MIPS Computer System on FPGA

Presented by:
Mostafa Abd El-Aziz

Supervisor:
Prof. Dr. Layla Abou Hadid

Presented to:
Alexandria University
Faculty of Engineering
Computer and Systems Engineering Department

July 2016

Abstract

Acknowledgements

```
            ,----------------,              ,---------,
        ,-----------------------,          ,"        ,"|
      ,"                      ,"|        ,"        ,"  |
     +-----------------------+  |      ,"        ,"    |
     |  .-----------------.  |  |     +---------+      |
     |  |                |  |  |     | -==----'|      |
     |  |  WE LOVE MIPS  |  |  |     |         |      |
     |  |  QUAF> _       |  |  |/----|`---=    |      |
     |  |                |  |  |   ,/|==== ooo |      ;
     |  |                |  |  |  // |(((( [33]|    ,"
     |  `-----------------'  |," .;'| |((((     |  ,"
     +-----------------------+  ;;  | |         |,"
        /_)______________(_/  //'   | +---------+
   ___________________________/___  `,
  /  oooooooooooooooo  .o.  oooo /,   \,"-----------
 / ==ooooooooooooooo==.o.  ooo= //   ,`\--{)B     ,"
/_==__==========__==_ooo__ooo=_/'   /___________,"
`-----------------------------'
```

Contents

Abstract ii

Acknowledgements iii

1 Introduction 1
- 1.1 Overview . . . 1
- 1.2 Scope of Project . . . 3
- 1.3 Overview of Document . . . 3

2 Related Work 6

3 Architecture 8
- 3.1 Top-Level Components . . . 8
- 3.2 Central Processing Unit . . . 10
  - 3.2.1 Introduction . . . 10
    - 3.2.1.1 Load and Store Instructions . . . 11
    - 3.2.1.2 ALU Instructions . . . 12
    - 3.2.1.3 Jump and Branch Instructions . . . 15
    - 3.2.1.4 Miscellaneous Instructions . . . 17
    - 3.2.1.5 Control Instructions . . . 18
    - 3.2.1.6 Instruction Encoding Summary . . . 19
    - 3.2.1.7 Addressing Modes . . . 21
    - 3.2.1.8 Programmer-Visible Pipeline Effects . . . 21
  - 3.2.2 Registers . . . 22
    - 3.2.2.1 General-Purpose Registers . . . 23
    - 3.2.2.2 LO and HI Registers . . . 23
    - 3.2.2.3 Control Registers . . . 24
  - 3.2.3 Memory Management . . . 29
    - 3.2.3.1 MIPS Logical Address Space . . . 31
    - 3.2.3.2 Translation Look-aside Buffer (TLB) . . . 32
    - 3.2.3.3 Cache Memory . . . 37
  - 3.2.4 Exception Handling . . . 37
    - 3.2.4.1 Exception Cases . . . 38
    - 3.2.4.2 Exception Vectors . . . 39
    - 3.2.4.3 Exception Protocol . . . 40
    - 3.2.4.4 Precise Exceptions . . . 41
- 3.3 System Bus . . . 41
- 3.4 Bus Decoder . . . 43

3.5 Memory Interface 43
3.6 VGA Interface 44
3.6.1 Character Memory 45
3.6.2 Attribute Memory 46
3.6.3 Font Memory 48
3.6.4 Special Registers 50
3.7 Interrupt Controller 51
3.8 Programmable Timer 52
3.9 Keyboard Controller 53

4 Implementation 56

5 Software 58

6 Results 60

7 Conclusion 62
7.1 Future Work 62

A Instruction Listing 64

List of Figures

1.1 Overview of the computer system . . . 2
1.2 System layers . . . 2

3.1 Top-level components of the system. . . . 9
3.2 MIPS-I instruction formats . . . 19
3.3 Instructions encoded by opcode field . . . 20
3.4 Instructions encoded by function field when opcode field = *SPECIAL* . . 20
3.5 Instructions encoded by rt field when opcode = *REGIMM* . . . 21
3.6 All possible addressing modes . . . 21
3.7 Conventional names/uses of MIPS registers . . . 23
3.8 Control registers . . . 24
3.9 Format of Index register . . . 24
3.10 Format of EntryLo register . . . 25
3.11 Format of BadVaddr register . . . 25
3.12 Format of EntryHi register . . . 26
3.13 Format of Status register . . . 26
3.14 Format of Cause register . . . 28
3.15 ExcCode values with their meaning . . . 28
3.16 Format of EPC register . . . 28
3.17 Memory access levels . . . 29
3.18 Process logical address spaces and mappings to physical addresses . . . 30
3.19 Address translation using page table . . . 33
3.20 Address translation using two-level page table scheme . . . 34
3.21 Format of TLB entry . . . 35
3.22 Excetion vectors for various exception types . . . 40
3.23 Memory map. . . . 43
3.24 VGA decoding mechanism. . . . 44
3.25 VGA color codes. . . . 46
3.26 Hello world example: text. . . . 47
3.27 Hello world example: colors. . . . 48
3.28 Hello world example: results on screen. . . . 48
3.29 Shape of letter 'A' in font memory. . . . 49
3.30 VGA special registers. . . . 50
3.31 IRQ lines and their connected devices. . . . 51
3.32 Interrupt controller registers. . . . 51
3.33 Programmable timer registers. . . . 53
3.34 PS/2 keyboard scan codes . . . 54
3.35 Keyboard controller registers . . . 54

Chapter 1

Introduction

The past few years have witnessed great development in the area of field programmable gate array (FPGA) devices. The high flexibility and performance of FPGAs make them a suitable choice for systems with high constraints. The main purpose of our project is to build a simple computer system using a cheap FPGA device.

1.1 Overview

In this project we shall use FPGA technology to implement a general-purpose computer machine based on MIPS-I instruction set architecture. We define the term 'general-purpose computer' as a computer that should be able to perform most common computing tasks, like editing documents, performing calculations, and programming.

The system consists of several components, like CPU, memory, input devices, and output devices, all connected together to build up a general-purpose computer. The hardware part of the machine is implemented using the FPGA chip. The software part is a set of programs stored in the memory of the system.

The CPU is an implementation for MIPS-I instruction set architecture with a 5-stage pipeline. CPU, along with other hardware components, is implemented in VHDL (VHSIC Hardware Description Language) while software is written in MIPS assembly and C programming language. There is also an emulator for the machine written in C language for GNU/Linux.

Figure 1.1 shows the basic conception of the proposed system. The system has a CRT/LCD screen used to output information to the user, and a keyboard used for input. Power is regulated using an AT-compatible power supply. The FPGA board contains an FPGA chip, which is automatically configured with our hardware design on start up.

The FPGA board also contains other components, including a ROM chip, a RAM chip, LEDs, switches, VGA connector, and a PS/2 keyboard connector. Our hardware design (that lies inside the FPGA chip) is responsible for interfacing all these components. The software part of our machine is stored on the ROM chip.

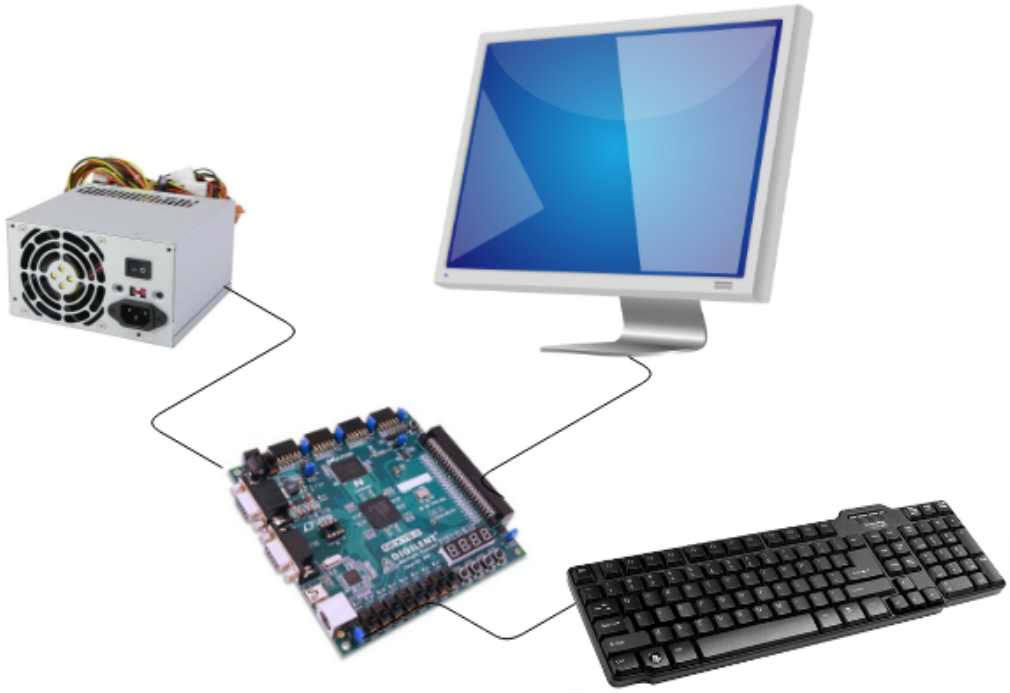


Figure 1.1: Overview of the computer system

The system actually consists of several layers: hardware, firmware, and software. User interacts with applications that runs over the operating system. Figure 1.2 shows the layered architecture of the system.

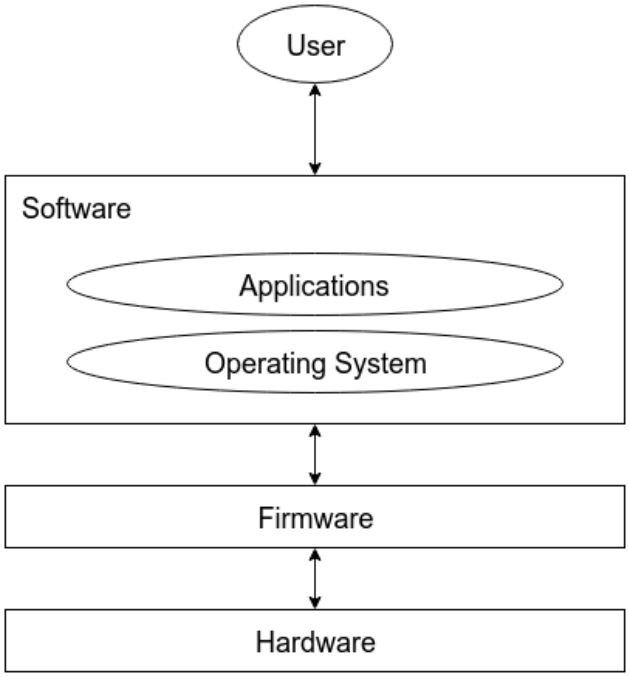


Figure 1.2: System layers

When the computer boots up, CPU starts executing firmware instructions that are stored on the ROM chip on the board. Firmware initializes the hardware components and loads the operating system into RAM. The OS boots up the machine and provides a user-friendly interface to the end user.

1.2 Scope of Project

As we mentioned in the overview section, the goal of this project is to build a general-purpose computer system based on MIPS architecture using FPGA technology. Users shall communicate with the system using a command-line interface. Since the machine is 'programmable', in the sense that end user can write programs, compile, and execute them, the machine can be used for any purpose. However, because the resources of the FPGA are very limited, the machine doesn't have a graphical user interface.

1.3 Overview of Document

In this document, we describe the architecture and the implementation of the proposed computer system in details. First, we shall define two keywords that will be used throughout the document:

- **Architecture**: The behaviour of the processor as seen by the programmer. This includes the instruction set, their opcodes and formats, architectural registers, interrupt/exception protocols, and other behavioral concerns that will be seen later in the section. The architecture describes all that can be abstracted from the implementation of the processor. Other authors may call this an instruction set architecture (ISA) or macro-architecture.

- **Micro-architecture**: How the architecture of the processor is implemented in switching circuits. We use this keyword to refer to our VHDL implementation for the architecture described in this section. There may exist several different implementations for the same architecture. The only restriction on the micro-architecture of a processor is to behave exactly as described by the architecture, but the implementation itself doesn't matter. For example, the architecture might guarantee that the 'effect' of the execution of a set of instructions will be in order, however, the processor might execute the instructions in an out of order fashion (OoOE) yet effectively behaves the same as described by the architecture.

We use the word 'architecture' to refer to the architecture of the processor and the whole system. On the other hand, the word 'implementation' is used to refer to the micro-architecture of the processor and the whole system. Indeed, there are two implementations for the system: one using FPGA, and the other is emulator software.

