

ATGP_RISC-V: Automation of Test Generator using Pluggy for RISC-V Architecture

B Madhavan
madhavansankar007@gmail.com
http://orcid.org/0000-0001-6119-9074

A Kamerish
kamerish160@gmail.com
http://orcid.org/0000-0001-8637-8191

R Manimegalai
drmm@psgitech.ac.in
http://orcid.org/0000-0003-1398-4080

Department of Computer Science and Engineering
PSG Institute of Technology and Applied Research, Coimbatore, India.

Abstract— *The reduced instruction set computing (RISC) architecture is a free and open Instruction Set Architecture (ISA), which enables a new era of processor innovation through open standard collaboration. It directly challenges several well-established processor families such as intel x-86, Motorola 68k processor. To thrive an RISC-V ecosystem, the core suppliers need an independent verification solution to ensure that their designs are compliant with the ISA specification. Verification of RISC-V designs become challenging due to their optional features, implementation flexibility, and provisions for customer extensions. Hence, a thorough verification is essential to compete successfully against the established processor families. Automation is the key for reducing the time taken for the processor verification. This paper provides a way to develop an automated tool ATGP_RISC-V, which uses the same arguments to run all instruction generators. This helps in verifying the processor in an efficient way by reducing the time taken to manually compare the test results.*

Keywords—*RISC-V, instruction, exceptions, VCS, verification, testbed, functional verification, instruction set architecture.*

I Introduction

Functional verification [7] is an important aspect of the integrated circuit (IC) design cycle. It is vital that the design is functionally verified and any potential bug is eliminated at an early stage of processor design. Processor verification is done in various methods and stages. Functional verification is the task of verifying the logic design completely with the rules of specification. Formal verification helps to mathematically check the technical requirements. Intelligent verification is an automation to adapt the testbench changes in the transfer level code. Emulation is the process of building a version of design using programmable logic. Various techniques of functional verification are as follows; static verification is done to reduce the verification effort at the Register Transfer Logic (RTL) level. Static verification is the process of

verifying the design against some predefined rules without knowing the actual working of design.

Techniques involved in static verification [9] can contain any of the following like reset domain crossing, clock domain crossing, lint, formal verification and static checks.

Functional simulation is done to verify the individual Intellectual Property (IP) or the individual blocks of the IC. Functional simulation is the process of verifying the functional behaviour of a design by simulating it in software. Field Programmable Gate Array (FPGA) prototyping is the process of verifying the functionality of the system (IC) on FPGAs and it is done to verify that the design operates as expected when it is driven with live data. Emulation is also called as pre-silicon validation. Emulation is done on a hardware device to verify the functionality of the system. An emulator can handle both RTL designs and system-level designs. Emulation uses live data to find issues in the system level design. Universal Verification Methodology (UVM) has a well-defined testbench structure with predefined set of coding guidelines [12]. It provides a System Verilog Base Class Library (BCL) for building advanced reusable verification component. Intellectual Property (IP) is highly complex and it takes time to verify it completely.

The RISC-V architecture is proposed and designed in the University of California, Berkeley [2]. The RISC-V process has a wide range of flexibility such as variable width size with three word-widths 32, 64 and 128 bits. It also has plenty of base parts with additionally added extension for the ease of both developer and user. The RISC-V processor operates in one of the following modes:

1. U-Mode, which stands for user mode, which is the lowest privileged level.
2. S-Mode, which stands for Supervisor Mode.
3. M-Mode, which stands for machine mode and has the highest privilege.

Typically, user-level applications will execute in user Mode. Whenever the application wants an OS service, it will make a system call. The OS code that

handles this call will execute in Supervisor Mode, i.e., at a higher privileged level. Upon return to the application, the mode will be lowered back to user mode. The RISC-V architecture has 17 extensions and the standard extensions are specified to work with all of the standard bases [1] and with each other without conflict. Some of the extensions are listed below

- i) M denotes the extension for Integer multiplication and division
- ii) A denotes the extension for atomic Instruction
- iii) F denotes the extension for single precision floating point
- iv) D denotes the extension for double precision Floating point
- v) G denotes the shorthand for some of the extensions like base
- vi) C denotes the extension for compressed instructions.
- vii) Q denotes the extension for quad-precision Floating point.
- viii) L denotes the extension for decimal floating-point
- ix) B denotes the extension for bit manipulation
- x) J denotes the extension for dynamically translated languages
- xi) T denotes the extension for transactional memory
- xii) P denotes the extension for packed SIMD instructions
- xiii) V denotes the extension for vector operations
- xiv) N-denotes the extension for user-level interrupts and high level interrupts
- xv) H denotes an extension for the level interrupts in H-standard

