

# The Capstone-RISC-V Instruction Set Reference

## Table of Contents

|                                                           |    |
|-----------------------------------------------------------|----|
| 1. Introduction                                           | 5  |
| 1.1. Goals                                                | 5  |
| 1.2. Major Design Elements                                | 5  |
| 1.3. Capstone-RISC-V ISA Overview                         | 6  |
| 1.4. Capstone Variants                                    | 6  |
| 1.5. Assembly Mnemonics                                   | 7  |
| 1.6. Notations                                            | 7  |
| 1.7. Bibliography                                         | 7  |
| 2. Programming Model                                      | 9  |
| 2.1. Capabilities                                         | 9  |
| 2.2. Extension to General-Purpose Registers               | 11 |
| 2.3. Extension to Other Registers                         | 13 |
| 2.4. Extension to Memory                                  | 14 |
| 2.5. Added Registers                                      | 14 |
| 2.6. Instruction Set                                      | 16 |
| 3. Capability Manipulation Instructions                   | 17 |
| 3.1. Cursor, Bounds, and Permissions Manipulation         | 17 |
| 3.2. Type Manipulation                                    | 20 |
| 3.3. Dropping                                             | 21 |
| 3.4. Revocation                                           | 22 |
| 4. Memory Access Instructions                             | 24 |
| 4.1. <i>Pure Capstone</i>                                 | 24 |
| 4.2. <i>TransCapstone</i>                                 | 25 |
| 5. Control Flow Instructions                              | 27 |
| 5.1. Jump to Capabilities                                 | 27 |
| 5.2. Domain Crossing                                      | 28 |
| 5.3. A World Switching Extension for <i>TransCapstone</i> | 30 |
| 6. Control State Instructions                             | 33 |
| 6.1. Capability CSR (CCSR) Manipulation                   | 33 |
| 7. Adjustments to Existing Instructions                   | 34 |
| 7.1. Memory Access Instructions                           | 34 |
| 7.2. Control Flow Instructions                            | 37 |
| 8. Interrupts and Exceptions                              | 40 |
| 8.1. Exception and Exit Codes                             | 40 |

|                                                             |    |
|-------------------------------------------------------------|----|
| 8.2. Pure Capstone .....                                    | 41 |
| 8.3. TransCapstone .....                                    | 42 |
| 9. Memory Consistency Model .....                           | 45 |
| Appendix A: Debugging Instructions (Non-Normative) .....    | 46 |
| A.1. World Switching .....                                  | 46 |
| A.2. Exception Handling .....                               | 46 |
| Appendix B: Instruction Listing .....                       | 48 |
| B.1. Debugging Instructions .....                           | 48 |
| B.2. Capstone Instructions .....                            | 48 |
| B.3. Extended RV64I Memory Access Instructions .....        | 50 |
| Appendix C: Assembly Code Examples .....                    | 53 |
| Appendix D: Abstract Binary Interface (Non-Normative) ..... | 54 |

Contributors to this document include (in alphabetical order): Jason Zhijingcheng Yu, Mingkai Li

**Version Information:** Draft version. Refer to the commit hash.

# 1. Introduction

The Capstone project is an effort to explore the design of a new CPU instruction set architecture that achieves multiple security goals including memory safety and isolation with one unified hardware abstraction.

## 1.1. Goals

The ultimate goal of Capstone is to unify the numerous hardware abstracts that have been added as extensions to existing architectures as afterthought mitigations to security vulnerabilities. This goal requires a high level of flexibility and extensibility of the Capstone architecture. More specifically, we aim to support the following in a unified manner.

### **Exclusive access**

Software should be guaranteed exclusive access to certain memory regions if needed. This is in spite of the existence of software traditionally entitled to higher privileges such as the OS kernel and the hypervisor.

### **Revocable delegation**

Software components should be able to delegate authority to other components in a revocable manner. For example, after an untrusted library function has been granted access to a memory region, the caller should be able to revoke this access.

### **Dynamically extensible hierarchy**

The hierarchy of authority should be dynamically extensible, unlike traditional platforms which follow a static hierarchy of hypervisor-kernel-user. This makes it possible to use the same set of abstractions for memory isolation and memory sharing regardless of where a software component lies in the hierarchy.

### **Safe context switching**

A mechanism of context switching without trusting any other software component should be provided. This allows for a minimal TCB if necessary in case of a highly security-critical application.

## 1.2. Major Design Elements

The Capstone architecture design is based on the idea of capabilities, which are unforgeable tokens that represent authority to perform memory accesses and control flow transfers. Capstone extends the traditional capability model with new capability types including the following.

### **Linear capabilities**

Linear capabilities are guaranteed not to alias with other capabilities. Operations on linear capabilities maintain this property. For example, linear capabilities cannot be duplicated. Instead, they can only be moved around across different registers or between registers and memory. They can hence enable safe exclusive access to memory regions. Capabilities that do not have this property are called *non-linear* capabilities.

## Revocation capabilities

Revocation capabilities cannot be used to perform memory accesses or control flow transfers. Instead, they convey the authority to revoke other capabilities. Each revocation capability is derived from a linear capability and can later be used to revoke (i.e., invalidate) capability derived from the same linear capability. This mechanism enables revocable and arbitrarily extensible chains of delegation of authority.

## Uninitialised capabilities

Uninitialised capabilities convey write-only authority to memory. They can be turned into linear capabilities after the memory region has been "initialised", that is, when the whole memory region has been overwritten with fresh data. Uninitialised capabilities enable safe initialisation of memory regions and prevent secret leakage without incurring extra performance overhead.

# 1.3. Capstone-RISC-V ISA Overview

While Capstone does not assume any specific modern ISA, we choose to propose a Capstone extension to RISC-V due to its open nature and the availability of toolchains and simulators.

The Capstone-RISC-V ISA is an RV64I extension that makes the following types of changes to the base architecture:

- Each general-purpose register is extended to 129 bits to accommodate 128-bit capabilities.
- New instructions for manipulating capabilities are added.
- New instructions for memory accesses using capabilities are added.
- New instructions for control flow transfers using capabilities are added.
- New instructions for machine state control are added.
- Semantics of a small number of existing instructions are changed to support capabilities.
- Semantics of interrupts and exceptions are changed to support capabilities.

# 1.4. Capstone Variants

In addition to Capstone, which is referred to as *Pure Capstone* in the Capstone-RISC-V ISA, we propose a variant of Capstone, called *TransCapstone*. While memory accesses and control flow transfers are only possible using capabilities in *Pure Capstone*, *TransCapstone* fuses capabilities with privilege levels and virtual memory found in traditional architectures, which allows for a smooth transition from existing architectures to Capstone.

The following types of changes are made to *Pure Capstone* to obtain *TransCapstone*:

- The physical memory is partitioned into [two disjoint regions](#), one exclusively for accesses through capabilities and the other exclusively for accesses through the virtual memory.
- Software components are allowed to run in either of the 2 *worlds*, i.e., the *normal world* and the *secure world*.
  - The *normal world* follows the traditional privilege levels, allows both capability-based accesses and virtual memory accesses, and is therefore compatible with existing softwares.

- The *secure world* follows the *Pure Capstone* design, limits memory accesses to capability-based accesses and provides the security guarantees of Capstone.
- A [world switching mechanism](#) is added to support the secure switching between the 2 *worlds*.
- Semantics of a small number of *Pure Capstone* instructions are changed to support the *normal world* and the *secure world*.
- Semantics of interrupts and exceptions are extended to support the *normal world* and the *secure world*.

Table 1. Memory Accesses in TransCapstone

| World               | MMU | Capabilities |
|---------------------|-----|--------------|
| <i>Normal World</i> | Yes | Yes          |
| <i>Secure World</i> | -   | Yes          |

## 1.5. Assembly Mnemonics

Each Capstone-RISC-V instruction is given a mnemonic prefixed with **CS..** In contexts where it is clear we are discussing Capstone-RISC-V instructions, we will omit the **CS.** prefix for brevity.

In assembly code, the list of operands to an instruction is supplied following the instruction mnemonic, with the operands separated by commas, in the order of **rd**, **rs1**, **rs2**, **imm** for any operand the instruction expects.

## 1.6. Notations

When specifying the semantics of instructions, we use the following notations to represent the type of each operand:

**I**

Integer register.

**C**

Capability register.

**S**

Sign-extended immediate.

**Z**

Zero-extended immediate.

## 1.7. Bibliography

The initial design of Capstone has been discussed in the following paper:

- [Capstone: A Capability-based Foundation for Trustless Secure Memory Access](#) by Jason Zhijingcheng Yu, Conrad Watt, Aditya Badole, Trevor E. Carlson, Prateek Saxena. In *Proceedings*

*of the 32nd USENIX Security Symposium. Anaheim, CA, USA. August 2023.*

## 2. Programming Model

The Capstone-RISC-V ISA has extended part of the machine state, including both some registers and the memory, to enable the storage and handling of capabilities.

### 2.1. Capabilities

#### 2.1.1. Width

The width of a capability is 128 bits. We represent this as  $CLEN = 128$  and  $CLENBYTES = 16$ . Note that this does not affect the width of a raw address, which is  $XLEN = 64$  bits, or equivalently,  $XLENBYTES = 8$  bytes, same as in RV64I.

#### 2.1.2. Fields

Each capability has the following architecturally-visible fields:

Table 2. Fields in a capability

| Name   | Range           | Description                                                                                                                                                                    |
|--------|-----------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| valid  | $0..1$          | Whether the capability is valid: $0$ = invalid, $1$ = valid                                                                                                                    |
| type   | $0..6$          | The type of the capability: $0$ = linear, $1$ = non-linear, $2$ = revocation, $3$ = uninitialised, $4$ = sealed, $5$ = sealed-return, $6$ = exit                               |
| cursor | $0..2^{XLEN}-1$ | Not applicable when $type = 4$ (sealed), $type = 5$ (sealed-return), or $type = 6$ (exit). The memory address the capability points to (to be used for the next memory access) |
| base   | $0..2^{XLEN}-1$ | Not applicable when $type = 6$ (exit). The base memory address of the memory region associated with the capability                                                             |
| end    | $0..2^{XLEN}-1$ | Not applicable when $type = 4$ (sealed), $type = 5$ (sealed-return), or $type = 6$ (exit). The end memory address of the memory region associated with the capability          |

| Name  | Range | Description                                                                                                                                                                                                                                                                                                                         |
|-------|-------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| perms | 0..7  | Not applicable when <code>type = 4</code> (sealed), <code>type = 5</code> (sealed-return), or <code>type = 6</code> (exit). One-hot encoded permissions associated with the capability: 0 = no access, 1 = execute-only, 2 = write-only, 3 = write-execute, 4 = read-only, 5 = read-execute, 6 = read-write, 7 = read-write-execute |
| async | 0..1  | Only applicable when <code>type = 4</code> (sealed) or <code>type = 5</code> (sealed-return). Whether the capability is sealed asynchronously: 0 = synchronously, 1 = asynchronously                                                                                                                                                |
| reg   | 0..31 | Only applicable when <code>type = 5</code> (sealed-return). The index of the general-purpose register to restore the capability to                                                                                                                                                                                                  |

The range of the `perms` field has a partial order  $\leq_p$  defined as follows:

```

<=p = {
  (0, 0), (0, 1), (0, 2), (0, 3), (0, 4), (0, 5), (0, 6), (0, 7),
  (1, 1), (1, 3), (1, 5), (1, 7),
  (2, 2), (2, 3), (2, 6), (2, 7),
  (3, 3), (3, 7),
  (4, 4), (4, 5), (4, 6), (4, 7),
  (5, 5), (5, 7),
  (6, 6), (6, 7),
  (7, 7)
}

```

We say a capability `c` aliases with a capability `d` if and only if the intersection between `[c.base, c.end)` and `[d.base, d.end)` is non-empty.

For two revocation capabilities `c` and `d` (i.e., `c.type = d.type = 2`), we say `c <_t d` if and only if

- `c` aliases with `d`
- The creation of `c` was earlier than the creation of `d`

In addition to the above fields, an implementation also needs to maintain sufficient metadata to test the  $<_t$  relation. It will be clear that for any pair of aliasing revocation capabilities, the order of their creations is well-defined.

## Note

The **valid** field is involved in **revocation**, where it might be changed due to a **revocation operation** on a different capability. A performant implementation, therefore, may prefer not to maintain the **valid** field inline with the other fields.

Implementations are free to maintain additional fields to capabilities or compress the representation of the above fields, as long as each capability fits in **CLEN** bits. It is not required to be able to represent capabilities with all combinations of field values in a compressed representation, as long as the following conditions are satisfied:

- For load and store instructions that move a capability between a register and memory, the value of the capability is preserved.
- The resulting capability values of any operation are not more powerful than when the same operation is performed on a Capstone-RISC-V implementation without compression. More specifically, if an execution trace is valid (i.e., without exceptions) on the compressed implementation, then it must also be valid on the uncompressed implementation. For example, a trivial yet useless compression would be to store nothing and always return a capability with **valid** = 0 (TODO: double-check this claim).

## Note

For different types of capabilities, a specific subset of the fields is used. The table below summarises the fields used for each type of capabilities.

