

**ĐẠI HỌC BÁCH KHOA THÀNH PHỐ HỒ CHÍ MINH**

**Khoa Điện- Điện tử**

**Bộ Môn Điện tử**



# **BÁO CÁO BÀI TẬP LỚN**

**Môn học: Kỹ thuật số nâng cao**

GVHD: TS. Trần Hoàng Linh

SVTH: Lê nguyên Gia Thịnh

MSSV: 1613347

**HỒ CHÍ MINH, THÁNG 11 NĂM 2020**

## **Mục lục:**

1	Tổng quan về kiểu số thực Floating-point: .....	4
1.1	Giới thiệu: .....	4
1.2	Định dạng và phân bố số Floating-point: .....	4
1.3	Tiêu chuẩn IEEE cho số Floating-point: .....	4
2	Chuyển đổi số Floating-point và các trường hợp ngoại lệ: .....	5
2.1	Chuyển đổi số Floating-point: .....	5
2.2	Các trường hợp ngoại lệ: .....	6
3	Các phép toán cơ bản với số Floating-point: .....	6
3.1	Phép cộng/trừ: .....	6
3.2	Phép nhân/chia: .....	6
4	Thực hiện trên phần cứng: .....	6
4.1	Bộ cộng/trừ: .....	<b>Error! Bookmark not defined.</b>
4.1.1	Phương pháp sử dụng bộ cộng Carry Lookahead Adder: .....	7
4.1.2	Phương pháp sử dụng bộ cộng Full Adder: .....	7
4.2	Bộ nhân chia nhân/chia: .....	7
4.3	Bộ chuyển đổi đầu vào đầu ra từ Floating Point sang Decimal: .....	8
5	Một số kết quả mô phỏng: .....	9
5.1	Bộ cộng/trừ: .....	9
5.2	Bộ nhân: .....	10
5.3	Bộ chia: .....	11
5.4	Bộ chuyển định dạng: .....	12
5.5	Bộ lấy căn bậc hai: .....	14
6	Đánh giá và kết luận	
7	Phụ lục: .....	13

Hình 1. Định dạng thông thường của số FP .....	4
Hình 2. Khoảng phụ và các giá trị đặc biệt của số FP.....	4
Hình 3. Tiêu chuẩn IEEE định dạng số FP .....	5
Hình 4. Phép cộng/trừ .....	<b>Error! Bookmark not defined.</b>
Hình 5. Phép nhân.....	<b>Error! Bookmark not defined.</b>
Hình 6. Phép chia .....	<b>Error! Bookmark not defined.</b>
Hình 7. Phép biến đổi dấu chấm động sang thập phân .....	<b>Error! Bookmark not defined.</b>
Hình 8. Carry Lookahead Adder .....	7
Hình 9. Full Adder .....	7
Hình 13. Waveform bộ cộng/trừ (phương pháp 1).....	9
Hình 14. Waveform bộ cộng/trừ (phương pháp 2).....	9
Hình 15. Schematic Carry Lookahead Adder .....	10
Hình 16. Schematic Full Adder.....	10
Hình 17. Waveform bộ nhân (phương pháp 1) .....	10
Hình 18. Waveform bộ nhân (phương pháp 2) .....	11
Hình 19. Schematic bộ nhân Baugh - Wooley .....	12
Hình 16. Schematic bộ nhân từng phần .....	10
Hình 17. Waveform bộ chia .....	10
Hình 18. Schematic bộ chia.....	11
Hình 19. Waveform bộ chuyển đổi thập phân.....	12
Hình 20. Schematic bộ chuyển đổi thập phân.....	10
Hình 17. Waveform bộ lấy căn bậc 2.....	10
Hình 18. Schematic bộ lấy căn bậc 2 .....	11

Bảng 1. Hệ thống số.....	4
Bảng 2. Một vài đặc điểm của tiêu chuẩn IEEE 754-2008 .....	5
Bảng 3. Các trường hợp chuyển đổi.....	8

## 1 Tổng quan về kiểu số thực Floating-point:

### 1.1 Giới thiệu:

Không có hệ thống số nào có thể biểu diễn tất cả số thực.

Một vài hệ thống có thể sử dụng cho tập con của số thực được liệt kê trong Bảng 1.