The book is organized in chapters, summarized in the following lines:

- **Architecture**: Explains the architecture of the whole system, and the architecture of system components, including CPU, system bus, memory, and I/O devices.

- **Implementation**: Describes how the architecture is implemented on FPGA using VHDL language.
- **Software**: Shows the architecture of the firmware and the operating system.
- **Results**: Screenshots and sample runs for the system.
- **Conclusion**: Brief conclusion for our work and suggestions for future work.

Chapter 2

Related Work

Chapter 3

Architecture

In this chapter we introduce the high-level design of the machine we propose in this project. We use the word 'architecture' to refer to the programming model of the machine (the design of the machine as seen by the programmer), as we've already explained in chapter 1.

First, we will show the major *architectural* components of the system. Next, we describe every component in a separate section. We refer to our machine using the word 'system', since it can be seen as a set of interacting components whose behaviour can be described using black box model. This chapter is dedicated for getting into this black box; in order to explore the architecture (the behaviour and the interface) of every component by itself.

3.1 Top-Level Components

Similar to the majority of modern digital computers, our computer consists of a centeral processing unit, connected to memory and I/O devices through system bus. The I/O devices include a circuit for interfacing PS/2 keyboard (called keyboard controller), a programmable timer, interrupt controller, and a video graphics adapter (VGA) component.

Figure 3.1 shows the architectural components of the system (yet it hides many details). The components are summarized in the following points:

- **Central Processing Unit (CPU)**: The CPU is actually the bus master. It has the capability of driving the system bus at any time. CPU fetches/writes data from/to memory and I/O devices through system bus. When CPU needs to access a word in memory (or an I/O device), CPU puts the address corresponding to the word in memory or the target I/O device on address bus. The device shall respond to the access request by reading/writing data from/to data bus.
- **System Bus**: The bus through which all components of the system are connected. The bus consists of three sub-buses:
  - Address bus (32 lines): always driven by the CPU. Contains the address of the data word being read/written by CPU.

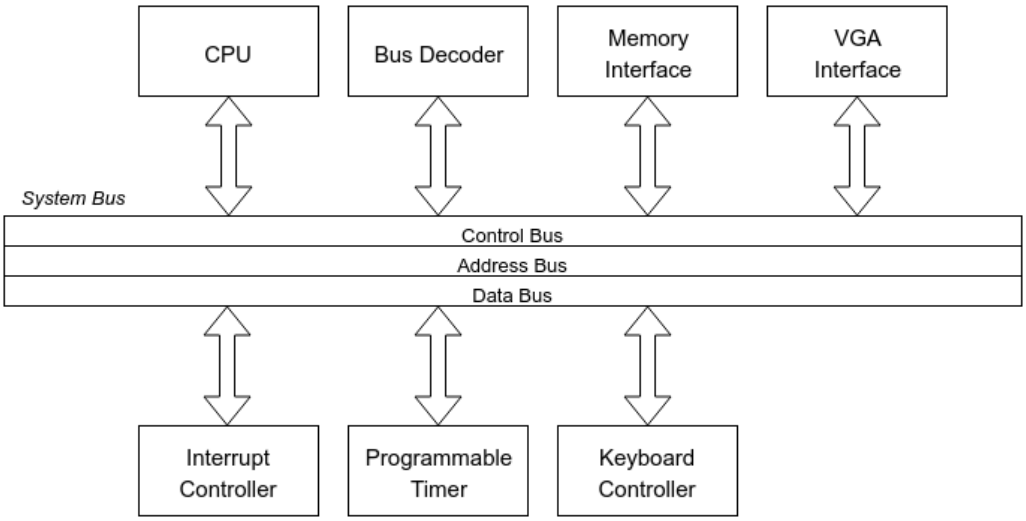


Figure 3.1: Top-level components of the system.

  - Data bus (32 lines): holds the data during an access cycle.
  - Control bus: set of control signals controlled by the processor and address decoder during an access cycle.

- **Bus Decoder**:

  This component keeps listening to the address bus to control which device should respond to read/write request issued by CPU. The decoder takes the most significant bits of the address as input, and controls a number of chip select (CS) signals for all I/O devices in the system. A device whose CS signal is active is the one that should respond to CPU.

- **Memory Interface**:

  This is a circuit that works as an interface between our system and a CMOS pseduo-static 16MB RAM chip used as main memory. Another ROM chip is also connected to the back side of the circuit. The circuit can be considered as a translation between the language of our system bus and the pinout of the memory chips.

- **VGA Interface**:

  This is an interface between our system and a VGA-compatible connector. It is used to provide a method of displaying text data on a CRT/LCD screen that supports VGA interface.

- **Interrupt Controller**:

  Provides a protocol for interrupting the processor by various I/O devices. Several interrupt lines are fed into the interrupt controller. The output of the controller is an interrupt signal that is delivered to the processor. When interrupt service routine (ISR) is executed, software should ask the interrupt controller for more information about the interrupt source.

- **Programmable Interval Timer**:

  Used by the operating system to keep track of the time of day. Can be programmed to trigger an interrupt every specific interval configured by the OS. Details of the architecture of the interval timer will be seen later in this chapter.

- **Keyboard Controller**:

  Provides an interface between the system and an external PS/2 keyboard that is used as an input device. It is also connected to the interrupt controller through an interrupt signal (to interrupt the CPU whenever a key is pressed down).

3.2 Central Processing Unit

The CPU module implements MIPS-I instruction set architecture along with additional instructions that control the behavior of the processor. In this section we present detailed description for the programming interface of the CPU, followed by a description for the implementation of that interface.

3.2.1 Introduction

CPU instruction (also called order, operation) is the basic unit of execution. The main task of a processor is to execute a sequence of instructions (i.e, program) which shall have effect on the state of the processor and devices connected with the processor.

Following Von Neumann's model, the program is stored in memory along with variables which are accessible and modifiable by the program. Memory is an array of words, where a word in MIPS is 32 bits. Every instruction in MIPS is encoded in a single word that contains the opcode and the operands of the instruction.

Every word in memory has an address, that also means that every instruction in memory has an address called 'instruction address'. CPU maintains a register that stores the instruction address of next instruction to fetch, called program counter (PC). When an instruction is fetched, program counter is increased by 4 (number of bytes in word) without touching lower 2 bits (they are always zero), a behaviour which implies that:

- Unless there is a branch, a jump, or an exception, instructions are always executed in sequence, as opposed to machines with one-plus-one addressing scheme like RPC-4000 where a program could be scattered over the drum.
- Instructions should always be stored in 4-byte aligned words.

It is important to note that MIPS-I instruction set is a proper subset of the instruction set of our processor; meaning that the processor supports additional instructions and registers that control the behaviour of the processor as will be seen below.

CPU instructions are divided into the following groups:

1. Load and Store

2. ALU
3. Jump and Branch
4. Miscellaneous
5. Control

3.2.1.1 Load and Store Instructions

This group holds all instructions related to memory access, either read/load or write/store accesses. Every instruction word contains the following fields:

- **Opcode**: specifies that this instruction is a memory instruction, along with whether this instruction is a byte, half-word, or a word instruction, and whether this instruction is an aligned/unaligned, load/store, and signed/unsigned instruction (6 bits).
- **Source/destination register**: If the instruction is a load instruction, this is the address of the register where the value loaded from memory will be stored. If the instruction is a store instruction, this is the address of the register where the value stored to memory will be loaded. Since there are 32 data registers specified by MIPS ISA, this field takes a value from 0 to 31 (5 bits).
- **Base register**: address of data register that contains the base memory address. Just like source/destination register field, this field takes a value from 0 to 31 (5 bits).
- **Index immediate**: A 16-bit signed integer that is extended to 32-bit signed integer then added to the value of the base register. Result is the memory address that will be accessed.

This instruction group has a total of 12 instructions, each instruction has a unique value for the opcode field:

1. **LB**: Load Byte
2. **LBU**: Load Byte Unsigned
3. **SB**: Store Byte
4. **LH**: Load Halfword
5. **LHU**: Load Halfword Unsigned
6. **SH**: Store Halfword
7. **LW**: Load Word
8. **SW**: Store Word
9. **LWL**: Load Word Left
10. **LWR**: Load Word Right

11. **SWL**: Store Word Left
12. **SWR**: Store Word Right

The detailed behaviour of those instructions is described in the appendix.

As described above, the target memory address is calculated by adding the extended 32-bit signed immediate to the value of the base register, forming a 32-bit memory address. This implies that memory address space of MIPS-I instruction set is 4GB. There are two characteristics that apply to the address field:

- **Alignment**: All instructions (except LWL, LWR, SWL, and SWR) are aligned. This means that LW and SW target address must be a multiple of 4, while LH, LHU, and SH target address must be a multiple of 2.
- **Endianness**: Our processor is little endian; the least significant byte of a word in memory is the byte having the lowest address, and the most significant byte is the byte having the highest address.

3.2.1.2 ALU Instructions

This group contains all instructions related to arithmetic and logical operations. We divide them into 4 categories, according to how they are encoded:

1. ALU instructions with an immediate operand
2. ALU instructions with three operands
3. Shift instructions
4. Multiply and divide instructions.

We call the last three groups together as 'R-TYPE ALU instructions' for reasons illustrated below. In the next lines we describe the instruction groups in detail.

- **ALU Instructions With Immediate Operand:**

  The reader may expect that the first group (ALU instructions with immediate operand) is encoded exactly the same way as memory instructions. Both groups are included in the same encoding class (called I-TYPE) as will be seen later in the encoding subsubsection. Just like memory instructions, an ALU instruction with an immediate operand shall consist of the following fields:

  - **Opcode**: specifies that this instruction is an ALU with immediate operand instruction, along with the type of operation. (6 bits)
  - **Source register**: except for LUI, all operations in this group are binary operations, thus this field contains the address of the register that contains the value of the first parameter of the binary operation. For LUI instruction, this field must be zero. (5 bits)

- **Destination register**: address of data register where the result of the operation will be stored. (5 bits)
- **Index immediate**: except for ANDI, ORI, XORI, and LUI instructions, this field is a 16-bit signed integer that is extended to 32-bit signed integer then used as the second parameter for the binary operation. For ANDI, ORI, and XORI, this field is a 16-bit unsigned integer that is extended to 32-bit unsigned integer then used as the second parameter for the binary operation. For LUI, this field is just copied into the highest 16 bits of the destination register.

Following is a listing for all ALU with immediate operand instructions specified by MIPS-I ISA, with every instruction having a unique value for the opcode field:

1. **ADDI**: Add Immediate Word
2. **ADDIU**: Add Immediate Unsigned Word
3. **SLTI**: Set on Less Than Immediate
4. **SLTIU**: Set on Less Than Immediate Unsigned
5. **ANDI**: And Immediate
6. **ORI**: Or Immediate
7. **XORI**: Exclusive Or Immediate
8. **LUI**: Load Upper Immediate

The detailed behaviour of those instructions is described in the appendix.

- **R-TYPE ALU Instructions:**

The remaining three groups of ALU instructions (ALU instructions with three operands, shift instructions, and multiply/divide instructions) are all encoded the same way which is shown below. All instructions encoded that way are called R-TYPE instructions. We will summarize all encoding classes later in this section.

- **Opcode**: always zero. (6 bits)
- **Source register 1**: except for SLL, SRL, SRA, SLLV, SRLV, SRAV, MFHI, MTHI, MFLO, and MTLO instructions, this field contains the address of the register that holds the value of the first parameter for the binary operation. For SLLV, SRLV, and SRAV, this field contains the address of the register that holds the amount of shift. For SLL, SRL, SRA, MFHI, and MFLO instructions, this field must be zero. For MTHI and MTLO instructions, this field contains the address of the register that holds the value that will be moved into the HI/LO register. (5 bits)
- **Source register 2**: except for SLL, SRL, SRA, SLLV, SRLV, SRAV, MFHI, MTHI, MFLO, and MTLO instructions, this field contains the address of the register that holds the value of the second parameter for the binary operation.

For SLL, SRL, SRA, SLLV, SRLV, and SRAV instructions, this field contains the address of the register that holds the value to be shifted. For MFHI, MTHI, MFLO, and MTLO instructions, this field must be zero. (5 bits).

- **Destination register**: except for MULT, MULTU, DIV, DIVU, MFHI, MTHI, MFLO, and MTLO instructions, this field contains the address of the data register where the result of the binary operation or shift operation will be stored. For MFHI and MFLO instructions, value stored in LO/HI register will be moved into the register addressed by this field. For MULT, MULTU, DIV, DIVU, MTHI, and MTLO instructions, this field must be zero, since the result of these instructions is stored directly in LO register, HI register, or both. (5 bits)
- **Shift amount**: for SLL, SRL, and SRA instructions, this field simply contains the amount of shift. For all other instructions, this field must be zero. (5 bits)
- **Function code**: An extension to the opcode field, contains a unique code that defines the instruction to be performed. (6 bits).