Table 3. Fields used for each type of capabilities

| Type          | type | valid | cursor | base | end | perms | async | reg |
|---------------|------|-------|--------|------|-----|-------|-------|-----|
| Linear        | 0    | Yes   | Yes    | Yes  | Yes | Yes   | -     | -   |
| Non-linear    | 1    | Yes   | Yes    | Yes  | Yes | Yes   | -     | -   |
| Revocation    | 2    | Yes   | Yes    | Yes  | Yes | Yes   | -     | -   |
| Uninitialised | 3    | Yes   | Yes    | Yes  | Yes | Yes   | -     | -   |
| Sealed        | 4    | Yes   | -      | Yes  | -   | -     | Yes   | -   |
| Sealed-return | 5    | Yes   | -      | Yes  | -   | -     | Yes   | Yes |
| Exit          | 6    | Yes   | -      | -    | -   | -     | -     | -   |

## 2.2. Extension to General-Purpose Registers

The Capstone-RISC-V ISA extends each of the 32 general-purpose registers, so it contains either a

capability or a raw **XLEN**-bit integer. The type of data contained in a register is maintained and confusion of the type is not allowed, except for **x0/c0** as discussed below. In assembly code, the type of data expected in a register operand is indicated by the alias used for the register, as summarised in the following table.

| Index | XLEN-bit integer | Capability        |
|-------|------------------|-------------------|
| 0     | <b>x0/zero</b>   | <b>c0/cnull</b>   |
| 1     | <b>x1/ra</b>     | <b>c1/cra</b>     |
| 2     | <b>x2/sp</b>     | <b>c2/csp</b>     |
| 3     | <b>x3/gp</b>     | <b>c3/cgp</b>     |
| 4     | <b>x4/tp</b>     | <b>c4/ctp</b>     |
| 5     | <b>x5/t0</b>     | <b>c5/ct0</b>     |
| 6     | <b>x6/t1</b>     | <b>c6/ct1</b>     |
| 7     | <b>x7/t2</b>     | <b>c7/ct2</b>     |
| 8     | <b>x8/s0/fp</b>  | <b>c8/cs0/cfp</b> |
| 9     | <b>x9/s1</b>     | <b>c9/cs1</b>     |
| 10    | <b>x10/a0</b>    | <b>c10/ca0</b>    |
| 11    | <b>x11/a1</b>    | <b>c11/ca1</b>    |
| 12    | <b>x12/a2</b>    | <b>c12/ca2</b>    |
| 13    | <b>x13/a3</b>    | <b>c13/ca3</b>    |
| 14    | <b>x14/a4</b>    | <b>c14/ca4</b>    |
| 15    | <b>x15/a5</b>    | <b>c15/ca5</b>    |
| 16    | <b>x16/a6</b>    | <b>c16/ca6</b>    |
| 17    | <b>x17/a7</b>    | <b>c17/ca7</b>    |
| 18    | <b>x18/s2</b>    | <b>c18/cs2</b>    |
| 19    | <b>x19/s3</b>    | <b>c19/cs3</b>    |
| 20    | <b>x20/s4</b>    | <b>c20/cs4</b>    |
| 21    | <b>x21/s5</b>    | <b>c21/cs5</b>    |
| 22    | <b>x22/s6</b>    | <b>c22/cs6</b>    |
| 23    | <b>x23/s7</b>    | <b>c23/cs7</b>    |
| 24    | <b>x24/s8</b>    | <b>c24/cs8</b>    |
| 25    | <b>x25/s9</b>    | <b>c25/cs9</b>    |
| 26    | <b>x26/s10</b>   | <b>c26/cs10</b>   |
| 27    | <b>x27/s11</b>   | <b>c27/cs11</b>   |
| 28    | <b>x28/t3</b>    | <b>c28/ct3</b>    |
| 29    | <b>x29/t4</b>    | <b>c29/ct4</b>    |

| Index | XLEN-bit integer | Capability |
|-------|------------------|------------|
| 30    | x30/t5           | c30/ct5    |
| 31    | x31/t6           | c31/ct6    |

`x0/c0` is a read-only register that can be used both as an integer and as a capability, depending on the context. When used as an integer, it has the value `0`. When used as a capability, it has the value `{ valid = 0, type = 0, cursor = 0, base = 0, end = 0, perms = 0 }`. Any attempt to write to `x0/c0` will be silently ignored (no exceptions are raised).

In this document, for `i = 0, 1, ..., 31`, we use `x[i]` to refer to the general-purpose register with index `i`.

## 2.3. Extension to Other Registers

### 2.3.1. Program Counter

The following changes are made to the program counter (`pc`):

- *Pure Capstone*: The program counter (`pc`) is extended to contain a capability only.
- *TransCapstone*: Similar to the general-purpose registers, the program counter (`pc`) is also extended to contain a capability or an integer.

**During the instruction fetch stage, an exception is raised when any of the following conditions are met:**

- *Pure Capstone*
  - `x[pc].valid` is `0` (invalid).
  - `x[pc].type` is neither `0` (linear) nor `1` (non-linear).
  - `x[pc].cursor` is not aligned to `4`.
  - `2 <= p x[pc].perms` does not hold.
  - `x[pc].cursor` is not in the range `[pc.base, pc.end - 4]`.
- *TransCapstone*
  - `cwld` is `1` (secure world) and any of the conditions for *Pure Capstone* are met.
  - `cwld` is `0` (normal world) and `x[pc]` is not an integer.
  - `cwld` is `0` (normal world) and any of the conditions for RV64I are met.

**If no exception is raised:**

- *Pure Capstone*: The instruction pointed to by `x[pc].cursor` is fetched and executed. The `x[pc].cursor` is then incremented by `4` (i.e., `x[pc].cursor += 4`).
- *TransCapstone*:
  - `cwld` is `1` (secure world): Same as *Pure Capstone*.
  - `cwld` is `0` (normal world): The instruction pointed to by `x[pc]` is fetched and executed. The

$x[pc]$  is then incremented by 4 (i.e.,  $x[pc] += 4$ ).

## 2.4. Extension to Memory

The memory is addressed using an  $XLEN$ -bit integer at byte-level granularity. In addition to raw integers, each  $CLEN$ -bit aligned address can also store a capability. The type of data contained in a memory location is maintained and confusion of the type is not allowed.

In *Pure Capstone*, the memory can only be accessed through capabilities.

| Address Space   | Access Method |
|-----------------|---------------|
| $[0, 2^{XLEN})$ | Capabilities  |

In *TransCapstone*, the physical memory is divided into two disjoint regions: the *normal memory* and the *secure memory*. While the *normal memory* is only accessible through MMU (Memory Management Unit), the *secure memory* can only be accessed through capabilities.

| Memory Region | Address Space                      | Access Method |
|---------------|------------------------------------|---------------|
| Normal Memory | $[0, SBASE) \cup [SEND, 2^{XLEN})$ | MMU           |
| Secure Memory | $[SBASE, SEND)$                    | Capabilities  |

## 2.5. Added Registers

The Capstone-RISC-V ISA adds the following registers:

Table 4. Additional Registers in Capstone-RISC-V ISA

| Capstone Variant | Additional Registers |               |                                                                        |
|------------------|----------------------|---------------|------------------------------------------------------------------------|
| Pure Capstone    | Mnemonic             | CCSR encoding | Description                                                            |
|                  | ceh                  | 0x000         | The sealed capability for the exception handler                        |
|                  | cih                  | 0x001         | The sealed capability for the interrupt handler                        |
|                  | cinit                | 0x010         | The initial capability covering the entire address space of the memory |

| Capstone Variant | Additional Registers |               |                                                                                                                                                 |
|------------------|----------------------|---------------|-------------------------------------------------------------------------------------------------------------------------------------------------|
| TransCapstone    | Mnemonic             | CCSR encoding | Description                                                                                                                                     |
|                  | ceh                  | 0x000         | The sealed capability for the exception handler                                                                                                 |
|                  | cinit                | 0x010         | The initial capability covering the entire address space of the secure memory                                                                   |
|                  | emode                | 0x011         | The encoding mode of the system. 0 = integer encoding mode, 1 = capability encoding mode                                                        |
|                  | cwrlld               | -             | The currently executed world. 0 = normal world, 1 = secure world                                                                                |
|                  | normal_pc            | -             | The program counter for the normal world before the secure world is entered                                                                     |
|                  | normal_sp            | -             | The stack pointer for the normal world before the secure world is entered                                                                       |
|                  | switch_reg           | -             | The index of the general-purpose register used when switching worlds. This register needs to be stored during context switching in normal world |
|                  | switch_cap           | -             | The capability used to store contexts when switching worlds. This register needs to be stored during context switching in normal world          |
|                  | exit_reg             | -             | The index of the general-purpose register for receiving the exit code when exiting the secure world                                             |

Some of the registers only allow capability values and have special semantics related to the system-wide machine state. They are referred to as *capability control state registers (CCSRs)*. Under their respective constraints, CCSRs can be manipulated using [CCSR manipulation instructions](#).

The manipulation constraints for each CCSR are indicated below.

Table 5. Manipulation Constraints for CCSRs

| Mnemonic | Read                                                                          | Write                                                                         |
|----------|-------------------------------------------------------------------------------|-------------------------------------------------------------------------------|
| ceh      | No constraint                                                                 | No constraint                                                                 |
| cih      | The original content must be an invalid capability ( <code>valid = 0</code> ) | The original content must be an invalid capability ( <code>valid = 0</code> ) |
| cinit    | No constraint                                                                 | Silently ignored                                                              |

## Note

`ceh` and `cih` should be handled differently. `ceh` is about the functionality of a domain only. A domain should be allowed to set `ceh` for itself. That also means it needs to be switched when switching domains. `cih` is about the functionality of the system, which should normally only be set once. To prevent any domain from setting `cih`, we require the original content of `cih` to be invalid for an attempt to change it to succeed.

## Note

`cinit` is a capability CSR used to provide the initial capability in the system. During system initialisation, `x[cinit]` is set by the hardware to a linear capability as follows:

Table 6. Initial Capability of `x[cinit]`

| Variant       | type       | cursor | base  | end    | perms                  | valid     |
|---------------|------------|--------|-------|--------|------------------------|-----------|
| Pure Capstone | 1 (linear) | 0      | 0     | 2^XLEN | 7 (read-write-execute) | 1 (valid) |
| TransCapstone | 1 (linear) | SBASE  | SBASE | SEND   | 7 (read-write-execute) | 1 (valid) |

CCSR manipulation instructions can be used to read the initial capability of `x[cinit]` and store it in one of the general-purpose registers. `x[cinit]` is set to `x[cnull]` after the first operation and any attempt to write to `x[cinit]` is silently ignored.

## 2.6. Instruction Set

The Capstone-RISC-V instruction set is based on the RV64I instruction set. The (uncompressed) instructions are fixed 32-bit wide, and laid out in memory in little-endian order. In the encoding space of the RV64I instruction set, Capstone-RISC-V instructions occupies the "custom-2" subset, i.e., the opcode of all Capstone-RISC-V instructions is `0b1011011`.

Capstone-RISC-V instruction encodings follow two basic formats: R-type and I-type, as described below (more details are also provided in the [RISC-V ISA Manual](#)).

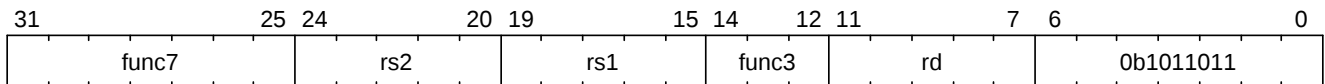


Figure 1. R-type instruction format

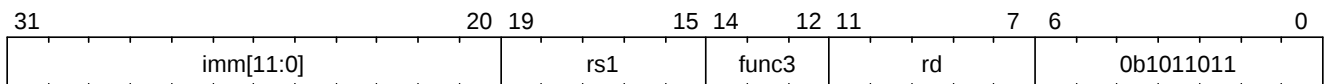


Figure 2. I-type instruction format

R-type instructions receive up to three register operands, and I-type instructions receive up to two register operands and a 12-bit-wide immediate operand.

## 3. Capability Manipulation Instructions

Capstone provides instructions for creating, modifying, and destroying capabilities. Note that due to the guarantee of provenance of capabilities, those instructions are the *only* way to manipulate capabilities. In particular, it is not possible to manipulate capabilities by manipulating the content of a memory location or register using other instructions.

### 3.1. Cursor, Bounds, and Permissions Manipulation

#### 3.1.1. Capability Movement

Capabilities can be moved between registers with the `MOVC` instruction.

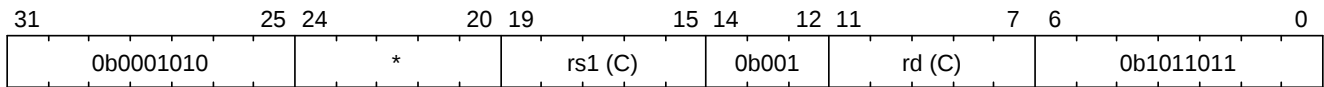


Figure 3. `MOVC` instruction format

An exception is raised when any of the following conditions are met:

- `x[rs1]` is not a capability

**If no exception is raised:** If `rs1 = rd`, the instruction is a no-op. Otherwise, write `x[rs1]` to `x[rd]`, and if `x[rs1]` is neither a non-linear capability (i.e., `type != 1`) nor an exit capability (i.e., `type != 6`), write `x[null]` to `x[rs1]`.

#### 3.1.2. Cursor Increment

The `CINCOFFSET` and `CINCOFFSETIMM` instructions increment the *cursor* of a capability by a give amount (offset).