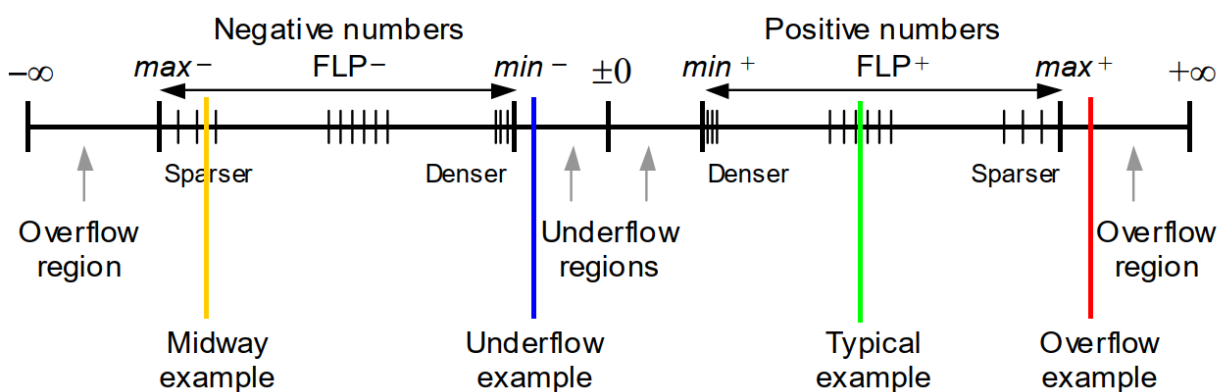
**Bảng 1. Hệ thống số**

Hệ thống số	Biểu diễn	Tính chất
Fixed-point	$\pm w.f$	Độ chính xác thấp và khoảng biểu diễn ít
Rational	$\pm p/q$	Khó tính toán
Floating-point	$\pm s \times b^e$	Phổ biến nhất
Logarithmic	$\pm \log_b x$	Bị giới hạn trường hợp so với Floating-point

### 1.2 Định dạng và phân bố số Floating-point:

Sign	Exponent	Fraction
$\pm$	Số nguyên có dấu, nhưng thường được dùng như số nguyên không dấu để biểu diễn giá trị bằng cách cộng một lượng bias.	Biểu diễn như một số fixed-point, thường được chuẩn hóa bằng cách dịch dấu chấm.

*Hình 1. Định dạng thông thường của số FP*

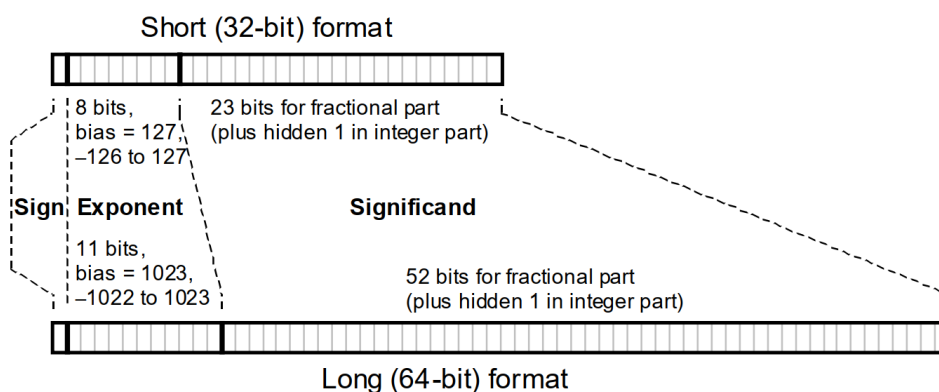


*Hình 2. Khoảng phủ và các giá trị đặc biệt của số FP*

### 1.3 Tiêu chuẩn IEEE cho số Floating-point:

Tiêu chuẩn IEEE đầu tiên cho việc tính toán số nhị phân floating-point được thông qua năm 1985.

Năm 2008, sau vài năm thảo luận, một bản sửa đổi tiêu chuẩn IEEE cho số floating-point được xuất bản gọi tắt là IEEE 754-2008.



*Hình 3. Tiêu chuẩn IEEE định dạng số FP*

Bảng 2. Một vài đặc điểm của tiêu chuẩn IEEE 754-2008

Stt	Đặc điểm	Single FP	Double FP
1	Word width (bits)	32	64
2	Significand bits	23+1 hidden	52+1 hidden
3	Significand range	$[1, 2 \cdot 2^{-23}]$	$[1, 2 \cdot 2^{-52}]$
4	Exponent bits	8	11
5	Exponent bias	127	1023
6	Zero	$e + \text{bias} = 0, f = 0$	$e + \text{bias} = 0, f = 0$
7	Denormal	$e + \text{bias} = 0, f \neq 0$	$e + \text{bias} = 0, f \neq 0$
8	Infinity	$e + \text{bias} = 255, f = 0$	$e + \text{bias} = 2047, f = 0$
9	Not-a-number	$e + \text{bias} = 255, f \neq 0$	$e + \text{bias} = 255, f \neq 0$
10	min	$2^{-126}$	$2^{-1022}$
11	max	$2^{128}$	$2^{1024}$

## 2 Chuyển đổi số Floating-point và các trường hợp ngoại lệ:

### 2.1 Chuyển đổi số Floating-point:

Chuyển đổi từ số fixed-point sang floating-point.

Chuyển đổi giữa các định dạng floating-point.