With a unique function code for every instruction (and an opcode of 0), ALU instructions with 3 operands group includes the following instructions:

1. **ADD**: Add Word
2. **ADDU**: Add Unsigned Word
3. **SUB**: Subtract Word
4. **SUBU**: Subtract Unsigned Word
5. **SLT**: Set on Less Than
6. **SLTU**: Set on Less Than Unsigned
7. **AND**: And operation
8. **OR**: Or operation
9. **XOR**: Exclusive Or operation
10. **NOR**: Nor operation

The third group (shift instructions) contains the following instructions:

1. **SLL**: Shift Word Left Logical
2. **SRL**: Shift Word Right Logical
3. **SRA**: Shift Word Right Arithmetic
4. **SLLV**: Shift Word Left Logical Variable
5. **SRLV**: Shift Word Right Logical Variable
6. **SRAV**: Shift Word Right Arithmetic Variable

Finally, following is a listing for all instructions in the last group (multiply/divide instructions):

1. **MULT**: Multiply Word

2. **MULTU**: Multiply Unsigned Word
3. **DIV**: Divide Word
4. **DIVU**: Divide Unsigned Word
5. **MFHI**: Move From HI
6. **MTHI**: Move To HI
7. **MFLO**: Move From LO
8. **MTLO**: Move To LO

For detailed description for the behaviour of all of the instructions listed above, please refer to the appendix.

3.2.1.3 Jump and Branch Instructions

MIPS-I architecture defines a number of operations that control the path of execution. Jump instruction shall unconditionally modify PC register, while branch instructions modify PC register only when the condition specified within the instruction word is tautology.

Jump and branch instructions are divided into four groups:

1. Jump Instructions Jumping Within a 256MB Region
2. Jump Instructions to Absolute Address
3. PC-Relative Conditional Branch Instructions Comparing 2 Registers
4. PC-Relative Conditional Branch Instructions Comparing Against Zero

Next we illustrate every group in detail.

- **Jump Instructions Jumping Within 256MB Region:**

  The way the instructions of the first group (jump instructions jumping within 256MB region) are encoded is different from I-TYPE and R-TYPE. It's called J-TYPE and consists of the following fields:

  - **Opcode**: specifies that the instruction is a jump instruction of J-TYPE, along with whether the instruction should alter the link register or not. (6 bits)
  - **Instruction Index**: a 26-bit integer that is multiplied by 4 to form the lowest 28-bit of the new value of PC register. The highest 4 bits are taken from the old value of PC, that's why the instructions of this group are said to jump within 256MB region (4GB/16).

  This group only contains two instructions:

  1. **J**: Jump
  2. **JAL**: Jump and Link

- **Jump Instructions to Absolute Address:**

  Second group (jump instructions to absolute address) consists of instructions similar to instructions of the first group, yet have the ability to jump to any memory location, not just locations inside the current 256MB region. These instructions are encoded as R-TYPE like most ALU instructions, thus the instruction word shall consist of the following fields:

  - **Opcode**: always zero. (6 bits)
  - **Source register 1**: contains address of the register that holds the target 32-bit address. (5 bits)
  - **Source register 2**: must be zero. (5 bits)
  - **Destination register**: for JALR instruction, this field contains the address of the register where return address shall be stored. For JAL instruction, this field should be zero.
  - **Shift amount**: must be zero. (5 bits)
  - **Function code**: An extension to the opcode field, contains a unique code that defines the instruction to be performed (only jump or jump with link). (6 bits).

  Analogous to the first group, this group consists of these two instructions:

  1. **JR**: Jump Register
  2. **JALR**: Jump and Link Register

- **Conditional Branch Instructions Comparing 2 Registers:**

  The remaining two groups are concerned with conditional branching. The third group consists of instructions that compare two registers, then jump if the comparison yields a specific result, while the last group compare one register against zero, and jump if the comparison yields a specific result.

  The third group is just encoded as I-TYPE (like ALU immediate instructions and load/store instructions). The fields of the instruction word are interpreted as following:

  - **Opcode**: specifies that this instruction is a conditional branch instruction comparing 2 registers, along with the type of comparison to be performed. (6 bits)
  - **Source register**: this field contains the address of the register that contains the value that will be used as the first operand for comparison. (5 bits)
  - **Destination register**: except for BLEZ and BGTZ instructions, this field contains the address of the register that contains the value that will be used as the second operand for comparison. For BLEZ and BGTZ instructions, this field must be zero. (5 bits)

- **Index immediate**: a 16-bit signed integer that is extended to a 32-bit signed integer, then added to PC to form the effective address. Thus this group of instructions are 'PC-relative'.

The group consists of the following instructions:

1. **BEQ**: Branch on Equal
2. **BNE**: Branch on Not Equal
3. **BLEZ**: Branch on Less Than or Equal to Zero
4. **BGTZ**: Branch on Greater Than Zero

- **Conditional Branch Instructions Comparing Against Zero:**

The fourth group is also encoded as I-TYPE instructions, however, the fields shall have special values and meanings:

- **Opcode**: always one. (6 bits)
- **Source register**: this field contains the address of the register that contains the value that will be compared against zero. (5 bits)
- **Destination register**: used as an extension to the opcode field. This field shall contain a unique identifier for the operation to be performed. (5 bits)
- **Index immediate**: a 16-bit signed integer that is extended to a 32-bit signed integer, then added to PC to form the effective address. Thus this group of instructions are 'PC-relative'.

Finally, this last group consists of the following instructions:

1. **BLTZ**: Branch on Less Than Zero
2. **BGEZ**: Branch on Greater Than or Equal Zero
3. **BLTZAL**: Branch on Less Than Zero and Link
4. **BGEZAL**: Branch on Greater Than or Equal Zero and Link

For detailed description for the behaviour of all of the instructions listed above, please refer to the appendix.

3.2.1.4 Miscellaneous Instructions

This group of instruction consists of instructions used to trigger exceptions by software. In MIPS-I ISA, this type of instructions is called 'exception instructions'.

When the CPU encounters one of the instructions in this group, it pauses execution of the current program and transfers control to the exception handler. This behaviour is similar to software interrupts in x86 architecture.

The format of this class of instructions is similar to R-TYPE, with the 'code' field replacing source register, target register, destination register, and shift amount fields:

- **Opcode**: always zero. (6 bits)
- **Code**: available for use as a software parameter to be examined by the exception handler by loading the instruction word into a register.
- **Function code**: An extension to the opcode field, contains a unique code that defines the instruction to be performed (only break or system call). (6 bits).

This group consists of the following instructions:

1. **SYSCALL**: System Call
2. **BREAK**: Breakpoint

For detailed description for the behaviour of all of the instructions listed above, please refer to the appendix.

3.2.1.5 Control Instructions

MIPS-I architecture provides an ability to define coprocessor instructions in the instruction address space. The architecture can contain up to 4 groups of coprocessor instructions. MIPS-I assigns system control purposes to coprocessor 0, thus we allocate coprocessor 0 instructions for system control. Coprocessor instructions are specified by system designer, not MIPS-I ISA itself.

In this document, we shall use the keyword 'control instructions' to refer to this group of instructions. Control instructions are R-TYPE instructions with the following fields:

- **Opcode**: always 16. (6 bits)
- **Source register 1**: used to specify type of instruction. (5 bits)
- **Source register 2**: for MFC0 and MTC0 instructions, this field specifies the address of the general-purpose register whose value will be transferred to a control register, or altered by the value of the control register specified by 'destination register' field. For other instructions, this field must be zero. (5 bits)
- **Destination register**: for MFC0 and MTC0 instructions, this field specifies the address of the control register whose value will be transferred to or altered by the general-purpose register specified in the previous field. For other instructions, this field must be zero. (5 bits)
- **Shift amount**: must be zero. (5 bits)
- **Function code**: used, along with source register 1 field, to specify the type of the instruction. (6 bits)

The following list shows the instructions contained in this group:

1. **MFC0**: Move From Control Register
2. **MTC0**: Move To Control Register

3. **TLBR**: Read TLB Entry at Index
4. **TLBWI**: Write TLB Entry at Index
5. **RFE**: Return From Exception

For detailed description for the behaviour of all of the instructions listed above, please refer to the appendix.

3.2.1.6 Instruction Encoding Summary

As seen above, all instructions are encoded using either R-TYPE, I-TYPE, or J-TYPE format. Figure 3.2 summarizes the three formats and their corresponding fields. Note that field use/meaning depends on the encoded instruction itself, not the format.

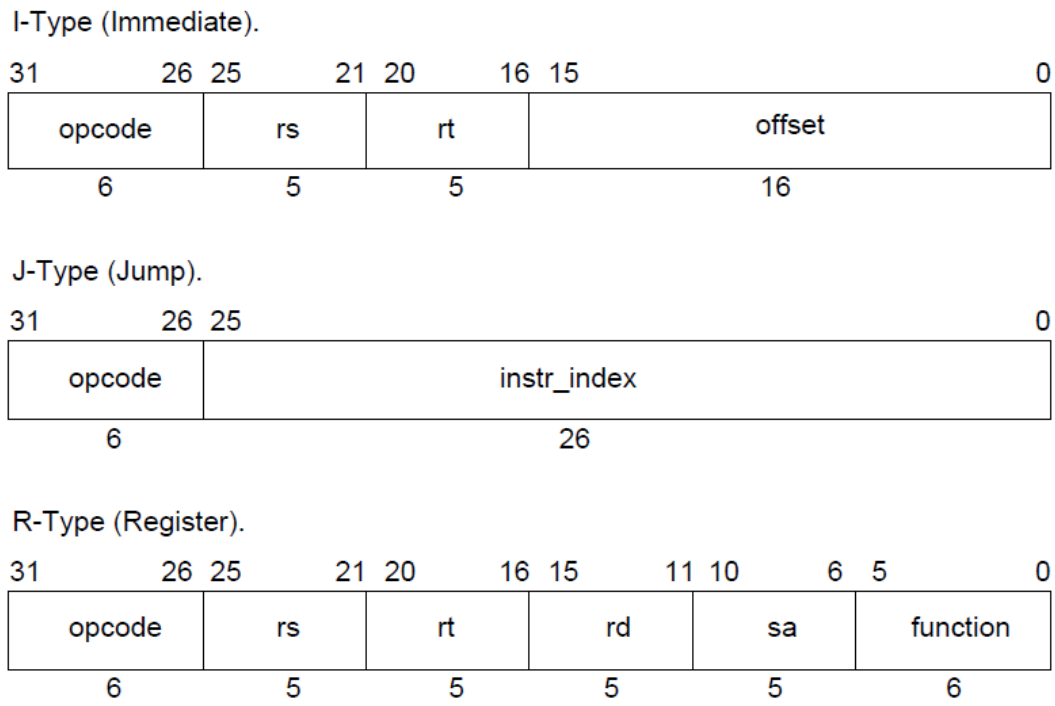
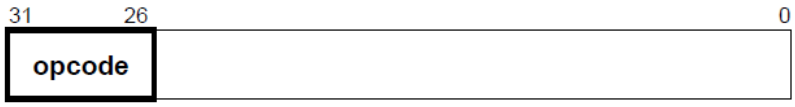


Figure 3.2: MIPS-I instruction formats

In all J-TYPE and I-TYPE instructions (except branch instructions comparing one register against zero), the opcode field is used to specify instruction type. This class of instructions is called 'instructions encoded by opcode field'. All possible values for opcode field under this class are shown in figure 3.3, along with valid mnemonics.



opcode bits 28..26 Instructions encoded by **opcode** field.

bits 31..29		0 000	1 001	2 010	3 011	4 100	5 101	6 110	7 111
0	000	*SPECIAL* δ	*REGIMM* δ	J	JAL	BEQ	BNE	BLEZ	BGTZ
1	001	ADDI	ADDIU	SLTI	SLTIU	ANDI	ORI	XORI	LUI
2	010	*COP0* δ,π	*COP1* δ,π	*COP2* δ,π	*COP3* δ,π,κ	*	*	*	*
3	011	*	*	*	*	*	*	*	*
4	100	LB	LH	LWL	LW	LBU	LHU	LWR	*
5	101	SB	SH	SWL	SW	*	*	SWR	*
6	110	*	LWC1 π	LWC2 π	LWC3 π,κ	*	*	*	*
7	111	*	SWC1 π	SWC2 π	SWC3 π,κ	*	*	*	*

Figure 3.3: Instructions encoded by opcode field

All R-TYPE instructions (including control instructions) are encoded using funct field, where opcode field is constant. For control instructions, the opcode field is equal to *COP*0 constant, while opcode field of remaining instruction is equal to *SPECIAL* constant. R-TYPE instructions with opcode equal to *SPECIAL* constant form another third class, called 'instructions encoded by function field when opcode field = *SPECIAL*'. Figure 3.4 shows members of this class.

