

ĐẠI HỌC QUỐC GIA TP.HỒ CHÍ MINH

TRƯỜNG ĐẠI HỌC BÁCH KHOA

KHOA ĐIỆN – ĐIỆN TỬ

BỘ MÔN ĐIỆN TỬ

-----oOo-----



BÀI TẬP LỚN

KỸ THUẬT SỐ NÂNG CAO

GVHD: TS. TRẦN HOÀNG LINH

SVTH: NGUYỄN DƯƠNG PHÚC TÀI

MSSV: 1713012

TP. HỒ CHÍ MINH, THÁNG 07 NĂM 2020

Mục lục:

1	Tổng quan về kiểu số thực Floating-point:	5
1.1	Giới thiệu:	5
1.2	Định dạng và phân bố số Floating-point:	5
1.3	Tiêu chuẩn IEEE cho số Floating-point:	5
2	Chuyển đổi số Floating-point và các trường hợp ngoại lệ:	7
2.1	Chuyển đổi số Floating-point:	7
2.2	Các trường hợp ngoại lệ:	7
3	Các phép toán cơ bản với số Floating-point:	8
3.1	Phép cộng/trừ:	8
3.2	Phép nhân/chia:	8
4	Thực hiện trên phần cứng:	9
4.1	IP Core cộng/trừ:	9
4.1.1	Phương pháp sử dụng bộ cộng Carry Lookahead Adder:	10
4.1.2	Phương pháp sử dụng bộ cộng Full Adder:	10
4.2	IP Core nhân/chia:	11
4.3	IP Core chuyển đổi đầu vào đầu ra:	12
5	Một số kết quả mô phỏng:	13
5.1	IP Core cộng/trừ:	13
5.2	IP Core nhân:	14
5.3	IP Core chia:	14
5.4	IP Core chuyển định dạng:	14
6	Phụ lục Schematic:	14

Hình 1. Định dạng thông thường của số FP	5
Hình 2. Khoảng phụ và các giá trị đặc biệt của số FP.....	5
Hình 3. Tiêu chuẩn IEEE định dạng số FP.....	6
Hình 4. Phép cộng/trừ	8
Hình 5. Phép nhân	8
Hình 6. Phép chia	8
Hình 7. Data path IP Core cộng/trừ.....	9
Hình 8. Carry Lookahead Adder	10
Hình 9. Full Adder	10
Hình 10. Data path IP Core nhân	11
Hình 11. Data path IP Core chia	12
Hình 12. Data path IP Core chuyển đổi	12
Hình 13. Waveform IP Core cộng/trừ	13
Hình 14. Schematic Full Adder.....	13
Hình 15. Schematic Carry Lookahead Adder	13
Hình 16. Waveform IP Core nhân (1)	14
Hình 17. Waveform IP Core nhân (2)	14
Hình 18. Waveform IP Core chia.....	14
Hình 19. Waveform IP Core chuyển định dạng	14

Bảng 1. Hệ thống số.....	5
Bảng 2. Một vài đặc điểm của tiêu chuẩn IEEE 754-2008	7
Bảng 3. Các trường hợp chuyển đổi.....	12

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.