II Related Work

Imperas developed a tool named Open Virtual Platform (OVP) for RISC [13]. This simulator is a multiprocessor emulator used to run unchanged production binaries of the target hardware. The riscvOVPSim simulator implements the full and complete functionality of the RISC-V foundation's user and privilege specifications. It has the ability to simulate nearly a billion instructions in a very short period of time. Ivannikov institute for system programming has developed a tool for processor verification [5]. This work presents MicroTESK, a tool that automates the construction of test program generators for microprocessors. MicroTESK for RISC-V is an Instruction Stream Generator (ISG) aimed at functional verification of RISC-V microprocessors. A constructed generator consists of the core that implements architecture-independent generation methods and the model that holds information required to generate tests for the corresponding architecture

[10]. MicroTESK extracts this information from formal specifications of the instruction set architecture to get the assembly format of the instructions in order to build the coverage model of the instruction set architecture. The coverage model is used to construct the instruction set simulator which is used as a reference model. Test programs are generated from test templates, describing the program's structural and behavioural properties. Onespın solutions has developed tools for RISC coverage test. Coverage test technology is a common software testing technology, which is the basic requirement of software testing. The coverage analysis can quantify the completeness of the test vector. Onespın [8] translates functional requirements in a formal and simulation executable format. It captures entire circuit transactions in a concise way similar to timing diagrams [6]. Yang, Yawen and Zhou have developed a tool which uses onespın technology to compare two kinds of mainstream coverage analysis techniques, code coverage and functional coverage. The covered code, uncovered code and the software defect in the test results are analyzed. Simple-uvms is a verification methodology that is developed based on UVM (Universal Verification Methodology) standard library. It builds a reference model through the aspect-oriented paradigm. It also generates the high functional coverage test cases based on the knowledge at the transaction level.

RISC-V DV

RISC-V DV is an UVM based open-source instruction generator for RISC processor verification written in system verilog [4]. The workflow of RISC-V is shown in Fig-3. It supports RV64IMAFDC, RV32IMAFDC instruction set and supported privileged modes are user mode, machine mode, and supervisor mode. It is built in Python on top of System Verilog. Python will take care of the instruction generation by using System Verilog files. List of instructions needed for all available extensions is predefined in the System Verilog files. Target files are written in System Verilog to specify the ISA about which instruction set is used for generating instructions. There is a reasonable amount of YAML files included in the working of a RISC-V DV instruction generator. These files specify a set of rules for executing an instruction generator. We specify the rules and the mode of occurrence of instruction in the yaml file via an attribute called gen-opts. Other functions like instruction count (no of instructions in the output file), no of fragments into which the output should be divided, no of branches in the instruction, number of exceptions (including user-defined and naturally occurring exception) generated per 100 instruction

generation and more. *VCS* is an RTL (Register transfer logic) simulator. It is a tool used for compiling verilog source code into object (.o) files. *VCS* invokes the C compiler (cc, GCC) to create an executable file which simulates your design. *Spike*, which is also called as Instruction Set Simulator (ISS) implements a functional model of one or more RISC-V harts. *Spike* supports the verbose mode to decode the generated instructions.

III ATGP-RISC-V: The Proposed Methodology for Test Program Generation

The flow of the proposed methodology for test program generation is illustrated in Figure. 1.

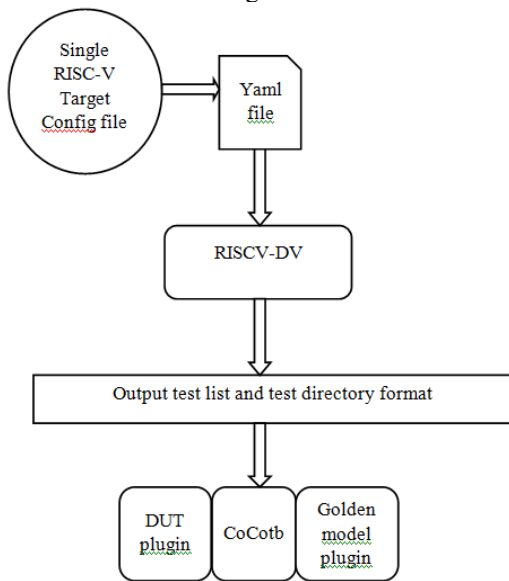


Figure 1. The Proposed Verification Methodology for RISC-V Processor

First the target file, which is the file on which the whole program will work on will be chosen using runtime interface. Then the yaml file corresponding to the target config file will be chosen. The files with .yaml extensions are the files which specifies the nature of the input and the input values for the system. These yaml files has a specific standard syntax associated with it. This file is used by the assembler and it decodes this file with that prescribed syntax. Since this file is used for supplying the input values, at-most care should be taken during its fabrication. *RISCV-DV* is the tool which is involved in the test program generation. The output test-list consists of two stages called assembly unit test and torture testing. In assembly unit test [11], sanitary testing occurs where the basic functionality of the assembly code is tested. In torture test, we have a mix of logical sequences which is executed upon the output file from the tool. This makes the output look more attractive. Design under test (DUT) [3] is a product which performs testing of a file. The DUT testing is classified into initial testing and life cycle testing. In initial testing, the first phase of the design is tested. In life cycle testing, review is made in the later stages. This report also shows the errors in the assembly code, and ways to correct it. After successful testing, the code with the corresponding output is made into a single executable file called the golden model plugin. This workflow is shown in Figure. 1.

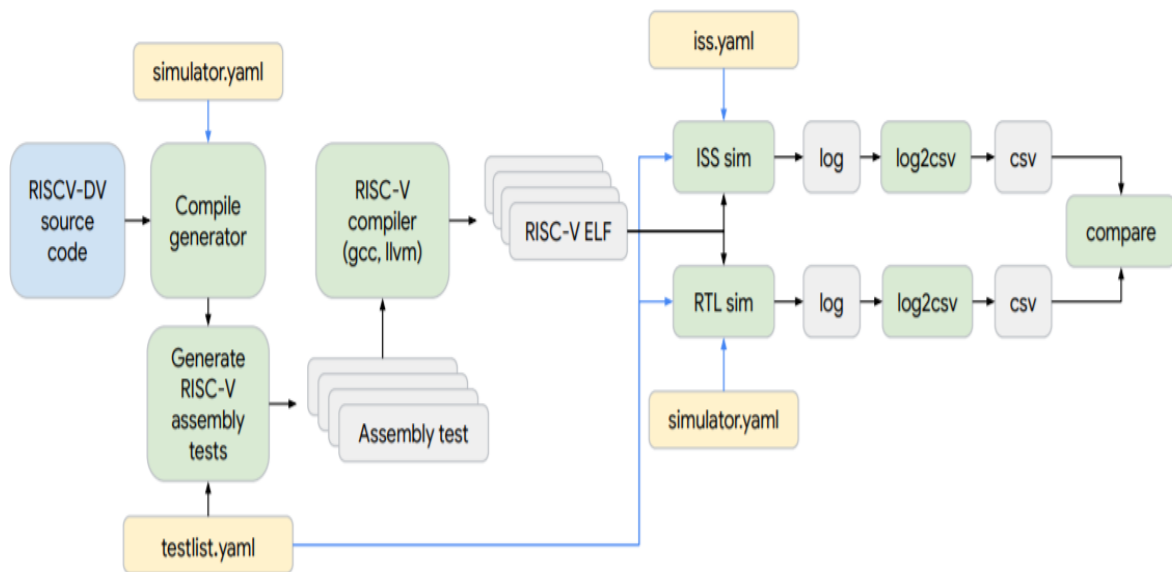


Figure 2. Workflow of RISC-V DV

A. Instructions

This paper focuses on generating instructions for the RISC-V processor in an optimal manner. The proposed methodology takes a seed value from the user or assigns a random seed value for the generator to generate instructions. The RISC-V ISA has four base instruction set namely RV32I, RV64I, RV32E, and RV128I. The base specifies instructions, control flow, registers with their sizes, memory and addressing logic manipulation, and ancillaries. The overall workflow of RISC-V DV is shown in Figure. 2. The initial phase is executing the file named simulator.yaml which simulates the code into the RISC-V core. Then it moves to generate assembly test instruction. The type of instruction to be generated is referred from the file testlist.yaml. This testlist is fed from Instruction Set Simulator (ISS). The same testlist is also fed from a Register Transfer Level (RTL) in parallel to the ISS. ISS and RTL generates log, log2csv, and CSV files. The assembly tests are fed into the compiler and an elf file will be generated to act as the supplier for ISS and RTL. Note that the ISS and RTL files are used in generating assembly test reports and to compare the spawned assembly codes. Simulation of both the results are compared for accuracy of the generated assembly code instruction. This workflow is shown in Figure. 2.

B. Exceptions

At the time of execution of instructions, an exception may occur. There are many types of exceptions which may occur in runtime. But in general, a total of fourteen exceptions are defined: Three memory exceptions are defined for memory access namely fetch, load, store/AMO, and nine misaligned address, page fault (address translation), access fault (physical memory attributes and protection), illegal instruction, CSR access rights violations, non-existing or reserved opcodes and encodings, other instructions in unprivileged mode (call/return), breakpoint (fetch, load, store of debugged address) , environment call, three separate exceptions based on originating mode (M, S, U).

Exceptions are classified into two types. They are synchronous and asynchronous exceptions. Synchronous exceptions are listed below:

- instruction_access_fault
- illegal_instruction
- breakpoint
- load_address_misaligned
- load_access_fault

- ecall_mmode
- ecall_umode, ecall_smode.

Asynchronous exceptions i.e. interrupts, are listed below:

- timer interrupt
- software interrupt
- external interrupt.

There are exception handlers which handles the exception when generated. An exception can be handled or ignored. A trap occurs when an exception is handled, The trap processing involves a transfer of control from a place where the trap has occurred to a trap handler routine. The trap handler routine is shown in Table 2. Trap processing consists of some hardware operations, such as modifying a couple of hardware lags, saving the PC, and effecting a transfer of control to the first instruction of the trap handler routine. These exception handlers maintain the flow of the program execution and prevents the program from sudden termination. Exception handler handles the exception and return back to the same place where the exception is generated and so, order of execution is also maintained.

Table-2: List of Exception Handlers

Exceptions handlers	Function
trap_load_access_fault	There is no file in the desired folder
illegal_instr_handler	For Unimplemented and illegal instructions
trap_misaligned_exception	If the target address is not 4-byte aligned
trap_instruction_access_fault	Access to this user is not valid
mmode_instr_handler	Trying to operate from u-mode or s-mode
store_fault_handler	When the store instruction is misaligned
load_fault_handler	When the load instruction is misaligned
instr_fault_handler	When the given instruction results into an error
ebreak_handler	When control is not properly transferred to the debugging environment
mtvec_handler	when mtvec has a writable value
pt_fault_handler	When the exception relates to an error

Exceptions are rare events that are triggered by the hardware and force the processor to execute an exception handler. There are exception handlers in the host file which gets invoked when instruction generation overrides the bounds given by the user. These exception handlers ensure that the execution process is not interrupted. There are many kind of handlers like `trap_illegal_access_exception` which gets generated when user account changes during the execution of the instruction. `load_access_fault_handler` is called to manage the scenario where the source file is missing from the given destination. `ebreak_handler` is standard used by GCC-compiler to stop the execution flow and return back into the debugger. This also marks the code that should not get executed during the program flow. Another main purpose of `ebreak` is that it supports semihosting. This is a type of host which alters the execution environment such that it has a debugger initialised inside it that provides an alternate service to an interface which is around `ebreak`.

IV Implementation for Test Program Generation

We start the execution of the program from the file named `run.py` in the command line. This python file can get the following parameters in the command line (`instr_cnt`, `num_of_sub_program`, `no_fence`, `no_data_page`, `boot_mode`, `no_csr_instr`). These generator options can also be given in the `testlist.yaml`. `Testlist` contains type of tests that can be executed on a particular given data. The test instruction is passed using a specific language. The python file finds the appropriate test specified in the test list YAML file and searches for the target in the target directory. If no test is specified in the command line then all the test gets executed. Custom targets can be added in the RISC-V DV framework to generate instructions for a specific extension but the targets created should be included in the `run.py` python file. `Run.py` file is the root file for all the process occurring in the system. It contains the list of test generators and codes for handling the runtime environment according to the test input. First it calls the `setup_parser` function which is used for setting the arguments which is to be used in runtime. Then, it calls `get_generator_cmd` where it looks into the file `testlist.yaml` which is given as input. `Run.py` then peaks into the template file and finds the matching simulator. If no simulator is found, then it returns

with the message stating that it cannot find the specified vcs simulator. After finding the matching test and target specified, it compiles the testlist file with the specified test and converts the information into binary (machine readable) format. If compilation is successful, then it invokes `do_simulate` function where simulation starts for the output generation. The seed value is created in simulation phase. This seed value is used to retrieve (or) regenerate the same report which is executed at that time when seed value is generated. The output may vary time to time and the seed value acts as the identity for every simulation. Then the `gen` function comes into role where it instructs the compiler to setup the compile and simulation command for the generator. Then it compiles and runs the instruction generator. It follows with `iss` simulation, compilation, and generation. The next stage is the compilation stage where we generate three types of files namely `asm`, `elf`, and `binary`. `asm` file is the output in Hexadecimal format. The generated instructions are stored in the directory specified in the `run.py` file. This can also be manipulated to set a custom output directory. The output directory contains `asm-test` folder which contains `binary`, `object` and `source` file. Each output also contains a seed value used to generate the same type of instructions.

A. Automated Test Program generation using Python-pluggy

Pluggy is the solidified core of plugin management and also a hook calling system for `pytest`. This executes as a part of normal program execution, and also enhances certain features of it. This promotes the flexibility of the user to greater heights where user gains privilege to expand and modify the characteristics of the host program. This is achieved by installing a plugin for that program. The basic motive of `pluggy` is to split the given code into fragments so that dependency of user on the output increases efficiently. `Pluggy`, in the core form can be classified as internal and external `pluggy` modules. The internally created `pluggy` hold the rules defined by the developer which should be followed to operate on it. The external `pluggy` is the area for the user to create an own `pluggy` which is a derived version of `pluggy` based on the rules defined by the developer. `Pluggy` is also used for intergerating various different types of components under one roof. Here we have intergerated `riscv ovp-sim`, `aapg`, instruction generation and exception handler when any error occurs during the generation of instruction.

B. Overview of Pluggy Components

The Host program contains hook functions and their implementation which is part of the program. ATPG will implement the prescribed hooks. It also participates in the execution of the program whenever these implementations are approaching the host. ATPG connects host and plugins by hook implementations, hook specifications and also by hook callers. Hook implementations are taken from the registered plugins. Hook specifications defines call signatures given by the host. Hook caller acts as a call loop which is triggered at appropriate positions in the program. The triggers are invoked in the host during the implementations and the results are collected. The host.py file contains the list of plugins available and moves to the specified area by getting a choice as input from the user. Pluggy contains the host program and the necessary components that host uses in it. Both the pluggy and host has its own path setting variable. It is generally advised that plugin and the host to be installed in the same environment to avoid unnecessary errors. The readme file in the pluggy clearly defines the users on how to work with pluggy. It helps users in difficult times like when there occurs an error which cannot be understood by the user.

C. RISC-V DV Pluggy

The overview of the files present in RISC-V DV pluggy is shown in figure-4. This pluggy contains the main file for instruction generation and the files associated with the set of instructions. The pluggy created here is termed as Automated Test Program Generator (ATPG). The RISC-V DV pluggy contains an input file which specifies the characteristics of the output file and other optional parameters such as number of instructions, amount of exceptions, jump instructions and more. The init.py file starts the code simulation which contains the input filename and the exact location of the file in the ATPG. The readme file in ATPG contains a short description about the contents in it and the mode of operation done on it. It is the sensitive part of the setup because the user operates on the ATPG tool based on the understanding they get from this file. The instruction generation starts from the triggering done in the file named instruction. Here, the input file is processed and the desired attributes are selected. Then the output directory is locked where the results are getting stored with the filename which gives the information about the generated output. The output location will not be reside in the

setup since the generated files depend on the user and so any changes made are reflected only in user's local directory. ATPG uses VCS simulator and spike tools which should be pre-installed in the hardware where the execution takes place. A seed value generated for each execution of a test file. Seed value acts as a unique id for the generated test. Even when the same test is executed after some time, the seed value for the second test gets changed. This seed value, if specified in the input module will generate the same instruction as the previous one.

