

論文2002-39TE-2-1

고속 Floating Point Unit 설계

(A Design of High Speed Floating Point Unit)

吳 幸 洙 *

(Haeng-soo Oh)

요 약

부동소수점 시스템은 IEEE754 표준을 따른다. 부동소수점 방식의 가산기를 2의 보수를 사용하지 않고 1의 보수를 사용함으로써 간단히 나타낼 수 있다. 즉 이 시스템은 단지 반전을 함으로써 간단하고 빠른 연산을 수행할 수 있도록 하였다. 새롭게 설계된 가산기의 연산속도 향상을 위해 53bit의 캐리 선택 가산기를 사용하였다. 본 논문에서는 연산속도 향상을 위한 고성능의 효율적인 마이크로프로세서 시스템을 위한 부동소수점 가산 장치를 설계하였다.

Abstract

Floating point unit system follows IEEE 754 Standard. In this paper, we used 1's complement system instead of 2's complement to practice the arithmetic. By converting we enable this system to compute simply and fast. To improve the speed of newly design adder, we used a transformation Carry selector adder of 53 bits. In paper, a design of floating point unit high efficiency micro processor system about for high speed.

Keyword: High Speed Floating Point Unit.

I. 서 론

요즘의 고성능 마이크로 프로세서 시스템은 고속의 부동 소수점 연산을 요구한다. 현재 사용되고 있는 부동 소수점 unit내에는 숫자 계산을 위한 ALU, 승산기, 제산기가 포함되어 있으며 ALU와는 별도 unit으로 설계되어 있다. 삼각함수 계산 unit도 따로 되어있으며 승산기와 제산기를 사용하여 계산한다.

DSP 칩은 연산 처리 방식에 따라 고정 소수점(fixed point) DSP와 부동 소수점(floating point) DSP로 구분할 수 있다. 수치적 자료의 표현에 있어서, 부동 소수점

표현 방식은 고정 소수점 표현 방식보다 자릿수를 많이 줄일 수 있고, 같은 수의 비트를 표현할 경우 더 큰 정수 혹은 작은 소수를 나타낼 수 있다. 또한 기억장치의 공간을 적게 차지하고, 연산회로를 포함한 중앙처리 장치 내에서의 자료처리 및 이동이 용이하며, 한정된 수의 비트를 가지고 자료 표현의 정밀도를 높일 수 있다. 부동 소수점 방식의 가산기를 실시간 영상신호처리 혹은 음성신호처리 분야와 같은 고속의 변환 속도를 요하는 곳에 사용할 경우, 부동 소수점 방식 가산기의 수행속도는 매우 중요한 문제가 된다.^[1-2] 이런 문제점에 착안하여 본 논문에서는 기존방식에서의 2의 보수를 사용한 산술 연산 방식에 비하여 음수를 간단하고 빠르게 변환할 수 있도록 1의 보수 방식을 사용하여, 연산 속도 향상을 위한 53bit의 캐리 선택 가산기(carry selector adder)를 변형한 새로운 가산기를 설계하였다.

* 正會員, 東岡大學 電子情報科

(Department of Electronic and Information Engineering)

接受日字:2001年9月27日, 수정완료일:2002年6月10日

II. 이론적 배경

IEEE 부동소수점은 단정도와 배정도라는 두 개의 부동소수점 수로 기술된다.

단정도와 배정도의 수는 그림 (1)과 같이 다음의 3가지로 구성되어있다.

- (1) 1bit 부호 s
- (2) Bias된 지수(e) $e = E - bias$
- (3) 가수 $f = b_1 b_2 \dots b_{p-1}$

단정도의 경우 p(the number of significant bit)에 23bit를 저장하지만 hidden bit라는 1bit를 포함하여 24bit, 배정도의 경우는 53bit이다.