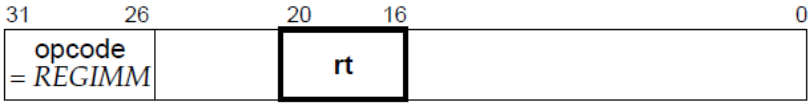
function bits 2..0 Instructions encoded by **function** field when opcode field = SPECIAL.

bits 5..3		0 000	1 001	2 010	3 011	4 100	5 101	6 110	7 111
0	000	SLL	*	SRL	SRA	SLLV	*	SRLV	SRAV
1	001	JR	JALR	*	*	SYSCALL	BREAK	*	*
2	010	MFHI	MTHI	MFLO	MTLO	*	*	*	*
3	011	MULT	MULTU	DIV	DIVU	*	*	*	*
4	100	ADD	ADDU	SUB	SUBU	AND	OR	XOR	NOR
5	101	*	*	SLT	SLTU	*	*	*	*
6	110	*	*	*	*	*	*	*	*
7	111	*	*	*	*	*	*	*	*

Figure 3.4: Instructions encoded by function field when opcode field = *SPECIAL*

Finally, branch instructions comparing one register against zero are encoded using rt field, while opcode field hold a constant value equal to *REGIMM* constant (register-immediate). This class, shown in figure 3.5, is called 'instructions encoded by the rt field

when opcode field = *REGIMM*'.



rt		bits 18..16		Instructions encoded by the rt field when opcode field = REGIMM.					
bits 20..19		0	1	2	3	4	5	6	7
		000	001	010	011	100	101	110	111
0	00	BLTZ	BGEZ	†	†	†	†	†	†
1	01	†	†	†	†	†	†	†	†
2	10	BLTZAL	BGEZAL	†	†	†	†	†	†
3	11	†	†	†	†	†	†	†	†

Figure 3.5: Instructions encoded by rt field when opcode = *REGIMM*

3.2.1.7 Addressing Modes

Figure 3.6 summarizes all possible addressing modes along with instructions that support them:

Addressing Mode	Corresponding Instructions
Register, [Base Register]+Immediate Offset	Load/Store Instructions
Register, Register, Immediate	ALU Instructions with Immediate Operand
Register, Register, Register	ALU Instructions with Three Operands
	SLLV, SRLV, and SRA
Register, Register, Shift Amount	SLL, SRL, and SRA
Register, Register	MULT, MULTU, DIV, and DIVU
Register	MFHI, MTHI, MFLO, and MTLO
Absolute Address Within 256MB Segment	Jump Instructions Jumping Within 256MB Region
Register Absolute	Jump Instructions to Absolute Address
Register, Register, PC-Relative Immediate	Branch Instructions Comparing 2 Registers
Register, PC-Relative Immediate	Branch Instructions Comparing Against Zero
Immediate Code	Exception Instructions
Register, Control Register	MFC0 and MTC0
Implicit	TLBR, TLBWI, and RFE

Figure 3.6: All possible addressing modes

We note that every instruction has no more than one single addressing mode, thus MIPS doesn't need an addressing mode field in instruction word.

3.2.1.8 Programmer-Visible Pipeline Effects

Several effects related to the behaviour of the pipeline should be stated clearly so the programmer can be aware of the exact performance parameters of the running program:

- **Delayed branching**: Since the effective address is not available immediately after fetching a jump/branch instruction, a delay slot should be inserted after the

jump/branch instruction. The delay instruction is always fetched and executed after the jump/branch instruction. The effective address becomes ready after fetching the delay instruction and next instruction would be the instruction at the new PC.

- **Load data not available**: If a memory load instruction is followed by an ALU instruction that requires the data acquired by the load instruction (addresses the same register), a data hazard will occur since ALU must wait until MEM unit is done fetching data from memory. This shall delay the execution of the ALU instruction by one delay slot (stalls the pipeline for one slot).
- **Consequent control instructions**: A control instruction may stall if there is another control instruction in the pipeline, to prevent data hazards and simultaneous access to TLB and stuff.
- **Branch instruction operands not ready**: If a branch instruction depends on result of the previous instruction (or second previous instruction), the pipeline will stall until data is ready.
- **Multiply/divide instructions**: These four special instructions take more than one CPU clock cycle to complete, therefore they might stall the pipeline.
- **Cache misses**: The pipeline shall stall if the instruction to be fetched or data specified by load instruction is not in cache. Furthermore, store instructions would always stall the pipeline since the cache system is write-through. This means that an instruction will take more than one CPU clock cycle if a cache miss happens.
- **Unaligned memory access**: This class of instructions shall take more than one CPU clock cycles because atomic access to unaligned word is not possible.

Other than the cases specified above, the pipeline will have no hazards/stalls and every instruction will take no more than one clock cycle delay.

3.2.2 Registers

A register is a small amount of memory that can be accessible faster than memory/cache. Processor provides direct register access through 'register' addressing mode.

MIPS architecture defines number of registers to be used as general-purpose registers by the running software. MIPS architecture also defines two special-purpose registers called LO and HI registers.

In addition to registers defined by MIPS standard, our processor contains special registers that control the behaviour of the processor, called 'control registers'. Surprisingly, control registers are accessible through control instructions.

3.2.2.1 General-Purpose Registers

Strictly speaking, MIPS architecture doesn't make any assumptions about the use of the 32 general-purpose registers. It's up to the programmer to use the register the way she likes, and the hardware will just perform the operations on those registers as specified by the programmer. However, there are two exceptions to this:

1. Register 0 is always zero, no matter what value the programmer writes to it.
2. Register 31 is used as a link register (register containing return address from subroutine). Some instructions implicitly address this register.

Although MIPS architecture doesn't make any assumptions on the use of general-purpose registers (except for register 0 and 31), people used to use those registers according to conventions listed in figure 3.7. The figure shows conventional names and uses associated with the registers. Available assemblers and compilers for MIPS architecture are expected to follow the same conventions.

Reg#	Name	Description
0	zero	Always zero
1	at	Used by assembler ($at = assembler\ temporary$)
2-3	v0-v1	Value returned by subroutine
4-7	a0-a3	Subroutine parameters ($a\# = argument\#$)
8-15	t0-t7	Used by subroutines without saving ($t\# = temporary\#$)
24-25	t8-t9	
16-23	s0-s7	Must be saved before use ($s\# = subroutine_var\#$)
26-27	k0-k1	Used by exception handler
28	gp	Global pointer
29	sp	Stack pointer
30	s8/fp	9th register variable/frame pointer
31	ra	Return address

Figure 3.7: Conventional names/uses of MIPS registers

3.2.2.2 LO and HI Registers

Along with the 32 registers defined above, two special registers are defined by the MIPS architecture to store results of MULT, MULTU, DIV, and DIVU instructions. Since these four instructions produce 64-bit results, they need two destination registers. Since MIPS architecture doesn't define more than one destination register for an instruction (for hardware considerations), the result should first stored in an intermediate 64-bit register (or two parallel 32-bit registers) then moved to the GPRs in two steps.

The intermediate registers are called LO and HI registers. Each one of them is 32-bit wide, so you can see them as one 64-bit register which is can be accessed in halfs. LO and HI can be read using MFLO and MFHI instructions, each of which copies the value of LO or HI register into a general-purpose register. The HI and LO registers can also altered by MTLO and MTHI instructions.

3.2.2.3 Control Registers

Now we turn our attention to the last class of registers, registers that control the behaviour of the processor. These registers are not defined by MIPS architecture itself, yet a processor needs to have a set of registers to store control switches. These registers are analogous to CR0, CR2, CR3, and EFLAGS registers of x86 architecture.

Control registers are accessible only through MFC0 and MTC0 instructions, which transfer values between GPRs and control registers. Every control register is addressed by a special index. Figure 3.8 shows all the control registers of our processor and their uses.

Reg Indx	Name	Description
0	Index	Stores index of TLB entry to be read/written
2	EntryLo	The lowest half of TLB entry to be read/written
9	BadVaddr	Stores the memory address that caused TLB miss
10	EntryHi	The highest half of TLB entry to be read/written
12	SR	Status register
13	Cause	Exception type
14	EPC	Address of the instruction that caused an exception

Figure 3.8: Control registers

Every control register has a specific format with a special function for every field. Next we show the format of every control register with a detailed description for the fields and how they can be used.

- **Index Register:**

Index Register (0)				
Field bit(s):	31	30..14	13..8	7..0
Field name:	0	x	Index	x

Figure 3.9: Format of Index register

Detailed description for fields in figure 3.9:

  - **Index**: Before the operating system reads or modifies a TLB entry, it shall first write the index for the target TLB entry in this field using MTC0 instruction. After then, the OS issues a TLB read or TLB write using TLBR or TLBWI instruction. Index takes a value from 0 to 63, which implies that our TLB has 64 entries.

- **EntryLo Register:**

EntryLo Register (2)				
Field bit(s):	31..12	11..10	9	8..0
Field name:	PFN	0	V	0

Figure 3.10: Format of EntryLo register

Detailed description for fields in figure 3.10:

  - **V**: This field matches the valid bit (V) of the TLB entry to be read or written by the OS. Before the OS writes a TLB entry, it writes the valid bit of the entry in this field. When TLBWI instruction is issued, the V field of EntryLo register is copied into the V field of TLB entry addressed by Index register. On the other hand, when a TLBR instruction is executed, the V field of TLB entry addressed by Index register is copied into V field of EntryLo register; a read for EntryLo.V field (using MFC0 instruction) reveals what the value of the V field of the addressed TLB entry was when TLBR was issued.
  - **PFN**: This field matches the physical frame number (PFN) field of the TLB entry to be read or written by the OS. Before the OS writes a TLB entry, it writes the PFN of the entry in this field. When TLBWI instruction is issued, the PFN field of EntryLo register is copied into the PFN field of TLB entry addressed by Index register. On the other hand, when a TLBR instruction is executed, the PFN field of TLB entry addressed by Index register is copied into PFN field of EntryLo register; a read for EntryLo.PFN field (using MFC0 instruction) reveals what the value of the PFN field of the addressed TLB entry was when TLBR was issued.

- **BadVaddr Register:**

BadVaddr Register (9)	
Field bit(s):	31..0
Field name:	BadVaddr

Figure 3.11: Format of BadVaddr register

Detailed description for fields in figure 3.11:

  - **BadVaddr**: When a TLB miss happens, the memory address that causes the TLB miss is copied in this field. There are two cases here:
    1. If a TLB miss happens while fetching an instruction from memory, the address of the instruction is copied into BadVaddr field before the exception handler is called.

2. If the instruction is fetched successfully, but it happens that the instruction is a load/store instruction, and a TLB miss happens during fetching/writing data from/to memory, the effective address of the data element addressed by this instruction is placed in BadVaddr before the exception handler is called.

- **EntryHi Register:**

EntryHi Register (10)		
Field bit(s):	31..12	11..0
Field name:	VPN	0

Figure 3.12: Format of EntryHi register

Detailed description for fields in figure 3.12:

- **VPN**: This field matches the virtual page number (VPN) of the TLB entry to be read or written by the OS. Before the OS writes a TLB entry, it writes the VPN field of the entry in this field. When TLBWI instruction is issued, the VPN field of EntryHi register is copied into the VPN field of TLB entry addressed by Index register. On the other hand, when a TLBR instruction is executed, the VPN field of TLB entry addressed by Index register is copied into VPN field of EntryHi register; a read for EntryHi.VPN field (using MFC0 instruction) reveals what the value of the VPN field of the addressed TLB entry was when TLBR was issued.

- **Status Register:**

Status Register (12)						
Field bit(s):	31..5	4	3	2	1	0
Field name:	0	IEo	0	IEp	0	IEc

Figure 3.13: Format of Status register

Detailed description for fields in figure 3.13:

- **IEc**: Interrupt enable/disable flag. When software writes 1 to this field, interrupts are enabled. When software writes to this field, interrupts are disabled and do not trigger any exception. This field is reset to 0 by hardware on two events:
  1. CPU is reset.
  2. An exception has taken place.

This field is also modified by RFE instruction which copies the value of IEp field into IEc field.