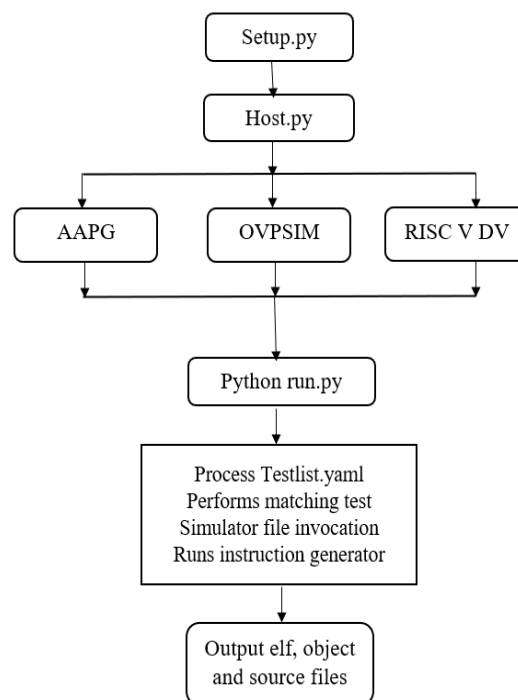


Figure 3: Pluggy Workflow

V Experimental Results

Components of ATPG contains setup file and the ATPG plugin folder. The plugin contains host, hookspecs and lib files. ATPG automates the tests program generation using the concept of pluggy. The components of ATPG is shown in Figure 4. Sample ebreak handler has been shown in Figure 5. This handler is invoked during the execution of the program when it encounters a trap named trap breakpoint. This is shown in figure 5.

```

    setup.py
    ATPG/ init_.py
    ATPG/hookspecs.py
    ATPG/host.py
    ATPG/lib.py
    ATPG.ieg-info/PKG-INFO
    ATPG.ieg-info/SOURCES.txt
    ATPG.ieg-info/dependency_links.txt
    ATPG.ieg-info/entry_points.txt
    ATPG.ieg-info/requirements.txt
    ATPG.ieg-info/top_level.txt
  
```

Figure 4. File Components of ATPG


```

ebreak_handler:
    csrr x10, mepc
    addi x10, x10, 4
    csrw mepc, x10
    lw x1, 4(x24)
    lw x2, 8(x24)
    lw x3, 12(x24)
    lw x4, 16(x24)
    lw x5, 20(x24)
    lw x6, 24(x24)
    lw x7, 28(x24)
    lw x8, 32(x24)
    lw x9, 36(x24)
    lw x10, 40(x24)
    lw x11, 44(x24)
    lw x12, 48(x24)
    lw x13, 52(x24)
    lw x14, 56(x24)
    lw x15, 60(x24)
    lw x16, 64(x24)
    lw x17, 68(x24)
    lw x18, 72(x24)
    lw x19, 76(x24)
    lw x20, 80(x24)
    lw x21, 84(x24)
    lw x22, 88(x24)
    lw x23, 92(x24)
    lw x24, 96(x24)
    lw x25, 100(x24)
    lw x26, 104(x24)
    lw x27, 108(x24)
    lw x28, 112(x24)
    lw x29, 116(x24)
    lw x30, 120(x24)
    lw x31, 124(x24)
    addi x24, x24, 124
    add x26, x24, zero
    csrrw x24, 0x340, x24
    mret
    
```

Figure 5. Sample code for ebreak handler

```

core 0: 0xffffffff80000482 (0x00008e65) c.and a2, s1
core 0: 0xffffffff80000484 (0x019903a3) sb s9, 7(s2)
core 0: 0xffffffff80000488 (0xff695303) lhu t1, -10(s2)
core 0: 0xffffffff8000048c (0xfeb90d23) sb a1, -6(s2)
core 0: 0xffffffff80000490 (0x00009002) c.ebreak
core 0: exception trap breakpoint, epc 0xffffffff80000490
core 0: tval 0xffffffff80000490
core 0: 0x0000000080010000 (0x0400006f) j pc + 0x40
core 0: 0xffffffff80010040 (0x34011173) csrrw sp, mscratch, sp
core 0: 0xffffffff80010044 (0x000d0133) add sp, s10, zero
core 0: 0xffffffff80010048 (0xf8410113) addi sp, sp, -124
core 0: 0xffffffff8001004c (0x0000c206) c.swsp ra, 4(sp)
    
```

Figure 6. Occurrence of an Exception

Figure 6 shows a part of the generated test program where the exception has been generated and a trap handler has been invoked. Now the flow of control will be transferred to exception handler shown in figure 5. After the successful execution of the trap handler, the control once again returns to the test program and continues the execution. ATPG, by using the concept of pluggy has automated the tedious process by running various test

consecutively. Hence, ATPG solves the problem in a robust way.