실제의 값(v)은 식 (1)과 같이 표시할 수 있다.

$$v = (-1)^s \cdot 2^E \cdot (1 \cdot f) \tag{1}$$

- (1) $e=2047, f \neq 0 \quad v = NaN$ (s를 고려 없이)
- (2) $e=2047, f=0 \quad v = (-1)^s \infty$
- (3) $0 < e < 2047 \quad v = (-1)^s \cdot 2^{e-1023} \cdot (1 \cdot f)$
- (4) $e=0, f \neq 0 \quad v = (-1)^s \cdot 2^{e-1023} \cdot (0 \cdot f)$
- (5) $e=0, f=0 \quad v = (-1)^s \cdot 0$ (zero)

단정도 부동 소수점의 경우 Bias의 값은 8bit의 exponent이므로 127, 배정도에서는 11bit이므로 1023이다. 본 논문에서 사용된 배정도의 경우 표시할 수 있는 가장 작은 수는 $v_{min} = \pm 2^{-1022}$ 이며 이보다 더 작은 수를 나타내기 위해서는 비정규(denormalized)수를 사용하여야 한다. 이외에도 확장 표준에 있어서 단정도에서는 32bit, 배정도에서는 64bit 이상을 사용하여 더 넓

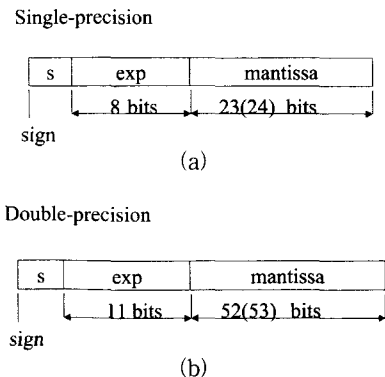


그림 1. IEEE 표준 단정도와 배정도 부동소수점
Fig. 1. Floating point of single and double.

은 숫자 범위를 표시할 수 있다.^[3-4]

III. 설 계

1. 구성도

보통의 부동소수점 DSP와 마찬가지로 기본적으로 IEEE 754를 따라 설계하였다. 부동 소수점 표현 방식을 이용한 가산기 회로의 입력 데이터 형태는 가수 부호 비트, 지수 비트, 그리고 가수 비트로 이루어져 있고, 초기 소수점(initial decimal point)은 가수의 첫 번째 비트 바로 앞으로 정한다. 부동소수점의 가산/감산기의 블록다이어그램은 그림 2와 같다.

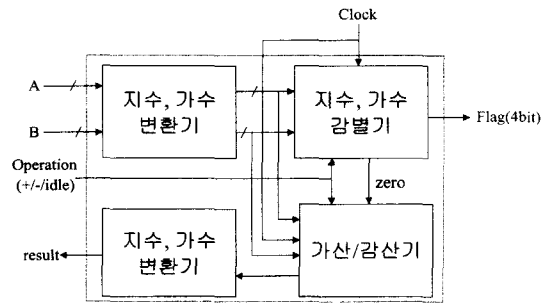


그림 2. 부동소수점 방식의 가산/감산기 블록도
Fig. 2. Adder/subtraction Block diagram of floating point.

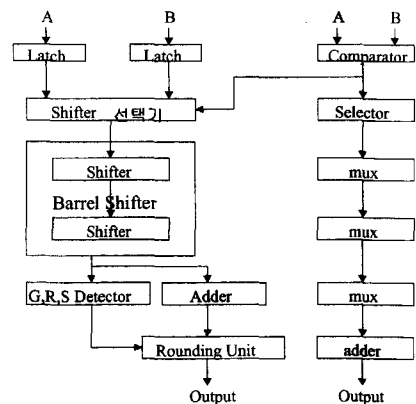


그림 3. 설계된 부동소수점 장치의 블록도
Fig. 3. Block diagram of designed floating point.

부동 소수점 방식 가산기의 동작원리는 우선 가산하고자 하는 두 수의 지수 크기를 비교하여 지수(exponent)의 차를 계산하고, 지수의 차이만큼, 크기가 작은 지수를 가진 수의 가수를 쉬프트(shift) 시켜서, 작은

수의 가수(mantissa)를 같게 만든 후에, 쉬프트하여 자리를 맞춘 후에 가수의 덧셈 또는 뺄셈을 수행하고 결과의 부호를 정한다. 마지막으로 반내림(rounding)과 정규화(normalization)를 수행하여 결과를 얻을 수 있다. 정규화 시킬 때 생기는 지수의 감소값을 감산하여 그 값을 지수 출력으로 한다. 위에서 설명한 알고리즘의 블록 다이어그램은 그림 3과 같다.