Chuyển đổi từ độ chính xác cao xuống độ chính xác thấp (làm tròn).

Tiêu chuẩn IEEE 754-2008 bao gồm năm chế độ làm tròn:

- Làm tròn gần nhất.
- Làm tròn về  $+\infty$ .
- Làm tròn về  $-\infty$ .

## 2.2 Các trường hợp ngoại lệ:

Một số chia cho 0.

Tràn trên (Overflow).

Tràn dưới (Underflow).

Các phép toán cho kết quả không tồn tại đều trả về kết quả NaN:

- Phép cộng:  $(+\infty) + (-\infty)$
- Phép nhân:  $0 \times \infty$
- Phép chia:  $0/0$  hoặc  $\infty/\infty$
- Phép lấy căn: toán hạng  $< 0$

## 3 Các phép toán cơ bản với số Floating-point:

### 3.1 Phép cộng/trừ:

Giả sử  $e_1 \geq e_2$ .

Thực hiện dịch để căn chỉnh phần exponent trước khi tính toán phần significand.

$$\begin{aligned}(\pm s_1 \times b^{e_1}) + (\pm s_2 \times b^{e_2}) &= (\pm s_1 \times b^{e_1}) + \left(\pm \frac{s_2}{b^{e_1-e_2}} \times b^{e_1}\right) \\ &= \left(\pm s_1 \pm \frac{s_2}{b^{e_1-e_2}}\right) \times b^{e_1} = \pm s \times b^e\end{aligned}$$

### 3.2 Phép nhân/chia:

$$(\pm s_1 \times b^{e_1}) \times (\pm s_2 \times b^{e_2}) = ((\pm s_1 \times s_2) \times b^{e_1+e_2})$$

$$\frac{\pm s_1 \times b^{e_1}}{\pm s_2 \times b^{e_2}} = \left(\pm \frac{s_1}{s_2}\right) \times b^{e_1-e_2}$$

## 4 Thực hiện trên phần cứng:

### Bộ cộng dấu chấm động

Các khâu tiến hành

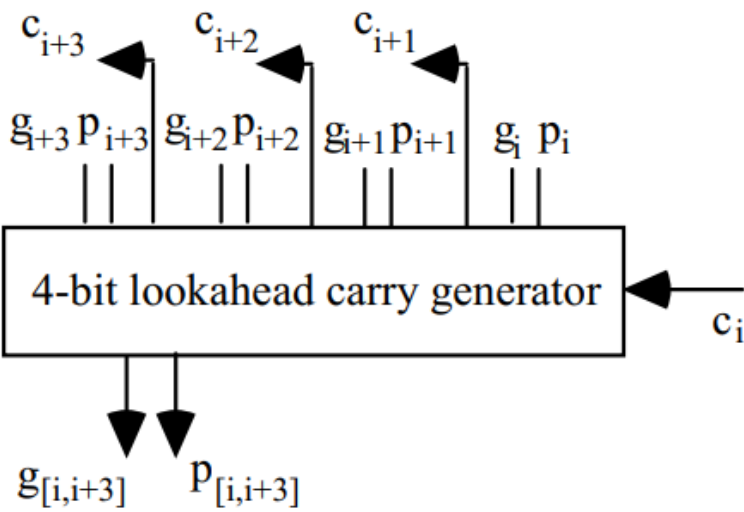
Gồm bốn khâu xử lý:

- So sánh phần Exponent của toán hạng vào.
- Dịch phần Exponent cho bằng nhau.
- Module cộng chính (thực hiện bằng hai phương pháp).
- Chuẩn hóa kết quả sau khi cộng.

Khi giá trị phần Exponent sau khi tính toán vượt khỏi miền biểu diễn thì xảy ra hiện tượng Overflow hoặc Underflow → Dịch trái hoặc dịch phải trong khâu xử lý chuẩn hóa.

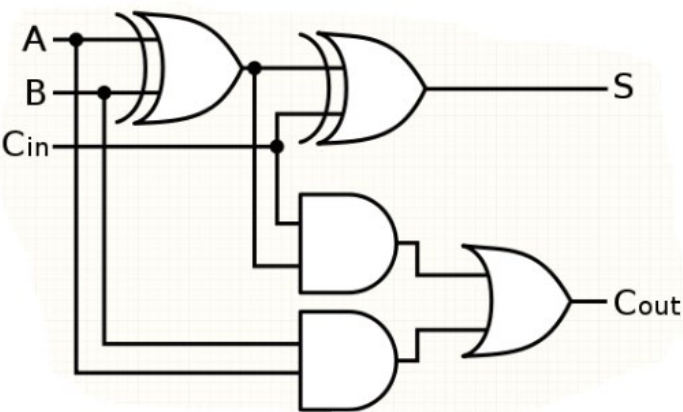
Các trường hợp ngoại lệ có thể được kiểm tra ở khâu mở gói định dạng đầu vào hoặc đóng gói định dạng kết quả ở đầu ra của bộ cộng.