It's important to note that when an exception happens and the value of IEc is reset to 0, the old value of IEc is not lost but rather is copied into IEp. Also, the old value of IEp is copied into IEo before IEc is copied into IEp. When exception handler returns, RFE instruction is executed and the value of IEp is copied back to IEc, and the value of IEo is copied back to IEp.

Actually, the three fields (IEc, IEp, and IEo) behave like a 3-level stack for the value of the interrupt flag. The top of the stack (IEc) is the current state of the interrupt flag. When an exception happens, a 0 is pushed into stack, causing interrupts to be disabled and the values of the three entries are shifted by 1 (IEp copied into IEo and IEc copied into IEp).

If a nested exception occurs again, the process repeats, thus IEo shall then hold the old value of the interrupt flag, IEp shall hold the value of the interrupt flag during handling the first exception, and IEc will be reset to 0.

When the exception handler returns, RFE instruction will cause the stack to pop; IEp will be copied back to IEc and IEo will be copied back to IEp, thus IEc returns back to its original state before the exception has occured.

This behaviour allows an exception to be nested into another exception without need to push status register into the software stack before the inner exception happens.

- **IEp**: As discussed above, this field functions as backup for IEc field during exception handling. The three fields (IEc, IEp, IEo) can be seen as a 3-level stack to store the value of the interrupt flag across exceptions.

  When an exception happens, the value of this field is copied into IEo field and the value of IEc is copied into this field before IEc is reset to 0. When an RFE instruction is executed, the value of this field is copied back to IEc and the value of IEo is copied into this field.

- **IEo**: As discussed above, this field functions as backup for IEp field during exception handling. The three fields (IEc, IEp, IEo) can be seen as a 3-level stack to store the value of the interrupt flag across exceptions.

  When an exception happens, the value of IEp is copied into this field before IEc is copied into IEp. When an RFE instruction is executed, the value of this field is copied back to IEp.

- **Cause Register:**

Cause Register (13)				
Field bit(s):	31	30..7	6..2	1..0
Field name:	BD	0	ExcCode	0

Figure 3.14: Format of Cause register

Detailed description for fields in figure 3.14:

- **ExcCode**: When an exception occurs, the type of the exception is copied into this field before the exception handler is called. Currently, only the types of exceptions shown in figure 3.15 are supported.

ExcCode Value	Description
0	Interrupt Exception
1	TLB Miss Exception

Figure 3.15: ExcCode values with their meaning

- **BD**: When an instruction causes exception (or gets interrutpted), this field is set to 1 if the instruction is a delay slot instruction, otherwise it is reset to 0; thus 'BD' means branch detected.

- **EPC Register:**

EPC Register (14)	
Field bit(s):	31..0
Field name:	EPC

Figure 3.16: Format of EPC register

Detailed description for fields in figure 3.16:

- **EPC**: When an instruction that is currently executed by the processor causes an exception (like TLB miss) or gets interrupted (by en external I/O device), the address of the instruction is copied into EPC field, then exception handler is called. If the instruction is a delay slot instruction, the address of the previous jump/branch instruction is copied into EPC field instead of the address of the delay slot instruction.

  This behaviour prevents software to misbehave when the exception handler returns (as the branch instruction needs to be executed again or the processor will keep fetching instructions that follow the delay slot no matter the branch is taken or not). This special case causes BD field of Cause register to set as

described above.

It's also important to note that before exception handler is called, the CPU makes sure that the instruciton doesn't have any effect on the architectural state of the processor (i.e, any effect caused by the instruction is undone).

3.2.3 Memory Management

Having discussed the instruction set and registers of our CPU, we now turn our attention to how the CPU translates memory accesses. When an instruction is fetched from memory, or when a load/store instruction requests data from memory, the memory address is first translated by means of translation look-aside buffer (TLB). The translated address is then used to look-up data in the cache which might then issue an access request to main memory if data is not found in cache. Figure 3.17 summarizes the process.

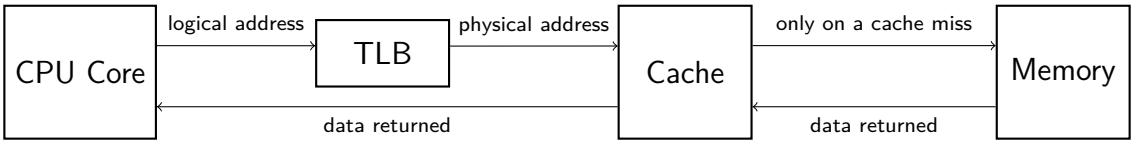


Figure 3.17: Memory access levels

The figure presents two different types of addresses: logical address and physical address. We define both terms as follows:

1. **Logical address**: the address of a memory location or an instruction from the prespective of the running program. This simply means that a program doesn't see memory as it is in reality. The running program doesn't need to know the real address of the memory item it is currently addressing, but the operating system must provide the mapping between the logical address of a memory item and its real location in memory. This mapping is used by translation circuitry that translates logical address to the real location before main memory is queried.

2. **Physical address**: this is the real location of the memory item or the instruction that is being accessed. Physical address is the address that is put on the bus when main memory access is established by the processor. Also, as will be seen later, data in cache is looked up by its physical address, not logical address.

The reason why there are two separate types of addresses and a translation circuitry is to allow every process to have its own address space by providing a distinct logical-to-physical address mapping for every running process. The process of translating a logical address to its corresponding physical address is called **address translation**.

Figure 3.18 illustrates how we can achieve process-level isolation using this technique. Every process has its own set of memory items: blue items belong to process 1, and red items belong to process 2. The OS needn't allocate the items in physical address space

in a specific order (they are scattered over memory) since the running program doesn't need to know the physical locations and doesn't need to make any assumptions about them.

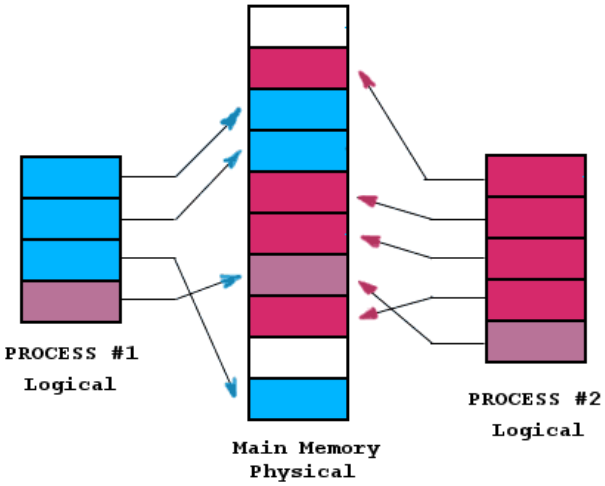


Figure 3.18: Process logical address spaces and mappings to physical addresses

The OS provides two separate mappings for both processes in the figure. Each mapping provides enough information needed to translate logical addresses of memory items belonging to the process to physical addresses. Each process only sees its own items in the logical address space, therefore a process can never access memory items of other processes because there is no mapping for them. By this mechanism, we have achieved the required isolation between processes: every process can only access the items allocated to it.

The OS might allow some memory items to be shared between more than one process. The purple item in the figure is seen in the logical address space of both process 1 and process 2. This implies that an item in main memory can have more than one logical address (for several mappings). However, every memory item can have no more than one physical address.

The process of allocating and organizing items in main memory is the responsibility of the OS. Also, the OS is responsible for providing the required mapping between logical and physical addresses. The OS provides the CPU with the mapping of the process that is currently being executed, and the CPU is responsible only for doing the translation on every memory access. Later in this section we show how the OS informs the CPU of the mapping.

It's important to note that some authors use terms like 'virtual address' and 'program address' to refer to logical addresses. We shall use the term 'logical address' throughout the document to prevent confusion with other terms and techniques, like 'virtual memory', which shouldn't be confused with 'logical address space', or equivalently, 'logical

memory'.

3.2.3.1 MIPS Logical Address Space

Since a logical address is 32 bits (as seen earlier), every process shall has a logical address space of 4GB. This 4GB address space is divided into four regions, each region is translated differently as will be seen below. The four regions are:

1. **kuseg** 0x00000000 - 0x7FFFFFFF (2GB): any address in this region is translated by the translation circuitry to its corresponding physical address using the mapping provided by the OS. This area usually contains the text and data of the running process.

2. **kseg0** 0x80000000 - 0x9FFFFFFF (512MB): addresses in this region are not translated using the mapping provided by the OS, but rather they are translated by stripping off the most significant bit. Thus, this region simply maps to the physical region: 0x00000000 - 0x1FFFFFFF. Operating system kernel can be simply put inside the lower 512MB zone of the main memory, and consequently it will appear inside kseg0. This implies that the kernel doesn't need to setup any mapping for itself, it is just mapped automatically to this region. This also implies that kernel memory area is shared among all processes.

3. **kseg1** 0xA0000000 - 0xBFFFFFFF (512MB): just like kseg0, addresses in this region also translate to the lower 512MB of physical address space. The only difference between kseg0 and kseg1 is that memory access of any address in kseg1 will skip the cache, so it is always uncached, while kseg0 is always cached. Access to memory-mapped I/O registers should always be done through this segment (to prevent caching). System designers must make sure that I/O registers have physical addresses in the low 512MB physical zone so that they can appear in kseg1.

4. **kseg2** 0xC0000000 - 0xFFFFFFFF (1GB): any address in this region is simply translated using the mapping provided by the operating system, this area is usually used to contain OS data structures which are maintained by the kernel.

We conclude that addresses in kuseg and kseg2 are translated the same way, using the mapping provided by the operating system, while addresses in kseg0 and kseg1 are translated by mapping them directly to the lower 512MB zone of physical memory. The only difference between kseg0 and kseg1 is that the former is cached while the later is uncached. The difference between kuseg and kseg2 is that kuseg is used to hold current process code and data, while kseg2 is used to hold kernel objects.

Having a static mapping for kseg0 and kseg1 has several advantages:

1. Processor reset vector is put in kseg1. Consequently, system initialization code (i.e, firmware) can start execution before translation circuitry is initialized. Caching also might not be wanted during system initialization, thus kseg1 is a safe place to start the system from.

2. Kernel is shared among all processes and doesn't need a specific mapping. This means that kernel initialization code can be executed before the translation circuitry is initialized. This also implies that a system call from process code running in kuseg can be handled in kseg0 without need to switch logical address space (i.e, without providing translation circuitry with another mapping). Microkernels can also use this area to put the tiny code that does the logical address switching.

When a process switch occurs, the OS must alter the translation circuitry with the mapping of the new process. The kernel can fix the mapping of kseg2 among all processes. Consequently, kernel objects can also be shared by all processes, allowing system calls to be handled without need to switching logical address space or alter the mapping.

3.2.3.2 Translation Look-aside Buffer (TLB)

Now we show how the operating system provides the CPU with a certain mapping for current process. Since there are 3GB of the address space (kuseg and kseg2 regions, see above) that can be dynamically mapped to physical memory, we need a data structure that describes how to map every logical address inside the two regions (kuseg and kseg2) into physical address.

One can think of an array of 4G elements, where the index of the array is the logical address, and the contents of an array element is the physical address that maps to the logical address associated with that entry. Notice that the third quarter of such an array would always be sparse, since kseg0 and kseg1 are not dynamically mapped as shown before.

Since a physical address is 32 bits, we need an array of 4Gx4B which is a 16GB data structure for every process. This is certainly unfeasible and we would need to think for another solution. Instead of mapping each logical address to a certain physical address, we can map every block of logical address to a block of physical addresses.

Using this technique, we just need an array with n elements, where n is number of blocks of the logical address space which is also equal to number of block of the address space. We call this scheme **'Paging'**, where a physical address is called a 'frame' and logical address block is called a 'page'.

Given a page size of 4KB, the 4GB address space could be divided into 1M pages, requiring an array of only 1M elements. The array is called a 'page table' and maps every page of the logical address space to a frame in the physical memory. A page table entry will simply contains the number of the frame it maps to, which is a 20-bit word (because we've got a total number of 1M frames).

Extra information might need to be stored along with frame number, like whether this page is mapped or not (a page can be unmapped if doesn't map to any physical address) as well as OS-specific flags. Hence we can extend the 20-bit entry to 32 bits: 12 bits for flags and 20 bits for frame number (FN). Hence we need an array of 1M*4B which is equal to 4MB page table for every process. This is much smaller than 16GB.

To translate a logical address, we divide the address into two parts:

1. Page number: the highest 20 bits of the address
2. Offset inside the page: the lowest 12 bits of the address.

Page number is then used to index the page table to get the corresponding page table entry. The frame number is extracted from the entry and concatenated with the 12-bit offset to form the actual physical address that this logical address maps to. Figure 3.19 illustrates this translation process.