54bit 가산기는 덧셈과 뺄셈을 할 수 있도록 설계되었다. 대부분의 컴퓨터의 ALU는 2의 보수 산술을 사용하여 뺄셈을 하나 본 논문에서는 1의 보수산술을 사용하였다. 그 이유는 IEEE 부동소수점 표준에 의하면 모든 숫자를 부호-절대값으로 나타내도록 되어있어 2의 보수를 사용하면 MSB가 1인 경우 부호-절대값으로 변환할 때 1을 빼주어야 하는 자리올림 전과 감산기(carry propagation subtracter)가 필요하기 때문이다. 이러한 감산기 회로는 복잡하고 계산이 느리다. 1의 보수 산술에서는 반전하는 것만으로 음수를 간단하고 빠르게 후보-절대값으로 변환할 수 있다. 그리고 수행 속도의 향상을 위해 모든 블록은 파이프라인으로 구성하였다.

파이프라인을 도입하여 설계된 부동소수점의 예약표는 표 1과 같다.

표 1. 설계된 부동소수점 장치의 예약표

Table 1. Schedule sign of designed floating point.

Component	Clock					
	1	2	3	4	5	6
Exponent Unit	○	○	○	○	○	○
Selector		○	○	○	○	○
Shifter 선택기		○	○	○	○	○
Shifter			○	○	○	○
Shifter				○	○	○
Adder					○	○
G. R. S Detector						○
Rounding unit						○
Adder						○

2. Barrel Shifter

Barrel shifter는 두 단계의 shifter로 설계되어 있으며 첫 번째 단계에서는 0,1,2,3,4,5,6,7bit까지 shift하고 두 번째 단계에서는 0,8,16,24,32,40,48bit를 shift하도록

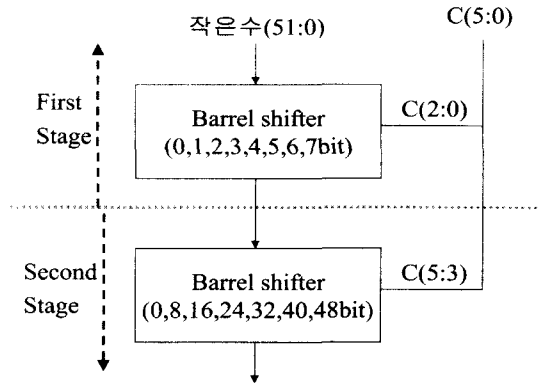


그림 4. 52bit barrel shifter의 블록도
Fig. 4. block-diagram on 52 bit barrel shifter.

설계되어 있다. 만약 첫 번째 단계와 두 번째 단계를 바꾸게 되면 guard, round, sticky bit detector 회로가 복잡해지므로 전자의 방식이 채택되었다.

그림 4는 52bit barrel shifter의 블록도이다.

3. Guard, Round, Sticky Bit Detector

Barrel shifter에서 shift한 후 53번째 bit를 Guard, 54번째로 bit를 Round, 그 나머지 bit에서 한 개라도 '1'이면 sticky bit를 '1'로 정한다. 이와 같이 구한 진리표값은 그림 5와 같다.

2개의 부동 소수점 숫자를 정확히 계산한 후 가수의 숫자가 표준보다 많아질 경우 표준과 같도록 하기 위해 반올림 또는 반내림하는 방법 중에 본 논문은 Round to nearest 즉, 가까운 쪽으로 반올림, 반내림하는 경우 barrel shift에서 guard, round, sticky bit을 검출하고, "100"상태 일 때는 shift를 하고 난 값의 최하위 비트가 "1"이면 "1"을 더하도록 설계하였다.

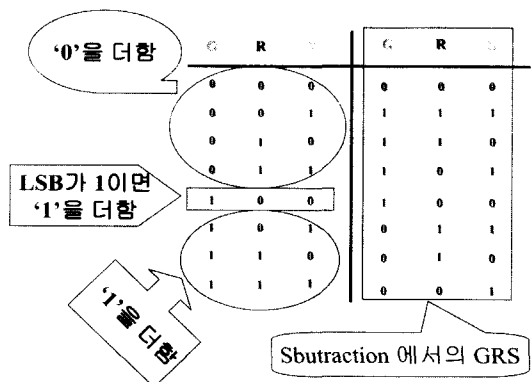


그림 5. GRS detector의 진리표
Fig. 5. Truth table of GRS detector.