VI Conclusion

The proposed RISC-V project defines and describes a standard instruction set architecture (ISA). RISC-V is an open-source specification for computer processor architecture. Based on the performance and the growing need for interoperability among vendors, it appears that the

RISC-V standard will gain more importance in future. RISC-V architecture holds good scope in the field of processors in the upcoming months. This work also focus on handling few pre-defined exceptions with instruction generation methods that results in minimising the no of errors. But the rate of exception generation is a major issue in these processors and it has been solved here by changing the exception handlers and also the user instruction generation methods. This paper deals with the tool which has been developed under the concept of python-pluggy. It automates the test program generation processes and reduces the human effort and time involved in processor verification. There are certain limitations in this work such that it cannot work on the errors other than the error described previously. It is also not capable of analysing the cause for errors. Also there is no automated program to automatically define the new errors. Error definition is still manual in this work.

References

- [1] Andrew Waterman et al., The RISC-V Instruction Set Manual, Volume I: UserLevel ISA, Version 2.1,online report available at <https://www2.eecs.berkeley.edu/Pubs/TechRpts/2016/EECS-2016-118.pdf>
- [2] Andrew Waterman et al.,The RISC-V Instruction Set Manual Volume II: Privileged Architecture Version 1.7, online report available at <http://www2.eecs.berkeley.edu/Pubs/TechRpts/2015/EECS-2015-49.html>.
- [3] Victor Jimenez Arador, S. in Verification Strategy for a RISC-V Core Design, in polytechnic University of Catalonia 2019, pp, 2-4.
- [4] Billie Thompson , riscv-dv, online tool available at <https://github.com/google/riscv-dv>.
- [5] M. Chupilko, A. Kamkin, A. Kotsynyak, A. Protsenko, S. Smolov and A. Tatarnikov, Test Program Generator MicroTESK for RISC-V, 19th International Workshop on Microprocessor and SOC Test and Verification (MTV), 2018, pp. 6-11.
- [6] Yang, Yawen & Zhou, Shan & Kong, Lu. Coverage Test Technology Based on ONESPIN Verification Platform. Journal of Physics: Conference Series. 2018, 1026.012004. pp. 5-6.
- [7] Xie, Z. & Wang, T. & Yong, S. & Chen, X. & Su, J. & Wang, X. A RISC CPU oriented reusable functional verification platform based on UVM. 2018, 50. 221-227. PP 2-3.
- [8] Yang, Yawen & Zhou, Shan & Kong, Lu. Coverage Test Technology Based on ONESPIN Verification Platform. Journal of Physics: Conference Series. 2018, 1026.012004.
- [9] Aijaz Fatima , The ABCs of functional verification techniques, online report available at <https://www.analogictips.com/what-are-abcs-of-functional-verification-techniques/>
- [10] Chupilko, Mikhail & Kamkin, Alexander & Kotsynyak, Artem & Tatarnikov, Andrei. (2017). MicroTESK: Specification-Based Tool for Constructing Test Program Generators. 217-220.
- [11] N. Gala, A. Menon, R. Bodduna, G. S. Madhusudan and V. Kamakoti, SHAKTI Processors: An Open-Source Hardware Initiative, 29th International Conference on VLSI Design (VLSID), 2016, pp. 7-8.
- [12] Gala, Neel & Madhusudan, Gs & George, Paul & Sahoo, Anmol & Menon, Arjun & Kamakoti, SHAKTI: An Open-Source Processor Ecosystem. Volume 2, issue 3, 2018. PP 4-5.
- [13] Gajendra Kumar Ranka, Dr. Manoj Kumar Jain , a validation of sim-a with ovpsim, Journal of Global Research in Computer Science Volume 2, No. 6, June 2011. pp. 2.