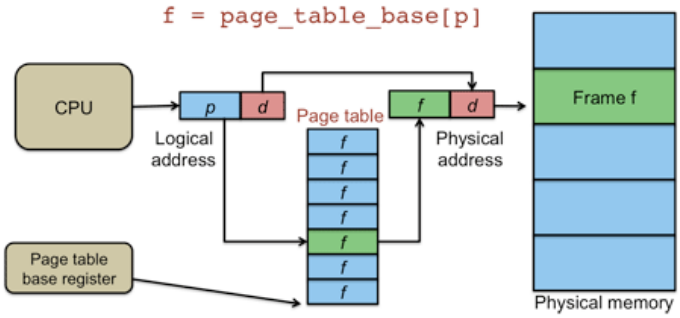


Figure 3.19: Address translation using page table

A page table of 4MB fo every process is still big, since when number of processes increase, a big portion of main memory will just be used to store page tables. Furthermore, a process won't usually need to map the whole 4GB address space. It can leave some pages unmapped, and only map pages where it needs to put its code and data. Actually, the usual case is that most pages are unmapped. And if virtual memory is not implemented, the total number of mapped pages across all page tables (across all processes) must be less than or equal to size of RAM (given that no frame is mapped to by more than one page), and RAM is usually much smaller than 4GB.

Since most entries in the page table are sparse (because most pages in the logical address space are unmapped), the OS can make use of this characteristic to make the size of the page table smaller. Thus two-level page tables where introduced. The big 1M-entry page table is divided into 1024 parts, with each part having 1024 entries. Each part is called a second-level page table. These second-level page tables could be scattered around memory. An 1024-entry table could be used to store the addresses of these 1024 2nd-level tables. Figure 3.20 shows how translation works under this two-level scheme.

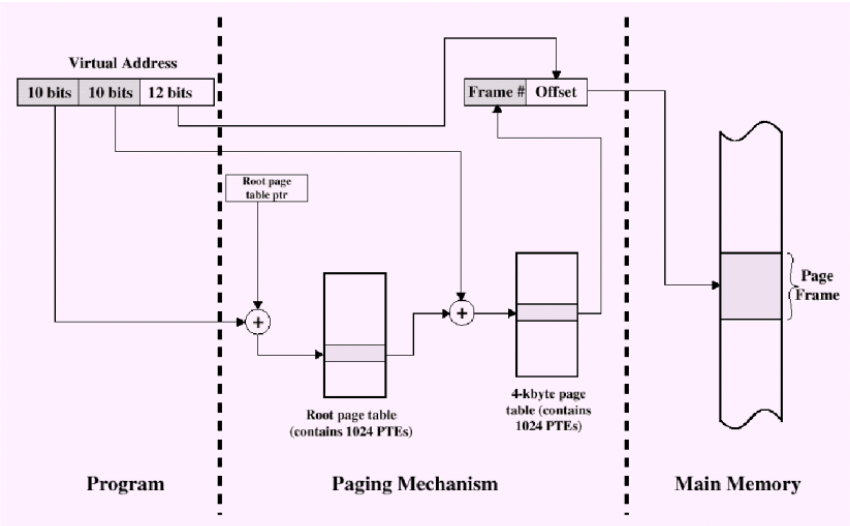


Figure 3.20: Address translation using two-level page table scheme

This way, if all the entries of a second-level page tables are unmapped, the whole second-level page table could be unmapped, setting a flag in the page directory that this page table doesn't exist at all, and all pages in this region are unmapped. Consequently, many second-level page tables wouldn't need to exist at all, reducing the size of allocated memory for page tables to a reasonable limit. However, if the process maps all of its logical memory, page table increases to 4MB plus the 4KB of the page directory.

This result implies that we cannot store the whole page table in the processor, since we would need 4MB of memory inside the CPU dedicated for page table. We would also need to re-initialize the whole 4MB on every process switch, a process that will take much time. That's why page tables are stored in memory and with a multi-level scheme to redcue their sizes.

So, how the CPU makes use of these page tables to do translation? The processor could simply be given the base physical address of the page directory that controls the current process. When a logical address is generated by core processor, the translation circuitry looks up the page directory then the corresponding 2nd-level page table to get frame number, as shown in the figure above. On a process switch, the CPU is given the page directory of next process.

This translation scheme requires two memory access for each translation, which is too much since memory access takes much time, so we need to cache page entries that are used frequently. Caching page entries will effectively reduce average translation time. This special-purpose cache is called **'translation look-aside buffer'**, or simply TLB. When a logical address is generated by core processor, TLB is queried for that logical address. If the entry of this address is found inside TLB, the physical address is returned immediately. If a TLB miss happens, the processor shall access main memory to get page entry.

The type of TLB described above is called **'hardware-managed TLB'**, where TLB is totally isolated from the OS and is managed by hardware. Another mechanism would be **'software-managed TLB'**, where TLB is controlled by the OS, and CPU doesn't need to know any information about page table. In a software-managed TLB, when a TLB miss happens, the CPU raises an exception, and it is the responsibility of the exception handler to look-up page tables and refill TLB with the missed entry.

Since our processor is designed for systems where resources are very limited (like FPGA), we chose to use software-managed TLB to simplify the implementation of the processor. On a process switch, the OS just needs to clear TLB from entries related to old process. After TLB is cleared, and when a logical address is generated, a TLB miss happens and exception handler is called. The OS queries page tables and refills TLB. Page tables are totally managed by the OS; the processor doesn't need to know anything about their structure and locations.

The TLB consists of 64 entries. Every entry is 64-bits wide with the structure illustrated by figure 3.21.

TLB Entry						
Field bit(s):	63..44	43..32	31..12	11..10	9	8..0
Field name:	VPN	0	PFN	0	V	0

Figure 3.21: Format of TLB entry

Fields are:

- **VPN**: the 20-bit logical page number to be translated.
- **PFN**: the 20-bit physical frame number this VPN maps to.
- **V**: valid bit. If this bit is set to 1, this is entry is valid and can be used for translation. If this bit is set to 0, this entry is not valid and won't be used for translation.

The reader can notice the similarity between TLB entry and EntryLo/EntryHi registers. The structure of the lower 32 bits of a TLB entry is identical to EntryLo register, and the structure of the upper 32 bits is identical to EntryHi register. EntryLo and EntryHi registers, along with Index register, are used to read and write TLB entries. For reading a TLB entry, the steps are:

1. The OS writes the index (0->63) of the TLB entry to be read into Index register using MTC0 instruction.
2. TLBR instruction is executed. The instruction causes the lower 32 bits of the TLB entry indexed by Index register to be copied into EntryLo register, and the upper 32 bits of the TLB entry to be copied into EntryHi register.
3. The operating system reads EntryLo register using MFC0 instruction. Result is the lower 32 bits of the target TLB entry.

4. The operating system reads EntryHi register using MFC0 instruction. Result is the upper 32 bits of the target TLB entry.

A similar sequence could be used for writing a TLB entry

1. The OS writes the index (0->63) of the TLB entry to be written into Index register using MTC0 instruction.
2. The operating system writes the lower 32 bits of the target TLB entry into EntryLo register using MTC0 instruction.
3. The operating system writes the upper 32 bits of the target TLB entry into EntryHi register using MTC0 instruction.
4. TLBWI instruction is executed. The instruction causes EntryLo register to be copied into the lower 32 bits of the TLB entry indexed by Index register, and EntryHi register to be copied into the upper 32 bits of the target TLB entry.

Now since the operating system has full access to TLB, paging can be implemented easily. Whenever a logical address is generated by the processor, the following steps happen:

1. The CPU reads TLB entry which has an index that is equal to bits 12 to 17 (inclusive) of the logical address.
2. If V flag of the read entry is 0, the CPU copies the logical address into BadVaddr register, and TLB miss exception happens. In this case translation stops. If V flag is 1, continue to step 3.
3. The VPN of the read field is compared with the highest 20 bits of the logical address. If both values are not equal, the logical address is copied into BadVaddr register and TLB miss exception happens. In this case translation stops. If both values are equal, continue to step 4.
4. The PFN of the read field is concatenated with the lower 12 bits of the generated logical address, the result is the physical address that will be used to access memory. Translation is done and the CPU continues work normally without exception.

Step 1 implies that our TLB is actually a **direct-mapped cache** where the index is computed using bits 12 to 17 of the logical address, and tag is VPN. Using this information, the OS can handle TLB miss exception as follows:

1. The OS reads the value of BadVaddr register using MFC0 instruction.
2. The OS then queries page tables to know the mapping of the address in BadVaddr register. If there is no mapping, the operating system may decide to stop the program or send an error signal for instance. If there is mapping, the operating system reads the physical frame number that this logical address maps to.
3. The operating system calculates the index of the TLB entry which should be modified to solve the miss. The index is actually equal to bits 12 to 17 (inclusive) of the logical address as explained above.

4. Now the OS has the index, VPN, and the corresponding PFN. The OS simply writes the index to Index register, writes VPN to EntryHi register, and writes PFN to EntryLo register (and sets V flag of EntryLo register to 1).

5. The OS executes TLBWI instruction. EntryLo and EntryHi registers are then copied into the TLB entry.

6. Exception handler returns and the instruction that causes the miss is executed again, but this time it doesn't cause a TLB miss and execution continues normally.

3.2.3.3 Cache Memory

After logical address is translated into physical address, physical address is used to access RAM. However, because memory access takes long time (i.e, has high penalty), a cache is put between TLB and RAM to fasten memory access time.

Since the pipeline of our processor might be fetching an instruction and handling memory access for a load/store instruction at the same time, two separate caches are employed: an instruction cache and a data cache. Both caches are direct-mapped caches. Every entry in the cache consists of three fields:

- **Valid flag**: If 1, data and tag fields are valid. Otherwise the entry is invalid.
- **Data field**: The 4-byte datum cached by this entry.
- **Tag field**: A 20-bit field storing the higher 20 bits of the physical address associated with this entry.

Every cache has 1024 entries, resulting in two 4KB caches. The process of looking up data in the cache is completely transparent from software. Bits 2 to 11 (inclusive) of the physical address are used as index. The tag of the entry at the calculated index is compared to the higher 20 bits of the physical address. If valid bit is 1 and comparison matches, data field of the entry is return to core processor. If valid bit or 0 or comparison doesn't match, the CPU initiates a memory cycle to bring data from RAM and updates the cache automatically.

Cache is write-through; any store instruction will cause data to be written to both cache and RAM. A cache miss (or a store instruction) causes the instruction to delay until data is brought from or written to memory. However, cache hit can be handled without any delay; pipeline doesn't stall and advance normally. Although caching is totally isolated from the programmer, the programmer can use her knowledge of how cache works to be able to write optimized programs that make full use of caching.

3.2.4 Exception Handling

In MIPS, an exception is a hardware event, trap, system call, or anything that interrupts the normal flow of the program and needs special handling. The same mechanism is used to handle all exceptions, which include:

- **Interrupts**: throughout this document, we will use the term 'interrupts' to refer only to those generated by hardware devices. Interrupts are external events that need immediate interaction by the processor. Interrupts can be enabled/disabled by software (refer to Status Register). If enabled, when an interrupt signal arrives to the processor, exception happens and the normal flow of the program is interrupted. Interrupts may include, for example, keystrokes, mouse button events, and disk I/O events.

- **System calls**: system calls are special calls that transfers the control to the operating system to execute some kernel-level subroutine then return back to user program and continue execution normally. System calls are handled by the processor as an exception trigger: when system call (or breakpoint) instruction is executed, an exception happens.

- **Page faults**: page faults need to be reported to operating system whenever they happen to fully support paging. When a page fault happens, the running program is interrupted and exception handler is called. Exception handler shall handle the fault then return back to the running program.

- **Program errors**: division erros, invalid opcodes, and invalid addresses are also forms of exceptions that cause the exception handler to be called. In this case the exception handler might, for example, call a specific signal handler in the faulty process or stop it completely.

Processor provide a mechanism to return from exception and resume execution of the interrupted process. It's the responsibility of the OS to store the context of the interrupted process at beginning of exception handling, so that when the exception handler returns back to the interrupted process, the process continues its work normally as if nothing has happened.

It's the responsibility of the processor to inform software with what kind of exception has actually happened, at which instruction the exception has happened, and any extra parameters needed. For example, if a program error has happened, the processor shall inform software of which instruction has caused the error. If a page fault happens, the processor should inform software of which instruction caused the fault, and of the logical address that couldn't be translated.

Up till now, we are talking about the concept of exceptions in general, without projecting the concept on our architecture. In the next lines we show how exceptions are designed in our processor.