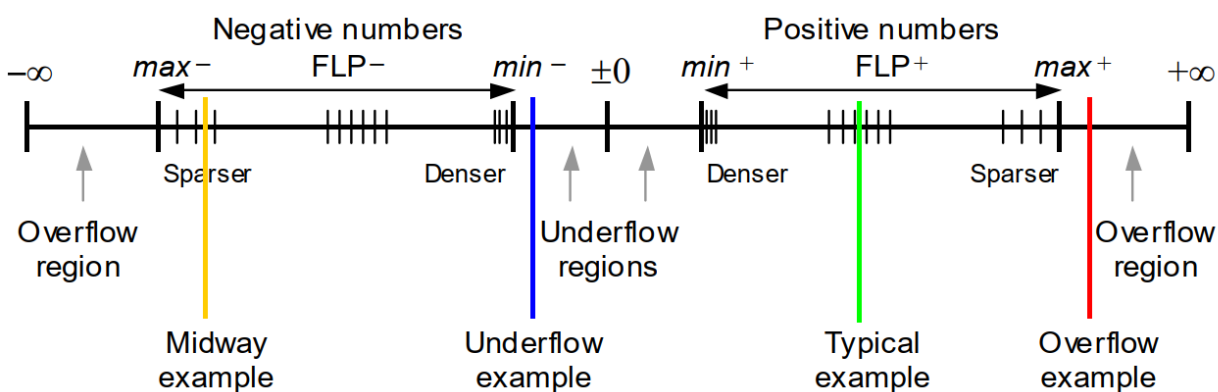
Bảng 1. Hệ thống số

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

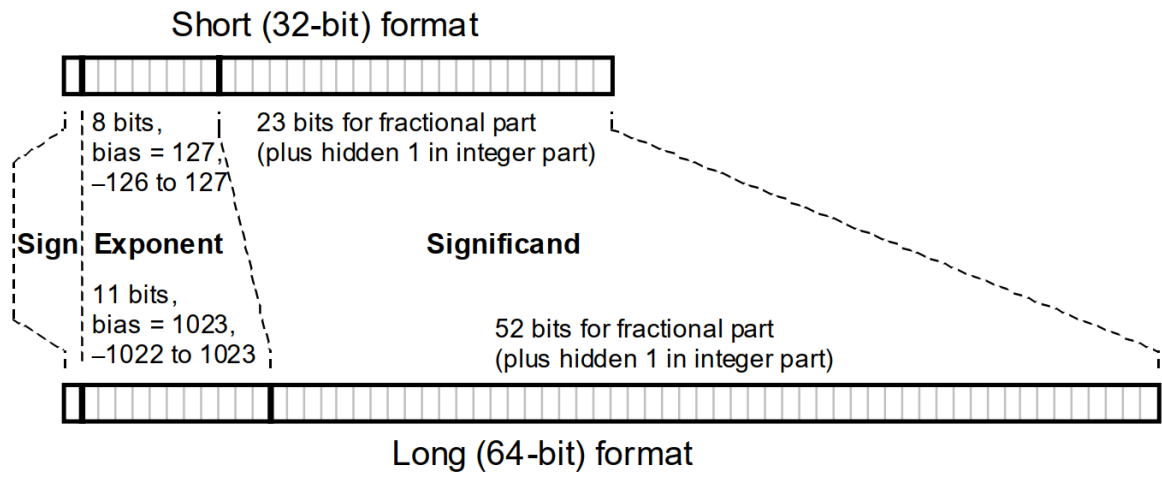


Hình 2. Khoảng phủ và các giá trị đặc biệt của số FP

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	$e+bias=0, f=0$	$e+bias=0, f=0$
7	Denormal	$e+bias=0, f \neq 0$	$e+bias=0, f \neq 0$
8	Infinity	$e+bias=255, f=0$	$e+bias=2047, f=0$
9	Not-a-number	$e+bias=255, f \neq 0$	$e+bias=2047, f \neq 0$
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 ∞/∞

- 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.