4. Leading One Detector와 Rounding

가산을 한 후의 값은 IEEE 표준 값이 아닐 수가 있으므로 좌로 shift하여 MSB에 '1'이 오도록 하여야 한다. 즉, MSB 쪽에서 몇 번째에서 '1'이 위치해 있는가를 추출하여 Barrel shifter에서 추출한 만큼 좌로 shift할 수 있도록 설계하였다.

5. 제안된 53bit 가산기

본 논문에서 제안된 가산기는 8bit의 캐리 선택 가산기를 이용하여 8bit 가산기를 11개, 4bit 가산기를 2개 사용하였으며, 가산기 이전에 mux를 통해 덧셈일 경우는 그냥 통과하고 뺄셈일 경우는 1의 보수를 취하여 다른 값과 더해진다.

만일 캐리 값이 '1'상태일 때만 '1'을 더 더해주면 된다. 그림 6과 같은 캐리 선택 가산기 방식을 이용한 회

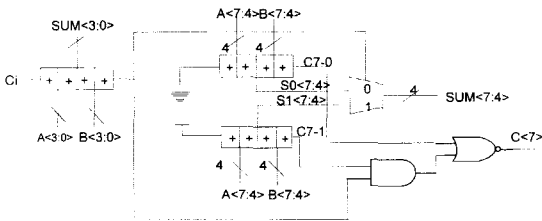


그림 6. 8bit 캐리 선택 가산기

Fig. 6. 8bit carry select adder.

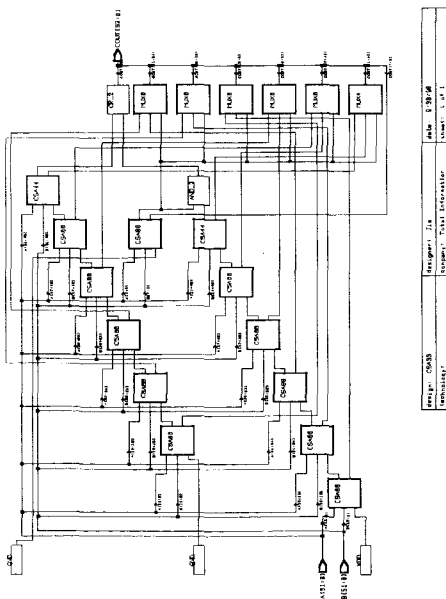


그림 7. 새롭게 디자인된 53bit 가산기

Fig. 7. The newly designed 53bit adder.

로를 구성하여 회로를 최소화하고, 그림 3의 가산기에서 가장 빠른 계산을 할 수 있는 가산기를 설계하였다.

그리고 클럭수를 줄여주기 위하여 leading one detector를 포함시켜 구성시켰다. 새롭게 디자인된 53bit 가산기는 그림 7과 같다. 전체 회로의 VHDL 코딩, 합성은 Synopsys 설계 툴로 수행하였고, 타이밍 시뮬레이션은 XCV50E로 수행하였다.

IV. 결과 및 검토

본 논문은 고성능 DSP 시스템에서 요구하는 고속의

표 2. 기존가산기와 제안된 가산기의 특성비교
Table 2. Comparative character of Existing adder and this paper adder.