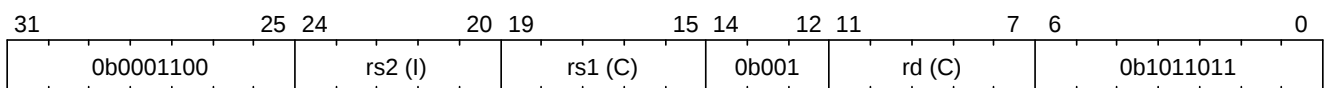


Figure 4. `CINCOFFSET` instruction format

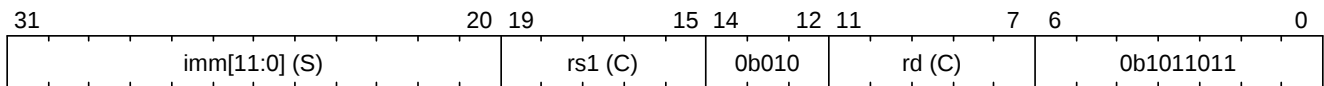


Figure 5. `CINCOFFSETIMM` instruction format

An exception is raised when any of the following conditions are met:

- `x[rs1]` is not a capability.
- `x[rs2]` is not an integer (for `CINCOFFSET`).
- `x[rs1]` does not have `type = 0` (linear) or `type = 1` (non-linear).

**If no exception is raised:** For `CINCOFFSET`, the offset is read from `x[rs2]`. For `CINCOFFSETIMM`, the offset is the 12-bit sign-extended immediate field `imm`. If the offset is `0`, the instructions are semantically equivalent to `MOVC rd, rs1`. Otherwise, the instructions are equivalent to an atomic

execution of `MOVC rd, rs1` followed by an increment of `x[rd].cursor` by the offset.

### 3.1.3. Cursor Setter

The `cursor` field of a capability can also be directly set with the SCC instruction.

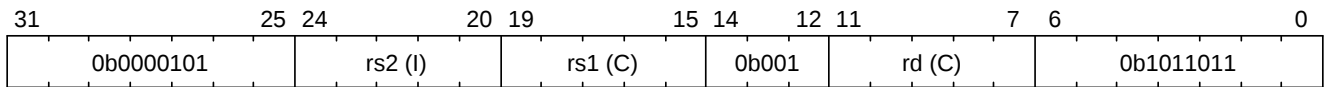


Figure 6. SCC instruction format

**An exception is raised if any of the following conditions are met:**

- `x[rs1]` is not a capability.
- `x[rs2]` is not an integer.
- `x[rs1]` does not have `type = 0` (linear) or `type = 1` (non-linear).

**If no exception is raised:** The instruction is equivalent to an atomic execution of `MOVC rd, rs1` followed by setting `x[rd].cursor` to `x[rs2]`.

### 3.1.4. Field Grab

The LCC instruction is used to read a field from a capability.

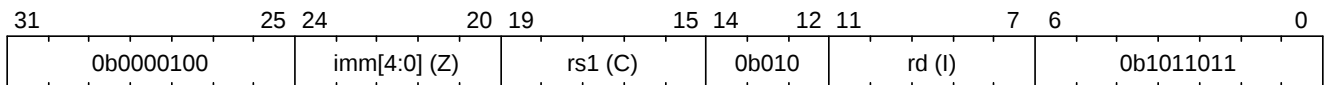


Figure 7. LCC instruction format

**An exception is raised if any of the following conditions are met:**

- `x[rs1]` is not a capability.
- `imm = 2` and `x[rs1]` has `type = 4` (sealed), `type = 5` (sealed-return), or `type = 6` (exit).
- `imm = 3` and `x[rs1]` has `type = 6` (exit).
- `imm = 4` and `x[rs1]` has `type = 4` (sealed), `type = 5` (sealed-return), or `type = 6` (exit).
- `imm = 5` and `x[rs1]` has `type = 4` (sealed), `type = 5` (sealed-return), or `type = 6` (exit).
- `imm = 6` and `x[rs1]` does not have `type = 4` (sealed) or `type = 5` (sealed-return).
- `imm = 7` and `x[rs1]` does not have `type = 5` (sealed-return).

**If no exception is raised:** If `imm > 7`, the instruction writes `x[zero]` to `x[rd]`. Otherwise, the instruction write different fields of `x[rs1]` to `x[rd]` according to the following table:

| imm | Content of <code>x[rd]</code> after LCC |
|-----|-----------------------------------------|
| 0   | <code>x[rs1].valid</code>               |
| 1   | <code>x[rs1].type</code>                |
| 2   | <code>x[rs1].cursor</code>              |
| 3   | <code>x[rs1].base</code>                |

| imm | Content of $x[rd]$ after LCC |
|-----|------------------------------|
| 4   | $x[rs1].end$                 |
| 5   | $x[rs1].perms$               |
| 6   | $x[rs1].async$               |
| 7   | $x[rs1].reg$                 |

### 3.1.5. Bounds Shrinking

The bounds (**base** and **end** fields) of a capability can be shrunk with the SHRINK instruction.

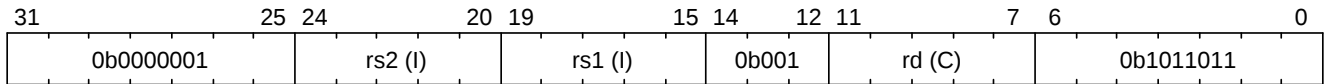


Figure 8. SHRINK instruction format

The instruction attempts to set the bounds of the capability  $x[rd]$  to  $[x[rs1], x[rs2]]$ .

**An exception is raised when any of the following conditions are met:**

- $x[rd]$  is not a capability.
- $x[rd].type$  is not 0, 1, or 3 (linear, non-linear, or uninitialised).
- $x[rs1]$  is not an integer.
- $x[rs2]$  is not an integer.
- $x[rs1] \geq x[rs2]$ .
- $x[rs1] < x[rd].base$  or  $x[rs2] > x[rd].end$ .

**If no exception is raised:** The instruction sets the bounds of the capability  $x[rd]$  to  $[x[rs1], x[rs2]]$ , i.e.,  $x[rd].base$  is set to  $x[rs1]$  and  $x[rd].end$  is set to  $x[rs2]$ .

### 3.1.6. Bounds Splitting

The SPLIT instruction can split a capability into two by splitting the bounds.

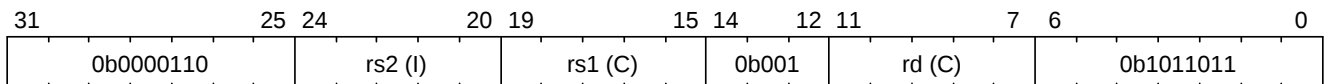


Figure 9. SPLIT instruction format

The instruction attempts to split the capability  $x[rs1]$  into two capabilities, one with bounds  $[x[rs1].base, x[rs2]]$  and the other with bounds  $[x[rs2], x[rs1].end]$ .

**An exception is raised when any of the following conditions are met:**

- $x[rs1]$  is not a capability.
- $x[rs1].valid$  is 0 (invalid).
- $x[rs1].type$  is neither 0 nor 1 (neither linear nor non-linear).
- $x[rs2]$  is not an integer.

- $x[rs2] \leq x[rs1].base$  or  $x[rs2] \geq x[rs1].end$ .

**If no exception is raised:** Set  $x[rs1].end$  to  $x[rs2]$ . A new capability is created with  $base = x[rs2]$  and the other fields equal to those of the original  $x[rs1]$ . The new capability is written to  $x[rd]$ .

### 3.1.7. Permission Tightening

The TIGHTEN instruction tightens the permissions ( $perms$  field) of a capability.

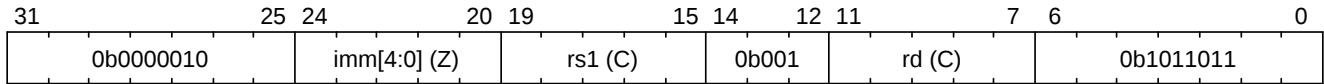


Figure 10. TIGHTEN instruction format

**An exception is raised when any of the following conditions are met:**

- $x[rs1]$  is not a capability.
- $x[rs1].type$  is not 0, 1, or 3 (linear, non-linear, or uninitialised).
- $x[rs1] \leq 7$  and  $x[rs1] \leq_p x[rd].perms$  does not hold.

**If no exception is raised:** The instruction is equivalent to:

1. If  $x[rs1] > 7$ , set  $x[rs1].perms$  to  $x[zero]$ . Otherwise, set  $x[rs1].perms$  to  $x[rs1]$
2. `MOVC rd, rs1`

## 3.2. Type Manipulation

Some instructions affect the  $type$  field of a capability.

### 3.2.1. Delinearisation

The DELIN instruction delinearises a linear capability.

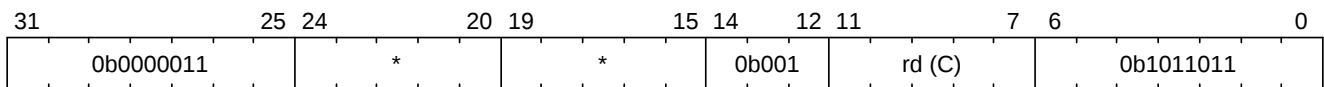


Figure 11. DELIN instruction format

**An exception is raised when any of the following conditions are met:**

- $x[rd]$  is not a capability.
- $x[rd].type$  is not 0 (linear).

**If no exception is raised:**  $x[rd].type$  is set to 1 (non-linear).

### 3.2.2. Initialisation

The INIT instruction transforms an uninitialised capability into a linear capability after its associated memory region has been fully initialised (written with new data).

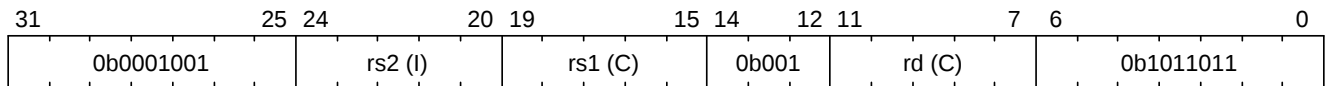


Figure 12. INIT instruction format

**An exception is raised when any of the following conditions are met:**

- $x[rs1]$  is not a capability.
- $x[rs1].type$  is not 3 (uninitialised).
- $x[rs1].cursor$  and  $x[rs1].end$  are not equal.
- $x[rs2]$  is not an integer.

**If no exception is raised:** The instruction is equivalent to:

1. Set  $x[rs1].type$  to 0 (linear) and  $x[rs1].cursor$  is set to  $x[rs2]$
2. `MOVC rd, rs1`

### 3.2.3. Sealing

The SEAL instruction seals a linear capability.

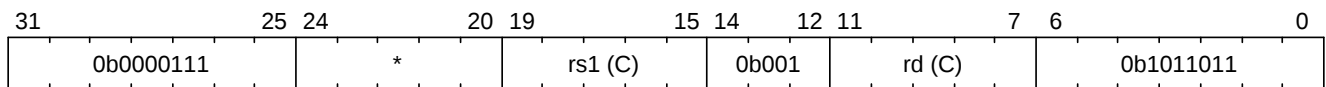


Figure 13. SEAL instruction format

**An exception is raised when any of the following conditions are met:**

- $x[rs1]$  is not a capability.
- $x[rs1].type$  is not 0 (linear).
- $x[rs1].perms$  is neither 6 (read-write) nor 7 (read-write-execute).
- The size of the memory region associated with  $x[rs1]$  is smaller than  $CLENBYTES * 33$  bytes. That is,  $x[rs1].end - x[rs1].base < CLENBYTES * 33$ .

**If no exception is raised:** The instruction is equivalent to:

1. Set  $x[rs1].type$  to 2 (sealed) and  $x[rs1].async$  to 0 (synchronous)
2. `MOVC rd, rs1`

## 3.3. Dropping

The DROP instruction invalidates a capability.

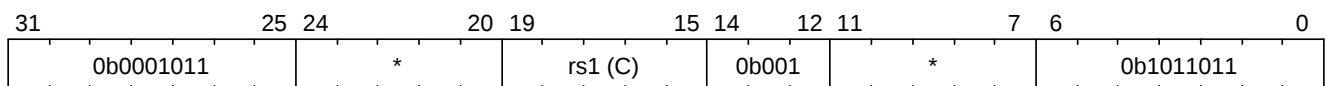


Figure 14. DROP instruction format

**An exception is raised when any of the following conditions are met:**

- `x[rs1]` is not a capability.
- `x[rs1].valid` is 0 (invalid).

If no exception is raised: `x[rs1].valid` is set to 0 (invalid).

## 3.4. Revocation

### 3.4.1. Revocation Capability Creation

The MREV instruction creates a revocation capability.

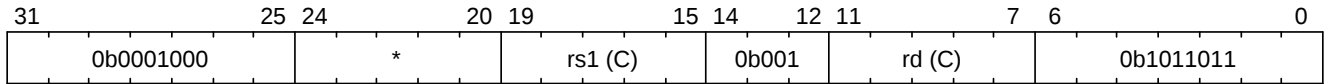


Figure 15. MREV instruction format

An exception is raised when any of the following conditions are met:

- `x[rs1]` is not a capability.
- `x[rs1].type` is not 0 (linear).
- `x[rs1].valid` is 0 (invalid).

If no exception is raised: A new capability is created in `x[rd]` with the same `base`, `end`, `perms`, `cursor` and `valid` fields as `x[rs1]`. The `type` field of the new capability is set to 2 (revocation).

### 3.4.2. Revocation Operation

The REVOKE instruction revokes a capability.

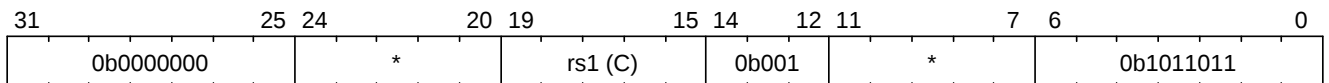


Figure 16. REVOKE instruction format

An exception is raised when any of the following conditions are met:

- `x[rs1]` is not a capability.
- `x[rs1].type` is not 2 (revocation).
- `x[rs1].valid` is 0 (invalid).

If no exception is raised:

For any capability `c` in the system (in either a register or memory location), `c.valid` is set to 0 (invalid) if any of the following conditions are met:

- `c.type` is not 2 (revocation), `c.valid` is 1 (valid), and `c` aliases with `x[rs1]`.
- `c.type` is 2 (revocation), `c.valid` is 1 (valid), and `x[rs1] < t c`.

`x[rs1].type` is set to 0 (linear) if, for any each invalidated capability `c`, at least one of the following

conditions are met:

- The type of `c` is non-linear (i.e., `c.type` is 1)
- $2 \leq_p c.perms$  does not hold

Otherwise, `x[rs1].type` is set to 3 (uninitialised), and `x[rs1].cursor` is set to `x[rs1].base`.

## 4. Memory Access Instructions

Capstone provides instructions to load and store capabilities from/to memory regions.

### 4.1. *Pure Capstone*

In *Pure Capstone*, two specific instructions (i.e., LDC and LTC) are used to load and store capabilities.

#### 4.1.1. Load Capabilities

The LDC instruction loads a capability from memory.

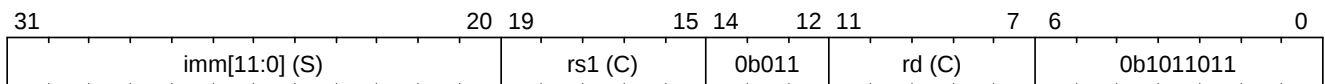


Figure 17. LDC instruction format

An exception is raised when any of the following conditions are met:

- $x[rs1]$  is not a capability.
- $x[rs1].type$  is neither 0 (linear) nor 1 (non-linear).
- $x[rs1].valid$  is 0 (invalid).
- $4 \leq x[rs1].perms$  does not hold.
- $x[rs1].cursor + imm$  is not in the range  $[x[rs1].base, x[rs1].end - CLENBYTES]$ .
- $x[rs1].cursor + imm$  is not aligned to CLEN bits.
- The data contained in the memory location  $[x[rs1].cursor + imm, x[rs1].cursor + imm + CLENBYTES)$  is not a capability.
- The `type` field of the capability being loaded is neither 1 (non-linear) nor 6 (exit), and  $x[rs1].perms$  is neither 6 (read-write) or 7 (read-write-execute).

**If no exception is raised:** Load the capability at the memory location  $[x[rs1].cursor + imm, x[rs1].cursor + imm + CLENBYTES)$  into  $x[rd]$ . If the capability being loaded is not a non-linear capability (i.e.,  $type \neq 1$ ) or an exit capability (i.e.,  $type \neq 6$ ), the data contained in the memory location  $[x[rs1].cursor + imm, x[rs1].cursor + imm + CLENBYTES)$  will be set to  $x[cnull]$ .

#### 4.1.2. Store Capabilities

The STC instruction stores a capability to memory.

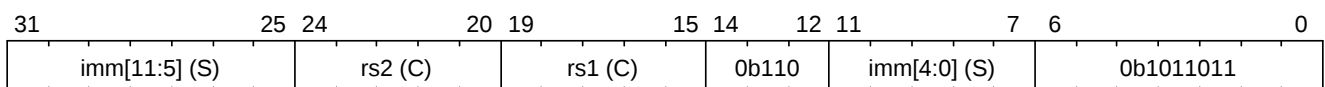


Figure 18. STC instruction format

An exception is raised when any of the following conditions are met:

- $x[rs1]$  is not a capability.

- `x[rs1].type` is not 0, 1, or 3 (linear, non-linear, or uninitialised).
- `x[rs1].type` is 3 (uninitialised) and `imm` is not 0.
- `x[rs1].valid` is 0 (invalid).
- `x[rs1].perms` is neither 6 (read-write) nor 7 (read-write-execute).
- `x[rs1].cursor + imm` is not in the range `[x[rs1].base, x[rs1].end - CLENBYTES]`.
- `x[rs1].cursor + imm` is not aligned to `CLEN` bits.
- `x[rs2]` is not a capability.

**If no exception is raised:** Store `x[rs2]` to the memory location `[x[rs1].cursor + imm, x[rs1].cursor + imm + CLENBYTES]`. If `x[rs1]` is an uninitialised capability (i.e., `x[rs1].type` is 3), `x[rs1].cursor` is set to `x[rs1].cursor + CLENBYTES`. If `x[rs2]` is not a non-linear capability (i.e., `type != 1`) or an exit capability (i.e., `type != 6`), `x[rs2]` will be set to `x[cnull]`.

## 4.2. TransCapstone

In *TransCapstone*, the LDC and STC instructions are extended to support loading and storing capabilities from/to the normal memory using raw addresses.

- In the secure world (i.e., `cwld` is 1), the LDC and STC instructions remain the same as in *Pure Capstone*.
- In the normal world (i.e., `cwld` is 0), the LDC and STC instructions behave differently in different *encoding modes*. When `emode` is 1 (capability encoding mode), the LDC and STC instructions behave the same as in *Pure Capstone*. When `emode` is 0 (integer encoding mode), the LDC and STC instructions are used to load and store capabilities from/to the normal memory using raw addresses.

### 4.2.1. Load Capabilities in integer encoding mode

When `cwld` is 0 (normal world) and `emode` is 0 (integer encoding mode), the LDC instruction loads a capability from the normal memory using raw addresses.

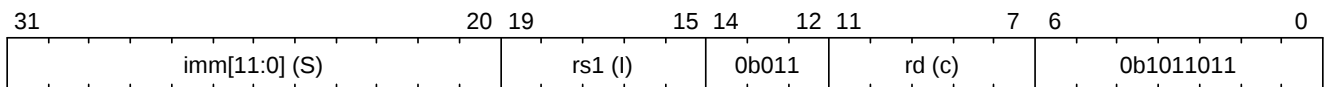


Figure 19. LDC instruction format

**An exception is raised when any of the following conditions are met:**

- `x[rs1]` is not an integer.
- `x[rs1] + imm` is not aligned to `CLEN` bits.
- `x[rs1] + imm` is in the range `[SBASE, SEND]`.
- The data contained in the memory location `[x[rs1] + imm, x[rs1] + imm + CLENBYTES]` is not a capability.

**If no exception is raised:** Load the capability at the memory location `[x[rs1] + imm, x[rs1] + imm + CLENBYTES]` into `x[rd]`. If the `type` field of the capability being loaded is neither 1 (non-linear) nor 6

(exit), the data contained in the memory location  $[x[rs1] + imm, x[rs1] + imm + CLENBYTES)$  will be set to  $x[cnull]$ .

### 4.2.2. Store Capabilities in integer encoding mode

When `cwld` is 0 (normal world) and `emode` is 0 (integer encoding mode), the STC instruction stores a capability to the normal memory using raw addresses.

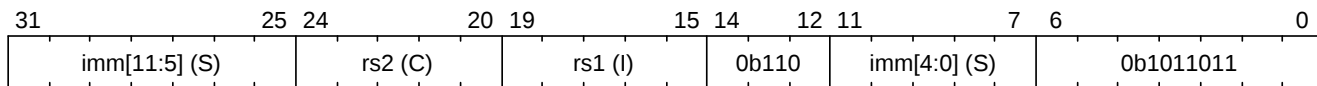


Figure 20. STC instruction format

**An exception is raised when any of the following conditions are met:**

- $x[rs1]$  is not an integer.
- $x[rs1] + imm$  is not aligned to `CLEN` bits.
- $x[rs1] + imm$  is in the range  $[SBASE, SEND)$ .
- $x[rs2]$  is not a capability.

**If no exception is raised:** Store  $x[rs2]$  to the memory location  $[x[rs1] + imm, x[rs1] + imm + CLENBYTES)$ . If the `type` field of the capability being stored is neither 1 (non-linear) nor 6 (exit),  $x[rs2]$  will be set to  $x[cnull]$ .

# 5. Control Flow Instructions

## 5.1. Jump to Capabilities

The CJALR and CBNZ instructions allow jumping to a capability, i.e., setting the program counter to a given capability, in a unconditional or conditional manner.

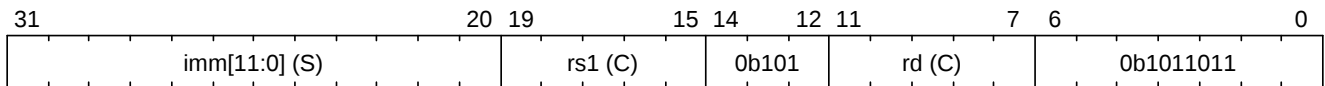


Figure 21. CJALR instruction format

**An exception is raised when any of the following conditions are met:**

*Pure Capstone*

- $x[rs1]$  is not a capability.

*TransCapstone*

- $cwrl d$  is 0 (normal world).
- $x[rs1]$  is not a capability.

**If no exception is raised:** Set  $x[pc]$  to  $x[rs1]$ , and  $x[pc].cursor$  to  $x[rs1].cursor + imm$ . Meanwhile, former  $x[pc]$  ( $x[pc].cursor$  is the address of the next instruction) is written to  $x[rd]$ .

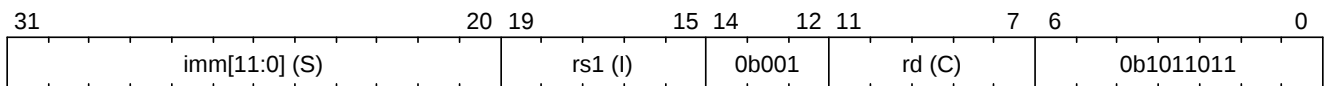


Figure 22. CBNZ instruction format

**An exception is raised when any of the following conditions are met:**

*Pure Capstone*

- $x[rd]$  is not a capability.
- $x[rs1]$  is not an integer.

*TransCapstone*

- $cwrl d$  is 0 (normal world).
- $x[rd]$  is not a capability.
- $x[rs1]$  is not an integer.

**If no exception is raised:** If  $x[rs1]$  is 0, the instruction is a no-op. Otherwise, set  $x[pc]$  to  $x[rs1]$ , and  $x[pc].cursor$  to  $x[rs1].cursor + imm$ .

## 5.2. Domain Crossing

*Domains* in Capstone-RISC-V are individual software compartments that are protected by a safe context switching mechanism, i.e., domain crossing. The mechanism is provided by the CALL and RETURN instructions.

### 5.2.1. CALL

The CALL instruction is used to call a sealed capability, i.e., to switch to another *domain*.

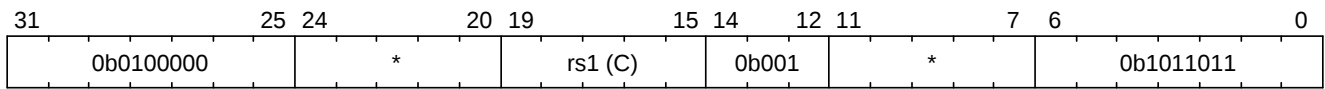


Figure 23. CALL instruction format

**An exception is raised when any of the following conditions are met:**

- *Pure Capstone*
  - $x[rs1]$  is not a capability.
  - $x[rs1].valid$  is 0 (invalid).
  - $x[rs1].type$  is not 4 (sealed).
  - $x[rs1].async$  is 1 (asynchronous).
- *TransCapstone*
  - $cworld$  is 0 (normal world).
  - Any of the conditions for *Pure Capstone* are met.

**If no exception is raised:**

1. Load the content at the memory location  $[x[rs1].base, x[rs1].base + CLENBYTES)$  to the program counter ( $pc$ ).
2. Load the content at the memory location  $[x[rs1].base + CLENBYTES, x[rs1].base + 2 * CLENBYTES)$  to  $ceh$ .
3. Load the content at the memory location  $[x[rs1].base + 2 * CLENBYTES, x[rs1].base + 3 * CLENBYTES)$  to  $csp$ .
4. Store the former  $pc$ ,  $ceh$  and  $csp$  values to the memory location  $[x[rs1].base, x[rs1].base + CLENBYTES)$ ,  $[x[rs1].base + CLENBYTES, x[rs1].base + 2 * CLENBYTES)$  and  $[x[rs1].base + 2 * CLENBYTES, x[rs1].base + 3 * CLENBYTES)$  respectively.
5. Set  $x[rs1].type$  to 5 (sealed-return),  $x[rs1].reg$  to  $rd$ , set  $x[rs1].async$  to 0 (synchronous), and write the resulting  $x[rs1]$  to the register  $cra$ .

### 5.2.2. RETURN

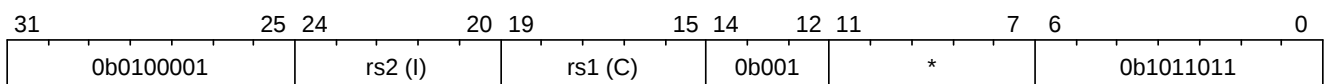


Figure 24. RETURN instruction format

An exception is raised when any of the following conditions are met:

- *Pure Capstone*
  - `x[rs1]` is not a capability.
  - `x[rs1].valid` is 0 (invalid).
  - `x[rs1].type` is not 5 (sealed-return).
  - `x[rs2]` is not an integer.
- *TransCapstone*
  - `cwrl` is 0 (normal world).
  - Any of the conditions for *Pure Capstone* are met.