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

Hình 4. Phép cộng/trừ

3.2 Phép nhân/chia:

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

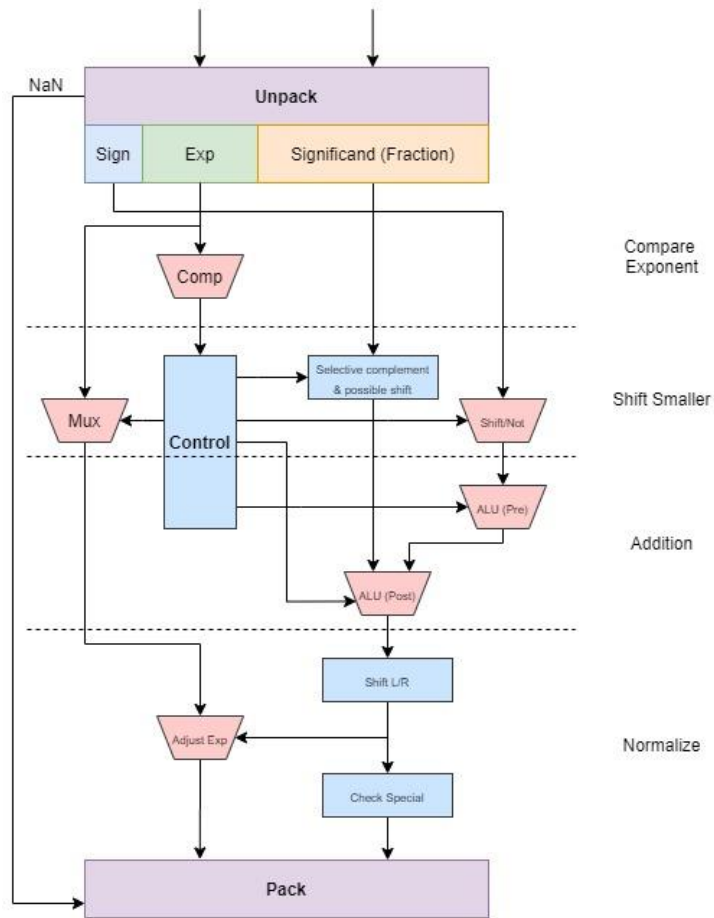
Hình 5. Phép nhân

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

Hình 6. Phép chia

4 Thực hiện trên phần cứng:

4.1 IP Core cộng/trừ:



Hình 7. Data path IP Core cộng/trừ

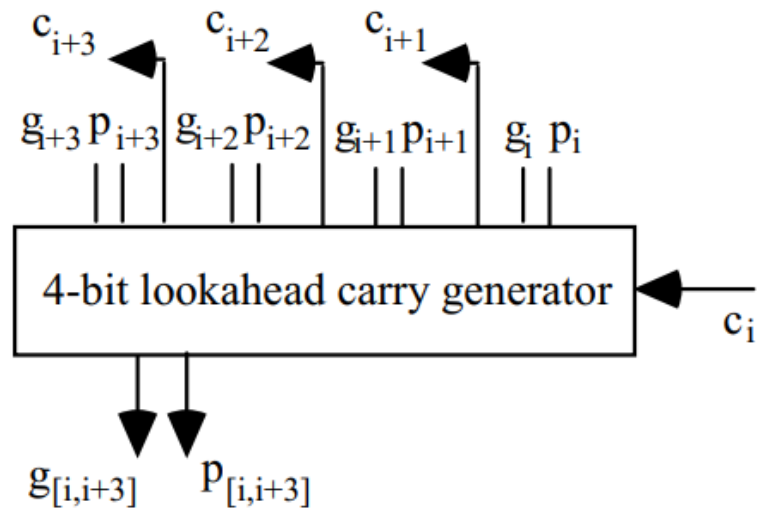
Luồng dữ liệu bao gồm bốn khâu xử lý: So sánh phần Exponent của toán hạng vào, Dịch phần Exponent cho bằng nhau, Module cộng chính có hai phương pháp thực hiện, Chuẩn hóa kết quả sau khi cộng.

Khi giá trị phần Exponent sau khi tính toán vượt khỏi miền biểu diễn thì xảy ra hiện tượng Overflow hoặc Underflow.

→ Hướng giải quyết: dịch trái hoặc dịch phải trong khâu xử lý chuẩn hóa.

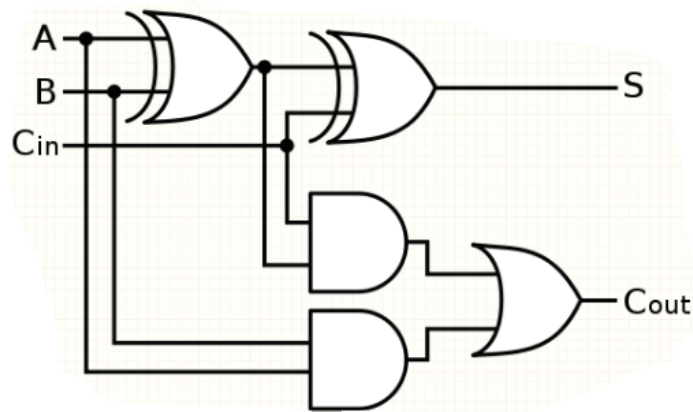
Các trường hợp ngoại lệ có thể được kiểm tra ở khâu mở gói định dạng đầu vào hoặc đóng gói định dạng kết quả ở đầu ra của IP Core.

