

STONE CUTTER

StoneCutter Language Specification

Tactical Computing Laboratories, LLC

Date: February 15, 2019
Revision: 0.1
Authors: Tactical Computing Labs
contact@tactcomplabs.com

Contents

1	Overview	6
2	StoneCutter Language	7
2.1	Overview	7
2.2	Syntactical Notes	8
2.3	Comments	9
2.4	Datatypes	10
2.5	Instruction Format Definitions	11
2.6	Register Class Definitions	12
2.7	Instruction Prototypes	13
2.8	Variable Definitions	15
2.9	Arithmetic Operations	16
2.10	Conditional Operations	17
2.10.1	Boolean Operators	17
2.10.2	Flow Control	18
2.11	Loop Operations	19
2.11.1	For Loops	19
2.11.2	While Loops	20
2.11.3	Do While Loops	21
3	Intrinsic Functions	22
3.1	Overview	22
3.2	Arithmetic Ininsics	23
3.2.1	CLZ	23
3.2.2	COMPRESS	24
3.2.3	COMPRESSM	25
3.2.4	CTZ	26
3.2.5	DOZ	27
3.2.6	EXTRACTS	28
3.2.7	EXTRACTZ	29
3.2.8	INSERTS	30
3.2.9	INSERTZ	31
3.2.10	MAJ	32
3.2.11	MAX	33
3.2.12	MERGE	34
3.2.13	MIN	35
3.2.14	NOT	36
3.2.15	POPCOUNT	37
3.2.16	REVERSE	38
3.2.17	ROTL	39
3.2.18	ROTR	40
3.2.19	SEXT	41
3.2.20	ZEXT	42
3.3	Memory Ininsics	43
3.3.1	LOAD	43
3.3.2	STORE	44
3.3.3	LOADELEM	45
3.3.4	STOREELEM	46
4	Appendix A: Sample StoneCutter Implementation	47
5	Appendix B: Intrinsic Function Table	53

List of Figures

1	StoneCutter Architecture	7
---	--------------------------	---

Listings

1	StoneCutter File Structure	8
2	StoneCutter Comments	9
3	Arbitrary Width Datatypes	10
4	Instruction Format Definition Syntax	11
5	Sample Instruction Format Definition	11
6	Register Class Definition Syntax	12
7	Sample Register Class Definition	12
8	Instruction Prototype Format	13
9	Instruction Prototype Format	14
10	Variable Definitions	15
11	Sample Variable Definitions	15
12	Flow Control Syntax	18
13	Sample If-Else Syntax	18
14	For Loop Syntax	19
15	Sample For Loop Syntax	19
16	While Loop Syntax	20
17	Sample While Loop Syntax	20
18	Do While Loop Syntax	21
19	Sample Do While Loop Syntax	21
20	Intrinsic Syntax	22
21	Sample StoneCutter	47

List of Tables

2	StoneCutter Datatypes	10
3	StoneCutter Instruction Format Field Types	11
4	StoneCutter Arithmetic Operations	16
5	StoneCutter Boolean Operations	17
6	StoneCutter Comparison Operations	19
7	CLZ Intrinsic	23
8	COMPRESS Intrinsic	24
9	COMPRESSM Intrinsic	25
10	CTZ Intrinsic	26
11	DOZ Intrinsic	27
12	EXTRACTS Intrinsic	28
13	EXTRACTZ Intrinsic	29
14	INSERTS Intrinsic	30
15	INSERTZ Intrinsic	31
16	MAJ Intrinsic	32
17	MAX Intrinsic	33
18	MERGE Intrinsic	34
19	MIN Intrinsic	35
20	NOT Intrinsic	36
21	POPCOUNT Intrinsic	37
22	REVERSE Intrinsic	38

23	ROTL Intrinsic	39
24	ROTR Intrinsic	40
25	SEXT Intrinsic	41
26	ZEXT Intrinsic	42
27	LOAD Intrinsic	43
28	STORE Intrinsic	44
29	LOADELEM Intrinsic	45
30	STOREELEM Intrinsic	46
31	StoneCutter Intrinsic	53

DRAFT

Revision History

Revision	Date	Author(s)	Description
0.1	02.15.2018	JLeidel	Initial public release

DRAFT

1 Overview

The StoneCutter language is utilized to develop the implementation of a single instruction within the System Architect design workflow. The StoneCutter language abstracts a large portion of traditional high level design languages such that the user may focus on the implementation details of a single instruction rather than the connectivity to the remainder of the infrastructure. The core features of the StoneCutter language infrastructure are noted as follows:

- **Compiled Language:** Unlike other HDL approaches, the StoneCutter language is, in fact, compiled. The core StoneCutter compiler infrastructure makes use of the LLVM compiler infrastructure for lexing, parsing, optimization and code generation. As a result, we have the ability to initiate traditional optimizing compiler passes, syntax tests and lexical analysis as other compiled languages. Further, we have the ability to craft tooling that integrates with the StoneCutter language in the same manner as traditional compiled languages.
- **Integration with CoreGen IR:** The StoneCutter language and associated tooling is architected in a manner that permits integration with the CoreGen intermediate representation (IR) [1]. This is done in two ways. First, users have the ability to write StoneCutter instruction definitions inline within the CoreGen IR. These `Impl` definitions are directly embedded within the overarching design. Second, the CoreGen IR can be utilized to verify the I/O architecture and instruction format for each StoneCutter instruction definition. This ensures that the prescribed instruction format in the CoreGen design is *verified* to be functionally correct prior to utilizing downstream synthesis tools.
- **C-Like Syntax:** Unlike other HDL approaches, the StoneCutter language utilizes a familiar syntactical structure that is designed to mimic traditional C procedural methods. Each instruction definition is contained within an effective function body with incoming arguments (registers). Arithmetic, boolean operations, conditional operations and loop structures all mimic traditional procedural C syntax. In this manner, the learning curve required to be productive with StoneCutter is minimized.
- **Support for Intrinsic:** Much in the same manner as traditional procedural languages such as C, StoneCutter has support for inline intrinsic operations. Intrinsic operations are designed to support optimized circuits for pathological operations such as sign/zero extension, multi-input operations (majority vote, etc) and special arithmetic operations. As the StoneCutter language continues to develop, we plan to augment the list of supported intrinsics.

The remainder of this language specification is organized as follows. Section 2 introduces the StoneCutter language and the associated syntax. Section 3 provides details associated with each of the currently supported intrinsics. Section 4 provides a sample set of StoneCutter instruction definitions. Section 5 provides a consolidated list of StoneCutter intrinsics.

2 StoneCutter Language

2.1 Overview

As mentioned above, the StoneCutter language is designed to take instruction definitions in a procedural syntax, optimize the instruction body and output each instruction in Chisel HDL. As shown in Figure 1, input files or buffers are parsed into an appropriate AST. Instruction intrinsics are recognized and annotated in the AST. Further, register classes and registers are identified and marked as special globals. The AST is then translated to standard LLVM intermediate representation (IR). The tools utilize standard LLVM optimization passes to optimize the instruction IR. The tools then execute a series of StoneCutter-specific IR passes against the optimized LLVM IR in order to ensure that the instruction design is correct per the prescribed instruction format(s). Once verified as correct, the instruction intrinsics are expanded inline and the tools generate Chisel HDL.

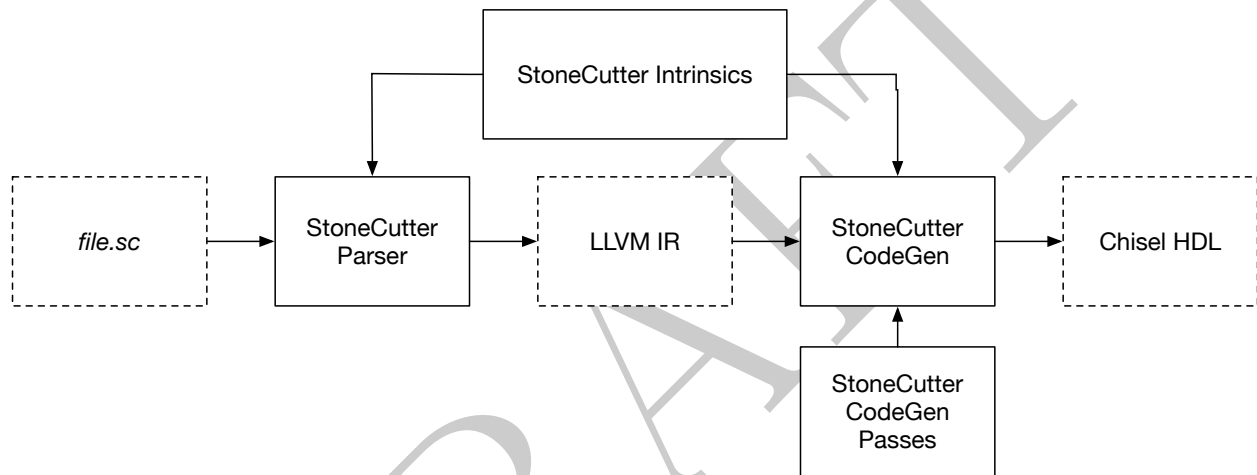


Figure 1: StoneCutter Architecture

Each StoneCutter input file is constructed in a specific, prescribed manner. The organization of the data members and the instruction definitions is done so specifically to elicit global and local hardware state within and across individual instructions. We find a sample of the StoneCutter file layout in Listing 1. StoneCutter input files begin with an optional comment block describing the contents of the file. Next, the individual register classes are defined for use by the instructions. These register classes should match the register class definitions in name and layout found in the respective CoreGen design input. Further information regarding the definition of register classes can be found in Section 2.6. The next portion of the file includes the actual instruction definitions. Each instruction definition is comprised of a *function*-like structure denoted with a `def` keyword. Each instruction takes some number of instruction arguments (I/O's) and contains private variables and arithmetic logic. Notice how the variable definitions are made near the top of the instruction definition. Much like the C language, this is required in StoneCutter in order to preemptively understand the scope of the temporary variables defined in the instruction. Sections 2.7-2.11 describe the rudimentary operations and the associated syntax that can be included in the instruction definition body.

```

1 #
2 # Sample StoneCutter File Layout
3 #
4
5 # Instruction Formats
6 instformat FORMAT1( FIELDTYPE FIELD1, FIELDTYPE FIELD2, ... )
7
8 # Register Class Definitions
9 regclass CLASS1( DATATYPE REG1, DATATYPE REG2, ... )
10 regclass CLASS2( DATATYPE REG3( DATATYPE SUBREG1, DATATYPE SUBREG2 ) )
11
12 # Instruction Definitions
13 def INST1( INPUT1 INPUT2 INPUT3 ) {
14     # Variable Definitions
15     DATATYPE VAR1, VAR2, VAR3
16     DATATYPE VAR4
17
18     # Instruction Body
19     VAR4 = INPUT2 + INPUT1
20     VAR1 = INPUT1 << 1
21     VAR2 = INPUT3 << 2
22     VAR3 = INPUT1 + INPUT2
23     INPUT1 = VAR3
24 }
25
26 def INST2( REG1 REG3 SUBREG1 ) {
27     # Instruction Body
28     REG1 = REG3 * SUBREG1
29 }

```

Listing 1: StoneCutter File Structure

2.2 Syntactical Notes

Prior to reading the remaining sections, we highly suggest users and readers understand the following syntactical notes. Understanding these notes will significantly reduce the time required to become productive in StoneCutter.

- **Semicolons:** Unlike the C language, StoneCutter does not require utilizing semicolons (;) to terminate an expression. All raw expressions are in static single assignment (SSA) form. This implies that each expression will be in the form of `TARGET = INPUT <OP> OUTPUT`. The only exception to this rule is when intrinsics are utilized. A call to an intrinsic function may be
- **Complement Operations:** The StoneCutter language contains all the standard arithmetic, boolean and logical operations except complements. The bitwise complement operator from C (~) is not supported and the boolean complement operator from C (!) is not supported. Complementing the result of a boolean operation and/or performing the bitwise complement of a variable or register must be performed using the NOT intrinsic. See Section 3.2.14 for more details.

2.3 Comments

Inline comments and comment text in StoneCutter must begin with the pound (#) sign. Comments may begin on new lines or inline with other code. Examples of using comments are shown in Listing 2.

```
1 # This is a stand alone comment
2 def inst0(RA RB RT) { # this is an inline comment
3     RT = RA + RB
4 }
```

Listing 2: StoneCutter Comments

DRAFT

2.4 Datatypes

Much in the same manner as traditional programming models such as C and C++, StoneCutter supports a common set of datatypes for intermediate variables and registers. However, unlike traditional programming models, hardware design languages are required to support datatypes in non-byte aligned types. In order to provide more hardware-centric support, StoneCutter supports traditional data types as well as arbitrary width signed and unsigned integer types. These types are documented in Table 2.

Table 2: StoneCutter Datatypes

Type	Width (in bits)	Description
bool	1	Boolean. Analogous to unsigned 1 bit integer (u1)
u8	8	Unsigned 8 bit integer. Analogous to uint8_t
u16	16	Unsigned 16 bit integer. Analogous to uint16_t
u32	32	Unsigned 32 bit integer. Analogous to uint32_t
u64	64	Unsigned 64 bit integer. Analogous to uint64_t
s8	8	Signed 8 bit integer. Analogous to int8_t
s16	16	Signed 16 bit integer. Analogous to int16_t
s32	32	Signed 32 bit integer. Analogous to int32_t
s64	64	Signed 64 bit integer. Analogous to int64_t
float	32	Single precision floating point
double	64	Double precision floating point
uN	N bits	Arbitrary unsigned integer of N bits
sN	N bits	Arbitrary signed integer of N bits

An example of defining arbitrary width integers is shown in Listing 3.

```

1  #-- unsigned 7 bit integer "foo"
2  u7 foo
3
4  #-- unsigned 1024 integer "bar"
5  u1024 bar
6
7  #-- signed 37 integer "foobar"
8  s37 foobar

```

Listing 3: Arbitrary Width Datatypes

2.5 Instruction Format Definitions

Instruction formats and their associated fields are special variables in the StoneCutter language. Each defined instruction field must reside within a respective instruction format. Instruction fields are marked in the IR with attributes in order to ensure that they are correctly associated with the correct instruction format. Further, for each field that is denoted as a register field, the associated register class must also be specified in order to correctly link the register read datapath to the correct register file. Each of the associated fields is annotated as a global variable such that it can be explicitly utilized in any instruction prototype or instruction body.

Table 3: StoneCutter Instruction Format Field Types

Field	Mnemonic	Read/Write Access	Description
Instruction Encodings	enc	Read-Only	Instruction encodings such as opcodes and function codes
Immediate Values	imm	Read-Only	Immediate values encoded in the instruction payload
Register Indices	reg[REGCLASS]	Read-Write	Register index encodings. Must include the respective REGCLASS

For each field in the instruction format, the respective field is designated with a field type and a field name. There are three types of instruction field types. First, instruction fields denoted as encoding fields are utilized to designate individual instruction encodings such as opcodes and function codes in RISC architectures. When utilizing these fields in the implementation of an instruction, the value contained within the field is utilized. For example, if an instruction defines a field `opc` that contains the values `0x0A` hex, reading from this field will provide the exact value `0x0A` hex. Second, immediate fields are immediate values encoded directly within the instruction payload. Reading from immediate fields from within the instruction payload will read the literal immediate from the instruction payload. Finally, register fields denote register indices that access a single register file. Reading from this field directly will not deliver the index. Rather, accessing a register field will read or write the value at the designated index of the target register file. For fields that are denoted as register class fields (`reg`), the associated register class (register file) must also be specified. We summarize the permissible instruction field types in Table 3. Further, the syntax for the an entire instruction format block is shown in Listing 4. Encoding and immediate fields are always marked as read-only. Only register fields may be written to.

```
1 instformat FName1( FIELDTYPE FIELD1, FIELDTYPE FIELD2, ... )
2 instformat FName2( reg[REGCLASS] FIELD1, ... )
```

Listing 4: Instruction Format Definition Syntax

We find an example instruction format definition for a simple RISC ISA with encodings for an opcode and function code as well as three register fields in Listing 5.

```
1 instformat RISC( enc opc, enc func, reg[GPR] RT, reg[GPR] RA, reg[GPR] RB )
```

Listing 5: Sample Instruction Format Definition

2.6 Register Class Definitions

Register classes and their associated registers are special variables in the StoneCutter language. Each defined register must reside within its respective register class. Registers are marked in the IR with attributes in order to ensure that they belong to the correct register class when being utilized in an instruction body. Each register is also annotated as a global variable such that it can be explicitly utilized in any instruction definition body.

```
1 regclass RCName1( DATATYPE RegName1, DATATYPE RegName2, ... )
2 regclass RCName2( DATATYPE RegName( DATATYPE SubReg1, DATATYPE SubReg2 ) )
```

Listing 6: Register Class Definition Syntax

We find the syntax for defining register classes and their associated registers in Listing 6. Each register class definition is marked with a `regclass` keyword followed by the register class name. Each of the registers defined within the register class are enclosed within parenthesis. Within the parenthesis, each register definition must be preceded by its respective datatype (Section 2.4) and the register name. Registers are separated by commas. For registers that have subfields, we may also explicitly annotate these within the register definition. For each register with subfields, the subfields are defined within parenthesis attached to the register definition (Line 2 in Listing 6). Each subfield must also contain its respective datatype. The combined number of bits for all subfields within a register definition must not exceed the total number of bits in that register.

We find an example register class definition for a simple RISC ISA in Listing 7.

```
1 regclass GPR( u64 r0, u64 r1, u64 r2, u64 r3, u64 r4, u64 r5, u64 r6,
2             u64 r7, u64 r8, u64 r9, u64 r10, u64 r11, u64 r12,
3             u64 r13, u64 r14, u64 r15, u64 r16, u64 r17, u64 r18,
4             u64 r19, u64 r20, u64 r21, u64 r22, u64 r23, u64 r24,
5             u64 r25, u64 r26, u64 r27, u64 r28, u64 r29, u64 r30,
6             r64 r31 )
7 regclass CTRL( u64 pc, u64 exc, u64 ne, u64 gt, u64 lt, u64 gte,
8              u64 lte, u64 sp, u64 fp, u64 rp )
```

Listing 7: Sample Register Class Definition

2.7 Instruction Prototypes

For each instruction defined in the StoneCutter syntax, we must construct an instruction prototype. The instruction prototype is an important and powerful feature of the StoneCutter language. The instruction prototype drives two major features in the downstream Chisel HDL produced by a compiled StoneCutter input. First, the prototype describes the attachment of the respective instruction implementation to the instruction crack logic generated by the CoreGen IR. The instruction mnemonic utilized in the instruction definition must match that of the instruction mnemonic defined in the CoreGen IR. In this manner, the combined StoneCutter and CoreGen IR infrastructure logic can match the instruction crack implementation and the instruction implementation (et al. ALU).

Second, the instruction definition argument list defines the set of standard I/O ports utilized by the instruction. Standard I/O ports are utilized to pipeline register read and write operations within an optimized pipeline. While it is entirely permissible to directly address registers or register files from within an instruction definition that are not defined in the argument list, the eventual downstream instruction implementation may require additional register read/write operations that induce pipeline stalls. Keep in mind that standard I/O ports such as the clock, register hazarding and stall signals are automatically instantiated.

```
1 def INSTNAME( ARG1 ARG2 ARG3 ... ) {
2 }
```

Listing 8: Instruction Prototype Format

As mentioned above, the instruction prototype includes two main structures: the instruction name and the instruction arguments (Listing 8). The instruction name must match the associated instruction name defined in the CoreGen IR. The instruction name is case sensitive. See the CoreGen IR specification for more information ???. The instruction argument list contains a set of register, instruction field or register class names that define the standard I/O functions for the optimized pipeline. The instruction argument list must match the set of register, instruction format fields or register file designators in the instruction format for the target instruction.

Register arguments are interpreted literally. In this manner, a register I/O is performed to the specific register index denoted by the target register. Register class arguments are interpreted logically. The register class arguments utilize the index specified in the assembled instruction payload to load or store from the respective register index in the target register file.

We see an example of a series of instruction prototypes in Listing 9. We utilize the register class and register definitions from Section 2.6. We define four instructions, `add`, `move`, `inc` and `sub`. The `add` instruction prototype contains a single argument, `GPR`, that defines I/O ports to/from the GPR register class. The `move` instruction utilizes both the `GPR` and the `CTRL` register classes. The `inc` instruction utilizes the explicit `pc` register as well as the `GPR` register class. Finally, the `sub` Notice that the arguments within the prototype are **not** separated by commas. Further, we see that the associated instruction definition body is contained within a set of brackets (`{}`).

```
1 instformat RISC( enc opc, enc func, reg[GPR] RT, reg[GPR] RA, reg[GPR] RB )
2 regclass GPR( u64 r0, u64 r1, u64 r2, u64 r3, u64 r4, u64 r5, u64 r6,
3             u64 r7, u64 r8, u64 r9, u64 r10, u64 r11, u64 r12,
4             u64 r13, u64 r14, u64 r15, u64 r16, u64 r17, u64 r18,
5             u64 r19, u64 r20, u64 r21, u64 r22, u64 r23, u64 r24,
6             u64 r25, u64 r26, u64 r27, u64 r28, u64 r29, u64 r30,
7             r64 r31 )
8 regclass CTRL( u64 pc, u64 exc, u64 ne, u64 gt, u64 lt, u64 gte,
9               u64 lte, u64 sp, u64 fp, u64 rp )
10
11 # add operation
12 def add( GPR ){
13 }
14
15 # move between GPR and CTRL
16 def move( GPR CTRL ){
17 }
18
19 # increment PC by GPR
20 def inc( pc GPR ){
21 }
22
23 # subtract RT = RA - RB
24 def sub( RT RA RB ){
25 }
```

Listing 9: Instruction Prototype Format

2.8 Variable Definitions

In addition to utilizing register classes and register fields as variables within an instruction definition, users may also specify local, temporary variables in the instruction definition body. These values may be utilized as temporary storage for complex, multi-stage operations, loop counters and other intermediate state. The StoneCutter compiler infrastructure will make every effort to minimize the overall hardware impact of these intermediate values to the size and timing of the target design. As a result, these temporary values may be optimized out and/or reused by adjacent instructions in the generated downstream Chisel HDL.

In the same manner as individual registers, variables are required to have a defined type. These types follow the same convention as defined in Section 2.4. Variable names may not collide with existing instruction format fields or register names. Much in the same manner as the C language, StoneCutter permits users to define multiple, independent variables using the same datatype. Variables must also be defined at the **top** of an instruction definition. This permits StoneCutter to sufficiently scope temporary variables across the body of an instruction. The format of variable definitions is noted in Listing 10. We provide an example of a series of variable definitions in Listing 11.

```

1 def INST( ... ){
2   DATATYPE name1
3   DATATYPE name2 = VALUE
4   DATATYPE name3, name4, ...
5
6   <instruction body>
7 }

```

Listing 10: Variable Definitions

```

1 instformat RISC( enc opc, enc func, reg[GPR] RT, reg[GPR] RA, reg[GPR] RB )
2
3 def add( RT RA RB ){
4   u64 var1, var2
5   if( RA == RB ){
6     var1 = 5
7     var2 = 6
8   }else{
9     var1 = 7
10    var2 = 998
11  }
12  RT = var1 & var2
13 }

```

Listing 11: Sample Variable Definitions

2.9 Arithmetic Operations

The StoneCutter language supports the standard set of arithmetic operations as defined by other C-like languages. However, given the nature of compiling the StoneCutter language to a target-specific circuit design, the StoneCutter compiler and tool chain enforces a number of idiosyncratic features when processing variables for arithmetic operations. First, the StoneCutter language scans the variables associated with the left hand side of a target arithmetic operation and automatically converts the operation to perform arithmetic using the width of the largest variable. Further, when storing the result of an operation to a target variable or register (right hand side of an operation), StoneCutter converts the resulting data to saturate the bit space of the target. The process of converting arithmetic operations and variables to larger bit space is performed using zero extension in order to avoid poisoning the numerical consistency of an operation. Users may override this behavior by utilizing sign and zero extended intrinsics (see Section 3.2).

We provide a summary of permissible arithmetic operations in Table 4.

Table 4: StoneCutter Arithmetic Operations

Operator	Example	Description
=	RT = RB	Assignment operation
+	RT = RA + RB	Add operation
-	RT = RA - RB	Subtract operation
*	RT = RA * RB	Multiplication operation
\	RT = RA \ RB	Division operation
%	RT = RA % RB	Modulo operation
&	RT = RA & RB	Bitwise <i>and</i> operation
	RT = RA RB	Bitwise <i>or</i> operation
^	RT = RA ^ RB	Bitwise <i>xor</i> operation
<<	RT = RA << RB	Shift left operation
>>	RT = RA >> RB	Shift right operation

NOTE: Please note that StoneCutter does not currently support the bitwise complement operation (\sim) and the logical not operation (!).

2.10 Conditional Operations

2.10.1 Boolean Operators

The StoneCutter language syntax supports the standard set of boolean operators as is defined by other C-like languages. The one exception being the logical not operation (!). StoneCutter does not support the logical not operation for boolean operations. Rather, users seeking to complement the result of a boolean operation of a singular boolean value must utilize the NOT intrinsic (Section 3.2.14).

We provide a summary of permissible boolean operations in Table 5.

Table 5: StoneCutter Boolean Operations

Operator	Example	Description
==	RA == RB	Logical equivalence
!=	RA != RB	Logical in-equivalence
<	RA < RB	Less than
>	RA > RB	Greater than
<=	RA <= RB	Less than or equal to
>=	RA >= RB	Greater than or equal to
&&	RA && RB	Logical and
	RA RB	Logical or

2.10.2 Flow Control

The StoneCutter language provides support for C-like conditional flow control using standard *if-else* clauses. The statements contained within these clauses must result in a boolean value that is interrogated in order to determine which body of code (circuit) to execute. Flow control operations are required to include a single *if* statement. They may optionally include a complementary *else* statement.

```
1 if( BOOLEAN OPERATION){
2   # Conditional body
3 }
4
5 if( BOOLEAN OPERATION){
6   # Conditional if body
7 }else{
8   # Conditional else body
9 }
```

Listing 12: Flow Control Syntax

We provide a set of example conditional flow control operations in Listing 13.

```
1 def add( RT RA RB ){
2   if( RA > RB ){
3     RB = RA + RB
4   }
5
6   if( RB < RT ){
7     RT = RB
8   }else{
9     RT = RT << 1
10  }
11 }
```

Listing 13: Sample If-Else Syntax

2.11 Loop Operations

2.11.1 For Loops

The first style of loop statement supported by the StoneCutter language is the *for* loop. The StoneCutter for loop is structured in a similar manner as the traditional C for loop. The loop is provided with a base case loop counter, a loop termination statement and an optional iterator trip step. The loop counter can be an existing variable or a newly defined temporary variable. Unlike other temporary variables, loop trip counters can be defined inline within the for loop structure. The loop termination statement includes one or more comparison statements that define the termination state of the loop using the loop trip counter. The optional loop trip counter iterator can be an immediate value or a variable. If this optional value is not specified, the iterator value is assumed to be "1". The for loop body is contained with brackets ({}) and may include any other set of StoneCutter statements. The for loop syntax is depicted in Listing 14.

```

1 for( LOOPTRIP = BASE; LOOPTRIP <COMPARATOR> TERMINATOR ) {
2   # for loop body
3 }
4
5 for( LOOPTRIP = BASE; LOOPTRIP <COMPARATOR> TERMINATOR; ITERATOR ) {
6   # for loop body
7 }

```

Listing 14: For Loop Syntax

The comparator operation utilized to terminate the loop is one of the standard variable comparison operations from C-like languages. We summarize these operators in Table 6.

Table 6: StoneCutter Comparison Operations

Operator	Description
<	Less Than
>	Greater Than
<=	Less Than Or Equal To
>=	Greater Than Or Equal To
==	Equal To
!=	Not Equal To

We provide a set of example for loops in Listing 15.

```

1 def add( RT RA RB ){
2   u64 var1
3   for( i = 1; i < RA ){
4     RB = RB + 1
5   }
6   for( j = 0; i < RB; 5 )
7     RT = RB | RT
8   }
9   for( var = 1024; i <= RT; RB ){
10    RT = var
11  }
12 }

```

Listing 15: Sample For Loop Syntax

2.11.2 While Loops

The second style of loop supported by the StoneCutter language is the *while* loop. The StoneCutter while loop is structured in a similar manner as the traditional C while loop. The loop is provided with a termination statement in the form of a comparator. It is up to the user to modify the variable utilized in the comparator statement such that a valid termination state can be reached. StoneCutter does not otherwise validate the potential for infinite while loops. The while loop body is contained within brackets ({}) and may include any other set of StoneCutter statements. The while loop syntax is depicted in Listing 16.

```
1 while( COMPARATOR STATEMENT ){
2   # while loop body
3 }
```

Listing 16: While Loop Syntax

We provide a set of example while loops in Listing 17.

```
1 def add( RT RA RB ){
2   u64 var1 = 1024
3   while( RA < RB ){
4     RT = RB << 1
5     RA = RA + 1
6   }
7   while( var1 >= RT ){
8     var1 = var1 - 5
9     RT = RT << 1
10  }
11 }
```

Listing 17: Sample While Loop Syntax

2.11.3 Do While Loops

The final style of loop supported by the StoneCutter language is the *do while* loop. The StoneCutter do while loop is structured in a similar manner as the traditional the traditional C do while loop. The loop is provided with an entry state provided by a *do* statement. The do statement is followed by a loop body contained within brackets (`{}`). The do while loop termination statement follows the last bracket and is contained in a *while* block similar in form as the while loop structure (Section 2.11.2). The while loop syntax is depicted in Listing 18.

```
1 do{
2   # do while loop body
3 }while( COMPARATOR STATEMENT )
```

Listing 18: Do While Loop Syntax

We provide a set of example do while loops in Listing 19.

```
1 def add( RT RA RB ){
2   u64 var1 = 1024
3   do{
4     RT = RB << 1
5     RA = RA + 1
6   }while( RA < RB )
7   do{
8     var1 = var1 - 5
9     RT = RT << 1
10  }while( var1 >= RT )
11 }
```

Listing 19: Sample Do While Loop Syntax

3 Intrinsic Functions

3.1 Overview

In the same manner as other C-like languages and compiler infrastructures, the StoneCutter language provides a notional set of builtin intrinsic functions. Much like other languages, these intrinsic functions provide pathological circuit functionality in a convenient and optimized package. However, unlike traditional strongly typed languages that provide unique intrinsic functions for disparate data types, the StoneCutter language provides single intrinsics that support all possible StoneCutter data types. In this manner, the StoneCutter compiler infrastructure interrogates the arguments of each intrinsic at compile time and constructs an optimized Chisel representation of the target operation using the user-specified types. As a result, users may utilize the forthcoming set of intrinsic operations with any target data types supported by the StoneCutter language.

StoneCutter intrinsics are classified as two disparate types: *Arithmetic* intrinsics and *Memory* intrinsics. Arithmetic intrinsics perform some notional permutation on the target input data. Arithmetic intrinsics generally require input and return the output as a unique data member (`OUTPUT = INTRINSIC(INPUT)`). Data members are not modified in place. Memory intrinsics are different in that they interact with the memory infrastructure generated by the CoreGen [1] code generation facilities such that the desired function unit may interact with external memories such as caches and off-chip memory units. Memory intrinsics are generally provide functionality such as *load* and *store* operations.

StoneCutter intrinsic operations may be utilized anywhere within the body of an instruction definition. The initial StoneCutter language parser *lowers* any encountered intrinsic operations to the equivalent of a function call. The StoneCutter code generator expands these intrinsic function calls into their full circuit descriptions using the target input and output types. In this manner, including intrinsics within the body of conditional flow control, loops and other basic blocks is both type safe and functionally sound.

Intrinsics can be utilized similar to C-style function calls. Intrinsic names are placed in the instruction body with the appropriate number of arguments. Intrinsic arguments are separated by commas. Intrinsics that return values can be utilized within other expressions and/or assigned to values (`RT = INTRIN(. . .)`). The syntax for utilizing intrinsics is depicted in Listing 20. A summary of the intrinsics is provided in Section 5.

```
1 def add( RT RA RB ){
2   u64 var1, var2
3   INTRIN( RB )
4   RB = INTRIN( RA )
5   RT = INTRIN( RA, RB )
6 }
```

Listing 20: Intrinsic Syntax

3.2 Arithmetic Intrinsics

3.2.1 CLZ

Table 7: CLZ Intrinsic

Mnemonic	<code>RT = CLZ (RA)</code>
Description	Counts the leading zeros (little endian)
Arguments	RA: Contains the value to count the leading zeros
Example	<pre>def INST(RA RB RT){ RT = CLZ(RA) }</pre>
Return	The number of leading zeros in the input value

DRAFT

3.2.2 COMPRESS

Table 8: COMPRESS Intrinsic

Mnemonic	<code>RT = COMPRESS (RA)</code>
Description	Bit compression. For the number of '1' values in RA, set the least significant bits to '1'. Zero extend the result.
Arguments	RA: Contains the value to compress
Example	<pre>def INST(RA RB RT){ RT = COMPRESS(RA) }</pre>
Return	The compressed version of RA

DRAFT

3.2.3 COMPRESSM

Table 9: COMPRESSM Intrinsic

Mnemonic	<code>RT = COMPRESSM(RA, RB)</code>
Description	Bit compression under mask. Perform a bitwise '&' operation of the input and bitmask. For the number of '1' values in the result, set the least significant bits to '1'. Zero extend the result.
Arguments	RA: Contains the value to compress RB: Contains the bitmask to select compression bits
Example	<pre>def INST(RA RB RT){ RT = COMPRESSM(RA, RB) }</pre>
Return	The compressed version of RA via the bitmask RB

DRAFT

3.2.4 CTZ

Table 10: CTZ Intrinsic

Mnemonic	<code>RT = CTZ (RA)</code>
Description	Counts the trailing zeros (little endian)
Arguments	RA: Contains the value to count the trailing zeros
Example	<pre>def INST(RA RB RT){ RT = CTZ(RA) }</pre>
Return	The number of trailing zeros in the input value

DRAFT

3.2.5 DOZ

Table 11: DOZ Intrinsic

Mnemonic	$RT = DOZ(RA, RB)$
Description	Performs "first grade subtraction." Returns $(RA - RB)$ IFF $RA \geq RB$. Returns 0 otherwise
Arguments	RA: Contains the left hand input value RB: Contains the right hand input value
Example	<pre>def INST(RA RB RT){ RT = DOZ(RA, RB) }</pre>
Return	If($RA \geq RB$) { return $RA - RB$ }else{ return 0 }

DRAFT

3.2.6 EXTRACTS

Table 12: EXTRACTS Intrinsic

Mnemonic	EXTRACTS (RA, RB, RC)
Description	Extract the target field starting a bit position RC from RB, store into RA and sign extend.
Arguments	RA: Contains the target output value RB: Input value to extract from RC: Starting bit position
Example	def INST(RA RB RC RT){ EXTRACTS(RA, RB, RC) }
Return	Nothing is returned

DRAFT

3.2.7 EXTRACTZ

Table 13: EXTRACTZ Intrinsic

Mnemonic	EXTRACTZ (RA, RB, RC)
Description	Extract the target field starting a bit position RC from RB, store into RA and zero extend.
Arguments	RA: Contains the target input value RB: Input value to extract from RC: Starting bit position
Example	def INST(RA RB RC RT){ EXTRACTZ(RA, RB, RC) }
Return	Nothing is returned

DRAFT

3.2.8 INSERTS

Table 14: INSERTS Intrinsic

Mnemonic	INSERTS (RA, RB, RC)
Description	Insert the field (RB) into RA starting at bit position RC and sign extend
Arguments	RA: Contains the target input value RB: Input value to insert RC: Starting bit position
Example	<pre>def INST(RA RB RC RT){ INSERTS(RA, RB, RC) }</pre>
Return	Nothing is returned

DRAFT

3.2.9 INSERTZ

Table 15: INSERTZ Intrinsic

Mnemonic	INSERTZ (RA, RB, RC)
Description	Insert the field (RB) into RA starting at bit position RC and zero extend
Arguments	RA: Contains the target input value RB: Input value to insert RC: Starting a bit position
Example	def INST(RA RB RC RT){ INSERTZ(RA, RB, RC) }
Return	Nothing is returned

DRAFT

3.2.10 MAJ

Table 16: MAJ Intrinsic

Mnemonic	<code>RT = MAJ(RA, RB, RC)</code>
Description	Performs a majority vote of the values RA, RB and RC. Returns true if at least two inputs are true; returns false otherwise
Arguments	RA: Input 1 RB: Input 2 RC: Input3
Example	<pre>def INST(RA RB RC RT){ RT = MAJ(RA, RB, RC) }</pre>
Return	Returns true if at least two inputs are true; returns false otherwise

DRAFT

3.2.11 MAX

Table 17: MAX Intrinsic

Mnemonic	<code>RT = MAX (RA, RB)</code>
Description	Returns the maximum value of RA and RB
Arguments	RA: Input 1 RB: Input 2
Example	<pre>def INST(RA RB RT){ RT = MAX(RA, RB) }</pre>
Return	<code>If(RA > RB){ return RA }else{ return RB }</code>

DRAFT

3.2.12 MERGE

Table 18: MERGE Intrinsic

Mnemonic	MERGE (RA, RB, RC)
Description	Selectively merge bits specified by the mask in RC from RB into RA
Arguments	RA: Contains the target data to merge into RB: Contains the value to selectively merge RC: Contains bitwise mask input
Example	<pre>def INST(RA RB RT){ MERGE(RA, RB, RC) }</pre>
Return	Nothing is returned

DRAFT

3.2.13 MIN

Table 19: MIN Intrinsic

Mnemonic	<code>RT = MIN (RA, RB)</code>
Description	Returns the minimum value of RA and RB
Arguments	RA: Input1 RB: Input2
Example	<pre>def INST(RA RB RT){ RT = MIN(RA, RB) }</pre>
Return	<code>If(RA < RB){ return RA }else{ return RB }</code>

DRAFT

3.2.14 NOT

Table 20: NOT Intrinsic

Mnemonic	<code>RT = NOT(RA)</code>
Description	Returns the bitwise complement of the input value. If the input value is a single bit boolean (<code>u1</code>), the logical complement is returned. If the input value is larger than a single bit, the bitwise complement is returned
Arguments	RA: Contains the input to be complemented
Example	<pre>def INST(RA RB RT){ RT = NOT(RA) }</pre>
Return	Bitwise complement of RA

DRAFT

3.2.15 POPCOUNT

Table 21: POPCOUNT Intrinsic

Mnemonic	<code>RT = POPCOUNT (RA)</code>
Description	Returns the population count of the input variable.
Arguments	RA: Contains the input value for the popcount
Example	<pre>def INST(RA RB RT){ RT = POPCOUNT(RA) }</pre>
Return	The number of positive (1) bits in the input value is returned. The output storage must contain sufficient space to storage the maximum popcount value (all positive bits)

DRAFT

3.2.16 REVERSE

Table 22: REVERSE Intrinsic

Mnemonic	RT = REVERSE (RA)
Description	Reverse the bit pattern of the input variable
Arguments	RA: Contains the input to be reversed
Example	<pre>def INST(RA RB RT){ RT = REVERSE(RA) }</pre>
Return	The reversed bit pattern of the input variable

DRAFT

3.2.17 ROTL

Table 23: ROTL Intrinsic

Mnemonic	<code>RT = ROTL(RA, RB)</code>
Description	Rotate the value in RA by RB bits to the left.
Arguments	RA: Contains the value to be rotated RB: Contains the number of bits to rotate by
Example	<pre>def INST(RA RB RT){ RT = ROTL(RA, RB) }</pre>
Return	The input value is rotated left by RB bits

DRAFT

3.2.18 ROTR

Table 24: ROTR Intrinsic

Mnemonic	<code>RT = ROTR(RA, RB)</code>
Description	Rotate the value in RA by RB bits to the right.
Arguments	RA: Contains the value to be rotated RB: Contains the number of bits to rotate by
Example	<pre>def INST(RA RB RT){ RT = ROTR(RA, RB) }</pre>
Return	The input value is rotated right by RB bits

DRAFT

3.2.19 SEXT

Table 25: SEXT Intrinsic

Mnemonic	<code>Rt = SEXT (RA)</code>
Description	Sign extend the input value
Arguments	RA: Contains the input value to sign extend
Example	<pre>def INST(RA RB RT){ RT = SEXT(RA) }</pre>
Return	The sign extended version of the input

DRAFT

3.2.20 ZEXT

Table 26: ZEXT Intrinsic

Mnemonic	ZEXT (RA, RB)
Description	Zero extend the input value
Arguments	RA: Contains the input value to zero extend
Example	def INST(RA RB RT){ RT = ZEXT(RA) }
Return	The zero extended version of the input

DRAFT

3.3 Memory Intrinsic

3.3.1 LOAD

Table 27: LOAD Intrinsic

Mnemonic	<code>RT = LOAD (RA)</code>
Description	Load the value at the address in RA. Attempt to derive the datatype
Arguments	RA: Contains the target address
Example	<pre>def INST(RA RB RT){ RT = LOAD(RA) }</pre>
Return	The data from the load operation

DRAFT

3.3.2 STORE

Table 28: STORE Intrinsic

Mnemonic	STORE (RA, RB)
Description	foo
Arguments	RA: Contains the data to be stored RB: Contains the target address
Example	<pre>def INST(RA RB RT){ STORE(RA, RB) }</pre>
Return	Nothing is returned

DRAFT

3.3.3 LOADELEM

Table 29: LOADELEM Intrinsic

Mnemonic	<code>RT = LOADELEM(RA, RB)</code>
Description	Load the value at the address in RA. The element size is contained in RB.
Arguments	RA: Contains the target address RB: Contains the size of the element
Example	<pre>def INST(RA RB RT){ RT = LOADELEM(RA, RB) }</pre>
Return	The data from the load operation

DRAFT

3.3.4 STOREELEM

Table 30: STOREELEM Intrinsic

Mnemonic	STOREELEM (RA, RB, RC)
Description	Stores the data in RA to the address at RB. The element size is contained in RC.
Arguments	RA: Contains the data to be stored RB: Contains the target address RC: Contains the size of the element
Example	def INST(RA RB RC RT){ STOREELEM(RA, RB, RC) }
Return	Nothing is returned

DRAFT

4 Appendix A: Sample StoneCutter Implementation

```
1  |-- StoneCutter source file for ISA=BasicRISC.ISA
2
3  # Instruction Formats
4  instformat Arith.if(reg[GPR] ra,reg[GPR] rb,reg[GPR] rt,enc opc,
5                      enc func,imm imm)
6  instformat ReadCtrl.if(reg[GPR] ra,reg[CTRL] rb,reg[GPR] rt,enc opc,
7                          enc func,imm imm)
8  instformat WriteCtrl.if(reg[GPR] ra,reg[GPR] rb,reg[CTRL] rt,enc opc,
9                          enc func,imm imm)
10
11 # Register Class Definitions
12 regclass GPR( u64 r0, u64 r1, u64 r2, u64 r3, u64 r4, u64 r5, u64 r6,
13              u64 r7, u64 r8, u64 r9, u64 r10, u64 r11, u64 r12, u64 r13,
14              u64 r14, u64 r15, u64 r16, u64 r17, u64 r18, u64 r19, u64 r20,
15              u64 r21, u64 r22, u64 r23, u64 r24, u64 r25, u64 r26, u64 r27,
16              u64 r28, u64 r29, u64 r30, u64 r31 )
17 regclass CTRL( u64 pc, u64 exc, u64 ne, u64 eq, u64 gt, u64 lt, u64 gte,
18               u64 lte, u64 sp, u64 fp, u64 rp )
19
20 # Instruction Definitions
21 # add
22 def add( ra rb rt imm )
23 {
24     rt = ra + rb
25 }
26
27 # sub
28 def sub( ra rb rt imm )
29 {
30     rt = ra - rb
31 }
32
33 # mul
34 def mul( ra rb rt imm )
35 {
36     rt = ra * rb
37 }
38
39 # div
40 def div( ra rb rt imm )
41 {
42     rt = ra / rb
43 }
44
45 # divu
46 def divu( ra rb rt imm )
47 {
48     rt = ra / rb
49 }
50
51 # sll
```

```
52 def sll( ra rb rt imm )
53 {
54   rt = ra << rb
55 }
56
57 # srl
58 def srl( ra rb rt imm )
59 {
60   rt = ra >> rb
61 }
62
63 # sra
64 def sra( ra rb rt imm )
65 {
66   rt = ra >> rb
67 }
68
69 # and
70 def and( ra rb rt imm )
71 {
72   rt = ra & rb
73 }
74
75 # or
76 def or( ra rb rt imm )
77 {
78   rt = ra | rb
79 }
80
81 # nand
82 def nand( ra rb rt imm )
83 {
84   rt = NOT(ra & rb)
85 }
86
87 # nor
88 def nor( ra rb rt imm )
89 {
90   rt = NOT(ra | rb)
91 }
92
93 # xor
94 def xor( ra rb rt imm )
95 {
96   rt = ra ^ rb
97 }
98
99 # cmp.ne
100 def cmp.ne( ra rb rt imm )
101 {
102   if( ra != rb ){
103     rt = 2
104   }else{
105     rt = 0
```



```
106   }
107 }
108
109 # cmp.eq
110 def cmp.eq( ra rb rt imm )
111 {
112   if( ra == rb ){
113     rt = 3
114   }else{
115     rt = 0
116   }
117 }
118
119 # cmp.gt
120 def cmp.gt( ra rb rt imm )
121 {
122   if( ra > rb ){
123     rt = 4
124   }else{
125     rt = 0
126   }
127 }
128
129 # cmp.lt
130 def cmp.lt( ra rb rt imm )
131 {
132   if( ra < rb ){
133     rt = 5
134   }else{
135     rt = 0
136   }
137 }
138
139 # cmp.gte
140 def cmp.gte( ra rb rt imm )
141 {
142   if( ra >= rb ){
143     rt = 6
144   }else{
145     rt = 0
146   }
147 }
148
149 # cmp.lte
150 def cmp.lte( ra rb rt imm )
151 {
152   if( ra <= rb ){
153     rt = 7
154   }else{
155     rt = 0
156   }
157 }
158
159 # lb
```

```
160 def lb( ra rb rt imm )
161 {
162   rt = SEXT(LOADELEM(ra+imm,8))
163 }
164
165 # lh
166 def lh( ra rb rt imm )
167 {
168   rt = SEXT(LOADELEM(ra+imm,16))
169 }
170
171 # lw
172 def lw( ra rb rt imm )
173 {
174   rt = SEXT(LOADELEM(ra+imm,32))
175 }
176
177 # ld
178 def ld( ra rb rt imm )
179 {
180   rt = LOADELEM(ra+imm,64)
181 }
182
183 # sb
184 def sb( ra rb rt imm )
185 {
186   STOREELEM(ra,rt+imm,8)
187 }
188
189 # sh
190 def sh( ra rb rt imm )
191 {
192   STOREELEM(SEXT(ra),rt+imm,16)
193 }
194
195 # sw
196 def sw( ra rb rt imm )
197 {
198   STOREELEM(SEXT(ra),rt+imm,32)
199 }
200
201 # sd
202 def sd( ra rb rt imm )
203 {
204   STOREELEM(SEXT(ra),rt+imm,64)
205 }
206
207 # lbu
208 def lbu( ra rb rt imm )
209 {
210   rt = ZEXT(LOADELEM(ra+imm,8))
211 }
212
213 # lhu
```

```
214 def lhu( ra rb rt imm )
215 {
216   rt = ZEXT(LOADELEM(ra+imm,16))
217 }
218
219 # lwu
220 def lwu( ra rb rt imm )
221 {
222   rt = ZEXT(LOADELEM(ra+imm,32))
223 }
224
225 # sbu
226 def sbu( ra rb rt imm )
227 {
228   STOREELEM(ZEXT(ra),rt+imm,8)
229 }
230
231 # shu
232 def shu( ra rb rt imm )
233 {
234   STOREELEM(ZEXT(ra),rt+imm,16)
235 }
236
237 # swu
238 def swu( ra rb rt imm )
239 {
240   STOREELEM(ZEXT(ra),rt+imm,32)
241 }
242
243 # not
244 def not( ra rb rt imm )
245 {
246   rt = NOT(ra)
247 }
248
249 # bra
250 def bra( ra rb rt imm )
251 {
252   pc = rt
253 }
254
255 # br
256 def br( ra rb rt imm )
257 {
258   pc = SEXT(pc + rt)
259 }
260
261 # cadd
262 def cadd( ra rb rt imm )
263 {
264   rt = ra + rb
265 }
266
267 # brac
```

```
268 def brac( ra rb rt imm )
269 {
270   if( ra == rb ){
271     pc = rt
272   }else{
273     pc = pc + 4
274   }
275 }
276
277 # brc
278 def brc( ra rb rt imm )
279 {
280   if( ra == rb ){
281     pc = pc + rt
282   }else{
283     pc = pc + 4
284   }
285 }
286
287 # ladd
288 def ladd( ra rb rt imm )
289 {
290   rt = ra + rb
291 }
292
293 # brr
294 def brr( ra rb rt imm )
295 {
296   pc = rt
297 }
```

Listing 21: Sample StoneCutter

5 Appendix B: Intrinsic Function Table

Table 31: StoneCutter Ininsics

Intrinsic	Inputs	Outputs	Description
CLZ	1	1	Count leading zero
COMPRESS	1	1	Bit compress
COMPRESSM	1	2	Bit compress under mask
CTZ	1	1	Count trailing zero
DOZ	1	2	First grade subtraction
EXTRACTS	0	3	Bit extract and sign extend
EXTRACTZ	0	3	Bit extract and zero extend
INSERTS	0	3	Bit insert and sign extend
INSERTZ	0	3	Bit insert and zero extend
MAJ	1	3	Majority vote
MAX	1	2	Max
MERGE	0	3	Merge under mask
MIN	1	2	Min
NOT	1	1	Bit complement
POPCOUNT	1	1	Population count
REVERSE	1	1	Bit reverse
ROTL	1	2	Rotate left
ROTR	1	2	Rotate right
SEXT	1	1	Signe extend
ZEXT	1	2	Zero extend
LOAD	1	1	Memory load
STORE	0	2	Memory store
LOAELEM	1	2	Memory element load
STOREELEM	0	3	Memory element store

References

- [1] J. Leidel and F. Conlon, “Coregen intermediate representation language specification,” 2019. [Online]. Available: <https://github.com/opensocsysarch/CoreGenIRSpec>

DRAFT