CSX64

Processor Specification

Contents

[Purpose 6](#_Toc508227278)

[Processor Details 7](#_Toc508227279)

[Registers 7](#_Toc508227280)

[The Stack 7](#_Toc508227281)

[Error Codes 7](#_Toc508227282)

[Main Memory 8](#_Toc508227283)

[Numeric Types 9](#_Toc508227284)

[Integer 9](#_Toc508227285)

[Unsigned 9](#_Toc508227286)

[Signed 9](#_Toc508227287)

[Floating-Point 10](#_Toc508227288)

[64-bit – Double Precision 11](#_Toc508227289)

[32-bit – Single Precision 13](#_Toc508227290)

[Machine Code 14](#_Toc508227291)

[Register Format 14](#_Toc508227292)

[Size Format 14](#_Toc508227293)

[Multiplier Format 14](#_Toc508227294)

[Memory Address Format 15](#_Toc508227295)

[Conditions 15](#_Toc508227296)

[Translating Machine Code 16](#_Toc508227297)

[Assembly Language 18](#_Toc508227298)

[CSX64 Assembly Language 18](#_Toc508227299)

[imm – Immediate 18](#_Toc508227300)

[Instant Imm 20](#_Toc508227301)

[r – Register 20](#_Toc508227302)

[m – Memory 21](#_Toc508227303)

[Symbols 22](#_Toc508227304)

[Predefined Symbols 22](#_Toc508227305)

[Labels 23](#_Toc508227306)

[Local Label 23](#_Toc508227307)

[Local Symbol 23](#_Toc508227308)

[Operation Syntax 23](#_Toc508227309)

[Directives 24](#_Toc508227310)

[Global 24](#_Toc508227311)

[Def 24](#_Toc508227312)

[Emit 24](#_Toc508227313)

[Assembly Process 24](#_Toc508227314)

[Operations 25](#_Toc508227315)

[NOP – No Operation 27](#_Toc508227316)

[STOP – Stop 28](#_Toc508227317)

[SYSCALL – System Call 29](#_Toc508227318)

[MOVcc – Conditional Move 30](#_Toc508227319)

[SWAP – Swap 31](#_Toc508227320)

[UX – Unsigned Extend 32](#_Toc508227321)

[SX – Sign Extend 33](#_Toc508227322)

[UMUL – Unsigned Multiply 34](#_Toc508227323)

[SMUL – Signed Multiply 35](#_Toc508227324)

[UDIV – Unsigned Divide 36](#_Toc508227325)

[SDIV – Signed Divide 37](#_Toc508227326)

[ADD – Add 38](#_Toc508227327)

[SUB – Subtract 39](#_Toc508227328)

[BMUL – Binary Multiply 40](#_Toc508227329)

[BUDIV – Binary Unsigned Divide 41](#_Toc508227330)

[BUMOD – Binary Unsigned Modulus 42](#_Toc508227331)

[BSDIV – Binary Signed Divide 43](#_Toc508227332)

[BSMOD – Binary Signed Modulus 44](#_Toc508227333)

[SL – Shift Left 45](#_Toc508227334)

[SR – Shift Right 46](#_Toc508227335)

[SAL – Shift Arithmetic Left 47](#_Toc508227336)

[SAR – Shift Arithmetic Right 48](#_Toc508227337)

[RL – Rotate Left 49](#_Toc508227338)

[RR – Rotate Right 50](#_Toc508227339)

[AND – Bitwise And 51](#_Toc508227340)

[OR – Bitwise Or 52](#_Toc508227341)

[XOR – Bitwise Exclusive Or 53](#_Toc508227342)

[CMP – Compare 54](#_Toc508227343)

[TEST – Logical Test 55](#_Toc508227344)

[INC – Increment 56](#_Toc508227345)

[DEC – Decrement 57](#_Toc508227346)

[NEG – Negate 58](#_Toc508227347)

[NOT – Bitwise Not 59](#_Toc508227348)

[ABS – Absolute Value 60](#_Toc508227349)

[CMPZ – Compare Zero 61](#_Toc508227350)

[LA – Load Address 62](#_Toc508227351)

[Jcc – Conditional Jump 63](#_Toc508227352)

[FADD – Floating-Point Add 64](#_Toc508227353)

[FSUB – Floating-Point Subtract 65](#_Toc508227354)

[FMUL – Floating-Point Multiply 66](#_Toc508227355)

[FDIV – Floating-Point Divide 67](#_Toc508227356)

[FMOD – Floating-Point Modulus 68](#_Toc508227357)

[FPOW – Floating-Point Power 69](#_Toc508227358)

[FSQRT – Floating-Point Square Root 70](#_Toc508227359)

[FEXP – Floating-Point Exponentiate 71](#_Toc508227360)

[FLN – Floating-Point Natural Logarithm 72](#_Toc508227361)

[FNEG – Floating-Point Negate 73](#_Toc508227362)

[FABS – Floating-Point Absolute Value 74](#_Toc508227363)

[FCMPZ – Floating-Point Compare Zero 75](#_Toc508227364)

[FSIN – Floating-Point Sine 76](#_Toc508227365)

[FCOS – Floating-Point Cosine 77](#_Toc508227366)

[FTAN – Floating-Point Tangent 78](#_Toc508227367)

[FSINH – Floating-Point Hyperbolic Sine 79](#_Toc508227368)

[FCOSH – Floating-Point Hyperbolic Cosine 80](#_Toc508227369)

[FTANH – Floating-Point Hyperbolic Tangent 81](#_Toc508227370)

[FASIN – Floating-Point Arcsine 82](#_Toc508227371)

[FACOS – Floating-Point Arccosine 83](#_Toc508227372)

[FATAN – Floating-Point Arctangent 84](#_Toc508227373)

[FATAN2 – Floating-Point Binary Arctangent 85](#_Toc508227374)

[FLOOR – Floor 86](#_Toc508227375)

[CEIL – Ceiling 87](#_Toc508227376)

[ROUND – Round 88](#_Toc508227377)

[TRUNC – Truncate 89](#_Toc508227378)

[FCMP – Floating-Point Compare 90](#_Toc508227379)

[FTOI – Floating-Point to Integer 91](#_Toc508227380)

[ITOF – Integer to Floating-Point 92](#_Toc508227381)

[PUSH – Push 93](#_Toc508227382)

[POP – Pop 94](#_Toc508227383)

[CALL – Call 95](#_Toc508227384)

[RET – Return 96](#_Toc508227385)

[BSWAP – Byte Swap 97](#_Toc508227386)

[BEXTR – Bitfield Extract 98](#_Toc508227387)

[BLSI – Binary Lowest-Set Isolate 99](#_Toc508227388)

[BLSMSK – Binary Lowest-Set Mask 100](#_Toc508227389)

[BLSR – Binary Lowest-Set Reset 101](#_Toc508227390)

[ANDN – Bitwise And Not 102](#_Toc508227391)

[GETF – Get Flags 103](#_Toc508227392)

[SETF – Set Flags 104](#_Toc508227393)

[LOOP – Loop 105](#_Toc508227394)

[FX – Floating-Point Extend 106](#_Toc508227395)

[Virtual Operating System 107](#_Toc508227396)

[Default CSX64 Specification 107](#_Toc508227397)

[Appendix 108](#_Toc508227398)

[Converting Binary Integers 108](#_Toc508227399)

[Proof of 2’s Complement 109](#_Toc508227400)

[The Heap 110](#_Toc508227401)

[Encoding IEEE 754 64-bit Floating-Point Values 111](#_Toc508227402)

# Purpose

The CSX64 virtual processor was designed primarily to be an educational tool for learning assembly language and low-level processor details. The processor acts as a 64-bit machine code interpreter with its own instruction set that includes hardware support for integral and floating-point computations. The instruction set was designed around the x86\_64 processor architecture to be as realistic as possible, while cutting out all of the hardware-dependent aspects of real machine code that can confuse people when trying to learn it.

The benefit of using the CSX64 virtual processor is that its instruction set and system calls are not platform dependent, and both the machine code programs and the interpreter itself are portable to any machine that supports the .NET framework. Additionally, CSX64 has a high degree of interoperability, being native to C# and thus compatible with all .NET languages, as well as being portable to C and C++ with minimal effort. Due to this, it’s incredibly easy to create extensions for the CSX64 processor to handle file IO, email messaging, video rendering, or anything else you can think of.

# Processor Details

## Registers

Registers are small storage cells built directly into a processor that are vastly faster than [main memory](#_Main_Memory) (RAM), but are also vastly more expensive per byte. Because of this price factor, there is not typically much room in a processor for storing data. However, due to their speed, they are the place in which data is manipulated by the processor. In most modern processors data is loaded from main memory (RAM) into registers, processed, and then moved back into main memory.

The CSX64 processor contains 16 64-bit general purpose registers (which will be referred to as R0-R15 in this manual) for processing both [integral](#_Integer) and [floating-point](#_Floating-Point) operations. The low 32, 16, and 8 bits of each 64-bit register can be used independently of the high bit portions as depicted below:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| High bits | | | | Low bits | | | |
| Qword (quadruple word) - 64 bits | | | | | | | |
|  |  |  |  | Dword (double word) - 32 bits | | | |
|  |  |  |  |  |  | Word - 16 bits | |
|  |  |  |  |  |  |  | Byte 8 bits |

It should be noted that these smaller subdivisions of the registers are still in fact just subdivisions of the same data. Thus, changing one segment of a register affects all segments of that register.

## The Stack

For all non-trivial programs, it is desirable to have a tool called function calling, where program execution jumps somewhere else, processes something, and then later returns to where it was called from to resume execution. For this purpose, a stack is required to manage how to return to where a function was called from.

A stack memory structure can be thought of as a literal stack of something. Given a stack of books, all you can do is either put another book on the top or take the top book off. However, like a stack of books, something can theoretically by taken from the middle, or even the bottom, but in both the memory structure and the analogy, this can easily end in disaster.

The stack is managed by [R15](#_Registers). Upon program initialization, R15 is set to the address of the top of the stack, which begins at the high side of the program’s available memory and grows downward. Because of this, R15 will always point to the most-recently-added item on the stack.

R15 may be modified directly, but care should be taken when doing so, as this could result in permanent damage to the stack structure.

## Error Codes

During program execution, the processor may encounter an error (such as division by zero). When this happens, the it will set an error code and immediately halt execution. Reviewing the error code emitted and any recent modifications to your program is the best way to diagnose the cause and correct the error.

With careful coding, all of these errors can be trivially avoided.

|  |  |
| --- | --- |
| **Error Code** | **Name** |
| 0 | None |
| 1 | Out of Bounds |
| 2 | Unhandled Syscall |
| 3 | Undefined Behavior |
| 4 | Arithmetic Error |

## Main Memory

Because the [registers](#_Registers) in the processor are much too small to hold any real amount of information, a larger storage base is required by any non-trivial program. Main memory (RAM) serves as this data reservoir, and is where the vast majority of your data should be held.

Upon program initialization, a page of memory is allocated to the program by the [virtual operating system](#_Virtual_Operating_System), enough to contain the contents of the program as well as an amount of additional space for the [stack](#_Stack) and [heap](#_Heap). The program’s contents are then loaded into this memory starting at address zero.

Data stored and retrieved from main memory are little-endian, meaning the least significant bytes come first in memory. This is opposed to big-endian systems, in which the most significant bytes come first. This can easily be conceptualized using base 10 digits:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Memory** | | | | **Value** | |
| **Low Address** | | **High Address** | | **Little-Endian** | **Big-Endian** |
| 4 | 1 | 9 | 0 | 914 | 4190 |
| 2 | 4 | 6 | 3 | 3642 | 2463 |
| 7 | 7 | 0 | 5 | 5077 | 7705 |
| 0 | 0 | 1 | 8 | 8100 | 18 |

You may be wondering why anyone in their right mind would pick little-endian over big-endian, as little-endian essentially reads the numbers backwards. You would, in fact, have a good argument. The truth is: there’s no clear winner, and even some systems today are still big-endian (though little-endian is the industry standard in modern computing, which is why CSX64 uses it).

One of the most compelling reasons to use little-endian systems is that it makes pointer type conversions be no-op (meaning, if given the address of a 64-bit integer, that same address can also be used to get the value as if it were a 32-bit value), which comes up in low-level programming more often than you’d think.

It should be noted that attempting to load or store memory beyond the page provided by the virtual operating system will result in an [Out of Bounds error](#_Error_Codes).

# Numeric Types

## Integer

The majority of processor operations involve the manipulation of integers of various widths (sizes). In fact, the integer is really the core of all modern computing, as all integers can be stored perfectly (i.e. no rounding) as a finite series of binary digits, which is perfect for a computer, as that’s all it’s available to anyway.

Now, the computer is a pure logic machine. It runs entirely on true or false. There is no in between. The reason behind this lies in manufacturing tolerances. At the heart of every processor is an intricate structure of electrical components. The simplest of such components for our purposes are logic gates (e.g. *and*, *or*, *xor*, *not*), which take in one or more electrical inputs and produce an output based on the voltages. For instance, an *and* gate outputs high voltage if all of its input voltages are high, and outputs low voltage otherwise.

So, what does it mean for the voltage to be high or low? Well, this is where the natural imperfection of the universe comes into play: there’s no strict definition. You could take two different logic gates, run the same voltages through all of the inputs, and ultimately get different results. This is because the gates are not perfectly identical: one might have slightly more resistance on its output or on one or more inputs due to flaws in the material and manufacturing.

Because of this, designing a processor to understand anything between purely on or off unavoidably results in it not being guaranteed to be repeatable (as well as vastly increasing the manufacturing cost due to the stricter specifications that would be required to even come close to pulling it off). And so, we use binary (a.k.a. 1’s and 0’s).

For information on how to convert integers to and from binary, see the [examples page on binary integers](#_Binary_Integers).

### Unsigned

Unsigned integers have no sign, hence the name. This is the simplest type of binary integer, as it’s exactly what you would expect: pure binary numbers. No tricks. Because they are unsigned, however, they can only range 0, 1, 2, …, meaning they mainly good for dealing with raw data or values you know absolutely cannot be negative (e.g. size of an array, temperature in kelvin, your age, etc.)

Whenever this criterium is met, it is usually better to use unsigned rather than signed values due to their simplicity and higher range, which will be elaborated on in the signed integer section

### Signed

Unlike unsigned integers, signed integers can be used to represent negative numbers. In most modern architectures (and indeed in the CSX64 architecture), this is done via an encoding mechanism known as 2’s complement.

In 2’s complement, the high bit (bit with the highest weight) is used to represent the sign of the number (1 for negative, 0 for positive). However, taking the negative is not as simple as just flipping the sign bit (and it’s good this is not the case, as you’re about to find out).

To take the negative of a signed integer, we apply the algorithm ~v + 1 where v is the value to negate and ~ is the bitwise not.

Because of the bitwise not, the high (colloquially, sign) bit is also flipped. As you can see, this means -0 = 0, which is a good thing (where this would not be the case had we simply flipped the sign bit).

Rather conveniently ([but really by definition](#_Proof_of_2’s)), 2’s complement results in a beautiful encoding scheme in which addition and subtraction are exactly the same process regardless of signage (i.e. signed or unsigned). Moreover, the low half of multiplication is also the same, regardless of signage, though the high half of the product (which is generally ignored in all but assembly/machine code) does differ for signed and unsigned values.

The consequence of not having a positive and negative zero is that there is one more negative number than positive numbers, meaning there is actually a single non-zero value for any width of signed integer whose negative is itself (e.x. for 8-bit signed integers: -128 is a valid negative number, but its negative is still -128, as 128 is not a valid positive number).

## Floating-Point

Integers are very simple, both to understand and to convert from “human script” (base 10) to binary. However, they are very limited numerically (e.x. 1 / 2 = 0). It would be nice to have decimal numbers as well, but this also has complications.

If we used a fixed-width integer to represent a binary number extending into negative powers of 2 (just as decimal does where 0.23 is really 2\*10-1 + 3\*10-2), we would have a structure that would either wouldn’t have many decimal places or would have a maximum value much smaller than its same-width integer would.

If we used a type that could expand or contract as needed to store the integral and decimal portions it would be very big (not to mention would need to be [allocated dynamically](#_Heap_1)), and thus would likely not fit in a register for processing. Because of this, operations involving it would be highly entangled with [main memory](#_Main_Memory), which would slow things down tremendously. And even then, what about numbers that repeated? Where should we cut them off?

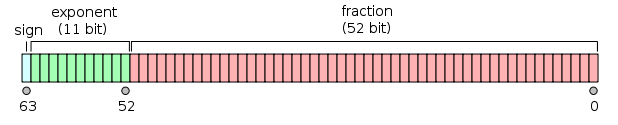
Conveniently, scientists have given us the answer with their inspirationally-titled *scientific notation*. If this entry in the CSX64 manual hasn’t phased you yet, you’re well accustomed to scientific notation already. The idea of scientific notation is to use a fixed with binary number with negative powers of 2 (as discussed above), but also multiplying it by a power of 2, which solves its issue of being limited in range compared to an integer of equal width. Another key aspect of scientific notation is that it is of the form a\*10b, where a is always in the range [1, 10). To scale the number around to other magnitudes, we simply change b, hence the name floating-point. Along with this comes the concept of significand digits (the idea that we will only have a couple of digits that actually define the number), which is perfect for our uses, as we simple can’t efficiently pull off variable significant digits.

And so, we come to the big question: how do we encode these supposed floating-point monstrosities?

### 64-bit – Double Precision

The processor contains hardware support for handling 64-bit (“double precision”) IEEE 754 floating-point values. There are plenty of utilities that will convert decimal numbers into their respective floating-point representations, but it’s occasionally beneficial to understand the underlying encoding scheme of the structure (e.x. say you wanted to extract what the power of 2 is without the overhead of taking its absolute value, then its logarithm, then casting it to an integer via truncation, as well as domain shifting due to not being able to take the logarithm of zero). If you understood its structure, it would simply be a right shift, a bitwise and, and a subtraction, all of which are vastly faster than logarithms or potential function calls.

Here’s a graphical representation of the data structure:



Sign – A single bit to represent the sign of the number (i.e. 1 for negative, 0 for positive).

Exponent – An 11-bit exponent encoded in excess-1023 (i.e. the “real” power + 1023). This 1023 value is known as the bias (b) of the representation, and would be different for different widths of floating-point representations.

Fraction – A 52-bit fractional portion that omits the first 1 (and is therefore technically 53 bits for the purpose of calculating precision) (e.x. 1.00101101 would be stored as 00101101).

Because the fraction portion is ordinarily assumed to have a hidden 1 in the front (which is not stored in the binary representation), it would be impossible to represent zero. Because of this, having an exponent of zero denotes a denormalized value, where there is not assumed to be a hidden 1. This also means that the true exponent is interpreted is -b+1 to account for this sudden loss of a high order significant digit (rather than the -b it would be under normalized interpretation: e-b), thus closing the gap between de/normalized values.

There is also a concept of infinity, which is represented as an exponent of all 1’s and a fraction of all zeros. The sign bit can be changed to represent positive and negative infinity.

Additionally, there is a concept of invalid arithmetic: NaN (Not a Number). These come in two varieties: QNaN (quiet), and SNaN (signaling). QNaN values are errorlessly produced by some floating-point operations (e.g. dividing zero by zero). The result of any arithmetic involving QNaN will also result in QNaN, thus safely propagating through a series of operations. SNaN however, will cause the floating-point module in a modern processor to emit an error when used in any operations (though this is usually “quietened” to a QNaN, thus propagating through arithmetic and not halting program execution).

A summary of 64-bit floating-point values is as follows, where x denotes either a 1 or a 0:

|  |  |  |  |
| --- | --- | --- | --- |
| **Floating Point Values** | | | |
| **Sign** | **Exponent (e)** | **Fraction (f)** | **Value** |
| 0 | 00…00 | 00…00 | +0 |
| 0 | 00…00 | 00…01  ⋮  11…11 | Positive Denormalized Real  0.f\*2(-b+1) |
| 0 | 00…01  ⋮  11…10 | xx…xx | Positive Normalized Real  1.f\*2(e-b) |
| 0 | 11…11 | 00…00 | +∞ |
| 0 | 11…11 | 00…01  ⋮  01…11 | SNaN |
| 0 | 11…11 | 1x…xx | QNaN |
| 1 | 11…11 | 1x…xx | QNaN |
| 1 | 11…11 | 00…01  ⋮  01…11 | SNaN |
| 1 | 11…11 | 00…00 | -∞ |
| 1 | 00…01  ⋮  11…10 | xx…xx | Positive Normalized Real  -1.f\*2(e-b) |
| 1 | 00…00 | 00…01  ⋮  11…11 | Positive Denormalized Real  -0.f\*2(-b+1) |
| 1 | 00…00 | 00…00 | -0 |

The positive and negative zero values can both be used to represent zero, but fortunately will compare equal under floating-point comparison. They can be thought of as special cases of denormalized real values (where the fractional portion is zero).

There are several well-defined special cases for the result of floating point operations, some of which are listed below. In general, if the result of an operation has a conceptual value, it will result in that value (e.x. adding infinities of the same sign will result in another infinity with the same sign). If it does not produce a logical result (e.x. adding infinities of opposite signs), it will result in QNaN.

|  |  |  |  |
| --- | --- | --- | --- |
| **Operation** | **Result** | **Operation** | **Result** |
| n / ±∞ | 0 | ∞ + ∞ | ∞ |
| ±∞ \* ±∞ | ±∞ | ∞ - -∞ |
| ±nonzero / ±0 | ±∞ | -∞ - ∞ | -∞ |
| ±finite \* ±∞ | ±∞ | -∞ + -∞ |
| ±0 / ±0 | QNaN | ∞ - ∞ | QNaN |
| ±∞ / ±∞ | QNaN | -∞ + ∞ |
| ±∞ \* ±0 | QNaN |  |  |

For more information about how to encode floating-point values, refer to [Encoding IEEE 754 64-bit Floating-Point Values](#_Encoding_IEEE_754) in the Appendix.

For more information on the structure of the IEEE 754 format, refer to:

<http://steve.hollasch.net/cgindex/coding/ieeefloat.html>.

### 32-bit – Single Precision

CSX64 also supports IEEE 754 32-bit floating-point computations. This format is identical to the above explained 64-bit format, except that there are 23 bits to encode the fractional portion (which still excludes the leading 1), and 8 bits to record the exponent (meaning the exponent bias is 127 instead of 1023).

All other special properties (e.g. infinity, NaN, denormalization) function identically to its 64-bit counterpart, and have the same formats as shown in the table above.

In the past, 32-bit floating point was used because it was faster than 64-bit, which made it the obvious choice for anything that didn’t need the extra precision or range. Nowadays however, the speed difference is negligible (in fact, 64-bit is sometimes actually faster than 32-bit, depending on the processor). The main reason to use 32-bit in modern times is just to save space. If you need an array of one million floating point values and you don’t need high precision, you might as well use 32-bit and save four million bytes. This also has the effect of making code smaller, meaning more can fit into a cache line, thus reducing cache misses and ultimately speeding up execution.

# Machine Code

As discussed elsewhere, a processor is essentially just a logic machine. Values are encoded in a binary, put through some logical manipulation algorithm (which may be incredibly complex), and ultimately transformed into some meaningful result. But how does the processor know how to execute a program? Well, just like numbers, programs themselves must somehow be encoded into binary. This is what’s known as machine code, and it lies at the heart of every computer. There is no programming language (if it can even be called such) closer to the hardware itself than machine code.

Machine code varies from processor to processor, which makes it very difficult to program in. However, The CSX64 processor is cross-platform, meaning any program you write in CSX64 machine code will be able to run on any machine via the interpreter.

Now, machine code is very technical, and most operations require certain key pieces of information to be in specific bit fields of the binary code. Below is an overview of the standard formats used by CSX64 binary.

## Register Format

Nearly all operations involve at least one of the [general-purpose registers](#_Registers). As there are 16 of such registers, they are designated by a 4-bit register address, where the registers are labeled 0-15. Any opcode format specification denoting a source or destination register (or abbreviated “reg”, “r”, “r1”, “r2”, etc.) uses this format unless otherwise specified.

## Size Format

Most operations involving registers perform their operation on a specific subdivision of the registers into 64, 32, 16, or 8-bit portions. These subdivisions are designated by a 2-bit size code. Any opcode format specification denoting size uses this format unless otherwise specified.

|  |  |  |
| --- | --- | --- |
| **Size Code** | **Size** | **Size Name** |
| 0 | 8 bits | byte |
| 1 | 16 bits | word |
| 2 | 32 bits | dword |
| 3 | 64 bits | qword |

## Multiplier Format

Some operations may specify one or more size multipliers (most notably memory addressing operations). These multipliers are designated by a 3-bit multiplier code. Any opcode format specification denoting a multiplier (or abbreviated “mult”, “m”, “m1”, “m2”, etc.) uses this format unless otherwise specified.

|  |  |
| --- | --- |
| **Multiplier Code** | **Multiplier Value** |
| 0 | 0 |
| 1 | 1 |
| 2 | 2 |
| 3 | 4 |
| 4 | 8 |
| 5 | 16 |
| 6 | 32 |
| 7 | 64 |

## Memory Address Format

In CSX64, [memory](#_Main_Memory) addresses have a width of 64 bits, and are always in a standard format:

[1: base][3: m1][1: -m2][3: m2] ([4: r1][4: r2]) ([64: imm])

If base is 1, imm is assumed to be provided.

If m1 or m2 is nonzero, r1 and r2 are assumed to be provided.

If -m2 is 1, r2 is first negated before being used in the address computation (but r2 is not modified).

The resulting address is thus imm + m1\*r1 + m2\*r2.

This documentation form is used for all CSX64 machine code specifications, where an item in brackets represents a labeled collection of bits. These bit fields are grouped into blocks of 8, 16, 32, or 64 bits, with spacing in between for ease of reading. Each field in a block specifies how many bits are used to represent it, and thus the format’s structure, with the left of each block being high bits.

Any item in parenthesis may be omitted depending on the other fields, as described by each operation’s machine code specification.

## Conditions

At the hardware level, conditionals are processed based on the value of a special flags [register](#_Registers). The flags register contains several 1-bit flags that represent different processor states. These flags are modified in different ways by many operations, but the following table generalizes the behavior of each flag under typical use. The top section lists “public” flags, which can be modified by the [SETF](#_SETF_–_Set) operation. The others are “private” flags, which cannot be modified by SETF (but may be modified by the virtual operating system).

|  |  |  |
| --- | --- | --- |
| **Flag** | **Name** | **Description** |
| ZF | Zero Flag | The last operation resulted in zero |
| PF | Parity Flag | The low 8 bits of the last operation had even parity |
| OF | Overflow Flag | The last operation resulted in arithmetic overflow |
| CF | Carry Flag | The last operation resulted in a carry or borrow from the high bit |
| SF | Sign Flag | The last operation resulted in a negative value |
| SMF | Slow Memory Flag | Indicates that memory access is artificially slowed |

Once these flags have been set by an operation, they can be used by any conditional operation. The following table lists all the conditionals used in CSX64:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Label** | **Alias** | **Name** | **Signage** | **Condition** |
| Z | E | Zero / Equal | N/A | Z = 1 |
| NZ | NE | Not Zero / Not Equal | Z = 0 |
| S |  | Sign | S = 1 |
| NS |  | Not Sign | S = 0 |
| P | PE | Parity / Parity Even | P = 1 |
| NP | PO | Not Parity / Parity Odd | P = 0 |
| O |  | Overflow | O = 1 |
| NO |  | Not Overflow | O = 0 |
| C |  | Carry | C = 1 |
| NC |  | Not Carry | C = 0 |
| A | NBE | Above | Unsigned | CF = 0 and ZF = 0 |
| AE | NB | Above or Equal | CF = 0 |
| B | NAE | Below | CF = 1 |
| BE | NA | Below or Equal | CF = 1 or ZF = 1 |
| G | NLE | Greater | Signed | ZF = 0 and SF = OF |
| GE | NL | Greater or Equal | SF = OF |
| L | NGE | Less | SF != OF |
| LE | NG | Less or Equal | ZF = 1 or SF != OF |

Because not all operations affect the flags in the same way, it is necessary to reference the documentation for any operation before using its resulting flags. In CSX64, if an operation’s documentation doesn’t state that it modifies a particular flag, the value of said flag is unchanged (though this is not the case on all modern processor architectures).

## Translating Machine Code

Translating your program’s intent into machine code is an easy (albeit ludicrously tedious) task. All you have to do is follow the format the specification indicates, putting the values in the correct positions.

For example, to add 17 to the 32-bit segment of R3, we would fill in values for [ADD](#_ADD)’s binary specification, giving us:

[8: 29] [4: 3][2: 2][2: 0] [16: 17]

Which would thus result in the final binary form (in hexadecimal):

1d 38 11 00

Remember that CSX64 is [little-endian](#_Main_Memory), meaning the low order bytes of a value come first.

Now, say we wanted to compare the 32-bit segment of R3 to an element of an array (of 32-bit integers) in memory that starts at address 200. Let’s also say that we want the index of the item in the array to be the current value of R7. We thus need to use [CMP](#_CMP_–_Compare)’s binary specification, which in turn uses the [address specification](#_Memory_Address_Format):

[8: 45] [4: 3][2: 2][2: 1]

[1: 1][3: 3][1: 0][3: 0] [4: 7][4: 0] [64: 200]

Which gives us the binary form (in hexadecimal):

2d 39 f0 70 c8 00 00 00 00 00 00 00

Now let’s say we want to jump to address 1000 if the values were equal (i.e. result was zero). We thus use [Jz](#_Jcc_–_Conditional)’s binary specification:

[8: 63] [1: 1][3: 0][1: 0][3: 0] [64: 1000]

Which gives us the binary form (in hexadecimal):

3f 80 e8 03 00 00 00 00 00 00

Congratulations: you’ve just gone through the process of translating your first few lines of machine code! Now, putting it all together, we add 17 to the 32-bit segment of R3, compare it to a value from an array in memory with base address 200 and index in R7, then jump somewhere if they were equal. The full resulting machine code is thus:

1d 38 11 00

2d 39 f0 70 c8 00 00 00 00 00 00 00

3f 80 e8 03 00 00 00 00 00 00

Which is essentially the same as the following in C:

R3 += 17;

if (R3 == arr[R7]) goto somewhere;

# Assembly Language

As you can imagine, [writing programs in machine code](#_Translating_Machine_Code) is completely impractical. It requires you to reference the binary specification for every single operation you want to perform, and makes you have to do the grunt work of translating everything into binary and putting things in the proper bit fields (not to mention making you painfully aware of [endianness](#_Main_Memory)). For all those reasons and many, many more, programs are never written in machine code except for academic purposes. Assembly language serves as the bridge between the programmer and pure machine code and is the closest programming “language” (i.e. human-readable) to the underlying hardware.

Assembly languages are tailored to specific processors (though most assembly languages contain many operations that go by the same name and have the same syntax to increase code reusability). In essence, assembly language is just a human-readable form of machine code that is translated into machine language by an assembler.

As an example of what assembly language might look like, the following is the CSX64 assembly language equivalent of the example in the above section on [translating machine code](#_Translating_Machine_Code):

add:32 $3, 17

cmp:32 $3, [200 + 4\*$7]

je [1000]

As you can see, we specified the same values for each instruction that we translated before. The difference is that this is readable – you can look at this and know what it’s doing.

# CSX64 Assembly Language

CSX64 comes with a thorough, built-in assembler designed loosely around the standard Intel syntax (with several major changes to appeal to modern users).

Before we describe the syntax of operations and assembler utilities, we need to go over some of the basics of CSX64 assembly language (much of which holds for all assembly language).

Firstly, there are three places a value can come from: packaged-in with an operation (imm), from a register (r), or from memory (m).

## imm – Immediate

An imm, otherwise known as an immediate, is a value that is placed directly inside an operation’s binary specification. In assembly language this is a literal number, such as 0, 7, or 3.14. These serve the purpose of providing values to [fill in the bit fields](#_Translating_Machine_Code) of an operation’s binary specification, as we did above.

However, an imm doesn’t have to just be a pure number, it may be an expression:

(1 + 3) \* 8.3

It may even involve symbolic names:

(a + b) / 2

In the most general sense, an imm is anything that doesn’t come from a register or memory.

The following table lists all the imm operators, in order of decreasing precedence, where grouped operators have equal precedence:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Usage** | **Name** | **Result** | **Result Type** | **Evaluation Order** |
| A \* B | Multiply | A \* B | Floating if A or B is floating, otherwise integral | Left-to-Right |
| A / B | Divide | A / B |
| A % B | Mod | Remainder of  A / B |
| A + B | Add | A + B | Floating if A or B is floating, otherwise integral | Left-to-Right |
| A - B | Subtract | A - B |
| A << B | Shift Left | A << B | Floating if A or B is floating, otherwise integral | Left-to-Right |
| A >> B | Shift Right | A >> B |
| A < B | Less | 1 if condition met, otherwise 0 | Integral | Left-to-Right |
| A <= B | Less or Equal |
| A > B | Great |
| A >= B | Great or Equal |
| A == B | Equal | 1 if condition met, otherwise 0 | Integral | Left-to-Right |
| A != B | Not Equal |
| A & B | Bitwise And | A & B | Floating if A or B is floating, otherwise integral | Left-to-Right |
| A ^ B | Bitwise Xor | A ^ B | Floating if A or B is floating, otherwise integral | Left-to-Right |
| A | B | Bitwise Or | A | B | Floating if A or B is floating, otherwise integral | Left-to-Right |
| A && B | Logical And\* | 1 if A and B are nonzero, otherwise 0 | Integral | Left-to-Right |
| A || B | Logical Or\* | 1 if A or B is nonzero, otherwise 0 | Integral | Left-to-Right |
| A ?? B | Null-Coalescing\* | A if A is nonzero, otherwise B | Type of the branch taken | Left-to-Right |
| A ? B : C | Ternary Conditional\* | B if A is nonzero, otherwise C | Type of the branch taken | Right-to-Left |

(\*) This operation supports short-circuiting evaluation.

There are also several unary operations, which are listed in the following table:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Usage** | **Name** | **Result** | **Result Type** | **Evaluation Order** |
| +A | Identity | A | Type of A | Right-to-Left |
| -A | Negative | -A |
| ~A | Bitwise Not | ~A |
| !A | Logical Not | 1 if A is zero, otherwise 0 | Integral |
| \*A | To Floating | A as floating | Floating |
| /A | To Integral | A as integral via truncation | Integral |

These operators may be chained in any way desired, with parenthesis being available for grouping if needed.

Furthermore, there are several ways to denote numeric literals, which are described below.

|  |  |  |
| --- | --- | --- |
| **Type** | **Description** | **Example** |
| Hexadecimal Integer | Prefaced by “0x” | 0xdeadbeef |
| Binary Integer | Prefaced by “0b” | 0b10011011 |
| Octal Integer | Prefaced by “0” | 0712 |
| Decimal Integer | Otherwise | 167 |
| Character (Integer) | A character enclosed in single quotes | ‘A’ |
| Decimal Floating-Point | Begins with a digit and contains a decimal point and/or e or E with an exponent (which may optionally have a sign) for exponential notation.  “∞” for infinity.  “NaN” for NaN (case sensitive). | 0.1  3.14  2.5e4  1.67e-11 |

### Instant Imm

For many operations, an imm is required to be “instant”. This simply means that its value must be calculatable at that time. For instance, if an imm has expression (a + b) / 2, it is not instant until both a and b are instant. By definition, all numeric literals are instant, and thus expressions consisting entirely of numeric literals are also instant.

## r – Register

A register value is simply a value that is used to refer to a register. As discussed in the section on [registers](#_Registers), CSX64 has 16 general-purpose registers, indexed 0-15. To refer to one of these registers, you designate the index of the register, preceded by “$”. For instance, $3 would refer to R3.

In fact, the index may be any [instant imm](#_Instant_Imm), with the only limitation being that compound expressions must be parenthesized. The following table summarizes these rules:

|  |  |
| --- | --- |
| **Expression** | **Referenced Register** |
| $3 | R3 |
| $abc | Depends on value of abc, which must be instant |
| $(3+1) | R4 |
| $3+1 | Error – compound register index must be parenthesized |
| $(abc\*3 + 1) | Depends on value of abc, which must be instant |
| $(a + b)/2 | Error – compound register index must be parenthesized |

Additionally, the evaluated instant imm used as the register index must be integral and in the range 0-15.

## m – Memory

A memory value is an expression that evaluates to the address of some value in memory. In CSX64 assembly language, addresses are enclosed in brackets “[…]” with the address expression inside. An address expression follows the same rules as any imm, with the added ability to reference as many as two registers to use in computing the address, as discussed in the section on [memory addresses](#_Memory_Address_Format).

Just like the machine code version of addressing (or rather because of), the registers used in the calculation must be connected by addition or subtraction to the rest of the expression, may be multiplied by an integral 0, 1, 2, 4, 8, 16, 32, or 64, and only one of the registers may be subtracted (i.e. negated before adding).

The assembler is capable of deducing the overall signedness of the register in the expression (e.g. if a register is connected by an even number of negations or subtractions, it understands that it is in fact not being negated). It should be noted that the assembler does not recognize ~v+1 as being equivalent to -v for deducing register signedness.

Additionally, the assembler is capable of applying algebra to put a register expression into a legal form if possible. (e.g. [8\*($1+4) + 64] is converted to [8\*$1 + 96]). It should be noted that this is only done for multiplication: other algebraic simplifications must be done by hand.

With all of these simplification rules in mind, register symbols in memory addresses may only be connected by addition, subtraction, unary plus, negation, and multiplication.

Certain imm expression ops offer short-circuiting evaluation. An important note on this is that registers may not be used in these operations. (e.g. [abc ? $0 : 0] is ill-formed even if abc is instant).

Finally, the assembler will combine multipliers of the same register:

(e.g. [6\*$5 - 8\*$6 - 2\*$5] is converted to [4\*$5 - 8\*$6] and is thus legal).

In fact, more than two registers may be specified, so long as only two of them have nonzero multipliers and only one of them is negated:

(e.g. [4\*$5 - 8\*$6 - 4\*$5 + 2\*$7] is converted to [-8\*$6 + 2\*$7] and is thus legal).

Traditionally, assembly doesn’t have a notion of variables or data types: everything is just data. However, it becomes necessary to have symbolic names for referencing different values. For instance, if your program needed an IP address in several places spread out over many files, giving it a symbolic name would make the code far less obscure. Additionally, it means changing the IP address would only mean having to change it in that one file, as opposed to having to hunt down every occurrence of it and change them by hand.

Another compelling argument is in addresses: would you rather refer to your function as “the one at address 300” or “the one named cross\_product”? And even more compellingly: how do you even know what the absolute address of something will be? If multiple files are used, the order of those files in the final executable may vary. The simplest solution is to simply have the assembler and linker take care of all the heavy lifting for you.

## Symbols

A symbol is any user-created name that has an associated imm (though most assemblers/linkers defined a few symbols for you, some of which you couldn’t come up with on your own, as they hold information about the resulting executable itself).

Symbols can be used in any imm expression, even before they’ve been defined (though that would also make them not [instant imms](#_Instant_Imm)). Symbols should be used anywhere you would otherwise put a value that you will likely need to reference elsewhere.

It is important to remember that assembly language does not have a notion of variables. Symbols are not memory locations to put something in. In fact, they don’t impact the resulting executable at all, and are simply a means of making assembly language more convenient.

All symbols must have a legal symbol name, which follows the same rules as most programming languages: all characters in the name must be alphanumeric or underscores, and the first character may not be a number.

### Predefined Symbols

The CSX64 assembler/linker define several symbols automatically, which are listed in the table below. It is undefined behavior to declare any of these symbols as global.

|  |  |
| --- | --- |
| **Symbol Name** | **Value** |
| \_\_time\_\_ | The time of the assembly process expressed as the number of 100 nanosecond intervals since January 1, 0001. |
| \_\_version\_\_ | The version of CSX64, where increasing values indicate newer versions. |
| \_\_pinf\_\_ | Floating-point positive infinity. |
| \_\_ninf\_\_ | Floating-point negative infinity. |
| \_\_nan\_\_ | Floating-point NaN. |
| \_\_fmax\_\_ | The maximum (finite) floating-point value. |
| \_\_fmin\_\_ | The minimum (finite) floating-point value. |
| \_\_fepsilon\_\_ | The smallest positive floating-point value. |
| \_\_pi\_\_ | The mathematical constant pi. |
| \_\_e\_\_ | The mathematical constant e. |
| \_\_prog\_end\_\_ | The address of the end of the executable program segment. |

## Labels

Labels are the tool assembly languages use to keep tabs on addresses. They are simply symbols that will refer to the address of something in the resulting executable. To do this, the assembler takes note of where the label was defined, and then adds a base address, which will be defined by the linker. Because of this, any expression involving a label is never instant.

Labels should be used when you need to refer to something’s address (e.g. function).

### Local Label

A local label is label that is associated with a non-local label. Because assembly language is prolific with labels, you can easily run into the problem of having clashing label names. Rather than naming your labels “loop\_top\_1”, “loop\_top\_2”, etc., you can give them local names.

A local label is tied to the most-recent non-local label and is only visible within that scope (i.e. the next non-local label definition will have a different scope, and thus not be able to access the other scopes’ local labels.

In CSX64, local labels are created and referenced by putting a period in front of the normal label name.

### Local Symbol

In addition to local labels, CSX64 assembly generalizes this locality to symbols as well. They function in exactly the same way, but refer to a normal imm, rather than strictly an address.

## Operation Syntax

CSX64 assembly uses a very standardized approach to assembly language. Each line is divided in the following way:

(label: label: ...) (op(:size) arg1, arg2, ...) #comment

The label section consists of zero or more labels, each of which may optionally be made local by appending a period to the front of the label name. The colon is not part of the label name. The labels must be separated from one another and from the operation by white space. The labels are introduced from left-to-right.

The operation section consists of zero or one operation and zero or one size parameter (which have different meaning depending on the operation). If the size parameter is not specified, it is defaulted to 64 (bits). The size parameter can be any instant imm, with the sole exception being that compound expressions (i.e. expressions involving any operators) must be parenthesized (this is in order to distinguish it from the argument section). Furthermore, the size parameter (or the operation if none is provided) must be separated from the argument section by white space.

The argument section consists of zero or more comma-separated arguments for the selected operation.

The optional comment section begins with a hash (#) and continues to the end of the line. It is ignored during assembly.

## Directives

CSX64 assembly has several pseudo-operations that act as operations but do not necessarily correspond to the generation of any binary code in the resulting executable. Primarily these directives offer a means of controlling some aspect of the assembly process.

### Global

The Global directive takes one or more symbol names and adds them to the export table (i.e. the symbol names that are accessible from other assembly files). A symbol can be made global before being defined. Local symbols cannot be made global.

### Def

The Def directive introduces a new symbol, where the new symbol name is the first argument and its expression is the second argument.

### Emit

The emit directive takes one or more imms and writes them directly into the binary file in-place from left to right. The size of the values written is determined by the size parameter. Additionally, if given a line of text enclosed in quotations, emit will write each character in the string as if it were a character literal.

## Assembly Process

Assemblers take an assembly file and create an object file. Object files are binary representations of the assembly file, with additional information pertaining to missing information (i.e. symbols that were not defined in that file, including the base of all labels) and the definitions of global symbols.

An object file itself is not executable: it needs to have all of its missing information filled in. To do this, a separate program known as the linker takes all the object files and cross-references them to fill in any missing symbol definitions. Ultimately, the linker puts all the object files together and creates the final executable.

The purpose of this is to have a system where only files that were modified need to be reassembled when any change is made. This of course also means that large files (e.g. library implementations) do not need to be assembled every time you write a program that references them. Furthermore, many software developers prefer to distribute their libraries as object files, as this simultaneously mitigates the assembly time of clients of the library and prevents them from having the source code (although the object file could be disassembled, all the comments and most symbol names would be gone forever, leaving the disassembled file to be a virtually illegible enigma.

# Operations

|  |  |
| --- | --- |
| **Name** | **Description** |
| [NOP](#_NOP_–_No) | No operation |
| [STOP](#_STOP_–_Stop) | Stops processor execution |
| [SYSCALL](#_SYSCALL_–_System) | Invokes virtual operating system utilities |
| [MOVcc](#_MOVcc_–_Conditional) | Conditional move |
| [SWAP](#_SWAP_–_Swap) | Swaps two values |
| [UX](#_UEXTEND_/_UX) | Unsigned (zero) extend |
| [SX](#_SEXTEND_/_SX) | Sign extend |
| [UMUL](#_UMUL_–_Unsigned) | Unsigned multiply |
| [SMUL](#_SMUL_–_Signed) | Signed multiply |
| [UDIV](#_UDIV_–_Unsigned) | Unsigned divide |
| [SDIV](#_SDIV_–_Signed) | Signed divide |
| [ADD](#_ADD) | Add |
| [SUB](#_SUB) | Subtract |
| [BMUL](#_BMUL_–_Binary) | Binary multiply |
| [BUDIV](#_BUDIV_–_Binary) | Binary unsigned divide |
| [BUMOD](#_BUMOD_–_Binary) | Binary unsigned mod |
| [BSDIV](#_BSDIV_–_Binary) | Binary signed divide |
| [BSMOD](#_BSMOD_–_Binary) | Binary signed mod |
| [SL](#_SL_–_Shift) | Shift left |
| [SR](#_SR_–_Shift) | Shift right |
| [SAL](#_SAL_–_Shift) | Shift arithmetic left |
| [SAR](#_SAR_–_Shift) | Shift arithmetic right |
| [RL](#_RL_–_Rotate) | Rotate left |
| [RR](#_RR_–_Rotate) | Rotate right |
| [AND](#_AND_–_Bitwise) | Bitwise AND |
| [OR](#_OR_–_Bitwise) | Bitwise OR |
| [XOR](#_XOR_–_Bitwise) | Bitwise XOR |
| [CMP](#_CMP) | Compare |
| [TEST](#_TEST_–_Logical) | Test |
| [INC](#_INC_–_Increment) | Increment |
| [DEC](#_DEC_–_Decrement) | Decrement |
| [NEG](#_NEG_–_Negate) | Negate |
| [NOT](#_NOT_–_Bitwise) | Bitwise not |
| [ABS](#_ABS_–_Absolute) | Absolute value |
| [CMPZ](#_CMPZ_–_Compare) | Compare to zero |
| [LA](#_LA_–_Load) | Load address |
| [Jcc](#_Jcc_–_Conditional) | Conditional jump |
| [FADD](#_FADD_–_Floating-Point) | Floating-point add |
| [FSUB](#_FSUB_–_Floating-Point) | Floating-point subtract |
| [FMUL](#_FMUL_–_Floating-Point) | Floating-point multiply |
| [FDIV](#_FDIV_–_Floating-Point) | Floating-point divide |
| [FMOD](#_FMOD_–_Floating-Point) | Floating-point mod |
| [POW](#_POW_–_Power) | Floating-point power |
| [SQRT](#_SQRT_–_Square) | Floating-point square root |
| [EXP](#_EXP_–_Exponentiate) | Floating-point exponentiate |
| [LN](#_LN_–_Natural) | Floating-point natural log |
| [FNEG](#_FNEG_–_Floating-Point) | Floating-point negate |
| [FABS](#_FABS_–_Floating-Point) | Floating-point absolute value |
| [FCMPZ](#_FCMPZ_–_Floating-Point) | Floating-point compare to zero |
| [SIN](#_SIN_–_Sine) | Floating-point sine |
| [COS](#_COS_–_Cosine) | Floating-point cosine |
| [TAN](#_TAN_–_Tangent) | Floating-point tangent |
| [SINH](#_SINH_–_Hyperbolic) | Floating-point hyperbolic sine |
| [COSH](#_COSH_–_Hyperbolic) | Floating-point hyperbolic cosine |
| [TANH](#_TANH_–_Hyperbolic) | Floating-point hyperbolic tangent |
| [ASIN](#_ASIN_–_Arcsine) | Floating-point arcsine |
| [ACOS](#_ACOS_–_Arccosine) | Floating-point arccosine |
| [ATAN](#_ATAN_–_Arctangent) | Floating-point arctangent |
| [ATAN2](#_ATAN2_–_Binary) | Floating-point binary arctangent |
| [FLOOR](#_FLOOR_–_Floor) | Floating-point floor |
| [CEIL](#_CEIL_–_Ceiling) | Floating-point ceiling |
| [ROUND](#_ROUND_–_Round) | Floating-point round |
| [TRUNC](#_TRUNC_–_Truncate) | Floating-point truncate |
| [FCMP](#_FCMP_–_Floating-Point) | Floating-point compare |
| [FTOI](#_FTOI_–_Floating-Point) | Floating-point to integer |
| [ITOF](#_ITOF_–_Integer) | Integer to floating-point |
| [PUSH](#_PUSH_–_Push) | Push value onto stack |
| [POP](#_POP_–_Pop) | Pop value from stack |
| [CALL](#_CALL_–_Call) | Call function |
| [RET](#_RET_–_Return) | Return from function |
| [BSWAP](#_BSWAP_–_Byte) | Swap byte order of a value |
| [BEXTR](#_BEXTR_–_Bitfield) | Extract a bitfield |
| [BSLI](#_BLSI_–_Binary) | Isolate the lowest-order set bit |
| [BLSMSK](#_BLSMSK_–_Binary) | Get a mask containing up to and including the lowest-order set bit |
| [BLSR](#_BLSR_–_Binary) | Clear the lowest-order set bit |
| [ANDN](#_ANDN_–_Bitwise) | Bitwise and with a negated value |
| [GETF](#_GETF_–_Get) | Get the flags register |
| [SETF](#_SETF_–_Set) | Set the flags register |
| [LOOP](#_LOOP_–_Loop) | Decrement register and jump if nonzero |
| [FX](#_FEXTEND_/_FX) | Floating-point extend |

## NOP – No Operation

OP Code:

0

Description:

Specifies no operation. Usually only used for timing purposes.

Usage:

NOP

Flags Affected:

None

Format:

[8: NOP]

## STOP – Stop

OP Code:

1

Description:

Terminates the program and yields to the virtual operating system.

Usage:

STOP

Flags Affected:

None

Format:

[8: STOP]

## SYSCALL – System Call

OP Code:

2

Description:

Causes program execution to temporarily yield to the [virtual operating system](#_Virtual_Operating_System) for it to perform some specialized process. The results of this operation are thus virtual operating system-specific. If the operating system fails to perform the syscall, this operation generates an [Unhandled Syscall error](#_Heap).

This operation is the only way to access potential language extensions such as file IO or internet access.

Usage:

SYSCALL

Flags Affected:

Virtual operating system-dependent

Format:

[8: SYSCALL]

## MOVcc – Conditional Move

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Condition** | **OP Code** | **Condition** | **OP Code** | **Condition** | **OP Code** | **Condition** | **OP Code** |
| None | 3 |  |  | Z | 12 | NZ | 13 |
| A | 4 | G | 8 | S | 14 | NS | 15 |
| AE | 5 | GE | 9 | P | 16 | NP | 17 |
| B | 6 | L | 10 | O | 18 | NO | 19 |
| BE | 7 | LE | 11 | C | 20 | NC | 21 |

Description:

Conditionally copies a value from some source to a destination.

Does not support memory-to-memory transfers (in accordance with modern processor architectures).

Usage:

MOVcc(:size) r, imm/r/m

MOVcc(:size) m, imm/r

Flags Affected:

None

Format:

[8: MOVcc] [4: dest][2: size][2: mode]

mode = 0: [size: imm]

dest 🡨 imm

mode = 1: [address]

dest 🡨 M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 src

mode2 = 1: [address]

M[address] 🡨 src

mode = 3: [size: imm] [address]

M[address] 🡨 imm

## SWAP – Swap

OP Code:

22

Description:

Exchanges two values.

Does not support memory-to-memory transfers.

Usage:

SWAP(:size) r, r/m

SWAP(:size) m, r

Flags Affected:

None

Format:

[8: swap] [4: r1][2: size][1:][1: mem]

mem = 0: [4:][4: r2]

r1 🡨 r2

r2 🡨 r1

mem = 1: [address]

r1 🡨 M[address]

M[address] 🡨 r1

## UX – Unsigned Extend

OP Code:

23

Description:

Zero extends a register to the specified size, truncating the value if converting to a smaller size, or filling the new high bits with 0 if converting to a larger size. Converting to the same size is no-op.

Usage:

UX(:to\_size) instant imm from\_size, r

Flags Affected:

None

Format:

[8: UX] [4: reg][2: from\_size][2: to\_size]

(to\_size) reg 🡨 zero extend (from\_size) reg

## SX – Sign Extend

OP Code:

24

Description:

Sign extends a register to the specified size, truncating the value if converting to a smaller size, or filling the new high bits with the sign bit if converting to a larger size. Converting to the same size is no-op.

Usage:

SX(:to\_size) instant imm from\_size, r

Flags Affected:

None

Format:

[8: SX] [4: reg][2: from\_size][2: to\_size]

(to\_size) reg 🡨 sign extend (from\_size) reg

## UMUL – Unsigned Multiply

OP Code:

25

Description:

Computes the full product of multiplying two unsigned values, which will have a width of twice the operand size. In the case of a 64-bit product, the high bits are stored in R1, while the low bits are stored in R0.

The first value to multiply must already by loaded into R0. The second value however, may be an immediate, register, or memory value.

Due to being forced to use R0 and R1, as well as the fact that the high bits of a product are usually not relevant for most tasks requiring multiplication, [BMUL](#_BMUL) may be a more suitable alternative.

Usage:

UMUL(:size) imm/r/m

Operation:

byte: word R0 🡨 byte R0 \* byte val

word: dword R0 🡨 word R0 \* word val

dword: qword R0 🡨 dword R0 \* dword val

qword: dqword R1:R0 🡨 qword R0 \* qword val

Flags Affected:

OF and CF are both set if the high bits of the product are non-zero, and cleared otherwise.

Format:

[8: UMUL] [4: reg][2: size][2: mode]

mode = 0: [size: imm]

mode = 1: use reg

mode = 2: [address]

## SMUL – Signed Multiply

OP Code:

26

Description:

Computes the full product of multiplying two signed values, which will have a width of twice the operand size. In the case of a 64-bit product, the high bits are stored in R1, while the low bits are stored in R0.

The first value to multiply must already by loaded into R0. The second value however, may be an immediate, register, or memory value.

Due to being forced to use R0 and R1, as well as the fact that the high bits of a product are usually not relevant for most tasks requiring multiplication, [BMUL](#_BMUL) may be a more suitable alternative.

Usage:

SMUL(:size) imm/r/m

Operation:

byte: word R0 🡨 byte R0 \* byte val

word: dword R0 🡨 word R0 \* word val

dword: qword R0 🡨 dword R0 \* dword val

qword: dqword R1:R0 🡨 qword R0 \* qword val

Flags Affected:

OF and CF are both set if there are significant bits in the high portion of the product (i.e. truncation yields a different value), and cleared otherwise.

SF is set if the resulting value is negative (i.e. the highest bit of the result is set).

Format:

[8: SMUL] [4: reg][2: size][2: mode]

mode = 0: [size: imm]

mode = 1: use reg

mode = 2: [address]

## UDIV – Unsigned Divide

OP Code:

27

Description:

Computes the full quotient and remainder of the division of two unsigned values. The numerator has twice the width of the operand size, and must already be loaded into R0 (with the high bits in R1 in the case of qword division). The denominator however, may be an immediate, register, or memory value.

Dividing by zero or producing a quotient that cannot be stored in the operand register produces an [arithmetic error](#_Heap).

Due to having to use R0 and R1, as well as the fact that you usually don’t need both the quotient and the remainder of a division, [BUDIV](#_BUDIV) or [BUMOD](#_BUMOD) may be a more suitable alternative.

Usage:

UDIV(:size) imm/r/m

Operation:

byte: word R0 / byte val

byte R0 🡨 quotient , byte R1 🡨 remainder

word: dword R0 / word val

word R0 🡨 quotient , word R1 🡨 remainder

dword: qword R0 / dword val

dword R0 🡨 quotient , dword R1 🡨 remainder

qword: dqword R1:R0 / qword val

qword R0 🡨 quotient , qword R1 🡨 remainder

Flags Affected:

CF is set if the resulting remainder is nonzero, and cleared otherwise.

Format:

[8: UDIV] [4: reg][2: size][2: mode]

mode = 0: [size: imm]

mode = 1: use reg

mode = 2: [address]

## SDIV – Signed Divide

OP Code:

28

Description:

Computes the full quotient and remainder of the division of two signed values. The numerator has twice the width of the operand size and must already be loaded into R0 (with the high bits in R1 in the case of 64-bit division). The denominator however, may be an immediate, register, or memory value.

Dividing by zero or producing a quotient that cannot be stored in the operand register produces an [arithmetic error](#_Heap).

Due to having to use R0 and R1, as well as the fact that you usually don’t need both the quotient and the remainder of a division, [BSDIV](#_BSDIV) or [BSMOD](#_BSMOD) may be a more suitable alternative.

If the numerator or denominator is negative, the remainder is platform-dependent.

Usage:

SDIV(:size) imm/r/m

Operation:

byte: word R0 / byte val

byte R0 🡨 quotient , byte R1 🡨 remainder

word: dword R0 / word val

word R0 🡨 quotient , word R1 🡨 remainder

dword: qword R0 / dword val

dword R0 🡨 quotient , dword R1 🡨 remainder

qword: dqword R1:R0 / qword val

qword R0 🡨 quotient , qword R1 🡨 remainder

Flags Affected:

CF is set if the resulting remainder is nonzero, and cleared otherwise.

SF is set if the resulting quotient is negative (i.e. highest bit of quotient is set), and cleared otherwise.

Format:

[8: SDIV] [4: reg][2: size][2: mode]

mode = 0: [size: imm]

mode = 1: use reg

mode = 2: [address]

## ADD – Add

OP Code:

29

Description:

Adds the source (second argument) to the destination (first argument).

Usage:

ADD(:size) r, imm/r/m

ADD(:size) m, imm/r

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

CF is set if the addition caused a carry out from the high bit, and cleared otherwise.

OF is set if the addition resulted in arithmetic under/overflow, and cleared otherwise.

Format:

[8: ADD] [4: dest][2: size][2: mode]

mode = 0: [size: imm]

dest 🡨 dest + imm

mode = 1: [address]

dest 🡨 dest + M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 dest + src

mode2 = 1: [address]

M[address] 🡨 M[address] + src

mode = 3: [size: imm] [address]

M[address] 🡨 M[address] + imm

## SUB – Subtract

OP Code:

30

Description:

Subtracts the source (second argument) from the destination (first argument).

SUB can and should be used in place of [CMP](#_CMP) if the result is needed.

Usage:

SUB(:size) r, imm/r/m

SUB(:size) m, imm/r

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

CF is set if the addition caused a borrow from the high bit, and cleared otherwise.

OF is set if the addition resulted in arithmetic under/overflow, and cleared otherwise.

Format:

[8: SUB] [4: dest][2: size][2: mode]

mode = 0: [size: imm]

dest 🡨 dest - imm

mode = 1: [address]

dest 🡨 dest - M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 dest - src

mode2 = 1: [address]

M[address] 🡨 M[address] - src

mode = 3: [size: imm] [address]

M[address] 🡨 M[address] - imm

## BMUL – Binary Multiply

OP Code:

31

Description:

Computes the multiplication of two signed or unsigned integers, keeping only the low half of the result. If the full double-width product is needed, [UMUL](#_UMUL) or [SMUL](#_SMUL) can be used instead.

Multiplies the destination (first argument) by the source (second argument).

Usage:

BMUL(:size) r, imm/r/m

BMUL(:size) m, imm/r

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

Format:

[8: BMUL] [4: dest][2: size][2: mode]

mode = 0: [size: imm]

dest 🡨 dest \* imm

mode = 1: [address]

dest 🡨 dest \* M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 dest \* src

mode2 = 1: [address]

M[address] 🡨 M[address] \* src

mode = 3: [size: imm] [address]

M[address] 🡨 M[address] \* imm

## BUDIV – Binary Unsigned Divide

OP Code:

32

Description:

Computes the quotient of division of two unsigned integers. If the division of a double-width value is needed, [UDIV](#_UDIV) can be used instead.

Divides the destination (first argument) by the source (second argument).

Causes an [arithmetic error](#_Heap) on division by zero.

Usage:

BUDIV(:size) r, imm/r/m

BUDIV(:size) m, imm/r

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

Format:

[8: BUDIV] [4: dest][2: size][2: mode]

mode = 0: [size: imm]

dest 🡨 quotient dest / imm

mode = 1: [address]

dest 🡨 quotient dest / M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 quotient dest / src

mode2 = 1: [address]

M[address] 🡨 quotient M[address] / src

mode = 3: [size: imm] [address]

M[address] 🡨 quotient M[address] / imm

## BUMOD – Binary Unsigned Modulus

OP Code:

33

Description:

Computes the remainder of division of two unsigned integers. If the division of a double-width value is needed, [UDIV](#_UDIV) can be used instead.

Divides the destination (first argument) by the source (second argument) and takes the remainder.

Causes an [arithmetic error](#_Heap) on division by zero.

Usage:

BUMOD(:size) r, imm/r/m

BUMOD(:size) m, imm/r

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

Format:

[8: BUMOD] [4: dest][2: size][2: mode]

mode = 0: [size: imm]

dest 🡨 remainder dest / imm

mode = 1: [address]

dest 🡨 remainder dest / M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 remainder dest / src

mode2 = 1: [address]

M[address] 🡨 remainder M[address] / src

mode = 3: [size: imm] [address]

M[address] 🡨 remainder M[address] / imm

## BSDIV – Binary Signed Divide

OP Code:

34

Description:

Computes the quotient of division of two signed integers. If the division of a double-width value is needed, [SDIV](#_SDIV) can be used instead.

Divides the destination (first argument) by the source (second argument).

Causes an [arithmetic error](#_Heap) on division by zero.

Usage:

BSDIV(:size) r, imm/r/m

BSDIV(:size) m, imm/r

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

Format:

[8: BSDIV] [4: dest][2: size][2: mode]

mode = 0: [size: imm]

dest 🡨 quotient dest / imm

mode = 1: [address]

dest 🡨 quotient dest / M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 quotient dest / src

mode2 = 1: [address]

M[address] 🡨 quotient M[address] / src

mode = 3: [size: imm] [address]

M[address] 🡨 quotient M[address] / imm

## BSMOD – Binary Signed Modulus

OP Code:

35

Description:

Computes the remainder of division of two signed integers. If the division of a double-width value is needed, [SDIV](#_SDIV) can be used instead.

Divides the destination (first argument) by the source (second argument) and takes the remainder.

Causes an [arithmetic error](#_Heap) on division by zero.

If the destination or source is negative, the result is platform-dependent (though taking the remainder of negative numbers is virtually never needed).

Usage:

BSMOD(:size) r, imm/r/m

BSMOD(:size) m, imm/r

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

Format:

[8: BSMOD] [4: dest][2: size][2: mode]

mode = 0: [size: imm]

dest 🡨 remainder dest / imm

mode = 1: [address]

dest 🡨 remainder dest / M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 remainder dest / src

mode2 = 1: [address]

M[address] 🡨 remainder M[address] / src

mode = 3: [size: imm] [address]

M[address] 🡨 remainder M[address] / imm

## SL – Shift Left

OP Code:

36

Description:

Shifts the destination (first argument) to the left by source (second argument) bits.

Shifting an n-bit value is performed in modulo-n.

Vacated bits are filled with 0.

The amount to shift by is given by an 8-bit unsigned integer, but because CSX64 is little-endian, higher-width values from memory may also be used without having to modify the address.

Usage:

SL(:size) r, imm/r/m 8

SL(:size) m, imm/r 8

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

Format:

[8: SL] [4: dest][2: size][2: mode]

mode = 0: [8: imm]

dest 🡨 dest << imm

mode = 1: [address]

dest 🡨 dest << (8)M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 dest << (8)src

mode2 = 1: [address]

M[address] 🡨 M[address] << (8)src

mode = 3: [8: imm] [address]

M[address] 🡨 M[address] << imm

## SR – Shift Right

OP Code:

37

Description:

Shifts the destination (first argument) to the right by source (second argument) bits.

Shifting an n-bit value is performed in modulo-n.

Vacated bits are filled with 0.

The amount to shift by is given by an 8-bit unsigned integer, but because CSX64 is little-endian, higher-width values from memory may also be used without having to modify the address.

Usage:

SR(:size) r, imm/r/m 8

SR(:size) m, imm/r 8

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

Format:

[8: SR] [4: dest][2: size][2: mode]

mode = 0: [8: imm]

dest 🡨 dest >> imm

mode = 1: [address]

dest 🡨 dest >> (8)M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 dest >> (8)src

mode2 = 1: [address]

M[address] 🡨 M[address] >> (8)src

mode = 3: [8: imm] [address]

M[address] 🡨 M[address] >> imm

## SAL – Shift Arithmetic Left

OP Code:

38

Description:

Shifts the destination (first argument) to the left by source (second argument) bits.

Shifting an n-bit value is performed in modulo-n.

Vacated bits are filled with 0.

This operation is identical to a logical left shift but is a distinct operation to specify intent.

The amount to shift by is given by an 8-bit unsigned integer, but because CSX64 is little-endian, higher-width values from memory may also be used without having to modify the address.

Usage:

SAL(:size) r, imm/r/m 8

SAL(:size) m, imm/r 8

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

Format:

[8: SAL] [4: dest][2: size][2: mode]

mode = 0: [8: imm]

dest 🡨 dest << imm

mode = 1: [address]

dest 🡨 dest << (8)M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 dest << (8)src

mode2 = 1: [address]

M[address] 🡨 M[address] << (8)src

mode = 3: [8: imm] [address]

M[address] 🡨 M[address] << imm

## SAR – Shift Arithmetic Right

OP Code:

39

Description:

Shifts the destination (first argument) to the right by source (second argument) bits.

Shifting an n-bit value is performed in modulo-n.

Vacated bits are filled with the sign bit.

For non-negative numbers, this operation is equivalent to a logical right shift.

Due to vacated bits being filled in with the sign bit, negative numbers will never reach zero by any amount of arithmetic right shifting.

The amount to shift by is given by an 8-bit unsigned integer, but because CSX64 is little-endian, higher-width values from memory may also be used without having to modify the address.

Usage:

SAR(:size) r, imm/r/m 8

SAR(:size) m, imm/r 8

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

Format:

[8: SAR] [4: dest][2: size][2: mode]

mode = 0: [8: imm]

dest 🡨 dest >> imm

mode = 1: [address]

dest 🡨 dest >> (8)M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 dest >> (8)src

mode2 = 1: [address]

M[address] 🡨 M[address] >> (8)src

mode = 3: [8: imm] [address]

M[address] 🡨 M[address] >> imm

## RL – Rotate Left

OP Code:

40

Description:

Rotates the destination (first argument) to the left by source (second argument) bits.

Shifting an n-bit value is performed in modulo-n.

The amount to shift by is given by an 8-bit unsigned integer, but because CSX64 is little-endian, higher-width values from memory may also be used without having to modify the address.

Usage:

RL(:size) r, imm/r/m 8

RL(:size) m, imm/r 8

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

Format:

[8: RL] [4: dest][2: size][2: mode]

mode = 0: [8: imm]

dest 🡨 dest <<< imm

mode = 1: [address]

dest 🡨 dest <<< (8)M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 dest <<< (8)src

mode2 = 1: [address]

M[address] 🡨 M[address] <<< (8)src

mode = 3: [8: imm] [address]

M[address] 🡨 M[address] <<< imm

## RR – Rotate Right

OP Code:

41

Description:

Rotates the destination (first argument) to the right by source (second argument) bits.

Shifting an n-bit value is performed in modulo-n.

The amount to shift by is given by an 8-bit unsigned integer, but because CSX64 is little-endian, higher-width values from memory may also be used without having to modify the address.

Usage:

RR(:size) r, imm/r/m 8

RR(:size) m, imm/r 8

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

Format:

[8: RR] [4: dest][2: size][2: mode]

mode = 0: [8: imm]

dest 🡨 dest >>> imm

mode = 1: [address]

dest 🡨 dest >>> (8)M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 dest >>> (8)src

mode2 = 1: [address]

M[address] 🡨 M[address] >>> (8)src

mode = 3: [8: imm] [address]

M[address] 🡨 M[address] >>> imm

## AND – Bitwise And

OP Code:

42

Description:

ANDs the destination (first argument) with the source (second argument).

AND can and should be used in place of [TEST](#_TEST) if the result is needed.

Usage:

AND(:size) r, imm/r/m

AND(:size) m, imm/r

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

Format:

[8: AND] [4: dest][2: size][2: mode]

mode = 0: [size: imm]

dest 🡨 dest & imm

mode = 1: [address]

dest 🡨 dest & M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 dest & src

mode2 = 1: [address]

M[address] 🡨 M[address] & src

mode = 3: [size: imm] [address]

M[address] 🡨 M[address] & imm

## OR – Bitwise Or

OP Code:

43

Description:

ORs the destination (first argument) with the source (second argument).

Usage:

OR(:size) r, imm/r/m

OR(:size) m, imm/r

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

Format:

[8: OR] [4: dest][2: size][2: mode]

mode = 0: [size: imm]

dest 🡨 dest | imm

mode = 1: [address]

dest 🡨 dest | M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 dest | src

mode2 = 1: [address]

M[address] 🡨 M[address] | src

mode = 3: [size: imm] [address]

M[address] 🡨 M[address] | imm

## XOR – Bitwise Exclusive Or

OP Code:

44

Description:

XORs the destination (first argument) with the source (second argument).

Usage:

XOR(:size) r, imm/r/m

XOR(:size) m, imm/r

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

Format:

[8: XOR] [4: dest][2: size][2: mode]

mode = 0: [size: imm]

dest 🡨 dest ^ imm

mode = 1: [address]

dest 🡨 dest ^ M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 dest ^ src

mode2 = 1: [address]

M[address] 🡨 M[address] ^ src

mode = 3: [size: imm] [address]

M[address] 🡨 M[address] ^ imm

## CMP – Compare

OP Code:

45

Description:

After this operation, the flags are set in such a way as to satisfy [compound conditionals](#_Flags), where the destination is the left operand and the source is the right operand.

This operation is identical to [SUB](#_SUB), but the result is discarded.

SUB can and should be used in place of CMP if the result is needed.

Usage:

CMP(:size) r, imm/r/m

CMP(:size) m, imm/r

Flags Affected:

As SUB

Format:

As SUB

## TEST – Logical Test

OP Code:

46

Description:

As [AND](#_AND), but the result is discarded.

AND can and should be used in place of TEST if the result is needed.

Usage:

TEST(:size) r, imm/r/m

TEST(:size) m, imm/r

Flags Affected:

As AND

Format:

As AND

## INC – Increment

OP Code:

47

Description:

Equivalent to using [ADD](#_ADD) with a value of 1.

Usage:

INC(:size) r/m

Flags Affected:

Equivalent to using ADD with a value of 1.

Format:

[8: INC] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 dest + 1

mem = 1: [address]

M[address] 🡨 M[address] + 1

## DEC – Decrement

OP Code:

48

Description:

Equivalent to using [SUB](#_SUB) with a value of 1.

Usage:

DEC(:size) r/m

Flags Affected:

Equivalent to using SUB with a value of 1.

Format:

[8: DEC] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 dest - 1

mem = 1: [address]

M[address] 🡨 M[address] - 1

## NEG – Negate

OP Code:

49

Description:

Negates the destination.

Usage:

NEG(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

Format:

[8: NEG] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 -dest

mem = 1: [address]

M[address] 🡨 -M[address]

## NOT – Bitwise Not

OP Code:

50

Description:

NOTs the destination.

Usage:

NOT(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

Format:

[8: NOT] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 ~dest

mem = 1: [address]

M[address] 🡨 ~M[address]

## ABS – Absolute Value

OP Code:

51

Description:

If the destination is negative, the destination is negated. Otherwise no-op.

Because 2's complement always contains exactly one negative number that cannot be expressed as a positive number, the result of this operation may be negative. For more information, see [Signed Integers](#_Signed). SF can be used to test for this occurrence.

Usage:

ABS(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

Format:

[8: ABS] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 abs(dest)

mem = 1: [address]

M[address] 🡨 abs(M[address])

## CMPZ – Compare Zero

OP Code:

52

Description:

Equivalent to using [CMP](#_CMP) with a value of zero.

Usage:

CMPZ(:size) r/m

Flags Affected:

Equivalent to using CMP with a value of zero.

Format:

[8: CMP0] [4: dest][2: size][1:][1: mem]

mem = 0:

cmp dest, 0

mem = 1: [address]

cmp M[address], 0

## LA – Load Address

OP Code:

53

Description:

Assigns to destination (first argument) a memory address (second argument).

Because of the [memory address format](#_Memory_Address_Format), this operation is capable of adding an immediate and two registers together simultaneously, where one register may be negative, and both may be multiplied by small powers of 2. This makes LA a powerful tool for computing integral arithmetic.

Usage:

LA m

Flags Affected:

None.

Format:

[8: LA] [4:][4: dest] [address]

dest 🡨 address

## Jcc – Conditional Jump

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Condition** | **OP Code** | **Condition** | **OP Code** | **Condition** | **OP Code** | **Condition** | **OP Code** |
| None | 54 |  |  | Z | 63 | NZ | 64 |
| A | 55 | G | 59 | S | 65 | NS | 66 |
| AE | 56 | GE | 60 | P | 67 | NP | 68 |
| B | 57 | L | 61 | O | 69 | NO | 70 |
| BE | 58 | LE | 62 | C | 71 | NC | 72 |

Description:

Conditionally jumps execution to the provided address.

These operations form the basis of all conditional branching.

Usage:

Jcc m

Flags Affected:

None.

Format:

[8: Jcc] [address]

if condition:

jmp address

## FADD – Floating-Point Add

OP Code:

73

Description:

Adds the source (second argument) to the destination (first argument).

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FADD(:size) r, imm/r/m

FADD(:size) m, imm/r

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FADD] [4: dest][2: size][2: mode]

mode = 0: [size: imm]

dest 🡨 dest + imm

mode = 1: [address]

dest 🡨 dest + M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 dest + src

mode2 = 1: [address]

M[address] 🡨 M[address] + src

mode = 3: [size: imm] [address]

M[address] 🡨 M[address] + imm

## FSUB – Floating-Point Subtract

OP Code:

74

Description:

Subtracts the source (second argument) from the destination (first argument).

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FSUB(:size) r, imm/r/m

FSUB(:size) m, imm/r

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FSUB] [4: dest][2: size][2: mode]

mode = 0: [size: imm]

dest 🡨 dest - imm

mode = 1: [address]

dest 🡨 dest - M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 dest - src

mode2 = 1: [address]

M[address] 🡨 M[address] - src

mode = 3: [size: imm] [address]

M[address] 🡨 M[address] - imm

## FMUL – Floating-Point Multiply

OP Code:

75

Description:

Multiplies the destination (first argument) by the source (second argument).

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FMUL(:size) r, imm/r/m

FMUL(:size) m, imm/r

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FMUL] [4: dest][2: size][2: mode]

mode = 0: [size: imm]

dest 🡨 dest \* imm

mode = 1: [address]

dest 🡨 dest \* M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 dest \* src

mode2 = 1: [address]

M[address] 🡨 M[address] \* src

mode = 3: [size: imm] [address]

M[address] 🡨 M[address] \* imm

## FDIV – Floating-Point Divide

OP Code:

76

Description:

Divides the destination (first argument) by the source (second argument).

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FDIV(:size) r, imm/r/m

FDIV(:size) m, imm/r

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FDIV] [4: dest][2: size][2: mode]

mode = 0: [size: imm]

dest 🡨 quotient dest / imm

mode = 1: [address]

dest 🡨 quotient dest / M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 quotient dest / src

mode2 = 1: [address]

M[address] 🡨 quotient M[address] / src

mode = 3: [size: imm] [address]

M[address] 🡨 quotient M[address] / imm

## FMOD – Floating-Point Modulus

OP Code:

77

Description:

Divides the destination (first argument) by the source (second argument) and takes the remainder.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

The flags are set in such a way that signed conditionals correctly compare the result to zero.

If the destination or source is negative, the result is platform-dependent (though taking the remainder of negative numbers is virtually never needed).

Usage:

FMOD(:size) r, imm/r/m

FMOD(:size) m, imm/r

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FMOD] [4: dest][2: size][2: mode]

mode = 0: [size: imm]

dest 🡨 remainder dest / imm

mode = 1: [address]

dest 🡨 remainder dest / M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 remainder dest / src

mode2 = 1: [address]

M[address] 🡨 remainder M[address] / src

mode = 3: [size: imm] [address]

M[address] 🡨 remainder M[address] / imm

## FPOW – Floating-Point Power

OP Code:

78

Description:

Computes a floating-point value raised to the power of another floating-point value.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FPOW(:size) r, imm/r/m

FPOW(:size) m, imm/r

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FPOW] [4: dest][2: size][2: mode]

mode = 0: [size: imm]

dest 🡨 pow(dest, imm)

mode = 1: [address]

dest 🡨 pow(dest, M[address])

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 pow(dest, src)

mode2 = 1: [address]

M[address] 🡨 pow(M[address], src)

mode = 3: [size: imm] [address]

M[address] 🡨 pow(M[address], imm)

## FSQRT – Floating-Point Square Root

OP Code:

79

Description:

Computes the square root of a floating-point value.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FSQRT(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FSQRT] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 sqrt(dest)

mem = 1: [address]

M[address] 🡨 sqrt(M[address])

## FEXP – Floating-Point Exponentiate

OP Code:

80

Description:

Computes e raised to a floating-point value.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FEXP(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FEXP] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 exp(dest)

mem = 1: [address]

M[address] 🡨 exp(M[address])

## FLN – Floating-Point Natural Logarithm

OP Code:

81

Description:

Computes the natural logarithm (base e) of a floating-point value.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FLN(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FLN] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 ln(dest)

mem = 1: [address]

M[address] 🡨 ln(M[address])

## FNEG – Floating-Point Negate

OP Code:

82

Description:

Computes the negative of a floating-point value.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FNEG(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FNEG] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 -dest

mem = 1: [address]

M[address] 🡨 -M[address]

## FABS – Floating-Point Absolute Value

OP Code:

83

Description:

Computes the absolute value of a floating-point value.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

Specifying sizes other than 64 or 32 bits results in an [Undefined Behavior Error](#_Heap).

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FABS(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FABS] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 abs(dest)

mem = 1: [address]

M[address] 🡨 abs(M[address])

## FCMPZ – Floating-Point Compare Zero

OP Code:

84

Description:

Equivalent to using [FCMP](#_FCMP) with a floating-point value of zero.

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FCMPZ(:size) r/m

Flags Affected:

Equivalent to using FCMP with a floating-point value of zero.

Format:

[8: FCMPZ] [4: dest][2: size][1:][1: mem]

mem = 0:

fcmp dest, 0.0

mem = 1: [address]

fcmp M[address], 0.0

## FSIN – Floating-Point Sine

OP Code:

85

Description:

Computes the sine of a floating-point value.

Angles are in radians.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

Specifying sizes other than 64 or 32 bits results in an [Undefined Behavior Error](#_Heap).

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FSIN(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FSIN] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 sin(dest)

mem = 1: [address]

M[address] 🡨 sin(M[address])

## FCOS – Floating-Point Cosine

OP Code:

86

Description:

Computes the cosine of a floating-point value.

Angles are in radians.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

Specifying sizes other than 64 or 32 bits results in an [Undefined Behavior Error](#_Heap).

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FCOS(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FCOS] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 cos(dest)

mem = 1: [address]

M[address] 🡨 cos(M[address])

## FTAN – Floating-Point Tangent

OP Code:

87

Description:

Computes the tangent of a floating-point value.

Angles are in radians.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

Specifying sizes other than 64 or 32 bits results in an [Undefined Behavior Error](#_Heap).

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FTAN(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FTAN] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 tan(dest)

mem = 1: [address]

M[address] 🡨 tan(M[address])

## FSINH – Floating-Point Hyperbolic Sine

OP Code:

88

Description:

Computes the hyperbolic sine of a floating-point value.

Angles are in radians.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

Specifying sizes other than 64 or 32 bits results in an [Undefined Behavior Error](#_Heap).

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FSINH(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FSINH] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 sinh(dest)

mem = 1: [address]

M[address] 🡨 sinh(M[address])

## FCOSH – Floating-Point Hyperbolic Cosine

OP Code:

89

Description:

Computes the hyperbolic cosine of a floating-point value.

Angles are in radians.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

Specifying sizes other than 64 or 32 bits results in an [Undefined Behavior Error](#_Heap).

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FCOSH(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FCOSH] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 cosh(dest)

mem = 1: [address]

M[address] 🡨 cosh(M[address])

## FTANH – Floating-Point Hyperbolic Tangent

OP Code:

90

Description:

Computes the hyperbolic tangent of a floating-point value.

Angles are in radians.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

Specifying sizes other than 64 or 32 bits results in an [Undefined Behavior Error](#_Heap).

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FTANH(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FTANH] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 tanh(dest)

mem = 1: [address]

M[address] 🡨 tanh(M[address])

## FASIN – Floating-Point Arcsine

OP Code:

91

Description:

Computes the inverse sine of a floating-point value.

Angles are in radians.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

Specifying sizes other than 64 or 32 bits results in an [Undefined Behavior Error](#_Heap).

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FASIN(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FASIN] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 asin(dest)

mem = 1: [address]

M[address] 🡨 asin(M[address])

## FACOS – Floating-Point Arccosine

OP Code:

92

Description:

Computes the inverse cosine of a floating-point value.

Angles are in radians.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

Specifying sizes other than 64 or 32 bits results in an [Undefined Behavior Error](#_Heap).

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FACOS(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FACOS] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 acos(dest)

mem = 1: [address]

M[address] 🡨 acos(M[address])

## FATAN – Floating-Point Arctangent

OP Code:

93

Description:

Computes the inverse tangent of a floating-point value.

Angles are in radians.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

Specifying sizes other than 64 or 32 bits results in an [Undefined Behavior Error](#_Heap).

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FATAN(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FATAN] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 atan(dest)

mem = 1: [address]

M[address] 🡨 atan(M[address])

## FATAN2 – Floating-Point Binary Arctangent

OP Code:

94

Description:

Computes the arctangent of two floating point values, where y = dest and x = val.

Angles are in radians.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

Specifying sizes other than 64 or 32 bits results in an [Undefined Behavior Error](#_Heap).

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FATAN2(:size) r, imm/r/m

FATAN2(:size) m, imm/r

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FATAN2] [4: dest][2: size][2: mode]

mode = 0: [size: imm]

dest 🡨 atan2 (dest, imm)

mode = 1: [address]

dest 🡨 atan2 (dest, M[address])

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 atan2 (dest, src)

mode2 = 1: [address]

M[address] 🡨 atan2 (M[address], src)

mode = 3: [size: imm] [address]

M[address] 🡨 atan2(M[address], imm)

## FLOOR – Floor

OP Code:

95

Description:

Computes the floor of a floating-point value.

Result is still floating point.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

Specifying sizes other than 64 or 32 bits results in an [Undefined Behavior Error](#_Heap).

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FLOOR(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: FLOOR] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 floor(dest)

mem = 1: [address]

M[address] 🡨 floor(M[address])

## CEIL – Ceiling

OP Code:

96

Description:

Computes the ceiling of a floating-point value.

Result is still floating point.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

Specifying sizes other than 64 or 32 bits results in an [Undefined Behavior Error](#_Heap).

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

CEIL(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: CEIL] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 ceil(dest)

mem = 1: [address]

M[address] 🡨 ceil(M[address])

## ROUND – Round

OP Code:

97

Description:

Computes the result of rounding a floating-point value to the nearest integer.

Result is still floating point.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

Specifying sizes other than 64 or 32 bits results in an [Undefined Behavior Error](#_Heap).

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

ROUND(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: ROUND] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 round(dest)

mem = 1: [address]

M[address] 🡨 round(M[address])

## TRUNC – Truncate

OP Code:

98

Description:

Rounds a floating-point value towards zero by removing the fractional portion.

Result is still floating point.

May produce special values based on inputs. Refer to [64-bit Floating-Point Numbers](#_64-bit) for more information.

Specifying sizes other than 64 or 32 bits results in an [Undefined Behavior Error](#_Heap).

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

TRUNC(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative, and cleared otherwise.

OF is cleared.

CF is set if the result is positive or negative infinity, and cleared otherwise.

PF is set if the result is NaN, and cleared otherwise.

Format:

[8: TRUNC] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 trunc(dest)

mem = 1: [address]

M[address] 🡨 trunc(M[address])

## FCMP – Floating-Point Compare

OP Code:

99

Description:

As [FSUB](#_FSUB), but the result is discarded.

The flags are set in such a way that signed conditionals correctly compare the result to zero.

Usage:

FCMP(:size) r, imm/r/m

FCMP(:size) m, imm/r

Flags Affected:

As FSUB.

Format:

As FSUB.

## FTOI – Floating-Point to Integer

OP Code:

100

Description:

Converts a floating-point value into a signed integer via truncation.

Result is integral.

Specifying sizes other than 64 or 32 bits results in an [Undefined Behavior Error](#_Heap).

If the floating-point value cannot be represented as a signed integer of the specified width after truncation (which includes large-magnitude numbers, both infinities, and all NaN values), the result is 10…0 in binary (i.e. the minimum value for a signed integer of that width). Thus, the result of FTOI can be compared to this value to test for a conversion failure (except in the single case where the conversion would correctly yield this value, though this chance is insignificant for most applications.

Usage:

FTOI(:size) r/m

Flags Affected:

None

Format:

[8: FTOI] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 int(dest)

mem = 1: [address]

M[address] 🡨 int(M[address])

## ITOF – Integer to Floating-Point

OP Code:

101

Description:

Converts a signed integer value into a floating-point value.

Result is floating-point.

Specifying sizes other than 64 or 32 bits results in an [Undefined Behavior Error](#_Heap).

Usage:

ITOF(:size) r/m

Flags Affected:

None

Format:

[8: ITOF] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 float(dest)

mem = 1: [address]

M[address] 🡨 float(M[address])

## PUSH – Push

OP Code:

102

Description:

Pushes a size byte value onto [the stack](#_The_Stack).

Usage:

PUSH(:size) imm/r

Flags Affected:

None

Format:

[8: PUSH] [4: src][2: size][1:][1: reg]

reg = 0: [size: imm]

push imm

reg = 1:

push src

## POP – Pop

OP Code:

103

Description:

Removes a size byte value from [the stack](#_Stack).

Usage:

POP(:size) r

Flags Affected:

None.

Format:

[8: POP] [4: dest][2: size][2:]

pop value

dest 🡨 value

## CALL – Call

OP Code:

104

Description:

Pushes the 64-bit address of the next instruction onto the stack and jumps execution to the specified address.

Usage:

CALL m

Flags Affected:

None.

Format:

[8: CALL] [address]

push next ins address

jmp address

## RET – Return

OP Code:

105

Description:

Removes a 64-bit value from the stack and jumps to that address.

Usage:

RET

Flags Affected:

None.

Format:

[8: RET]

pop address

jmp address

## BSWAP – Byte Swap

OP Code:

106

Description:

Reverses the byte order of the specified register, effectively translating between big/little endian. Using this operation on an 8-bit register is no-op.

Usage:

BSWAP(:size) r/m

Flags Affected:

None.

Format:

[8: BSWAP] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 reverse(dest)

mem = 1: [address]

M[address] 🡨 reverse(M[address])

## BEXTR – Bitfield Extract

OP Code:

107

Description:

Extracts a contiguous series of bits from a value. The size parameter determines the size of the bitfield source. The bitfield extracted is determined by an additional 16-bit argument, where the high 8 bits determine the position of the lowest bit of the bitfield (zero being the low bit of the source), and the low 8 bits determine the length (in bits) of the bitfield.

Both the position and the size of the bitfield are performed in modulo-n, where n is the size of the source (in bits).

Attempting to extract bits beyond the size of the source is not an error, and the overflowing section is ignored.

Usage:

BEXTR(:size) r, imm/r/m

BEXTR(:size) m, imm/r

Flags Affected:

None.

Format:

[8: BEXTR] [4: dest][2: size][2: mode]

mode = 0: [16: imm]

dest 🡨 extract(dest, imm)

mode = 1: [address]

dest 🡨 extract(dest, (16)M[address])

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 extract(dest, (16)src)

mode2 = 1: [address]

M[address] 🡨 extract(M[address], (16)src)

mode = 3: [16: imm] [address]

M[address] 🡨 extract(M[address], imm)

## BLSI – Binary Lowest-Set Isolate

OP Code:

108

Description:

Isolates the lowest-order set bit, leaving it in its current position. All other bits are cleared. If the source contains no set bits, the result is zero.

Usage:

BLSI(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

Format:

[8: BLSI] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 dest & -dest

mem = 1: [address]

M[address] 🡨 M[address] & -M[address]

## BLSMSK – Binary Lowest-Set Mask

OP Code:

109

Description:

Extracts a bitmask containing only the low-order bits up to and including the lowest-set bit. If the source is zero (i.e. no set bits), the result is all 1’s.

Usage:

BLSMSK(:size) r/m

Flags Affected:

None.

Format:

[8: BLSMSK] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 dest ^ (dest – 1)

mem = 1: [address]

M[address] 🡨 M[address] ^ (M[address] – 1)

## BLSR – Binary Lowest-Set Reset

OP Code:

110

Description:

Clears the lowest-order set bit of the source. If source contains no set bits, the result is zero.

Usage:

BLSR(:size) r/m

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

Format:

[8: BLSR] [4: dest][2: size][1:][1: mem]

mem = 0:

dest 🡨 dest & (dest – 1)

mem = 1: [address]

M[address] 🡨 M[address] & (M[address] – 1)

## ANDN – Bitwise And Not

OP Code:

111

Description:

Computes the bitwise and of two unsigned integers, with the second being inverted before use.

Usage:

ANDN(:size) r, imm/r/m

ANDN(:size) m, imm/r

Flags Affected:

ZF is set if the result is zero, and cleared otherwise.

SF is set if the result is negative (high bit set), and cleared otherwise.

PF is set if the result has even parity in the low 8 bits, and cleared otherwise.

Format:

[8: ANDN] [4: dest][2: size][2: mode]

mode = 0: [size: imm]

dest 🡨 dest & ~imm

mode = 1: [address]

dest 🡨 dest & ~M[address]

mode = 2: [3:][1: mode2][4: src]

mode2 = 0:

dest 🡨 dest & ~src

mode2 = 1: [address]

M[address] 🡨 M[address] & ~src

mode = 3: [size: imm] [address]

M[address] 🡨 M[address] & ~imm

## GETF – Get Flags

OP Code:

112

Description:

Loads the current flags data to the specified 64-bit register.

Usage:

GETF r

Flags Affected:

None.

Format:

[8: GETF] [4:][4: dest]

dest 🡨 flags

## SETF – Set Flags

OP Code:

113

Description:

Loads the flags register from the specified 64-bit register. Only the [public flags](#_Conditions) will be modified by this operation.

Usage:

SETF r

Flags Affected:

None

Format:

[8: SETF] [4:][4: src]

flags 🡨 src

## LOOP – Loop

OP Code:

114

Description:

Decrements the specified register (by 1, 2, 4, or 8) and jumps to the provided address if the result is non-zero. Used for simple count-based iteration.

Usage:

LOOP(:size) r, m, (imm size\_step)

Flags Affected:

None.

Format:

[8: LOOP] [4: reg][2: size][2: size\_step] [address]

reg 🡨 reg – size(size\_step)

if reg != 0:

jmp address

## FX – Floating-Point Extend

OP Code:

115

Description:

Extends (or contracts) a floating-point value.

Specifying sizes other than 64 or 32 bits results in an [Undefined Behavior Error](#_Heap).

Usage:

FX(:to\_size) imm from\_size, r

Flags Affected:

None.

Format:

[8: FX] [4: reg][2: from\_size][2: to\_size]

(to\_size) reg 🡨 floating extend (from\_size) reg

## SLP – Sleep

OP Code:

116

Description:

Causes the processor to enter a sleep mode, which in real processors dramatically cuts power usage and thus heat production. SLEEP takes an unsigned integer and skips that many operation cycles as no-op.

Usage:

SLP(:size) imm/r/m

Flags Affected:

None

Format:

[8: SLP] [4: reg][2: size][2: mode]

mode = 0: [size: imm]

mode = 1: use reg

mode = 2: [address]

# Virtual Operating System

The virtual operating system is the “platform” on which the CSX64 processor runs. In essence, it acts as an extension mechanism for the default processor behavior. Specification-compliant virtual operating systems are effectively dormant until called upon by the execution of a [SYSCALL](#_SYSCALL_–_System) operation.

By default, CSX64 offers only an empty virtual operating system, containing only the bare essentials for the processor to function.

A custom virtual operating system is a program written in C# (or any compatible language) that creates a class that inherits the base virtual operating system, overriding the Syscall() function to wrap the desired effects.

## Default CSX64 Specification

# Appendix

## Converting Binary Integers

Because all modern computing is built upon the foundation of binary integers, you’ll need to know how to encode them for the processor to understand. Whereas decimal is base 10 (and each digit position is 10 times bigger than the last), binary is base 2 (and each digit position is 2 times bigger than the last):

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Decimal** | | | | **Binary** | | | |
| 103 | 102 | 101 | 100 | 23 | 22 | 21 | 20 |
| 1000 | 100 | 10 | 1 | 8 | 4 | 2 | 1 |

To get the value of a decimal number, you take each digit and multiply by its weight (position), then add them together.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **410710** | | | **10112** | | |
| **Digit** | **Weight** | **Digit \* Weight** | **Digit** | **Weight** | **Digit \* Weight** |
| 4 | 1000 | 4000 | 1 | 8 | 8 |
| 1 | 100 | 100 | 0 | 4 | 0 |
| 0 | 10 | 0 | 1 | 2 | 2 |
| 7 | 1 | 7 | 1 | 1 | 1 |
|  |  | **Sum: 4107** |  |  | **Sum: 11** |

So now you can turn a binary number into decimal. A neat trick, but if you really want to get into the nitty-gritty of computers, you’ll need to be able to do the reverse (and get pretty fast at it). An easy method to do this is to divide by 2 and take the remainder, which will always be either a 0 or a 1. These remainders, when read backwards, will be the binary value:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **75310** | | | **17210** | | |
| **Value** | **Quotient** | **Remainder** | **Value** | **Quotient** | **Remainder** |
| 753 / 2 | 376 | 1 | 172 / 2 | 86 | 0 |
| 376 / 2 | 188 | 0 | 86 / 2 | 43 | 0 |
| 188 / 2 | 94 | 0 | 43 / 2 | 21 | 1 |
| 94 / 2 | 47 | 0 | 21 / 2 | 10 | 1 |
| 47 / 2 | 23 | 1 | 10 / 2 | 5 | 0 |
| 23 / 2 | 11 | 1 | 5 / 2 | 2 | 1 |
| 11 / 2 | 5 | 1 | 2 / 2 | 1 | 0 |
| 5 / 2 | 2 | 1 | 1 / 2 | 0 | 1 |
| 2 / 2 | 1 | 0 |  |  |  |
| 1 / 2 | 0 | 1 |  |  |  |
|  | **Result:** | **10 1111 00012** |  | **Result:** | **1010 11002** |

Alternatively, you could subtract powers of 2, keeping track of which ones fit. The moral of the story is practice, practice, practice. Knowing your powers of 2 is absolutely essential, especially in machine language and assembly.

## Proof of 2’s Complement

The seemingly magical properties of 2’s complement can be a bit hard to understand at first. For instance, why would signed and unsigned addition and subtraction be the same when using 2’s complement to represent signed integers? What genius madman, in a fevered binary dream of passion decided to apply bitwise not and incrementation?

Well, you’ll be happy to know that this property of working the same for both signed and unsigned addition (and subtraction, since that’s the same thing as adding a negative and we’re deriving a scheme for taking the negative) was actually the whole idea. Indeed, 2’s complement was defined to work with unsigned addition (and subtraction). It was no happy coincidence. Observe:

|  |  |
| --- | --- |
| v + -v = 0 | We begin with a simple question: v + what = 0 under unsigned addition |
| v + ~v + 1 = 0 | Well, v + ~v would be all 1’s, and adding 1 would overflow back to 0 |
| v + -v = v + ~v + 1 | Since these both equal 0, we can equate them |
| v + -v + -v = v + ~v + 1 + -v | We can then add -v to both sides |
| -v = ~v + 1 | From the first equation: v + -v = 0, and we have our answer: -v = ~v + 1 |

Consider this the computer science equivalent of being told there is no Santa Claus. It was algebra all along.

## The Heap

In some cases, it is necessary to create a new variable in memory that will outlast its position in the stack. For instance, you may need a function to allocate memory to represent a customer. Now, with only a standard stack this is an impossible task. If the function creates the variable on the stack, pushing and popping will destroy it. If it puts in in static variables or a pre-allocated array, there can only be as many as you created “slots” to put them in. But what if you don’t know how many “slots” to make?

This is the question that is answered by what’s known as dynamic memory allocation. Essentially, we create some pool of memory that can be allocated as needed, typically by a function call. The important thing is that the memory is not in static variables or on the stack. It is somewhere else: the heap. Typically, any modern language would provide a library function to do this, but you can absolutely make your own (potentially specialized) algorithm to do this.

Historically (and carried into modern systems), the heap begins just after the program segment and grows upwards similarly to a stack. This means that the stack and heap are on opposite sides of address space and grow towards one another. This means that either one of them can grow to fill potentially all of the provided space, so long as they don’t cross one another.

## Encoding IEEE 754 64-bit Floating-Point Values