4.1.1 Phương pháp sử dụng bộ cộng Carry Lookahead Adder:



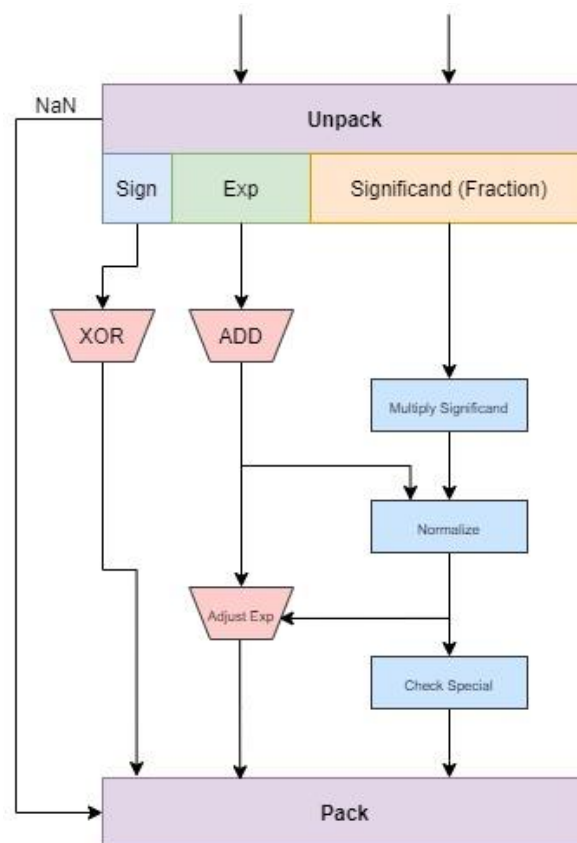
Hình 8. Carry Lookahead Adder

4.1.2 Phương pháp sử dụng bộ cộng Full Adder:

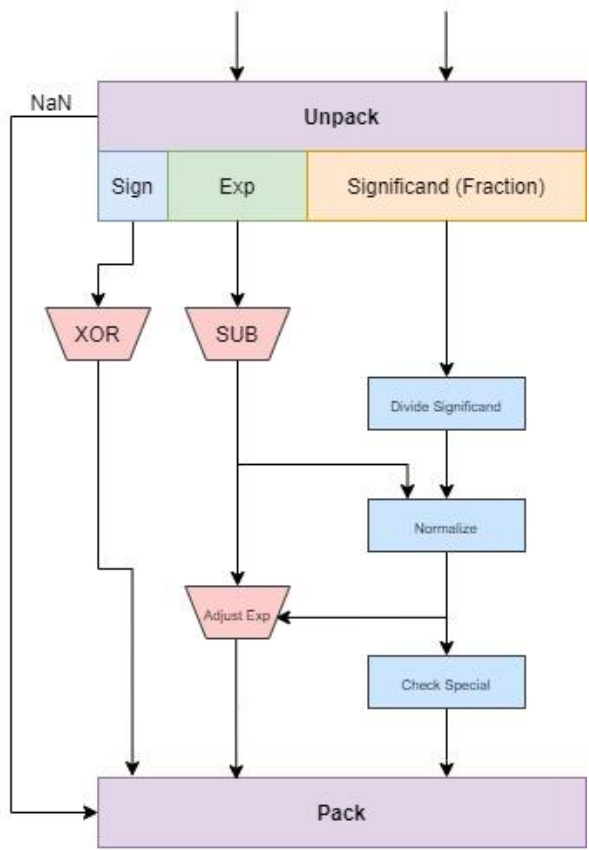


Hình 9. Full Adder

4.2 IP Core nhân/chia:

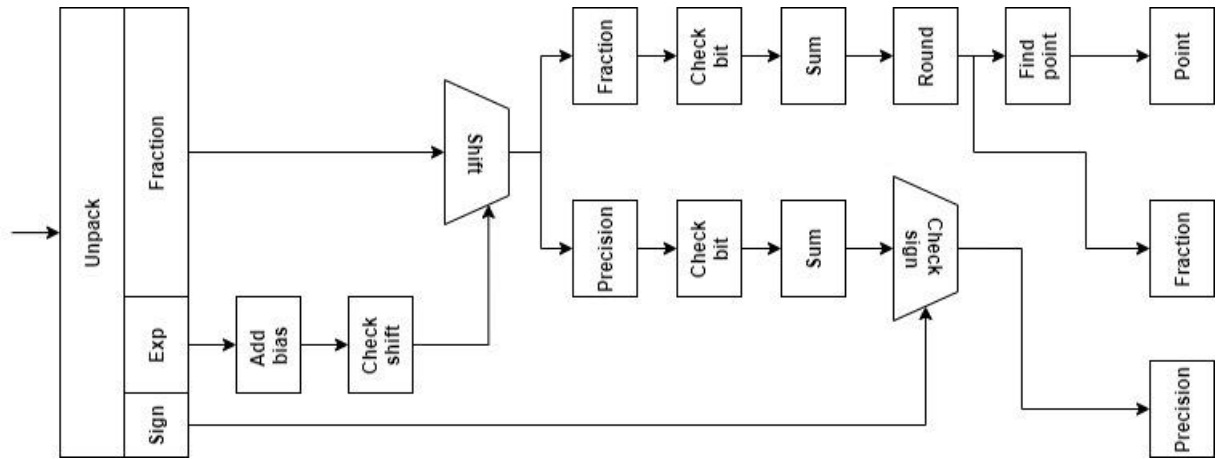


Hình 10. Data path IP Core nhân



Hình 11. Data path IP Core chia

4.3 IP Core chuyển đổi đầu vào đầu ra:



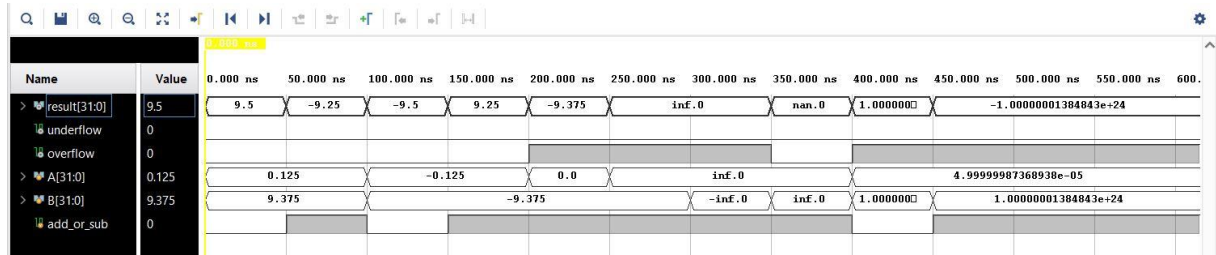
Hình 12. Data path IP Core chuyển đổi

Bảng 3. Các trường hợp chuyển đổi

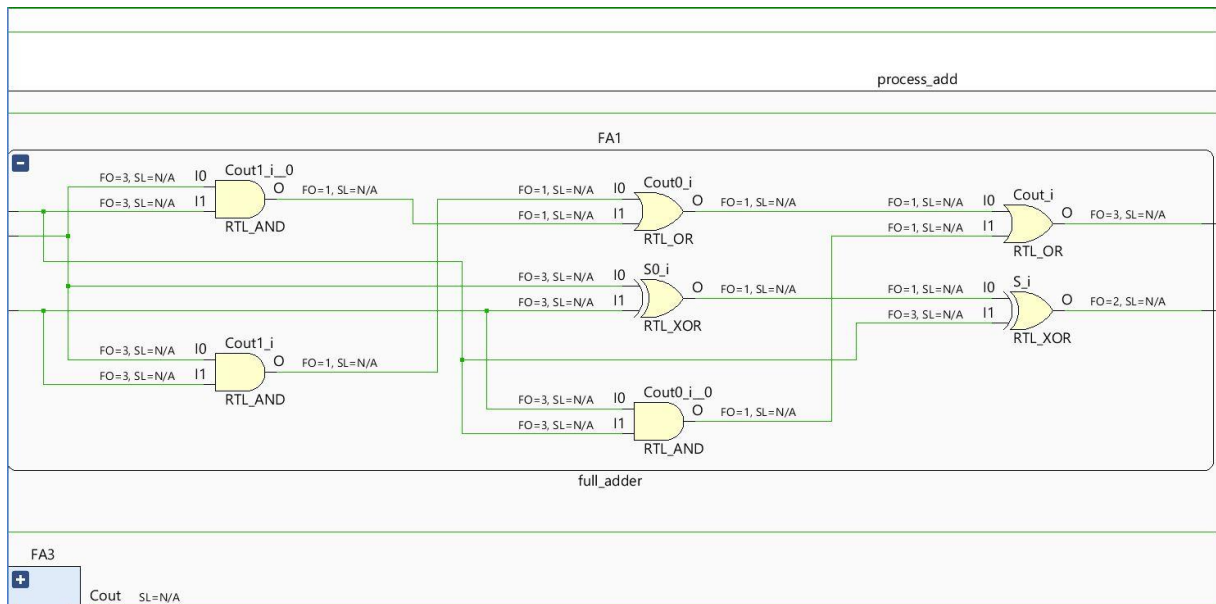
Stt	Shift Exponent	Sign Exponent	Phần nguyên	Phần thập phân
1	Giá trị shift > 23	(+)	Fraction	0
2	Giá trị shift > 23	(-)	0	0
3	Giá trị shift < 23	(+)	Shift fraction	Shift fraction
4	Giá trị shift < 23	(-)	0	Shift fraction