3.2.4.1 Exception Cases

Currently, these are all the events that cause our processor to raise an exception:

- **Interrupt**: Only if IEc flag of SR register is 1, if IRQ signal arrives to the processor, the processor sets ExcCode field of Cause register to 0, copies the address of instruction that is currently being executed into EPC register, then calls exception handler. We note that there is only one IRQ signal that is fed to the processor.

System designer should allow the IRQ signal to be shared by various I/O devices. Several techniques exist for sharing IRQ signals.

- **TLB miss**: When logical address fails to be translated (as descibed earlier), the processor sets ExcCode field of Cause register to 1, copies the logical address into BadVaddr register, and copies the address of the instruction that has caused the TLB miss into EPC register. The processor then calls the exception handler.
- **Reset signal**: Processor reset is also a form of exception, where the appropriate exception handler is called whenever this event happens. Reset doesn't cause the Cause register, EPC register, or any other registers to be modified. On RESET, software should re-initialize the processor and drives it to a known state.

It's important to note that when exception happens at some instruction, the processor makes sure that all previous instruction has done execution and their results has been commited successfuly before the exception handler is called. The instruction at which the exception has happened (the instruction whose address is copied to EPC) must not have any effect on the architectural state of the processor, so that when exception handler returns, the instruction can be re-executed without problems. We will later introduce the concept of precise exceptions.

3.2.4.2 Exception Vectors

Exception vector is the address where the exception handler is expected to reside. When exception happens, the processor performs a far jump to exception vector, starting from there the execution of the exception handler. Actually, one distinct exception vector could be assigned to every type of exception. In this case, there is an exception handler for every type of exception.

Exception vectors could be static or dynamic. For static exception vectors case, the vectors are predefined addresses that are hardwired in the processor. Static vectors can never change. Examples for architectures with static vectors include MIPS, 6502, and PIC architectures.

Dynamic vectors are set up by software. They could be stored inside one or more registers in the processor or could be stored in a table. The table will contains vectors for all possible exception types. Examples for architecture using dynamic vector mechanism include x86 and Zilog Z80. In some architectures, like 68K and PDP-11, the interrupting I/O device itself provides the processor with the interrupt vector when an interrupt is handled.

Since MIPS processors use static hard-wired exception vectors, our processor jumps to predefined addresses when exceptions happens. Figure 3.22 lists those predefined addresses and their associated exceptions.

Logical Address	Region	Physical Address	Exception Types
0xBFC00180	kseg1	0x1FC00180	All types of exception, except reset.
0xBFC00000	kseg1	0x1FC00000	Reset exceptions.

Figure 3.22: Excetion vectors for various exception types

We note that all exception handlers are located inside kseg1 memory region. This is important for reset vector to be located in kseg1 region as it is the only safe region to start the system from as explained before. Thus firmware location should be adjusted in memory (by system designer) such that first instruction of the firmware program matches reset vector.

The exception handler for other types of exception is also located in the uncached kseg1 region. This might be necessary during system initialization, but the OS later might want the exception handler to be called through kseg0 (instead of kseg1) so that it can be cached. A jump instruction could be inserted in the exception handler at kseg1 for that purpose.

3.2.4.3 Exception Protocol

Except for reset exceptions, the processor does the following when an exception happens:

1. Copy the address of the instruction that causes the exception into EPC register.
2. The 3-level stack inside SR register (explained before) is pushed and a 0 is inserted in IEc flag to disable interrupts.
3. Set the Cause register with appropriate ExcCode. If the exception is a TLB miss exception, the logical address is copied into BadVaddr register.
4. The processor transfers control to the exception handler.

The exception handler should go through these steps:

1. Store the state (context) of the interrupted program. EPC should also be stored.
2. Load Cause register and examine ExcCode field to determine the type of exception and call the appropriate handler.
3. Process the exception. This can alter the values of registers and change the state of the processor, that's why the context of the interrupted process is stored in first step.
4. Restore the state of the interrupted program.
5. Return from exception: this can be done by loading the value of the EPC register (that is stored in step 1) into one of the general purpose registers, then using JR instruction to jump back to the instruction referred to by EPC. RFE instruction should be inserted into the delay slot of the JR instruction so that the stack in the SR register is popped just before returning to the interrupted program (to enable interrupts again). This is the safest method to return from exception.

3.2.4.4 Precise Exceptions

We finally conclude our discussion with a note on the preciseness of exceptions. Due to pipelining, multiple exceptions may happen at the same time by various instructions in various stages. Furthermore, an exception might happen in an instruction where following instructions has already been executed and has changed the architectural state of the processor.

To handle these pipeline issues, exceptions need to be 'precise'. An exception is called precise if the saved processor state corresponds with the sequential model of the program execution, although the underlying hardware excutes instructions in parallel and out of order. In the case of precise exceptions, the OS knows well that all instructions before the instruction referred to by EPC have finished their execution successfully, and have already committed their results. The OS also knows that all instructions after EPC has never been executed and didn't effect the processor state. These two points is what we mean by saying that the saved processor state corresponds with the sequential model of execution.

Since our prcoessor implements precise exceptions, it guarantees the following:

- **Unambiguous proof of cause**: the values of EPC and Cause registers are exact and refer to the actual instruction that caused the exception and the actual type of exception. This wouldn't be the case on imprecise exceptions.
- **Exceptions are seen in instruction sequence**: as described above, precise exceptions guarantees that all instructions before EPC has finished their execution successfully. If several exceptions at several stages happen at the same time, the processor choose the earlier instruction of them to be put in EPC. When an exception happens at some stage of the pipeline, exception handler is not called until all instructions in the following stages finish execution and commit their results.
- **Subsequent instructions nullified**: all the stages before the stage that caused the exception are nullified and any change that happened by instructions in these stages is undone. This guarantees that all instructions after EPC can be restarted safely after exception handler returns.

3.3 System Bus

System bus is a set of signals that provide a communication mechanism between the processor, memory interface, and various I/O devices. As we have already explained, system bus consists of three sub-buses:

- **Address bus (32 lines)**: Contains the address of the data word being read/written by CPU. Since the address bus is 32 bits, the address space of our system has 4G addresses. The smallest addressible unit is 'byte'. Thus the whole address space is 4GB, with every address referring to a distinct byte in memory.

- **Data bus (32 lines)**: A bi-directional bus where CPU puts data word that is to be written to memory, or memory puts data word that is to be read by the processor. Since a data bus is 32 bits, the word size of our bus is 32 bits. Byte and half-word accesses are also supported.
- **Control bus**: This consists of a number of control signals that control the communication between the processor and I/O devices. Details on this will be seen below.

Address bus is always driven by the processor. No other devices are allowed to drive the bus (or contention may happen). Data bus is driven either by the processor (if data is to be written to memory) or by other devices (if data is to be read by the processor). Control bus signals are driven by various devices. Control bus supports these signals:

- **CLK**: clock signal to synchronize all components together. Generated by clock circuitry and fed into all components.
- **MEME**: memory enable signal, driven by CPU. When MEME changes from inactive level to active level, a memory access cycle starts.
- **RW**: read/write signal, driven by CPU. When inactive, CPU wishes to read something from memory. When active, CPU wishes to write.
- **DTYPE**: specify whether the CPU wants to read/write a single byte, a half-word, or a word.
- **IRQ**: IRQ signal controlled by interrupt controller. It is fed into the processor and changes to active level when interrupt happens. The CPU senses the inactive-to-active edge to detect interrupt requests.
- **IAK**: interrupt acknowledge signal, driven by CPU and fed into the interrupt controller. It changes to active level when CPU wants to acknowledge the interrupt signal.
- **IRQ#**: 8 IRQ lines controlled by various I/O devices are fed into the interrupt controller. When a device is interested in interrupting the processor for some reason, it activates its own IRQ line.
- **IAK#**: 8 IAK lines controlled by the interrupt controller and fed into various I/O devices. They are used to acknowledge the interrupt requests issued by devices.
- **CS#**: chip select lines fed into the I/O devices and driven by bus decoder. Bus decoder keeps listening to address bus. According to the address loaded on the address bus, a device is selected and its own CS signal changes to active level.

Now the sequence of a read/write cycle should be clear. First, the CPU puts the target address on the address bus, sets DTYPE and RW signals, then raises MEME to initiate the cycle. Bus decoder responds to MEME event by decoding the address and driving the corresponding CS signal to the active level.

Next, the device whose CS signal has been activated knows it should respond to the request. If RW is inactive (read cycle), the device puts data on the data bus. Otherwise, the device fetches data from data bus and processes it.

3.4 Bus Decoder

Bus decoder is a simple address decoder that selects an I/O device based on the state of the address bus. The decoder takes most significant bits of the address as input, and selects the I/O device based on the value of these bits. Thus, you can imagine the address space as a sequence of address blocks, where each block belongs to a specific I/O device.

It should be clear by now that all I/O registers are mapped in the same address space where memory is mapped. Therefore all I/O registers are 'memory-mapped registers' and no isolated I/O address space is supported.

Figure 3.23 shows the memory map of our system. The decoder shall activate the CS of a given device if and only if the address lies in the address range dedicated to that device.

Start Address	End Address	Size	Device
0x00000000	0x00FFFFFF	16MB	RAM
0x1E000000	0x1E003FFF	16KB	VGA
0x1E800000	0x1E800FFF	4KB	KBD (Keyboard Controller)
0x1E801000	0x1E801FFF	4KB	PIT (Programmable Interval Timer)
0x1E802000	0x1E802FFF	4KB	PIC (Programmable Interrupt Controller)
0x1F000000	0x1FFFFFFF	16MB	ROM

Figure 3.23: Memory map.

It is important to note that there is a separate CS signal for RAM and another separate CS signal for ROM. However, there is only one memory interface circuit for interfacing both RAM and ROM chips. The interface takes both CS signals as input, and selects whether to work with the RAM chip or ROM chip based on the level of the signals.

3.5 Memory Interface

RAM and ROM chips have pinouts and interfaces that are not compatible with the bus. Therefore there must be a circuit that converts between both interfaces. The chips might not support 4-byte accesses (as will be seen in the implementation chapter). Furthermore, control signals need to be translated to the controls of the chips.

The circuit takes the lower 24 bits of the address bus as input (this builds up to 16MB). CS_ROM and CS_RAM signals select whether the 24-bit address corresponds to RAM or ROM. The interface is designed that way because the available RAM and

ROM chips are actually 16MB each (as will be explained later in the implementation chapter).

From programmer's point of view, ROM is accessed the same way as RAM. If the address specified in LW/SW instruction corresponds to the address range of RAM, then RAM chip is activated. Otherwise, if it corresponds to the address range of ROM, then ROM chips is activated, and that's all. The implementation details of the circuit are hidden and do not affect the external behaviour seen by the programmer.

It is also important to note that writes to ROM (using SW instruction) are just ignored. Since the chips need time until data is ready on the data bus (or data has been written), the CPU should give the circuit enough time before it deactivates MEME signal.

3.6 VGA Interface

This component provides an interface between our system and any CRT/LCD screen with a VGA connector. VGA appears in the address space at the range 0x1E00000-0x1E03FFF (16KB). Internally, the chip contains four memory blocks of 2Kx9. This means that every block has 2K data units, and every unit is 9 bits wide.

It is important to note that all memory addresses mentioned in this section are physical addresses, refer to section 3.2.3 for more information on address translation in MIPS architecture.

A write to one of the addresses in the range 0x1E00000-0x1E03FFF selects a data unit in one (and only one) of the four RAM blocks. Any write should be peformed using SH (store half) instruction, and address should be always aligned to 2-byte boundary (i.e, lower 1 bit should always be zero) as specified by MIPS architecture for SH instruction.

When a halfword is written to VGA, only lower 9 bits are conisdered, and upper 6 bits are simply ignored. Figure 3.24 shows how the blocks are selected based on the value of the least significant 14 bits of the address.

Lower 14 Bits	Selected Block	Function
0xxxxxxxxxxx00	Block 0	Character Memory
0xxxxxxxxxxx10	Block 1	Attribute Memory
1xxxxxxxxxxx00	Block 2	Font Memory 1
1xxxxxxxxxxx10	Block 3	Font Memory 2

Figure 3.24: VGA decoding mechanism.

As you see, bit 0 is always 0, since writes are always 2-byte (half-word) writes. The memory block is selected based on the value of bit 1 and bit 13. This means that it is based on whether the address lies in the first half or the second half of VGA address

range, and whether the address (ignoring bit 0) is odd or even.

So, to summarize the table, we can say that if the address lies in the lower half of the VGA address range, then the address corresponds to block 0 or 1 (character and attribute memories). If the address, ignoring bit 0, is even, then select block 0. Otherwise (address is odd), select block 1.