**4.1.1 Phương pháp sử dụng bộ cộng Carry Lookahead Adder:**



Hình 4. Carry Lookahead Adder

**4.1.2 Phương pháp sử dụng bộ cộng Full Adder:**



Hình 5. Full Adder

**4.2 Bộ nhân/chia dấu chấm động:**

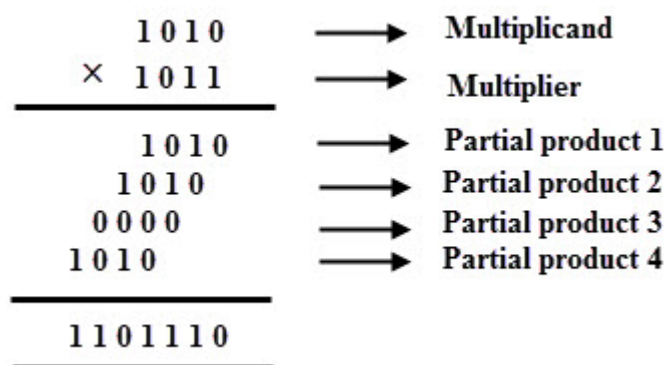
**4.2.2 Phương pháp nhân Baugh – Wooley**

$$\begin{array}{cccccccc}
 & & & & & \bar{a}_0 b_3 & a_0 b_2 & a_0 b_1 & a_0 b_0 \\
 & & & & \bar{a}_1 b_3 & a_1 b_2 & a_1 b_1 & a_1 b_0 & \\
 & & \bar{a}_2 b_3 & a_2 b_2 & a_2 b_1 & a_2 b_0 & & & \\
 & a_3 b_3 & a_3 \bar{b}_2 & a_3 \bar{b}_1 & a_3 \bar{b}_0 & & & & \\
 + & \bar{a}_3 & & & a_3 & & & & \\
 + & \bar{b}_3 & & & b_3 & & & & \\
 \hline
 & p_7 & p_6 & p_5 & p_4 & p_3 & p_2 & p_1 & p_0
 \end{array}$$



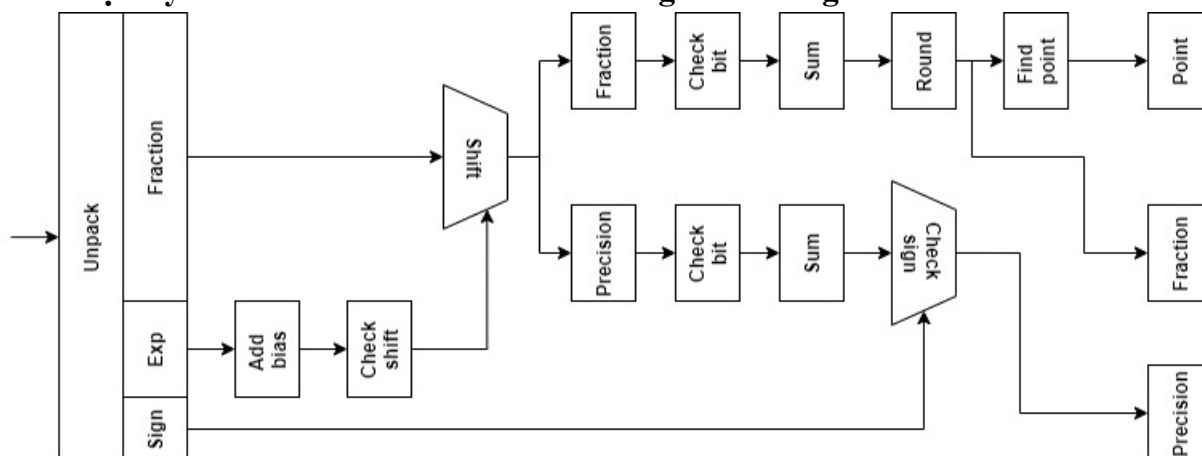
Hình 6. Bộ nhân Baugh – Wooley

#### 4.2.2 Phương pháp nhân từng bit :



Hình 7. Bộ nhân từng bit

#### 4.3 Bộ huyền đổi đầu vào đầu ra từ Floating Point sang Decimal:



Hình 8. Data path IP Core chuyển đổi

Bảng 3. Các trường hợp chuyển đổi

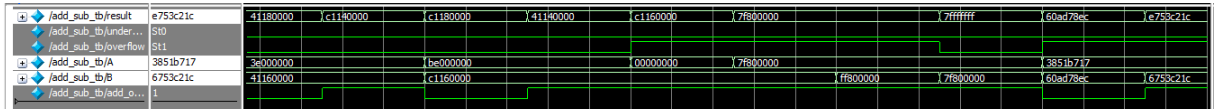
Stt	Shift Exponent	Sign Exponent	Phần nguyên	Phần thập phân
1	Giá trị shift > 23	(+)	Fraction	0
2	Giá trị shift > 23	(-)	0	0
3	Giá trị shift < 23	(+)	Shift fraction	Shift fraction
4	Giá trị shift < 23	(-)	0	Shift fraction

## 5 Một số kết quả mô phỏng:

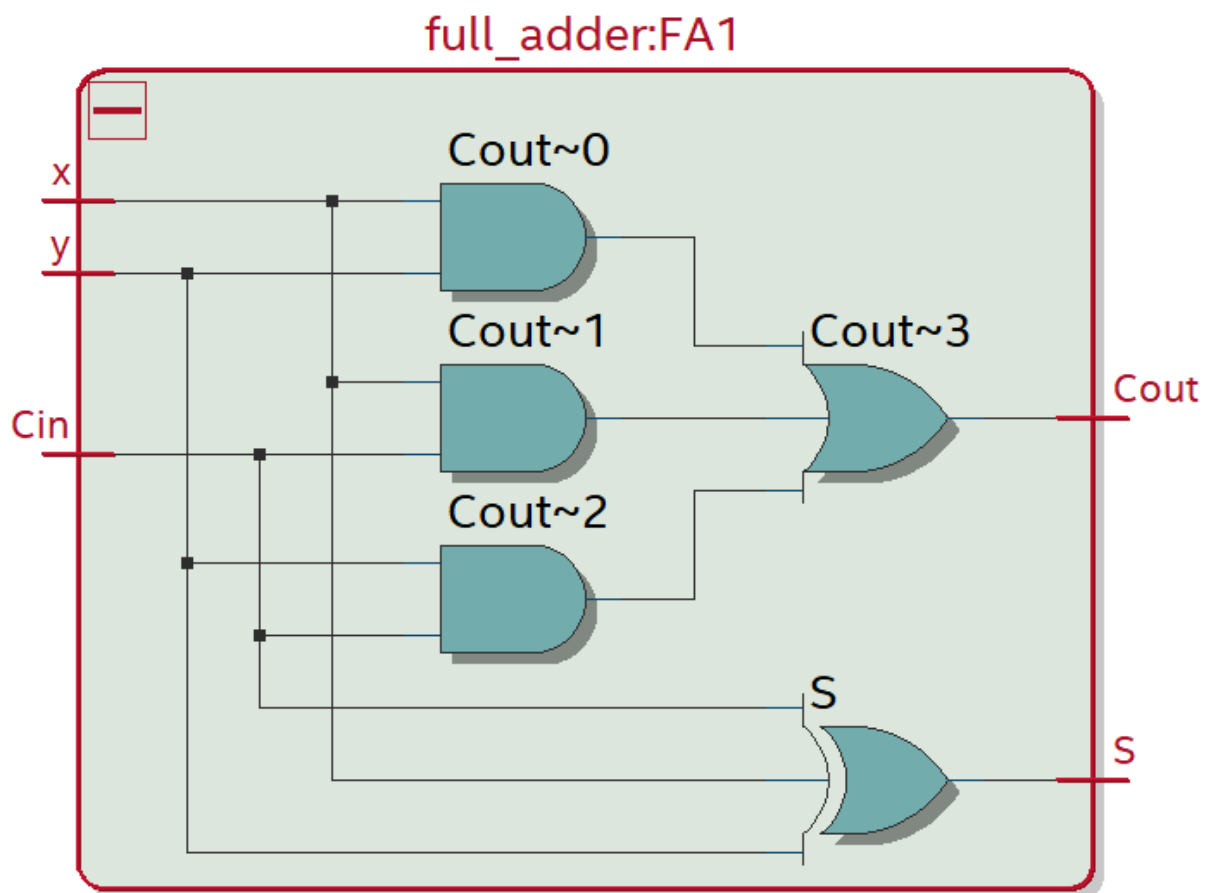
### 5.1 Bộ cộng/trừ:



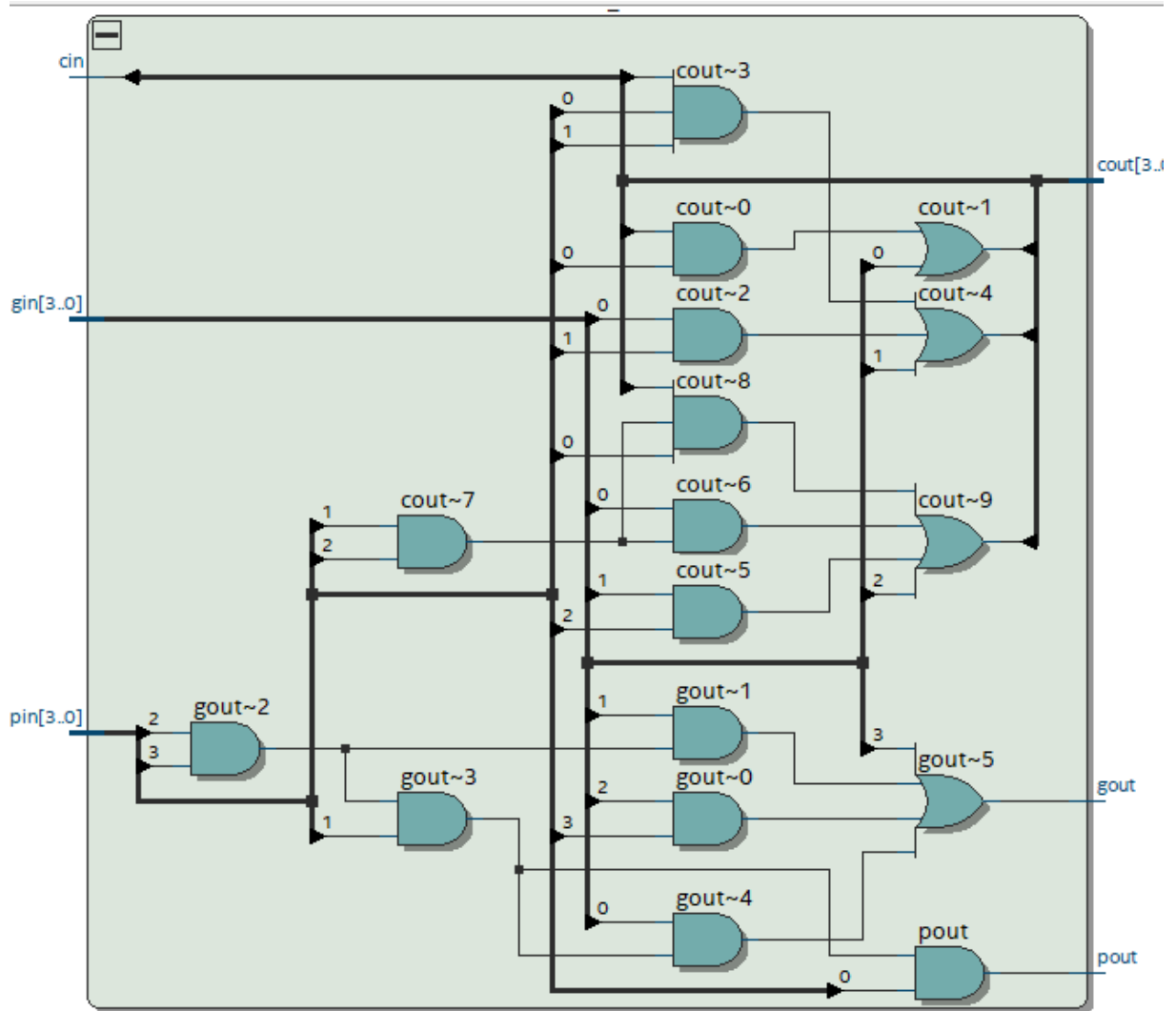
Hình 9. Waveform bộ cộng/trừ (phương pháp 1)



Hình 10. Waveform bộ cộng/trừ (phương pháp 2)