5 Một số kết quả mô phỏng:

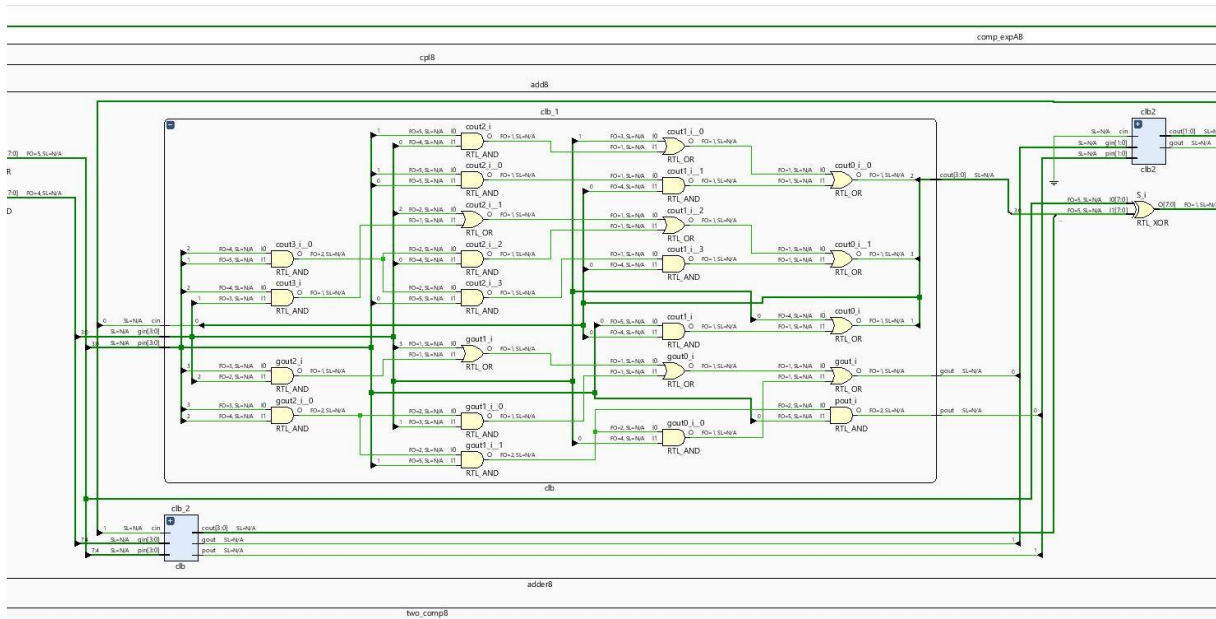
5.1 IP Core cộng/trừ:



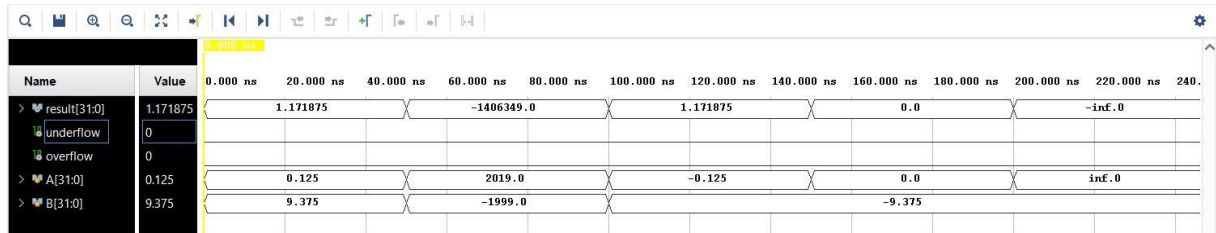
Hình 13. Waveform IP Core cộng/trừ



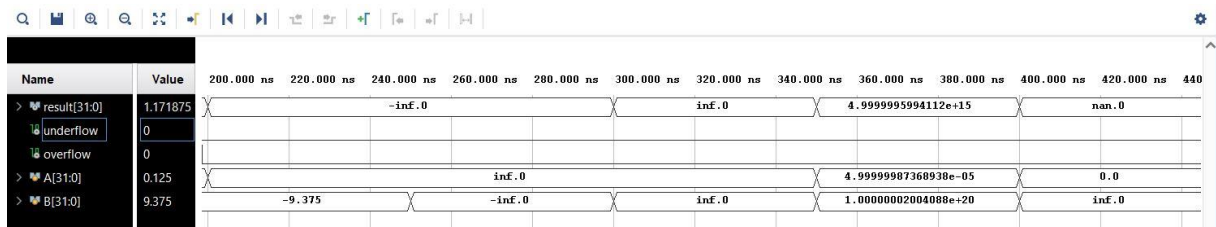
Hình 14. Schematic Full Adder



5.2 IP Core nhân:

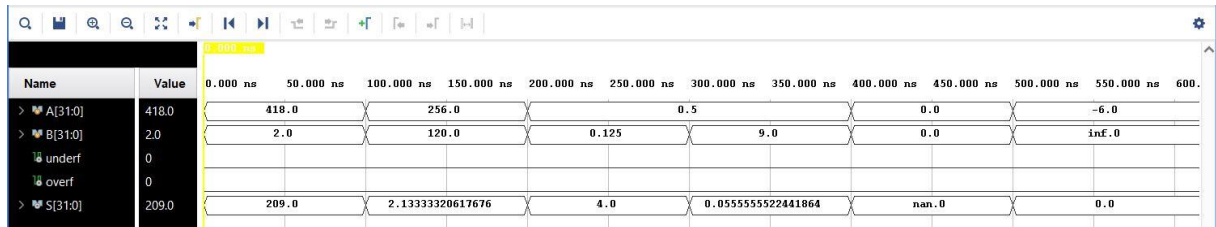


Hình 16. Waveform IP Core nhân (1)



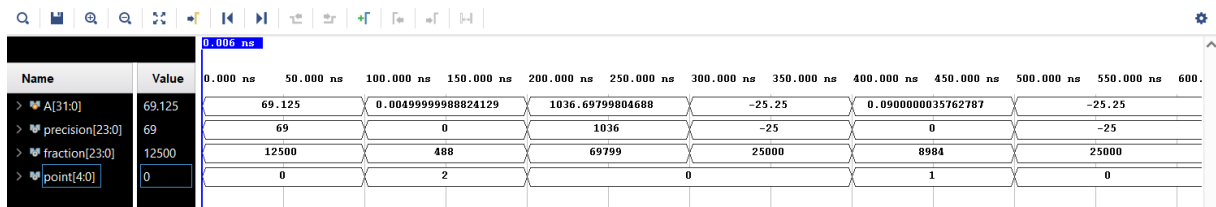
Hình 17. Waveform IP Core nhân (2)

5.3 IP Core chia:



Hình 18. Waveform IP Core chia

5.4 IP Core chuyển định dạng:



Hình 19. Waveform IP Core chuyển định dạng

6 Phụ lục Schematic: