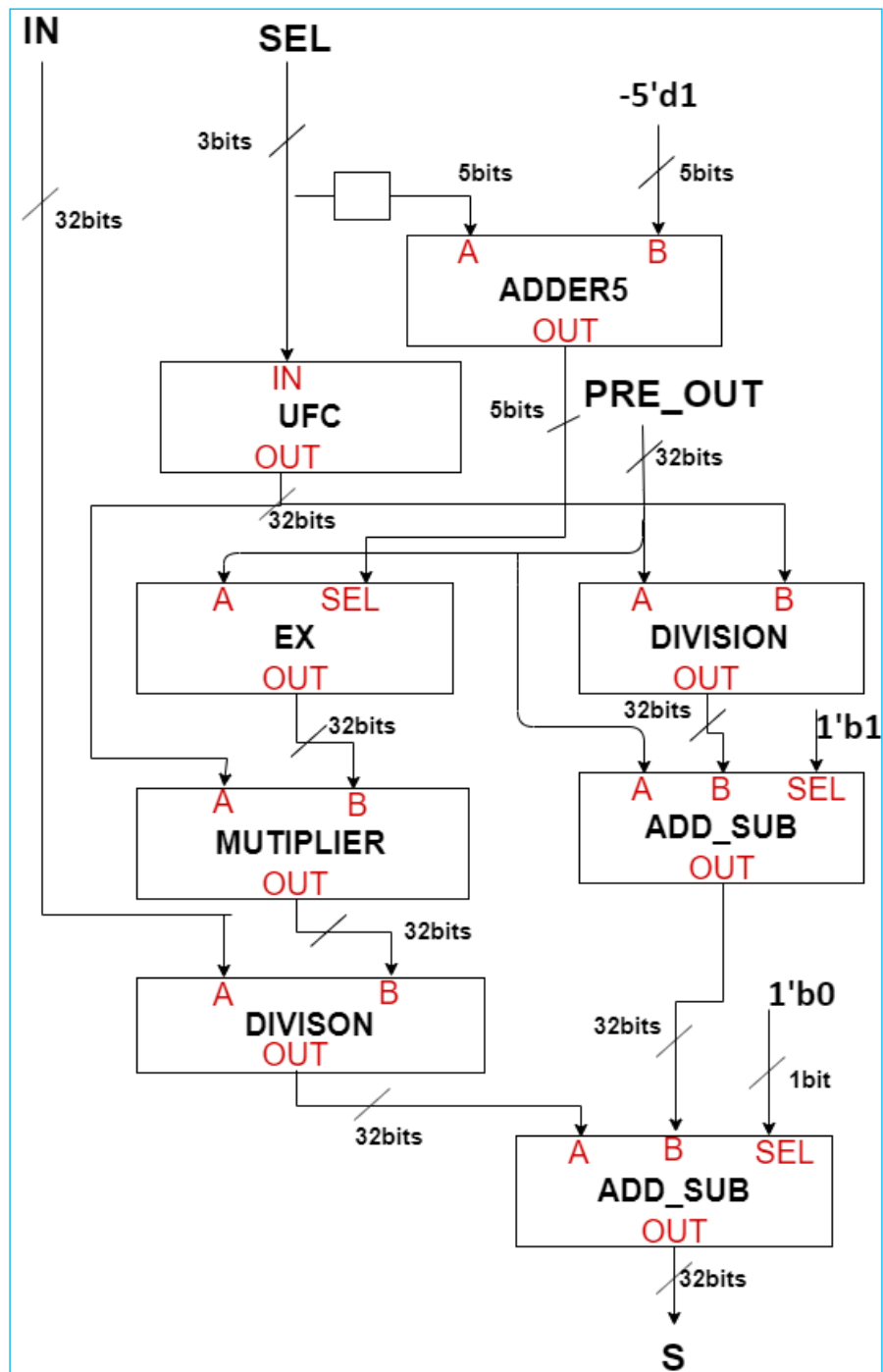


3. Module Unit

a. Mô tả:

$$Out = Pre_out - \frac{Pre_out}{n} + \frac{In}{Pre_out^{n-1}.n}$$

b. Mô hình chi tiết



4. Result_Rs

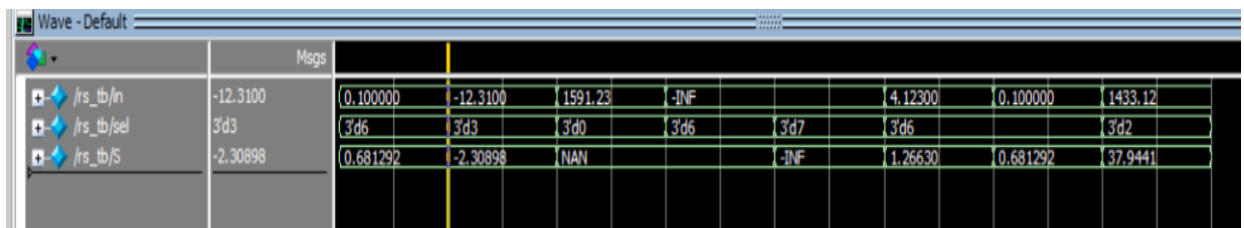
n	A	S
0	X	NaN
1	X	A
X	Inf	Inf
Odd	-Inf	-Inf
Event		NaN
X	0	0
X	NaN	NaN
Odd	x	ansS
Event	A<0	NaN
	A>0	ansS

5. Kết quả mô phỏng



Wave - Default	Msgs								
/rs_tb/in	942.124	(0.100000)	1423.42	942.124	-124.210	-96.6900	20.3100	25.7300	
/rs_tb/sel	3'd2	3'd0	3'd1	3'd2	3'd3	3'd4	3'd5	3'd7	
/rs_tb/S	30.6940	(NAN)	1423.42	30.6940	-4.98944	(NAN)	1.82617	1.59035	

Hình 1. Kết quả mô phỏng lần 1

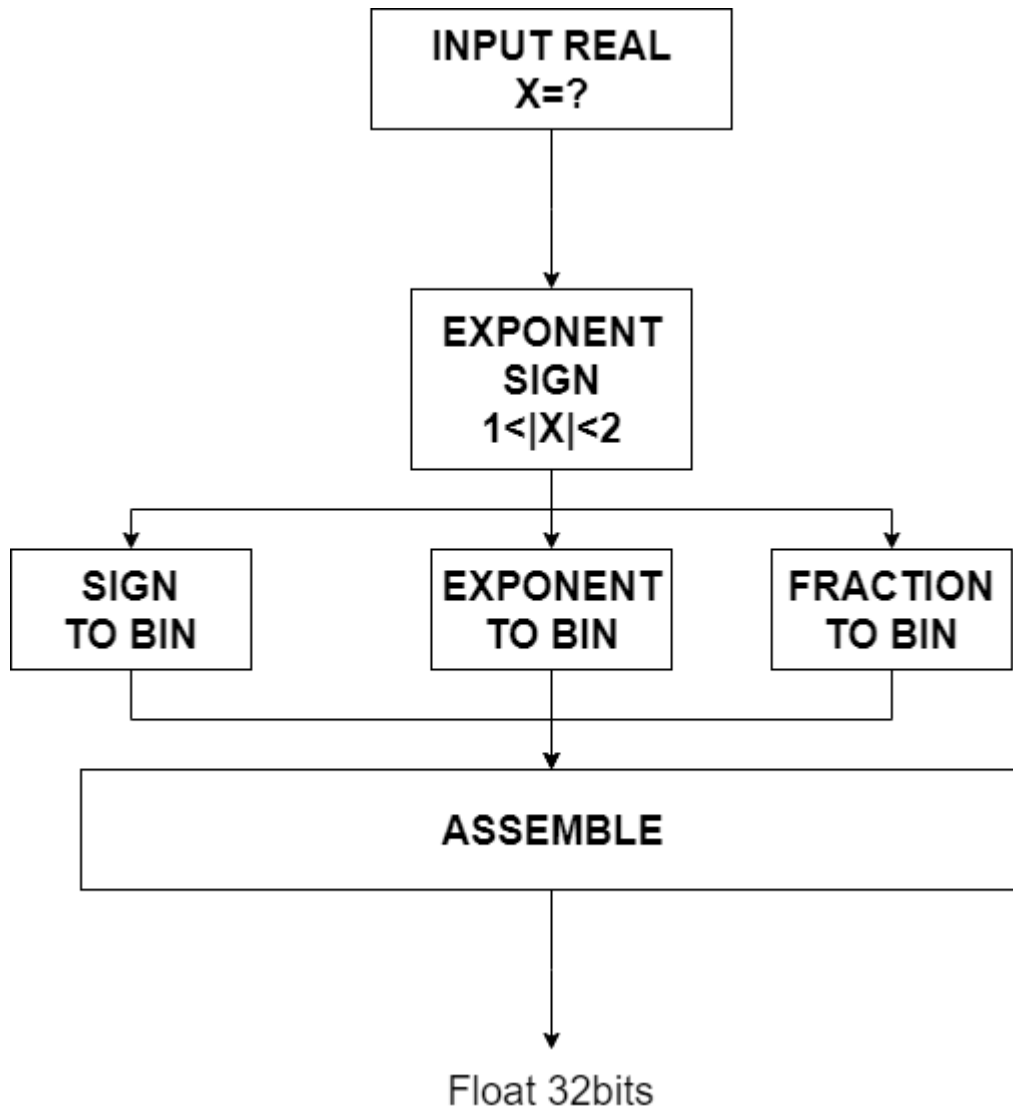


Wave - Default	Msgs								
/rs_tb/in	-12.3100	(0.100000)	-12.3100	1591.23	-INF		4.12300	0.100000	1433.12
/rs_tb/sel	3'd3	3'd6	3'd3	3'd0	3'd6	3'd7	3'd6		3'd2
/rs_tb/S	-2.30898	(0.681292)	-2.30898	(NAN)		-INF	1.26630	0.681292	37.9441