가산기 종류(53bit)	게이트수	시간(ns)
일반적 ripple carry adder	135	110.67
ripple carry(chain)	200	61.04
carry select adder	173	55.21
carry select adder(변환)	197	49.17
새롭게 만든 adder	204	44.15

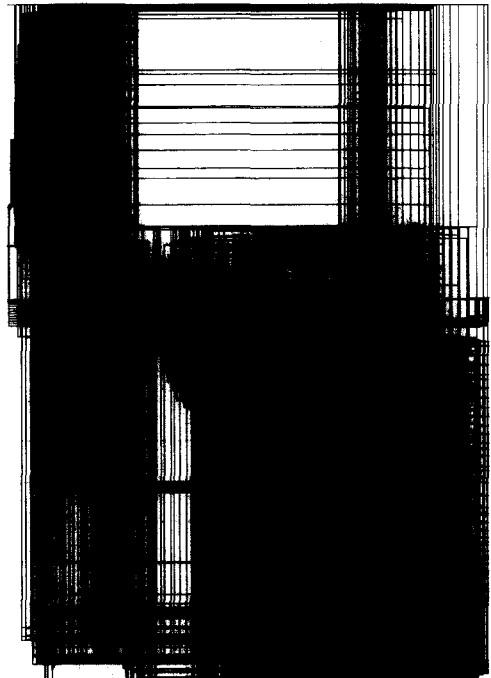


그림 8. 부동소수점 장치의 합성된 결과

Fig. 8. Synthesis result of floating point unit.

부동 소수점 unit를 설계하였다.

기존에는 2의 보수를 사용하여 산술을 행하였으나 반전하는 것만으로 음수를 간단하고 빠르게 변환할 수 있는 1의 보수 방식의 새롭게 디자인한 가산기를 사용함으로써 속도 향상을 얻을 수 있었다.

그리고 전체 회로의 VHDL 코딩, 합성은 Synopsys 설계 툴로 수행하였고, 타이밍 시뮬레이션은 XCV50E로 수행하였다.

동작주파수 23.3MHz로 작동하였을 때 기존의 가산기와 본 논문의 가산기의 비교 특성은 표 2와 같고, 본 논문의 동작주파수는 35MHz로 구동이 되었다.

본 논문에서 설계한 부동소수점 장치의 합성된 결과는 그림 8과 같다.

V. 결 론

본 논문에서는 DSP 구조 등의 고속의 변환 속도를 요하는 곳에 사용할 경우 문제가 되는 부동 소수점 방식 가산기의 수행속도를 빠르게 하는 고속 가산기를 설계하였다.

DSP 구조에서는 가산기를 필요로 하게 되는데 이 가산기의 연산 속도가 전체 시스템의 성능을 저하시킨다. 이를 위한 해결은 기존의 2의 보수 방식이 아닌 1의 방식을 취하여 부호-절대값 값을 빠르게 최소한의 회로로 계산하도록 하였으며, 캐리 선택 가산기(carry

select adder) 방식을 이용한 회로를 구성하여 회로를 최소화하고 가장 빠른 계산을 할 있는 가산기를 설계하였는데 기존 캐리 선택 가산기에 비해 면적비용이 1.2배정도 증가되고, 1.2배의 가산 속도를 향상시킬 수 있었다.

그리고 전체회로를 VHDL(VHSIC hardware description language)를 사용하여 설계하였기 때문에 전체적인 설계 시간을 단축할 수 있으며, 중요한 부분을 라이브러리로 등록하여 다른 회로의 설계에도 유용하게 사용할 수 있다.

참고문헌

- [1] Kim, Tae-Min, "Design of a low power floating point adder" 한국과학기술원 2000
- [2] Y. Shimazu, T.Kengaku, T.Fujiyama, E.Teraoka, T.Ohno, T.Tokuda, O.Tomisawa, and S. Tsujimichi, "A 50 MHz 24-b Floating-point DSP", ISSCC Dig. Tech. Papers, Feb 1989.
- [3] 이용식, "60MHz Clock 주파수의 IEEE 표준 Floating Point ALU", 전자공학회논문집, 제 28권, 1991
- [4] "IEEE Standard for Binary Floating-Point Arithmetic", Standard Committee of the IEEE Computer Society, 1985.

저 자 소 개

吳 幸 洙(正會員)

1980년 2월 : 조선대학교 전자공학과 졸업(공학사). 1982년 2월 조선대학교 대학원 전자공학과 졸업(공학석사). 1997년 2월 조선대학교 대학원 전기공학과 졸업(공학박사). 1984년 1월-1993년 2월 LG이노텍(주)헤드 설계과장. 1993년 3월-현재 동강대학 전자정보과 조교수. <주관심분야 : VLSI설계, 암호알고리즘>