If no exception is raised:

When `x[rs1].async = 0` (synchronous):

1. Load the content at the memory location `[x[rs1].base, x[rs1].base + CLENBYTES)` to the program counter (`pc`).
2. Load the content at the memory location `[x[rs1].base + CLENBYTES, x[rs1].base + 2 * CLENBYTES)` to `ceh`.
3. Load the content at the memory location `[x[rs1].base + 2 * CLENBYTES, x[rs1].base + 3 * CLENBYTES)` to `csp`.
4. Store the former `pc`, `ceh` and `csp` values to the memory location `[x[rs1].base, x[rs1].base + CLENBYTES)`, `[x[rs1].base + CLENBYTES, x[rs1].base + 2 * CLENBYTES)` and `[x[rs1].base + 2 * CLENBYTES, x[rs1].base + 3 * CLENBYTES)` respectively.
5. Set `x[rs1].type` to 4 (sealed), and write the capability to the register `x[x[rs1].reg]`.

When `x[rs1].async = 1` (asynchronous):

1. Load the content at the memory location `[x[rs1].base, x[rs1].base + CLENBYTES)` to the program counter (`pc`).
2. For `i = 1, 2, ..., 31`, load the content at the memory location `[x[rs1].base + i * CLENBYTES, x[rs1].base + (i + 1) * CLENBYTES)`, to `x[i]` (the `i`-th general-purpose register).
3. Write the former value of `pc`, with the `cursor` field replaced by `x[rs2]`, to the memory location `[x[rs1].base, x[rs1].base + CLENBYTES)`.
4. For `i = 1, 2, ..., 31`, store the content of `x[i]` (the `i`-th general-purpose register) to the memory location `[x[rs1].base + i * CLENBYTES, x[rs1].base + (i + 1) * CLENBYTES)`. When `i = rs1`, store the content of `cnull` instead to `[x[rs1].base + i * CLENBYTES, x[rs1].base + (i + 1) * CLENBYTES)`.
5. Set the `x[rs1].type` to 4 (sealed), and write the resulting `x[rs1]` to the register `ceh`.

### Note

When the `async` field of a sealed-return capability is 1 (asynchronous), some memory accesses

are granted by this capability. The following table shows the memory accesses granted by sealed and sealed-return capabilities in different scenarios.

Table 7. Memory accesses granted by sealed and sealed-return capabilities

| Capability type | <i>async</i> | Read                         | Write                        | Execute |
|-----------------|--------------|------------------------------|------------------------------|---------|
| Sealed          | 0            | No                           | No                           | No      |
| Sealed          | 1            | No                           | No                           | No      |
| Sealed-return   | 0            | No                           | No                           | No      |
| Sealed-return   | 1            | <i>cursor in [base, end)</i> | <i>cursor in [base, end)</i> | No      |

## 5.3. A World Switching Extension for *TransCapstone*

In *TransCapstone*, a pair of extra instructions, i.e., CAPENTER and CAPEXIT, is added to support switching between the secure world and the normal world. The CAPENTER instruction causes an entry into the secure world from the normal world, and the CAPEXIT instruction causes an exit from the secure world into the normal world.

The CAPENTER instruction can only be used in the normal world, whereas the CAPEXIT instruction can only be used in the secure world. In addition, the CAPEXIT instruction can only be used when an exit capability is provided. Attempting to use those instructions in the wrong world or without the required capability will cause an exception. The behaviours of these 2 instructions roughly correspond to the CALL and RETURN instructions respectively.

### 5.3.1. CAPENTER

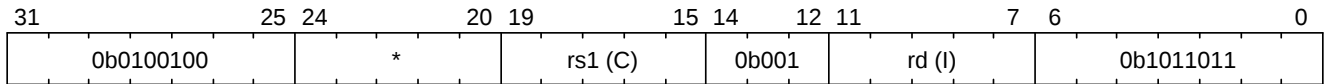


Figure 25. CAPENTER instruction format

An exception is raised when any of the following conditions are met:

- *cwrl*d is 1 (secure world).
- *x[rs1]* is not a capability.
- *x[rs1].valid* is 0 (invalid).
- *x[rs1].type* is not 4 (sealed).

If no exception is raised:

TODO: we need to find a way to access the content in the switch\_cap region

When *x[rs1].async* = 0 (synchronous):

1. Load the content at the memory location [*x[rs1].base*, *x[rs1].base* + *CLENBYTES*) to the program counter (*pc*).
2. Load the content at the memory location [*x[rs1].base* + *CLENBYTES*, *x[rs1].base* + 2 \* *CLENBYTES*).

CLENBYTES) to `ceh`.

3. Load the content at the memory location  $[x[rs1].base + 2 * CLENBYTES, x[rs1].base + 3 * CLENBYTES)$  to `csp`.
4. Store the former value of `pc` and `sp` to `normal_pc` and `normal_sp` respectively.
5. Set `x[rs1].type` to 5 (sealed-return), `x[rs1].async` to 0 (synchronous), and write the resulting `x[rs1]` to `switch_cap`.
6. Write `rs1` to `switch_reg`.
7. Create a capability of `type = 6` (exit) in `cra`.
8. Write `rd` to `exit_reg`.
9. Set `cwrlld` to 1 (secure world).

**When `x[rs1].async = 1` (asynchronous):**

1. Load the content at the memory location  $[x[rs1].base, x[rs1].base + CLENBYTES)$  to the program counter (`pc`).
2. Load the content at the memory location  $[x[rs1].base + CLENBYTES, x[rs1].base + 2 * CLENBYTES)$  to `ceh`.
3. For  $i = 1, 2, \dots, 31$ , load the content at the memory location  $[x[rs1].base + (i + 1) * CLENBYTES, x[rs1].base + (i + 2) * CLENBYTES)$ , to `x[i]` (the  $i$ -th general-purpose register).
4. Store the former value of `pc` and `sp` to `normal_pc` and `normal_sp` respectively.
5. Set `x[rs1].type` to 5 (sealed-return), `x[rs1].async` to 0 (synchronous), and write the resulting `x[rs1]` to `switch_cap`.
6. Write `rs1` to `switch_reg`.
7. Write `rd` to `exit_reg`.
8. Set `cwrlld` to 1 (secure world).

### Note

The `rd` register will be set to a value indicating the cause of exit when the CPU core exits from the secure world synchronously or asynchronously.

## 5.3.2. CAPEXIT

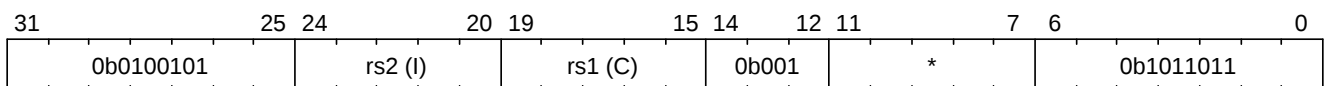


Figure 26. CAPEXIT instruction format

**An exception is raised when any of the following conditions are met:**

- `cwrlld` is 0 (normal world).
- `x[rs1]` is not a capability.

- `x[rs1].valid` is `0` (invalid).
- `x[rs1].type` is not `6` (exit).
- `x[rs2]` is not an integer.
- `switch_cap` is not a capability.
- `switch_cap.valid` is `0` (invalid).
- `switch_cap.type` is not `4` (sealed-return).
- `switch_cap.async` is `1` (asynchronous).

**If no exception is raised:**

1. Write the content of `normal_pc` and `normal_sp` to `pc` and `sp` respectively.
2. Write the former value of `pc`, with the `cursor` field replaced by `x[rs2]`, to the memory location `[switch_cap.base, switch_cap.base + CLENBYTES)`.
3. Write the former value of `ceh` and `csp` to the memory location `[switch_cap.base + CLENBYTES, switch_cap.base + 2 * CLENBYTES)` and `[switch_cap.base + 2 * CLENBYTES, switch_cap.base + 3 * CLENBYTES)` respectively.
4. Set `switch_cap.type` to `4` (sealed), `switch_cap.async` to `0` (synchronous), and write the resulting `switch_cap` to `x[switch_reg]`.
5. Set `x[exit_reg]` to `0` (normal exit).
6. Set `cwrlld` to `0` (normal world).

## 6. Control State Instructions

### 6.1. Capability CSR (CCSR) Manipulation

The CCSRRW instruction is used to read and write specified [capability CSRs](#) (CCSRs).

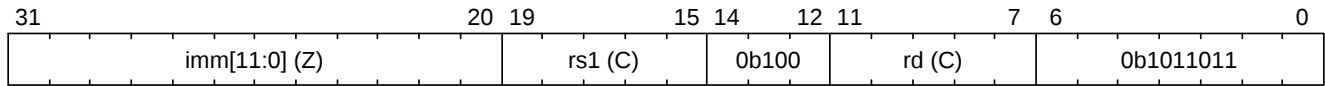


Figure 27. CCSRRW instruction format

**An exception is raised when any of the following conditions are met:**

- The immediate value `imm` does not correspond to the encoding of a valid capability CSR.
- `x[rs1]` is not a capability.

**If no exception is raised:**

1. Read from capability CSR

- If the [read constraint](#) is satisfied, the content of the capability CSR specified by the immediate value `imm` is written to `x[rd]`. If the current content of the capability CSR is neither a non-linear capability (i.e., `type != 1`) nor an exit capability (i.e., `type != 6`), it will be set to the content of `null`.
- Otherwise, `x[rd]` is set to the content of `null`.

2. Write to capability CSR

- If the [write constraint](#) is satisfied, `x[rs1]` is written to the capability CSR specified by the immediate value `imm`. If `x[rs1]` is neither a non-linear capability (i.e., `type != 1`) nor an exit capability (i.e., `type != 6`), it will be set to the content of `null`.
- Otherwise, the original current of the capability CSR is preserved.

## 7. Adjustments to Existing Instructions

For most existing instructions in RV64I, the adjustments are straightforward. Their behaviour is unchanged, and an exception is raised if any of the operands (i.e., `x[rs1]`, `x[rs2]` or `x[rd]`) is a capability. Apart from this operand constraint, the following instructions in RV64I are adjusted in Capstone:

- For memory access instructions, they are extended to use capabilities as addresses for memory access.
- For control flow instructions, they are slightly adjusted to support capability-aware control flow.

### 7.1. Memory Access Instructions

In RV64I, memory access instructions include load instructions (i.e., `lb`, `lh`, `lw`, `lbu`, `lhu`, `lwu`, `ld`), and store instructions (i.e., `sb`, `sh`, `sw`, `sd`). In RV64I, these instructions take an integer as a raw address, and load or store a value from/to this address. In Capstone, these instructions are extended to take a capability as an address.

#### 7.1.1. *Pure Capstone*

##### Load Instructions

In *Pure Capstone*, RV64I load instructions are modified to load integers of different sizes using capabilities.

##### Signed Load Instructions

`lb`, `lh`, `lw`, `ld` instructions are modified to load an integer in the size of byte, halfword, word, and doubleword (i.e., `XLENBYTES/8`, `XLENBYTES/4`, `XLENBYTES/2`, and `XLENBYTES` bytes), using capabilities respectively. The loaded integer is sign-extended to the size of the destination register (i.e., `XLEN` bits).

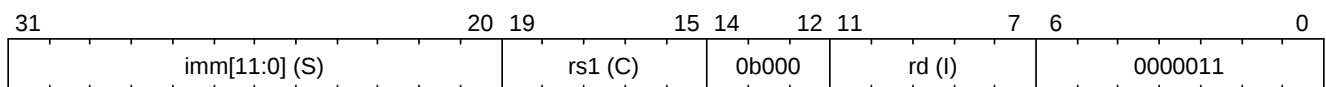


Figure 28. `lb` instruction format (I-type)

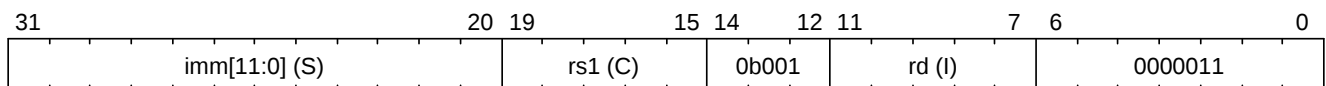


Figure 29. `lh` instruction format (I-type)

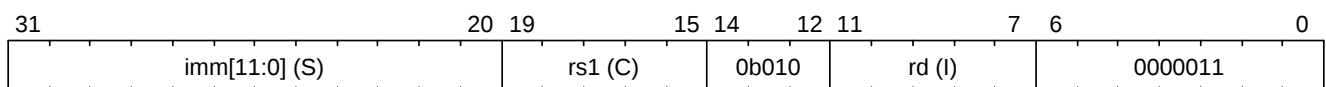


Figure 30. `lw` instruction format (I-type)

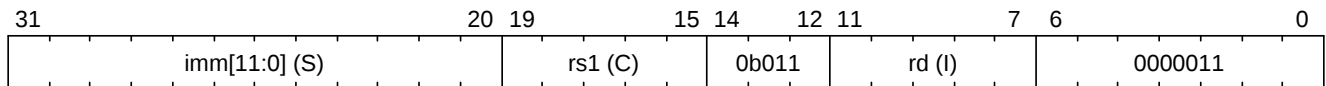


Figure 31. ld instruction format (I-type)

An exception is raised when any of the following conditions are met:

- $x[rs1]$  is not a capability.
- $x[rs1].type$  is neither 0 (linear) nor 1 (non-linear).
- $x[rs1].valid$  is 0 (invalid).
- $4 \leq x[rs1].perms$  does not hold.
- $x[rs1].cursor + imm$  is not in the range  $[x[rs1].base, x[rs1].end - size]$ , where  $size$  is the size (in bytes) of the integer being loaded.
- $x[rs1].cursor + imm$  is not aligned to the size of the integer being loaded.

**If no exception is raised:** The content at the memory location  $[x[rs1].cursor + imm, x[rs1].cursor + imm + size)$  is loaded as a signed integer to  $x[rd]$ , where  $size$  is the size of the integer being loaded (i.e.,  $XLENBYTES/8$ ,  $XLENBYTES/4$ ,  $XLENBYTES/2$ , and  $XLENBYTES$  bytes for `lb`, `lh`, `lw`, and `ld` respectively).

#### Unsigned Load Instructions

`lbu`, `lhu`, `lwu` instructions are modified to load an integer in the size of byte, halfword, and word (i.e.,  $XLENBYTES/8$ ,  $XLENBYTES/4$ , and  $XLENBYTES/2$  bytes), using capabilities respectively. The loaded integer is zero-extended to the size of the destination register (i.e.,  $XLEN$  bits).

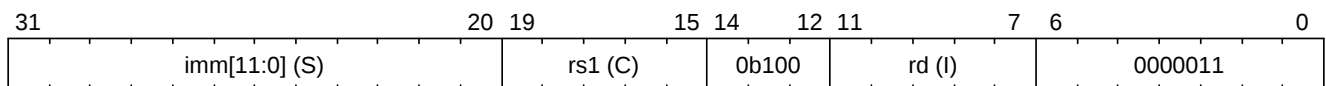


Figure 32. lbu instruction format (I-type)

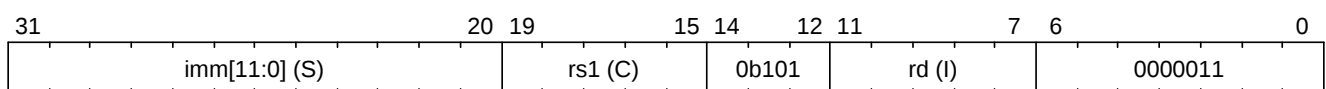


Figure 33. lhu instruction format (I-type)

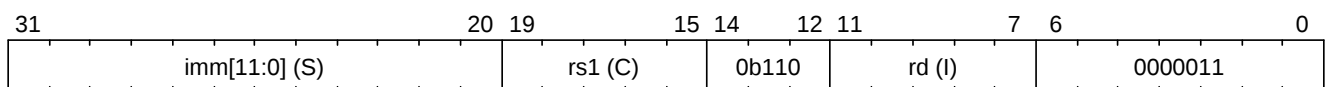


Figure 34. lwu instruction format (I-type)

An exception is raised when any of the following conditions are met:

- $x[rs1]$  is not a capability.
- $x[rs1].type$  is neither 0 (linear) nor 1 (non-linear).
- $x[rs1].valid$  is 0 (invalid).
- $4 \leq x[rs1].perms$  does not hold.
- $x[rs1].cursor + imm$  is not in the range  $[x[rs1].base, x[rs1].end - size]$ , where  $size$  is the size (in bytes) of the integer being loaded.

- $x[rs1].cursor + imm$  is not aligned to the size of the integer being loaded.

**If no exception is raised:** The content at the memory location  $[x[rs1].cursor + imm, x[rs1].cursor + imm + size)$  is loaded as an unsigned integer to  $x[rd]$ , where  $size$  is the size of the integer being loaded (i.e.,  $XLENBYTES/8$ ,  $XLENBYTES/4$ , and  $XLENBYTES/2$  bytes for `lbu`, `lhu`, and `lwu` respectively).

## Store Instructions

`sb`, `sh`, `sw`, `sd` instructions are modified to store an integer in the size of byte, halfword, word, and doubleword (i.e.,  $XLENBYTES/8$ ,  $XLENBYTES/4$ ,  $XLENBYTES/2$ , and  $XLENBYTES$  bytes), using capabilities respectively.

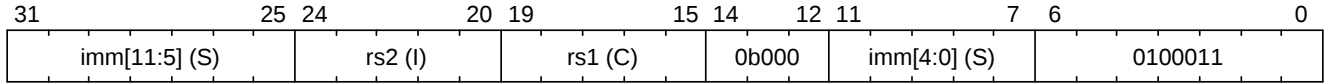


Figure 35. `sb` instruction format (S-type)

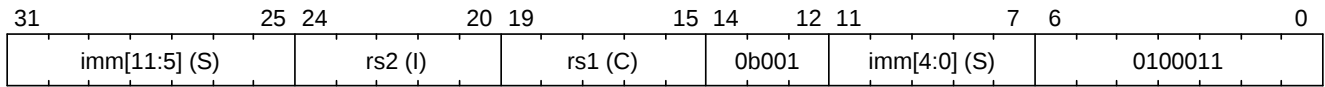


Figure 36. `sh` instruction format (S-type)

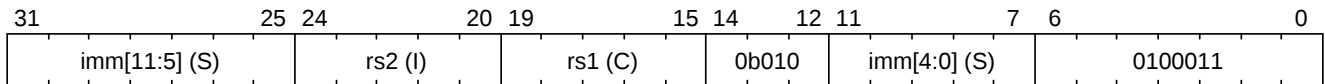


Figure 37. `sw` instruction format (S-type)

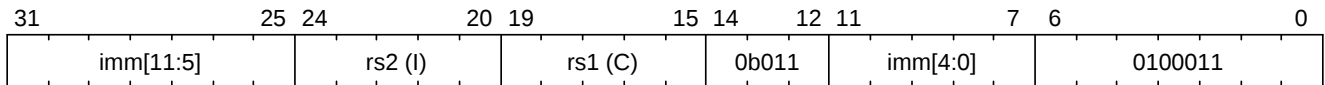


Figure 38. `sd` instruction format (S-type)

**An exception is raised when any of the following conditions are met:**

- $x[rs1]$  is not a capability.
- $x[rs1].type$  is not 0, 1, or 3 (linear, non-linear, or uninitialised).
- $x[rs1].type$  is 3 (uninitialised) and  $imm$  is not 0.
- $x[rs1].valid$  is 0 (invalid).
- $x[rs1].perms$  is neither 6 (read-write) nor 7 (read-write-execute).
- $x[rs1].cursor + imm$  is not in the range  $[x[rs1].base, x[rs1].end - size]$ , where  $size$  is the size (in bytes) of the integer being stored.
- $x[rs1].cursor + imm$  is not aligned to the size of the integer being stored.
- $x[rs2]$  is not an integer.

**If no exception is raised:**

1. The content of  $x[rs2]$  is stored as an integer to the memory location  $[x[rs1].cursor + imm, x[rs1].cursor + imm + size)$ , where  $size$  is the size of the integer being stored (i.e.,  $XLENBYTES/8$ ,  $XLENBYTES/4$ ,  $XLENBYTES/2$ , and  $XLENBYTES$  bytes for `sb`, `sh`, `sw`, and `sd` respectively).

2. If `x[rs1].type` is 3 (uninitialised), `x[rs1].cursor` is set to `x[rs1].cursor + size`.
3. The content in the `CLEN`-bit aligned memory location `[cbase, cend)`, which alias with memory location `[x[rs1].cursor + imm, x[rs1].cursor + imm + size)`, is set to integer type, where `cbase = (x[rs1].cursor + imm) & ~(CLENBYTES - 1)` and `cend = cbase + CLENBYTES`.

### 7.1.2. TransCapstone

In *TransCapstone*, RV64I memory access instructions behave the same as in *Pure Capstone* if `cwld` is 1 (secure world). However, if `cwld` is 0 (normal world), these instructions behave differently in different *encoding modes*.

- When `cwld` is 0 (normal world) and `emode` is 1 (capability encoding mode), these instructions behave the same as in *Pure Capstone*.
- When `cwld` is 0 (normal world) and `emode` is 0 (integer encoding mode), these instructions behave the same as in RV64I except that the following adjustments are made to these instructions:
  - An exception is raised if any of `x[rs1]`, `x[rs2]` and `x[rd]` contains a capability.
  - An exception is raised if the address to be accessed (i.e., `x[rs1] + imm`) is within the range `(SBASE - size, SEND)`, where `size` is the size (in bytes) of the integer to be loaded/stored.
  - For store instructions (i.e., `sb`, `sh`, `sw`, `sd`), the content in the `CLEN`-bit aligned memory location `[cbase, cend)`, which alias with memory location `[x[rs1] + imm, x[rs1] + imm + size)`, is set to integer type, where `cbase = (x[rs1] + imm) & ~(CLENBYTES - 1)` and `cend = cbase + CLENBYTES`.

#### Note

In Capstone-RISC-V, when using these instructions to access the memory location that does not contain an integer-type value, the result of the following operations is undefined:

- Load an integer from the memory location.
- Store an integer to the memory location and then load an integer from the rest of the `CLEN`-bit aligned memory location.

## 7.2. Control Flow Instructions

In RV64I, conditional branch instructions (i.e., `beq`, `bne`, `blt`, `bge`, `bltu`, and `bgeu`), and unconditional jump instructions (i.e., `jal` and `jalr`) are used to control the flow of execution. In Capstone, these instructions are adjusted to support the situation where the program counter is a capability.

### 7.2.1. Branch Instructions

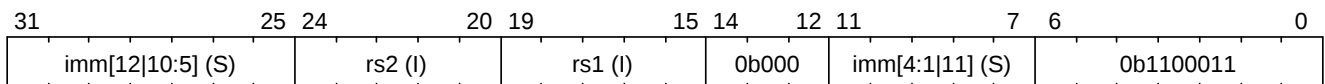


Figure 39. `beq` instruction format (B-type)

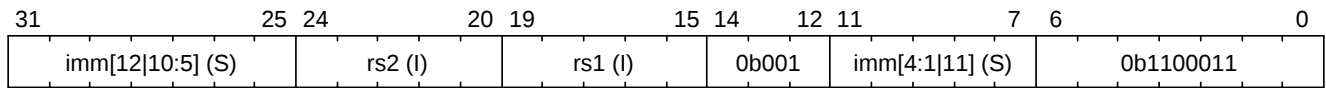


Figure 40. bne instruction format (B-type)

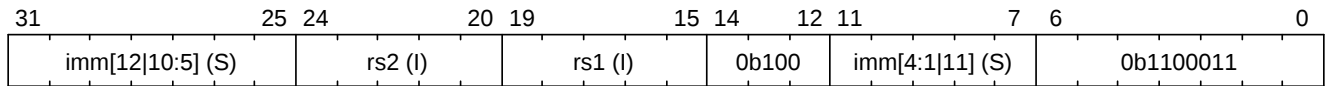


Figure 41. blt instruction format (B-type)

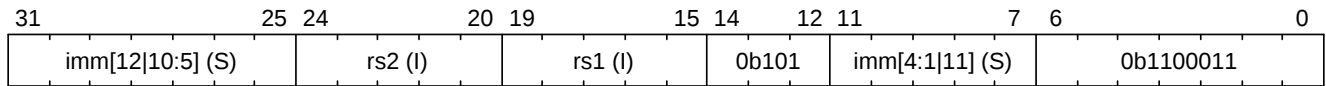


Figure 42. bge instruction format (B-type)

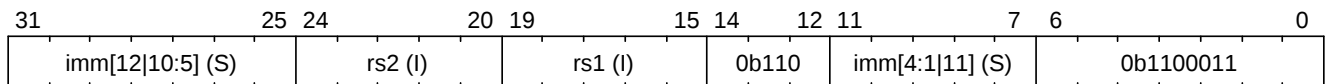


Figure 43. bltu instruction format (B-type)

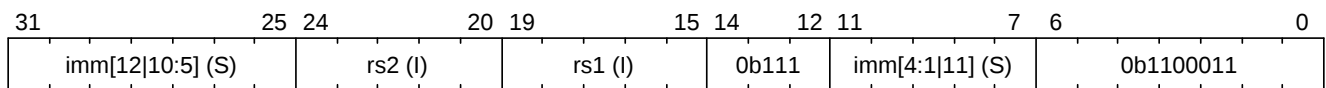


Figure 44. bgeu instruction format (B-type)

**The following adjustments are made to these instructions:**

*Pure Capstone*

- An exception is raised if  $x[rs1]$  or  $x[rs2]$  is a capability.
- $x[pc].cursor$ , instead of  $x[pc]$ , is changed by the instruction.

*TransCapstone*

- An exception is raised if  $x[rs1]$  or  $x[rs2]$  is a capability.
- When `cwd` is 1 (secure world),  $x[pc].cursor$ , instead of  $x[pc]$ , is changed by the instruction.

## 7.2.2. Jump Instructions

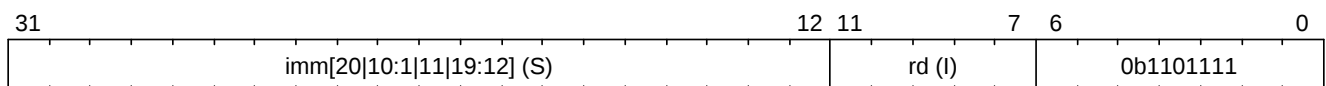


Figure 45. jal instruction format (J-type)

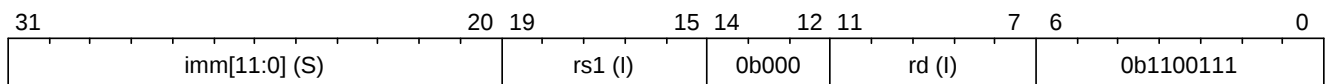


Figure 46. jalr instruction format (I-type)

**The following adjustments are made to these instructions:**

*Pure Capstone*

- An exception is raised if  $x[rs1]$  (if existed) or  $x[rd]$  is a capability.

- `x[pc].cursor`, instead of `x[pc]`, is changed by the instruction.
- `x[pc].cursor` (i.e., the address of the next instruction) is written to `x[rd]`.

### *TransCapstone*

- An exception is raised if `x[rs1]` (if existed) or `x[rd]` is a capability.
- When `cwld` is 1 (secure world), `x[pc].cursor`, instead of `x[pc]`, is changed by the instruction.
- When `cwld` is 1 (secure world), `x[pc].cursor` (i.e., the address of the next instruction) is written to `x[rd]`.

