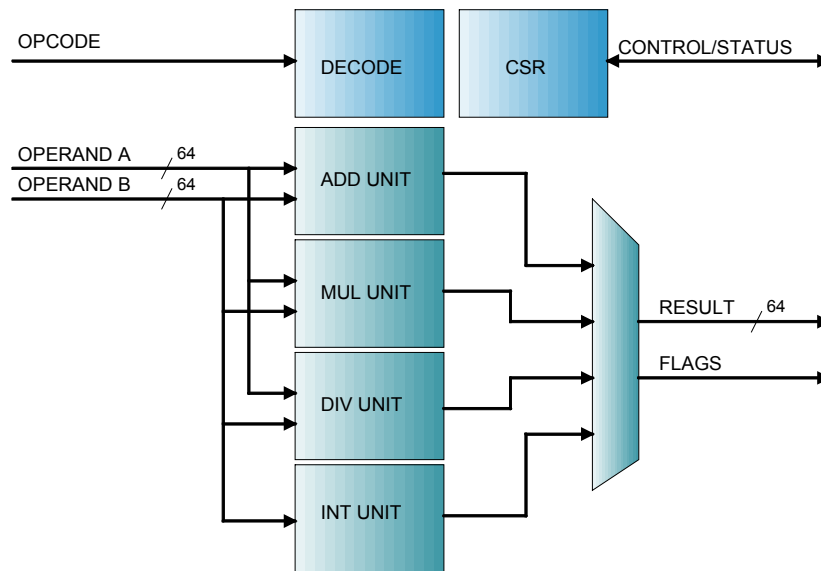# FPSDZ

IEEE-754 Floating Point Unit
Single/Double Precision with Flush-to-Zero Underflow

Overview

The FPSDZ is a co-processor unit providing floating-point computation compliant with the *ANSI/IEEE Std 754-1985, IEEE Standard for Binary Floating-Point Arithmetic* (IEEE-754 Standard). It is designed to provide high performance floating-point computation while minimizing die size and power.

The FPSDZ supports both single and double precision operands. The design's 3-stage execution pipeline combines high throughput with low latency, providing up to 250 MFLOPS on a 0.13u ASIC process. The host interface is clean and versatile, simplifying the interfacing to host processor pipelines.

Features

- IEEE-754 compliant (except underflow)
- Flush-to-Zero underflow implementation
- Single and double precision instructions
- 3-stage execution pipeline
- 1-cycle throughput for most instructions
- Instructions provided
    - Add/Subtract
    - Multiply
    - Divide
    - Remainder/Modulus
    - Square root
    - Floating Point Compare
    - Double ↔ Single Format Conversions
    - Floating Point ↔ Integer Conversions
    - Round to Integer
    - Absolute Value/ Negate
    - Move
- Flag outputs support conditional branching or conditional execution
- All IEEE rounding mode supported
- All IEEE exception flags supported
- Masked and unmasked exception control
- Control and status register

GB3

IEEE-754 Compliance

The FPSDZ is designed to provide a powerful floating-point capability while minimizing die size cost. To minimize unnecessary design size, some of the rarely used features of the IEEE specification are not implemented directly in the hardware design. The following IEEE-defined operations are not directly supported in FPSDZ hardware, but can be supported with software support:

- Gradual Underflow
- Denormal Numbers

In place of gradual underflow, the FPSDZ implements a flush-to-zero approach when underflow occurs. This feature allows the FPSDZ to maintain a one-cycle throughput in all operand cases, and minimizes design size.

Performance

Size: 85,000 NAND Gates
Timing: 200 MHz on 0.18u
        250 MHz on 0.13u

NOTE: The above performance data are estimates only, based on sample implementations using worst-case conditions. Achieved performance is highly dependent on the process technology, cell library, and synthesis tools used.

Instruction Timing

Instruction	Single Precision		Double Precision	
	Throughput	Latency	Throughput	Latency
Add, Subtract, Multiply, Compare, Round to Integer	1	3	1	3
Single/Double Format Conversions Integer Conversions	1	3	1	3
Absolute Value, Negate, Move	1	3	1	3
Divide	1 to 14	3 to 17	1 to 28	3 to 31
Square Root	1 to 13	3 to 16	1 to 27	3 to 30
Remainder, Modulus	1 to (e/2+2)	3 to (e/2+5)	1 to (e/2+2)	3 to (e/2+5)

Notes:
1) Divide, Square Root, Remainder, and Modulus are implemented with an "early out" algorithm, where the iterative calculations are stopped if the current remainder becomes zero.

2) e = operand exponent difference

GB3 Digital Systems
Phone: +1 (714) 731-5094
Fax: +1 (714) 731-7502
Email: sales@gb-3.com
Web: www.gb-3.com

GB3