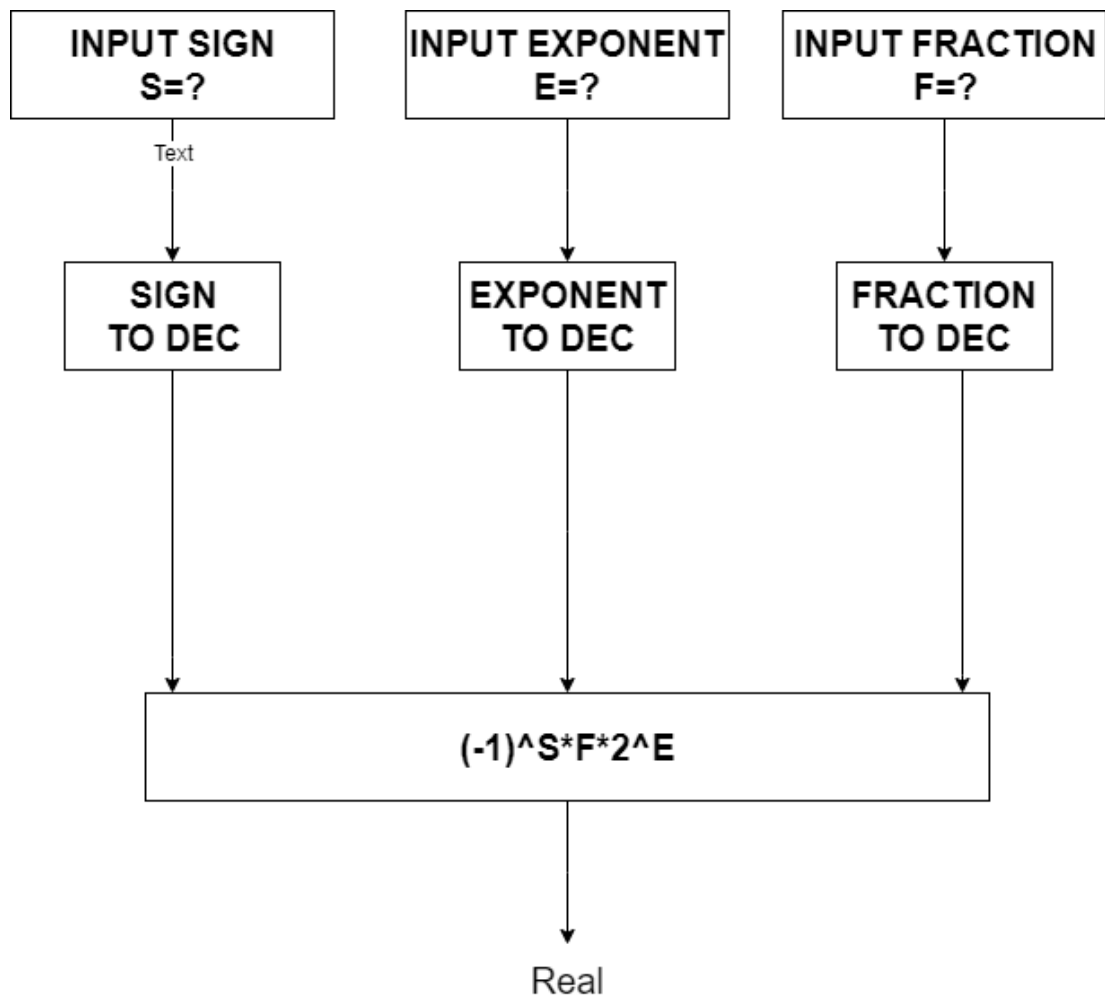
Hình 2. Kết quả mô phỏng lần 2

Chương 9: Chuyển Đổi Floating Point 32bits Trong Matlab

1. Chuyển từ số thực sang float 32bit



2. Chuyển từ số float 32bit sang số thực



3. Kết quả chạy trên matlab

```
Command Window

>> real2fp
Input Real: 0
Floating point 32bits: 0 | 00000000 | 000000000000000000000000
>> fp2real
Input floating point 32bits:
>> Bit sign (1bit): 0
>> Exponent (8bit): 00000000
>> Fraction (23bit): 000000000000000000000000
Real numbers: 0
>> real2fp
Input Real: 1235.314
Floating point 32bits: 0 | 10001001 | 00110100110101000001100
>> fp2real
Input floating point 32bits:
>> Bit sign (1bit): 0
>> Exponent (8bit): 10001001
>> Fraction (23bit): 00110100110101000001100
Real numbers: 1.235314e+03
>> real2fp
Input Real: -0.000231
Floating point 32bits: 1 | 01110010 | 11100100011100010010111
>> fp2real
Input floating point 32bits:
>> Bit sign (1bit): 1
>> Exponent (8bit): 01110010
>> Fraction (23bit): 11100100011100010010111
Real numbers: -2.310000e-04
fx >> |
```

KẾT LUẬN

- Thiết kế và mô phỏng thành công mô hình Floating Point Arithmetic Unit. Tính toán chính xác với bộ cộng trừ dùng giải thuật CLA, bộ nhân dùng giải thuật Double, chia dùng giải thuật Restoring Division.
- Chuyển thành công số floating point 32bits thành số real và ngược lại, mô phỏng được tất cả các số trên Matlab.
- Thiết kế và chạy thành công bộ Root Square of n dùng giải thuật Xấp xỉ.
- Thiết kế và chạy thành công bộ nhân dùng giải thuật Vedic
- Thiết kế và chạy thành công bộ cộng trừ dùng giải thuật CRA.
- Thiết kế và chạy thành công bộ lũy thừa dùng giải thuật Cổ Điển.
- Thiết kế và chạy thành công bộ Căn bậc 2 dùng giải thuật Xấp Xỉ.

TÀI LIỆU THAM KHẢO

- 1. DEVELOPING AN EFFICIENT IEEE 754 COMPLIANT FPU IN VERILOG (by RUBY DEV (108CS069)).**
- 2. DESIGN OF SINGLE PRECISION FLOAT ADDER (32-BIT NUMBERS) ACCORDING TO IEEE 754 STANDARD USING VHDL (by Dr. Roman Zálusky).**
- 3. A Hierarchical Verification of the IEEE-754 Table-Driven Floating-Point Exponential Function using HOL (by Amr Talaat Abdel-Hamid)**
- 4. Computer Arithmetic (by Reto Zimmermann).**