## 8. Interrupts and Exceptions

TODO: add support for nesting

### 8.1. Exception and Exit Codes

#### Note

For Pure Capstone, there is only one place where exception codes are relevant, which is the argument to pass to the exception handler domain. For TransCapstone, however, there are three places where we need to consider some form of exception codes:

1. (Handleable Exception) The argument to pass to the exception handler domain.
2. (Unhandleable Exception) The value returned to the CAPENTER instruction in the user process.
3. (Interrupt) The exception code that the OS sees.

The argument to pass to the exception handler domain will be in the register `a1`, and the `rd` operand of CAPENTER will be the exit code the user process receives.

The *exception code* is what the exception handler domain receives as an argument when an exception occurs on Pure Capstone or in TransCapstone secure world. It is an integer value that indicates what the type of the exception is. TransCapstone also has *exit codes*, which are the values returned to the CAPENTER instruction in case the exception cannot be handled in the secure world. We define the exception code and the exit code for each type of exception below. It aligns with the exception codes defined in RV64I, where applicable, for ease of implementation and interoperability.

Table 8. Exception codes and exit codes for Pure Capstone and TransCapstone secure world

| Exception                      | Exception code | TransCapstone exit code |
|--------------------------------|----------------|-------------------------|
| Instruction address misaligned | 0              | 1                       |
| Instruction access fault       | 1              | 1                       |
| Illegal instruction            | 2              | 1                       |
| Breakpoint                     | 3              | 1                       |
| Load address misaligned        | 4              | 1                       |
| Load access fault              | 5              | 1                       |
| Store/AMO address misaligned   | 6              | 1                       |
| Store/AMO access fault         | 7              | 1                       |
| Unexpected operand type        | 8              | 1                       |
| Invalid capability             | 9              | 1                       |

## Note

Currently, we use the same exit code `1` for all exception types to protect the confidentiality of the secure world execution.

## 8.2. Pure Capstone

For Pure Capstone, the handling of interrupts and exceptions is relatively straightforward. Regardless of whether the event is an interrupt or an exception, or what the type of the interrupt or exception is, the processor core will always transfer the control flow to the corresponding handler domain (specified in the `ceh` register for exceptions and the `cih` register for interrupts). The current context is saved and sealed in a sealed-return capability which is then supplied to the exception handler domain as an argument. When exception handling is complete, the exception handler domain can use the `RETURN` instruction to resume the execution of the excepted domain. This process resembles that of a `CALL-RETURN` pair, except that it is asynchronous, rather than synchronous, to the execution of the original domain.

TODO: specify what "panics" means here

TODO: specify what happens if any of the involved memory accesses fails

### 8.2.1. Handling of Interrupts

TODO: need to specify how to record cause of the interrupt

TODO: interrupt masking

TODO: record the pending interrupts

**The interrupt is ignored if any of the following conditions is met:**

- `cih` is not a capability.
- `cih.valid = 0` (invalid).
- `cih.type != 4` (sealed capability).

**Otherwise:**

1. Load the program counter `pc` from memory location `[cih.base, cih.base + CLENBYTES)`.
2. For `i = 1, 2, ..., 31`, load the content of `x[i]` from memory location `[cih.base + i * CLENBYTES, cih.base + (i + 1) * CLENBYTES)`.
3. Store the original program counter `pc` to the memory location `[cih.base + CLENBYTES, cih.base + 2 * CLENBYTES)`.
4. For `i = 1, 2, ..., 31`, store the *original* content of `x[i]` to memory location `[cih.base + i * CLENBYTES, cih.base + (i + 1) * CLENBYTES)`.
5. Set `cih.type` to `5` (sealed-return), `cih.reg` to `0` (asynchronous), and `cih.async` to `1` (asynchronous).

6. Write `cih` to the register `c1`.
7. Write the exception code to the register `x10`.

### 8.2.2. Handling of Exceptions

The CPU core panics if any of the following conditions is met:

- The `ceh` register does not contain a capability.
- The capability in `ceh` is invalid (`valid = 0`).
- The capability in `ceh` is not a sealed capability (`type != 4`).

Otherwise:

1. Load the program counter `pc` from memory location `[ceh.base, ceh.base + CLENBYTES)`.
2. For `i = 1, 2, ..., 31`, load the content of `x[i]` from memory location `[ceh.base + i * CLENBYTES, ceh.base + (i + 1) * CLENBYTES)`.
3. Store the original program counter `pc` to the memory location `[ceh.base + CLENBYTES, ceh.base + 2 * CLENBYTES)`.
4. For `i = 1, 2, ..., 31`, store the *original* content of `x[i]` to memory location `[ceh.base + i * CLENBYTES, ceh.base + (i + 1) * CLENBYTES)`.
5. Set `ceh.type` to 5 (sealed-return), `ceh.reg` to 0 (asynchronous), and `ceh.async` to 1 (asynchronous).
6. Write the content of `ceh` to the register `c1`.
7. Write the exception code to the register `x10`.

## 8.3. TransCapstone

TransCapstone retains the same interrupt and exception handling mechanisms for the normal world as in RV64I.

For the secure world in TransCapstone, the handling of interrupts and exceptions is more complex, and it becomes relevant whether the event is an interrupt or an exception.

For interrupts, in order to prevent denial-of-service attacks by the secure world, the processor core needs to transfer the control back to the normal world safely. The interrupt will be translated to one in the normal world that occurs at the CAPENTER instruction used to enter the secure world. Since interrupts are typically relevant only to the management of system resources, the interrupt should be transparent to both the secure world and the user process. In other words, the secure world will simply resume execution from where it was interrupted after the interrupt is handled by the normal-world OS.

For exceptions, we want to give the secure world the chance to handle them first. If the secure world manages to handle the exception, the normal world will not be involved. The end result is that the whole exception or its handling is not even visible to the normal world. If the secure world fails to handle an exception (i.e., when it would end up panicking in the case of Pure Capstone, such as when `ceh` is not a valid sealed capability), however, the normal world will take over. The exception

will not be translated into an exception in the normal world, but instead indicated in the exit code that the CAPENTER instruction in the user process receives. The user process can then decide what to do based on the exit code (e.g., terminate the domain in the secure world).

Below we discuss the details of the handling of interrupts and exceptions generated in the secure world.

### 8.3.1. Handling of Secure-World Interrupts

When an interrupt occurs in the secure world, the processor core directly saves the full context, scrubs it, and exits to the normal world. It then generates a corresponding interrupt in the normal world, and follows the normal-world interrupt handling process thereafter.

**If the content in `switch_reg` is a valid sealed capability:**

1. Store the current value of the program counter (`pc`) to the memory location [`switch_cap.base`, `switch_cap.base + CLENBYTES`).
2. For  $i = 1, 2, \dots, 31$ , store the content of `x[i]` to the memory location [`switch_cap.base + i * CLENBYTES`, `switch_cap.base + (i + 1) * CLENBYTES`).
3. Set `switch_cap.async` to 1 (asynchronous).
4. Write the content of `switch_cap` to the register `x[switch_reg]`.
5. Load the program counter `pc` and the stack pointer `sp` from `normal_pc` and `normal_sp` respectively.
6. Scrub the other general-purpose registers.
7. Set the `cwrlld` register to 0 (normal world).
8. Trigger an interrupt in the normal world.

**Otherwise:**

1. Write the content of `cnull` to `x[switch_reg]`.
2. Load the program counter `pc` and the stack pointer `sp` from `normal_pc` and `normal_sp` respectively.
3. Scrub the other general-purpose registers.
4. Set the `cwrlld` register to 0 (normal world).
5. Trigger an interrupt in the normal world.

Note that in this case, there will be another exception in the normal world when the user process resumes execution after the interrupt has been handled by the OS, due to the invalid `switch_cap` value written to the CAPENTER operand.

### 8.3.2. Handling of Secure-World Exceptions

When an exception occurs, the processor core first attempts to handle the exception in the secure world, in the similar way as in Pure Capstone. If this fails (`ceh` is not valid), the processor core saves the full context if it can and exits to the normal world with a proper error code.

**If the content in `ceh` is a valid sealed capability:**

1. Load the program counter `pc` from memory location `[ceh.base, ceh.base + CLENBYTES)`.
2. For  $i = 1, 2, \dots, 31$ , load the content of `x[i]` from memory location `[ceh.base + i * CLENBYTES, ceh.base + (i + 1) * CLENBYTES)`.
3. Store the original program counter `pc` to the memory location `[ceh.base + CLENBYTES, ceh.base + 2 * CLENBYTES)`.
4. For  $i = 1, 2, \dots, 31$ , store the *original* content of `x[i]` to memory location `[ceh.base + i * CLENBYTES, ceh.base + (i + 1) * CLENBYTES)`.
5. Set the `ceh.type` to 5 (sealed-return), and `ceh.async` to 0 (asynchronous).
6. Write the content of `ceh` to the register `c1`.
7. Write the exception code to the register `x10`.

Note that this is exactly the same as the handling of exceptions in Pure Capstone.

**Otherwise:**

**If the content in `switch_reg` is a valid sealed capability:**

1. Store the current value of the program counter (`pc`) to the memory location `[switch_cap.base, switch_cap.base + CLENBYTES)`.
2. For  $i = 1, 2, \dots, 31$ , store the content of the  $i$ -th general purpose to the memory location `[switch_cap.base + i * CLENBYTES, switch_cap.base + (i + 1) * CLENBYTES)`.
3. Set `switch_cap.async` to 1 (asynchronous).
4. Write the content of `switch_cap` to `x[switch_reg]`.
5. Load the program counter `pc` and the stack pointer `sp` from `normal_pc` and `normal_sp` respectively.
6. Write the exit code to `x[exit_reg]`.
7. Set the `cwrlld` register to 0 (normal world).

**Otherwise:**

1. Write the content of `cnull` to `x[switch_reg]`.
2. Load the program counter `pc` and the stack pointer `sp` from `normal_pc` and `normal_sp` respectively.
3. Write the exit code to `x[exit_reg]`.
4. Set the `cwrlld` register to 0 (normal world).

### Note

Compare this with [CAPEXIT](#). We require that CAPEXIT be provided with a valid sealed-return capability rather than use the latent capability in `switch_cap`. This allows us to enforce containment of domains in the secure world, so that a domain is prevented from escaping from the secure world when such a behaviour is undesired.

# 9. Memory Consistency Model

TODO

# Appendix A: Debugging Instructions (Non-Normative)

## A.1. World Switching

The instructions SETWORLD and ONPARTITION are related to world switching in TransCapstone.

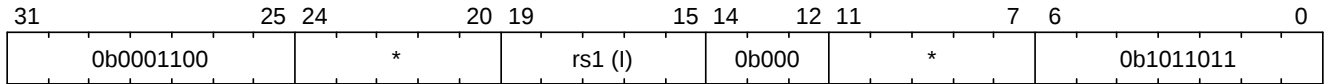


Figure 47. SETWORLD instruction format

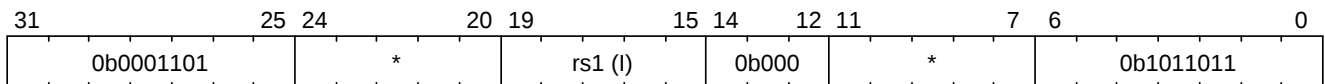


Figure 48. ONPARTITION instruction format

The instructions load their operands from the register  $x[rs1]$ , which expects an integer. SETWORLD directly sets the core to the specified world (0 for normal world and non-zero for secure world). The program counter will also be made into a capability or an integer correspondingly while retaining the `cursor` value. ONPARTITION switches on (non-zero) or off (0) the world partitioning checks in memory.

The instructions make it easy to set up the environment for testing either Pure Capstone or TransCapstone:

- Pure Capstone: secure world, world partitioning checks off
- TransCapstone: normal world, world partitioning checks on

## A.2. Exception Handling

The instructions SETEH and ONNORMALEH affect the behaviours of interrupt and exception handling.

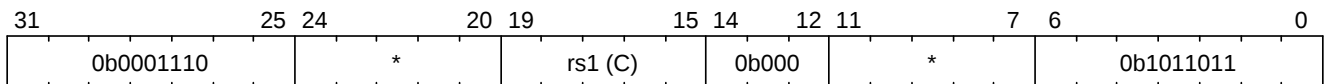


Figure 49. SETEH instruction format

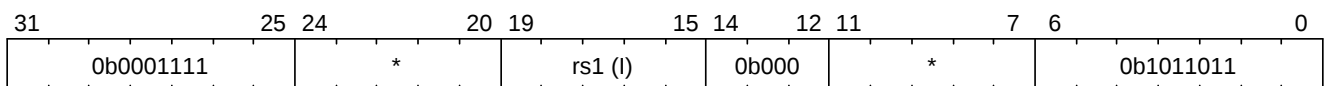


Figure 50. ONNORMALEH instruction format

The SETEH instruction sets the secure-world exception handler domain (i.e., `ceh`) to the specified capability  $x[rs1]$ . The ONNORMALEH instruction checks  $x[rs1]$  and switches on (non-zero) or off (0) normal world handling of secure-world exceptions. When this is on, an exception that occurs in the secure world will trap to the normal world first before being handled by the secure-world exception handler (`ceh`), which is the expected behaviour in TransCapstone. When it is off, the

exception will be directly handled by the secure-world exception handler, as is expected in Pure Capstone.