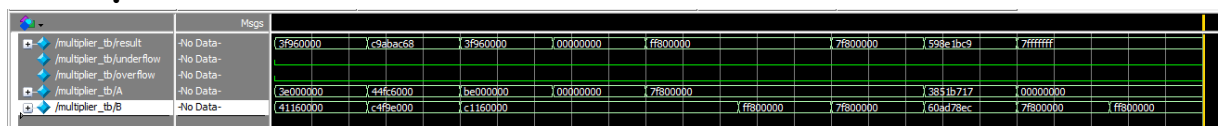


Hình 11. Schematic Full Adder

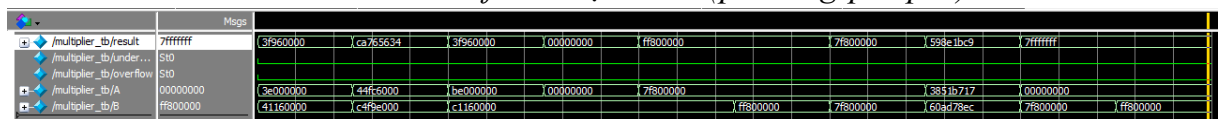


Hình 12. Schematic Carry Lookahead Adder

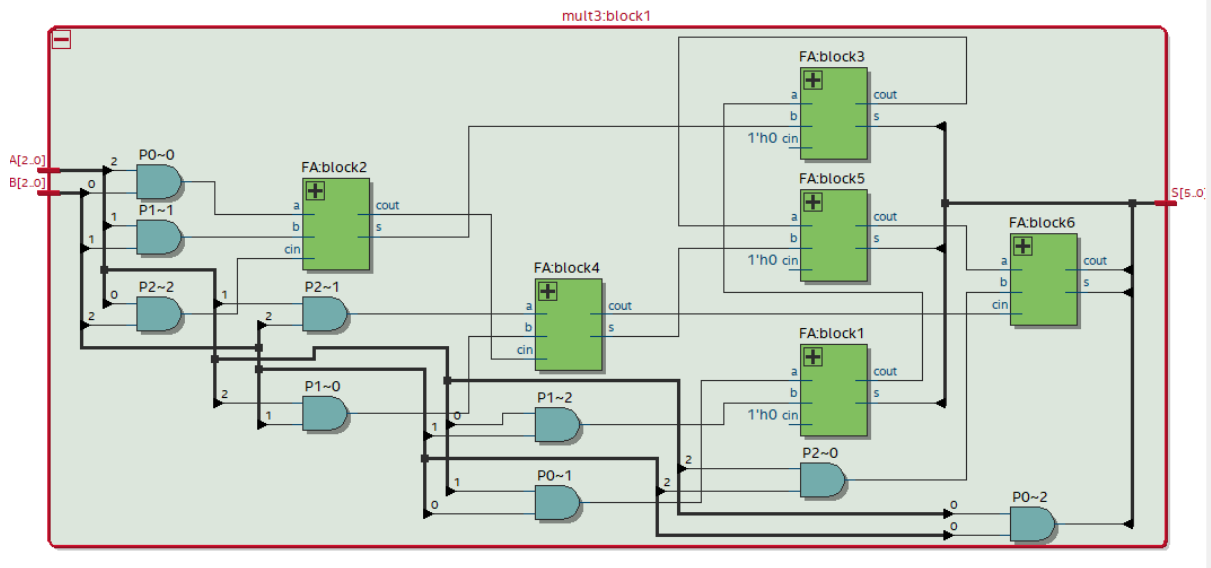
## 5.2 Bộ nhân:



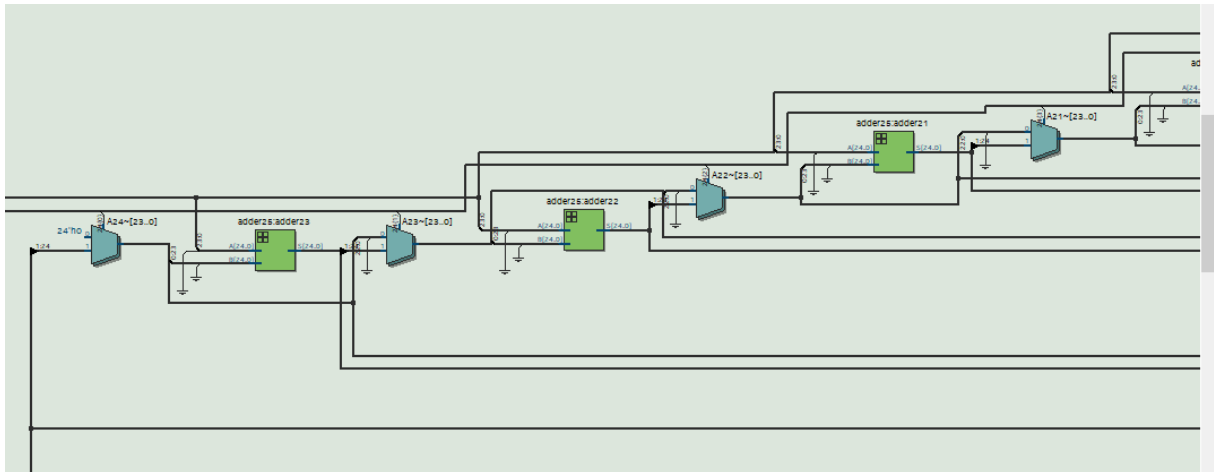
Hình 13. Waveform bộ nhân (phương pháp 1)



Hình 14. Waveform bộ nhân (phương pháp 2)

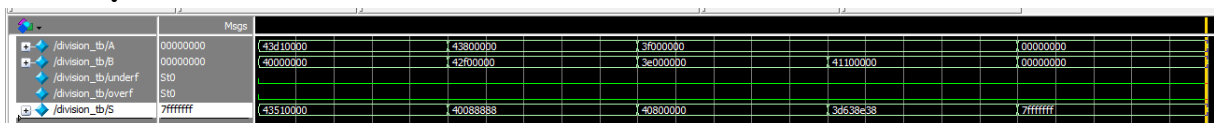


Hình 15. Schematic bộ nhân Baugh - Wooley (phương pháp 1)