# Appendix B: Instruction Listing

## B.1. Debugging Instructions

Table 9. Debugging instructions

| Mnemonic    | Format | Func3 | Func7   | rs1 | rs2 | rd | imm[11:0] | World | Variant |
|-------------|--------|-------|---------|-----|-----|----|-----------|-------|---------|
| QUERY       | R      | 000   | 0000000 | I   | -   | -  | -         | *     | *       |
| RCUPDATE    | R      | 000   | 0000001 | I   | -   | I  | -         | *     | *       |
| ALLOC       | R      | 000   | 0000010 | I   | -   | I  | -         | *     | *       |
| REV         | R      | 000   | 0000011 | I   | -   | -  | -         | *     | *       |
| CAPCREATE   | R      | 000   | 0000100 | -   | -   | C  | -         | *     | *       |
| CAPTYPE     | R      | 000   | 0000101 | I   | -   | C  | -         | *     | *       |
| CAPNODE     | R      | 000   | 0000110 | I   | -   | C  | -         | *     | *       |
| CAPPERM     | R      | 000   | 0000111 | I   | -   | C  | -         | *     | *       |
| CAPBOUND    | R      | 000   | 0001000 | I   | I   | C  | -         | *     | *       |
| CAPPRINT    | R      | 000   | 0001001 | I   | -   | -  | -         | *     | *       |
| TAGSET      | R      | 000   | 0001010 | I   | I   | -  | -         | *     | *       |
| TAGGET      | R      | 000   | 0001011 | I   | -   | I  | -         | *     | *       |
| SETWORLD    | R      | 000   | 0001100 | I   | -   | -  | -         | *     | T       |
| ONPARTITION | R      | 000   | 0001101 | I   | -   | -  | -         | *     | T       |
| SETEH       | R      | 000   | 0001110 | C   | -   | -  | -         | *     | T       |
| ONNORMALEH  | R      | 000   | 0001111 | I   | -   | -  | -         | *     | T       |

## B.2. Capstone Instructions

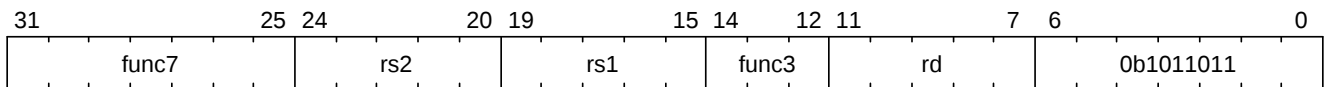


Figure 51. Instruction format: R-type

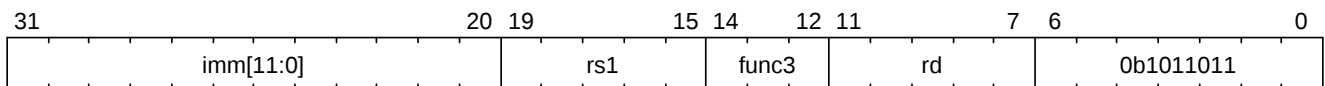


Figure 52. Instruction format: I-type

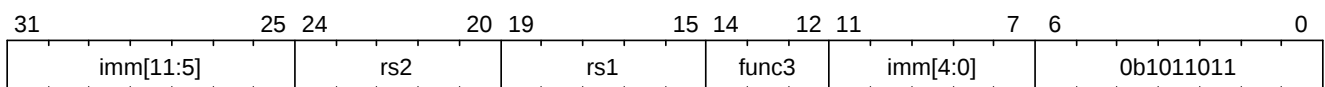


Figure 53. Instruction format: S-type

Table 10. Capability manipulation instructions

| Mnemonic      | Format | Func3 | Func7   | rs1 | rs2 | rd | imm[11:0] | World | Variant |
|---------------|--------|-------|---------|-----|-----|----|-----------|-------|---------|
| REVOKE        | R      | 001   | 0000000 | C   | -   | -  | -         | *     | *       |
| SHRINK        | R      | 001   | 0000001 | I   | I   | C  | -         | *     | *       |
| TIGHTEN       | R      | 001   | 0000010 | C   | Z   | C  | -         | *     | *       |
| DELIN         | R      | 001   | 0000011 | -   | -   | C  | -         | *     | *       |
| LCC           | R      | 001   | 0000100 | C   | Z   | I  | -         | *     | *       |
| SCC           | R      | 001   | 0000101 | I   | -   | C  | -         | *     | *       |
| SPLIT         | R      | 001   | 0000110 | C   | I   | C  | -         | *     | *       |
| SEAL          | R      | 001   | 0000111 | C   | -   | C  | -         | *     | *       |
| MREV          | R      | 001   | 0001000 | C   | -   | C  | -         | *     | *       |
| INIT          | R      | 001   | 0001001 | C   | I   | C  | -         | *     | *       |
| MOVC          | R      | 001   | 0001010 | C   | -   | C  | -         | *     | *       |
| DROP          | R      | 001   | 0001011 | C   | -   | -  | -         | *     | *       |
| CINCOFFSET    | R      | 001   | 0001100 | C   | I   | C  | -         | *     | *       |
| CINCOFFSETIMM | I      | 010   | -       | C   | -   | C  | S         | *     | *       |

Table 11. Memory access instructions

| Mnemonic | Format | emode | Func3 | Func7 | rs1 | rs2 | rd | imm[11:0] | World | Variant |
|----------|--------|-------|-------|-------|-----|-----|----|-----------|-------|---------|
| LDC      | I      | 0     | 011   | -     | I   | -   | C  | S         | N     | T       |
|          | I      | 1     | 011   | -     | C   | -   | C  | S         | N     | T       |
|          | I      | -     | 011   | -     | C   | -   | C  | S         | S     | T       |
|          | I      | -     | 011   | -     | C   | -   | C  | S         | -     | P       |
| STC      | S      | 0     | 100   | -     | I   | C   | -  | S         | N     | T       |
|          | I      | 1     | 100   | -     | C   | C   | -  | S         | N     | T       |
|          | I      | -     | 100   | -     | C   | C   | -  | S         | S     | T       |
|          | I      | -     | 100   | -     | C   | C   | -  | S         | -     | P       |

Table 12. Control flow instructions

| Mnemonic | Format | Func3 | Func7   | rs1 | rs2 | rd | imm[11:0] | World | Variant |
|----------|--------|-------|---------|-----|-----|----|-----------|-------|---------|
| CALL     | R      | 001   | 0100000 | C   | -   | -  | -         | S     | T       |
|          | R      | 001   | 0100000 | C   | -   | -  | -         | -     | P       |
| RETURN   | R      | 001   | 0100001 | C   | I   | -  | -         | S     | T       |
|          | R      | 001   | 0100001 | C   | I   | -  | -         | -     | P       |
| CJALR    | I      | 101   | -       | C   | -   | C  | S         | S     | T       |
|          | I      | 101   | -       | C   | -   | C  | S         | -     | P       |

| Mnemonic | Format | Func3 | Func7   | rs1 | rs2 | rd | imm[11:0] | World | Variant |
|----------|--------|-------|---------|-----|-----|----|-----------|-------|---------|
| CBNZ     | I      | 110   | -       | I   | -   | C  | S         | S     | T       |
|          | I      | 110   | -       | I   | -   | C  | S         | -     | P       |
| CAPENTER | R      | 001   | 0100010 | C   | -   | I  | -         | N     | T       |
| CAPEXIT  | R      | 001   | 0100011 | C   | I   | -  | -         | S     | T       |

Table 13. Control state instructions

| Mnemonic | Format | Func3 | Func7 | rs1 | rs2 | rd | imm[11:0] | World | Variant |
|----------|--------|-------|-------|-----|-----|----|-----------|-------|---------|
| CCSRRW   | I      | 111   | -     | C   | -   | C  | Z         | *     | *       |

## B.3. Extended RV64I Memory Access Instructions

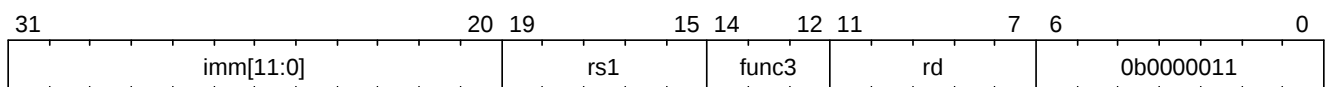


Figure 54. Instruction format: I-type

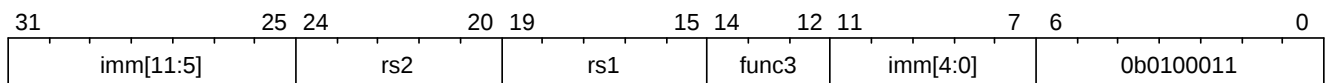


Figure 55. Instruction format: S-type

Table 14. Extended RV64I load instructions

| Mnemonic | Format | emode | Func3 | Func7 | rs1 | rs2 | rd | imm[11:0] | World | Variant |
|----------|--------|-------|-------|-------|-----|-----|----|-----------|-------|---------|
| lb       | I      | 0     | 000   | -     | I   | -   | I  | S         | N     | T       |
|          | I      | 1     | 000   | -     | C   | -   | I  | S         | N     | T       |
|          | I      | -     | 000   | -     | C   | -   | I  | S         | S     | T       |
|          | I      | -     | 000   | -     | C   | -   | I  | S         | -     | P       |
| lh       | I      | 0     | 001   | -     | I   | -   | I  | S         | N     | T       |
|          | I      | 1     | 001   | -     | C   | -   | I  | S         | N     | T       |
|          | I      | -     | 001   | -     | C   | -   | I  | S         | S     | T       |
|          | I      | -     | 001   | -     | C   | -   | I  | S         | -     | P       |
| lw       | I      | 0     | 010   | -     | I   | -   | I  | S         | N     | T       |
|          | I      | 1     | 010   | -     | C   | -   | I  | S         | N     | T       |
|          | I      | -     | 010   | -     | C   | -   | I  | S         | S     | T       |
|          | I      | -     | 010   | -     | C   | -   | I  | S         | -     | P       |
| ld       | I      | 0     | 011   | -     | I   | -   | I  | S         | N     | T       |
|          | I      | 1     | 011   | -     | C   | -   | I  | S         | N     | T       |
|          | I      | -     | 011   | -     | C   | -   | I  | S         | S     | T       |
|          | I      | -     | 011   | -     | C   | -   | I  | S         | -     | P       |

| Mnemonic | Format | emode | Func3 | Func7 | rs1 | rs2 | rd | imm[11:0] | World | Variant |
|----------|--------|-------|-------|-------|-----|-----|----|-----------|-------|---------|
| lbu      | I      | 0     | 100   | -     | I   | -   | I  | S         | N     | T       |
|          | I      | 1     | 100   | -     | C   | -   | I  | S         | N     | T       |
|          | I      | -     | 100   | -     | C   | -   | I  | S         | S     | T       |
|          | I      | -     | 100   | -     | C   | -   | I  | S         | -     | P       |
| lhu      | I      | 0     | 101   | -     | I   | -   | I  | S         | N     | T       |
|          | I      | 1     | 101   | -     | C   | -   | I  | S         | N     | T       |
|          | I      | -     | 101   | -     | C   | -   | I  | S         | S     | T       |
|          | I      | -     | 101   | -     | C   | -   | I  | S         | -     | P       |
| lwu      | I      | 0     | 110   | -     | I   | -   | I  | S         | N     | T       |
|          | I      | 1     | 110   | -     | C   | -   | I  | S         | N     | T       |
|          | I      | -     | 110   | -     | C   | -   | I  | S         | S     | T       |
|          | I      | -     | 110   | -     | C   | -   | I  | S         | -     | P       |

Table 15. Extended RV64I store instructions

| Mnemonic | Format | emode | Func3 | Func7 | rs1 | rs2 | rd | imm[11:0] | World | Variant |
|----------|--------|-------|-------|-------|-----|-----|----|-----------|-------|---------|
| sb       | S      | 0     | 000   | -     | I   | I   | -  | S         | N     | T       |
|          | S      | 1     | 000   | -     | C   | I   | -  | S         | N     | T       |
|          | S      | -     | 000   | -     | C   | I   | -  | S         | S     | T       |
|          | S      | -     | 000   | -     | C   | I   | -  | S         | -     | P       |
| sh       | S      | 0     | 001   | -     | I   | I   | -  | S         | N     | T       |
|          | S      | 1     | 001   | -     | C   | I   | -  | S         | N     | T       |
|          | S      | -     | 001   | -     | C   | I   | -  | S         | S     | T       |
|          | S      | -     | 001   | -     | C   | I   | -  | S         | -     | P       |
| sw       | S      | 0     | 010   | -     | I   | I   | -  | S         | N     | T       |
|          | S      | 1     | 010   | -     | C   | I   | -  | S         | N     | T       |
|          | S      | -     | 010   | -     | C   | I   | -  | S         | S     | T       |
|          | S      | -     | 010   | -     | C   | I   | -  | S         | -     | P       |
| sd       | S      | 0     | 011   | -     | I   | I   | -  | S         | N     | T       |
|          | S      | 1     | 011   | -     | C   | I   | -  | S         | N     | T       |
|          | S      | -     | 011   | -     | C   | I   | -  | S         | S     | T       |
|          | S      | -     | 011   | -     | C   | I   | -  | S         | -     | P       |

## Note

For instruction operands:

## I

Integer register

## C

Capability register

## S

Used as sign-extended immediate

## Z

Used as zero-extended immediate

-

Not used

### **For immediates:**

## S

Sign-extended

## Z

Zero-extended

-

Not used

### **For worlds:**

## N

Normal world

## S

Secure world

\*

Either world

### **For variants:**

## P

*Pure Capstone*

## T

*TransCapstone*

\*

Either variant

# Appendix C: Assembly Code Examples

TODO

# Appendix D: Abstract Binary Interface (Non-Normative)

TODO