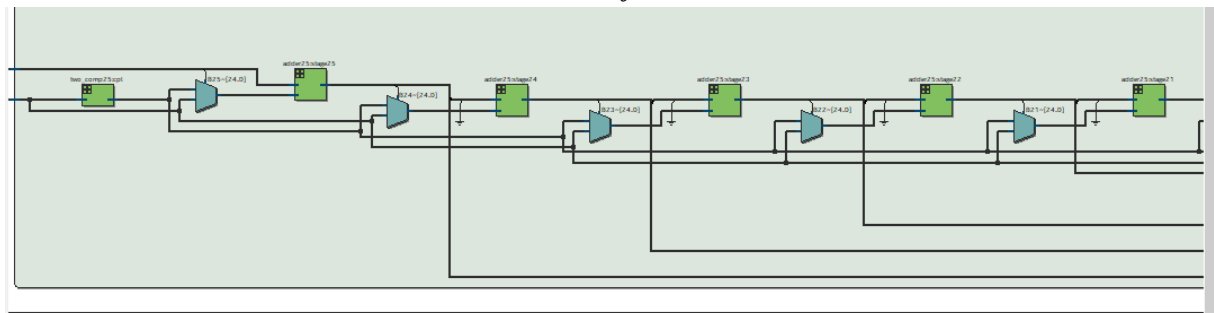


Hình 16. Schematic bộ nhân tự nhiên (phương pháp 2)

### 5.3 Bộ chia:

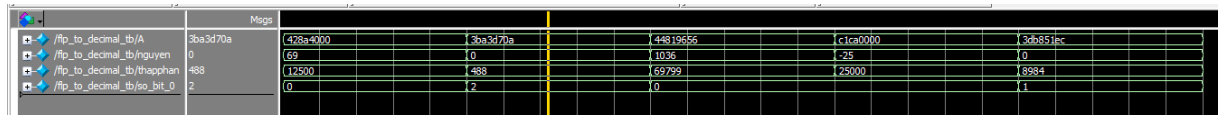


Hình 17. Waveform bộ chia

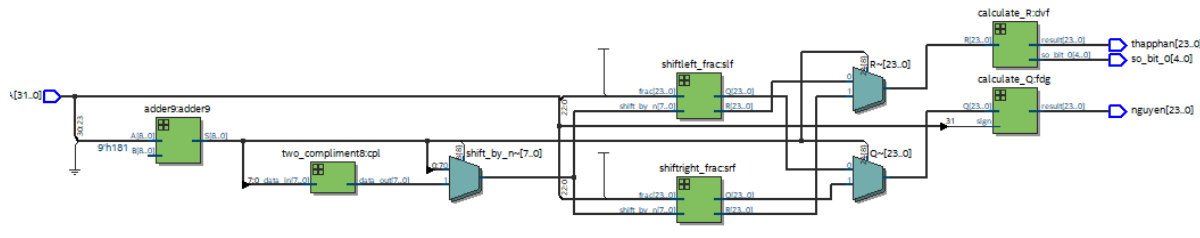


Hình 18. Schematic bộ chia

## 5.4 Bộ chuyển định dạng:

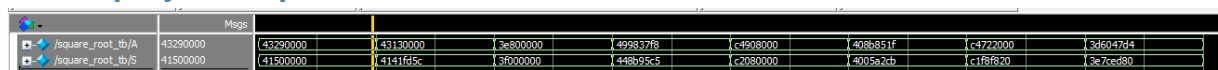


Hình 20. Waveform bộ chuyển định dạng



Hình 21. Schematic bộ chuyển định dạng

## 5.5 Bộ lấy căn bậc 2



Hình 22. Waveform bộ lấy căn bậc 2

## 6. Đánh giá và kết luận:

### Đánh giá:

Các module đã thực hiện đúng được các chức năng đã đề ra ban đầu.

Cả hai phương pháp được sử dụng cho bộ cộng và bộ nhân đều đạt được kết quả như mong muốn.

Do tính phức tạp của bộ chia nên hiện tại chỉ tiến hành thực hiện được bằng một phương pháp.

Đối với bộ chuyển đổi từ Floating Point sang Decimal. Hiện tại module chỉ có thể chuyển đổi được các số nằm trong phạm vi từ  $(10^{31}$  đến  $10^{-31})$ . Đối với các số có giá trị lớn hơn hiện tại module vẫn chưa thể đưa ra được đáp án chính xác.

Do giới hạn về thời gian và kiến thức nên hiện tại em chỉ có thể thực hiện được bộ lấy căn bậc hai của một số. Chưa thể tổng quát hóa lên thành lấy căn bậc n của số dấu chấm động.

### Kết luận:

Nhìn chung các module thiết kế đã đạt được hơn 90% các tiêu chí đã đề ra ban đầu. Các vấn đề còn lại có thể được giải quyết trong tương lai.

## **7. Phụ lục:**

Source code các module được đính kèm chung với bài báo cáo này!