On the other hand, if the address lies in the upper half of the address range, then the address corresponds to block 2 or 3 (font memories). If the address, ignoring bit 0, is even, then select block 2. Otherwise (address is odd), select block 3.

Now we know how a memory block is selected, the remaining part is to explain how a data unit (inside the selected memory block) is addressed. You can simply deduce that the remaining bits of the 14-bit address (the xxxxxxxxxxx) are used directly as an index for the memory block. Thus, 00000000000 will address the first data unit, where 00000000001 will address the second, and so on.

If you look closely, you will find that the number of 'x' bits is 11, which means that we can address up to 2048 (2^{11}) for each memory block. Every write is a 16-bit (half-word) write (with upper 6 bits ignored), thus we are dealing with 2Kx9 units.

The reader can notice the similarity between our design and the design of VGA programmer's interface of the IBM personal computer (the 0xB8000 text buffer). The basic difference is that our memory blocks are 9 bits wide (not 8), and that's the reason why writes are 16-bit writes, not 8-bit.

3.6.1 Character Memory

Our VGA module basically supports a 80x25 text mode. This means that screen is divided into 80 columns and 25 rows (2000 cells), with every cell containing an ASCII character. To draw text on screen, the operating system just needs to specify the contents of the 2000 cells.

Character memory (memory bank 0) consists of 2048 9-bit cells (from 0 to 2047). Ignoring last 48 cells, cells from 0 to 1999 are used to describe text on screen, with every cell corresponding to a character on screen.

To set the contents of character in row x (0-24), column y (0-79), you need to use the formula: $indx = x * 80 + y$, then you can then refer to table 3.24 which shows how to access the banks themselves.

For instance, to access row 2, column 1:

1. Calculate indx: $indx = 2 * 80 + 1 = 161$.
2. We know we want to access bank 0, thus lowest 14 bits of the address will be 0xxxxxxxxxxx00 where bit 0 is zero as usual, bit 1 and 13 are 00, (to indicate bank 0), and xxxxxxxxxxx is the index.

3. Substitute xxxxxxxxxxx with 161, so it becomes 00010100001. Therefore, the lowest 14 bits are: 00001010000100 (0x0284).

4. The full 32-bit address can be easily deduced by ORing the lowest 14 bits with 0x1E000000, thus the full address is 0x1E000284.

A write to 0x1E000284 will set the contents of cell 161 (which corresponds to row 2, column 1). If you write 'B' (0x42 in ASCII) to address 0x1E000284, you will see B shown on screen at third row (row 2) and second column (column 1). That way you can write any kind of ASCII text you want on screen. Note that bit 8 (highest bit) of a cell is always ignored, so only 256 ANSI symbols are supported.

3.6.2 Attribute Memory

Bank 1 is used exactly the same way as bank 0, except that it describes the colors, instead of the text itself. As we explained above, the screen is divided into 2000 cells, organized in 25 rows and 80 columns. Every cell has a corresponding 9-bit data cell in memory bank 0 containing ASCII character that is drawn on the cell. Also, every cell has another corresponding 9-bit data cell in memory bank 0 describing its color.

The format of the 9 bits is described as following:

- bits 0..3: color code of the ASCII character drawn in the cell (foreground).
- bits 4..7: color code of the background of the cell.
- bit 8: you better set it to zero.

As it is clear, there are only 16 colors (coded by numbers from 0 to 15). Color code are described in table 3.25.

Code	Color	Code	Color	Code	Color	Code	Color
0	Black	4	Red	8	Dark Gray	C	Light Red
1	Blue	5	Violet	9	Light Blue	D	Light Violet
2	Green	6	Brown	A	Light Green	E	Yellow
3	Cyan	7	Gray	B	Light Cyan	F	White

Figure 3.25: VGA color codes.

To set the color of the cell at row 2, column 1, we follow the same procedure we followed when we set the text of the same cell in previous section:

1. Calculate indx: $indx = 2 * 80 + 1 = 161$.

2. We know we want to access bank 1, thus lowest 14 bits of the address will be 0xxxxxxxxxxx10 where bit 0 is zero as usual, bit 1 and 13 are 01, (to indicate bank 1), and xxxxxxxxxxx is the index.

3. Substitute xxxxxxxxxxx with 161, so it becomes 00010100001. Therefore, the lowest 14 bits are: 00001010000110 (0x0286).

4. The full 32-bit address can be easily deduced by ORing the lowest 14 bits with 0x1E000000, thus the full address is 0x1E000286.

So, to set the text of that cell, we write ASCII code (using SH instruction) to 0x1E000284. To set the color, we write the attribute to 0x1E00286. As we mentioned before, bits 0..3 indicate the foreground color, where bits 4..7 indicate background. For a white foreground (color code F) and blue background (color code 1), we simply write 0x001F.

We conlude our discussion of character and attribute memory with a small example to explain how to print something on screen. In this example, we shall print 'Hello World!' at row 0 of screen, with blue background and white foreground. 'H' will be written at column 0, 'e' at column 1, 'l' at column 2, and so on.

We first type the text in character memory. Here we have 12 ASCII characters, which will be written into cells from 0 to 11 of character memory. In table 3.26, we calculate the address of every cell of the 12 cells, and we show the exact value that will be written in that cell using SH (store halfword) instruction.

Row	Column	Memory Address	Text	ASCII code
0	0	0x1E000000	H	0x0048
0	1	0x1E000004	e	0x0065
0	2	0x1E000008	l	0x006C
0	3	0x1E00000C	l	0x006C
0	4	0x1E000010	o	0x006F
0	5	0x1E000014		0x0020
0	6	0x1E000018	W	0x0057
0	7	0x1E00001C	o	0x006F
0	8	0x1E000020	r	0x0072
0	9	0x1E000024	l	0x006C
0	10	0x1E000028	d	0x0064
0	11	0x1E00002C	!	0x0021

Figure 3.26: Hello world example: text.

Next, we write the attributes of the 12 characters to attribute memory. Since we want all characters to have white foreground (F) and blue background (1), then color code is 0x1F. Table 3.27 shows the actual addresses that will be written.

Row	Column	Memory Address	Attributes	Color code
0	0	0x1E000002	White/Blue	0x001F
0	1	0x1E000006	White/Blue	0x001F
0	2	0x1E00000A	White/Blue	0x001F
0	3	0x1E00000E	White/Blue	0x001F
0	4	0x1E000012	White/Blue	0x001F
0	5	0x1E000016	White/Blue	0x001F
0	6	0x1E00001A	White/Blue	0x001F
0	7	0x1E00001E	White/Blue	0x001F
0	8	0x1E000022	White/Blue	0x001F
0	9	0x1E000026	White/Blue	0x001F
0	10	0x1E00002A	White/Blue	0x001F
0	11	0x1E00002E	White/Blue	0x001F

Figure 3.27: Hello world example: colors.

Addresses in tables 3.26 and 3.27 are calculated using the procedure explained in the previous section and this section. By writing the values to the addresses specified in the tables, 'Hello World' is printed on screen at row 0, as seen in figure 3.28.

Row/Col	0	1	2	3	4	5	6	7	8	9	10	11	12	...	79
0	H	e	l	l	o		W	o	r	l	d	!			
1															
2															
3															
4															
...															
24															

Figure 3.28: Hello world example: results on screen.

3.6.3 Font Memory

In previous sections we have shown how data stored in memory bank 0 (character memory) and memory bank 1 (attribute memory) are used to draw text on screen. The only remaining part is to describe the actual pixels inside every cell.

Every cell of the 2000 cells (on screen) consists of 16x9 pixels. The width of the cell is 9 pixels and the height of the cell is 16 pixels. Since the screen has 25 rows and 80 columns of cells, screen width is 80*9 which is equal to 720 pixels, and sreen height is 25*16 = 400. This results in a resolution of 720x400, which is the default screen resolution of many other systems, including IBM PC (during boot time).

Every ASCII character has a specific shape that is retrieved from bank 2 and bank 3 to be drawn on screen. For example, if cell at row 0, column 0 contains ASCII character

'A', VGA hardware looks for the shape of 'A' in bank 2 and bank 3. A data cell in VGA memory is 9 bits wide. Since every cell of screen has 16 rows and 9 columns of pixels, we can describe the shape of the cell using 16 9-bit data cells, with one bit for every pixel.

This means that coloring inside data cell is monochrome; a pixel is either 0 or 1. If the color of the pixel is 0, VGA chooses to draw background color described in the 'attribute' memory as we explained before. On the other hand, if pixel is 1, VGA chooses to draw the foreground color in that pixel.

For instance, consider the binary representation of letter 'A' in figure 3.29. The representation consists of 16 rows, with every row having 9 bits (the size of a data cell). The whole 16x9 bits describe the whole 16x9 pixels of any cell having 'A' in its character memory location, as we explained before.

```
row 0: 000000000 | 0x0000
row 1: 000000000 | 0x0000
row 2: 000010000 | 0x0010
row 3: 000111000 | 0x0038
row 4: 001101100 | 0x006C
row 5: 011000110 | 0x00C6
row 6: 011000110 | 0x00C6
row 7: 011111110 | 0x00FE
row 8: 011000110 | 0x00C6
row 9: 011000110 | 0x00C6
row A: 011000110 | 0x00C6
row B: 011000110 | 0x00C6
row C: 000000000 | 0x0000
row D: 000000000 | 0x0000
row E: 000000000 | 0x0000
row F: 000000000 | 0x0000
```

Figure 3.29: Shape of letter 'A' in font memory.

So if cell 0 of bank 0 contains ASCII code of letter 'A' (which is 0x0041), and cell 0 of bank 1 contains color code of white foreground and blue background (0x001F), VGA retrieves the shape of letter 'A' (0x0041 in ASCII) from bank 2 and bank 3. The retrieved data are the 16 9-bit words shown in figure 3.29.

Bits with value '0' describe pixels that will have the background color. Bits with value '1' describe pixels that will have foreground color. The final result is a matrix of 16x9 pixels drawn on screen, with letter 'A' drawn in white, over a blue background.

Bank 2 and 3 are called 'font memory', since they are used to describe the shapes of characters drawn on screen. Every ASCII character in the ASCII character space is described using 16 9-bit words stored in bank 2 and bank 3. Since we have 256 possible characters (remember, bit 8 of bank 0 is always ignored, so we only have 8 bits, from 0 to 7, to describe a character), then we need $256 * 16 = 4096$ words, i.e, we need two banks.

The odd rows of a given ASCII character (row 1, 3, 5, ...) are stored in bank 3, while even rows are stored in bank 2. First 8 words of bank 2 describe ASCII code 0, second 8 rows describe ASCII code 1, and so on... The same structure is used in bank 3. The OS loads font data on booting, then starts filling screen with ASCII characters by filling in bank 0 and 1.

3.6.4 Special Registers

As we mentioned before, VGA memory is mapped into 0x1E00000-0x1E03FFF. We described how to access any location in the four memory banks which all appear in this address region. However, there are several locations inside this memory region that are reserved for special purposes. These location are called 'special registers'.

Special registers are written using SH instruction, the same way writes to memory banks are performed. The only difference is that writes to sepcial register locations are handled using special circuits. Table 3.30 shows all special registers and their locations.

Memory Address	Register Name	Function
0x1E001FFA	ROW_BASE_REG	Screen scroll location
0x1E001FFC	CURSOR_ROW_REG	Cursor row location
0x1E001FFE	CURSOR_COL_REG	Cursor column location

Figure 3.30: VGA special registers.

It is important to note that all these registers are 8 bits wide, so bits 8-15 of the written half-word is always ignored. Now we shall summarize the purpose of each register in the following points:

- **ROW_BASE_REG**: Used to scroll up/down the screen. If this register is 0, screen is not scrolled, and everything is displayed normally as expected. Otherwise, this register shall contain the index of the row that will be displayed on the top of the screen. For example, if this register contains 10, then first row (the row on the top of screen) is row 10, next row is 11, and so on, until we reach row 24. Next rows will be 0, 1, up to 9. This helps when the OS needs to scroll up the screen by one row. The OS doesn't need to redraw the whole buffer; instead it can just increase this register by 1, and clear last row (the row at the bottom of the screen). Valid values for this register are 0 to 24.
- **CURSOR_ROW_REG**: This specifies the row number where the blinking cursor appears. Valid values are 0 to 24. Any other value will cause the cursor to be hidden.
- **CURSOR_COL_REG**: This specifies the column number where the blinking cursor appears. Valid values are 0 to 79. Any other value will cause the cursor to be hidden.

3.7 Interrupt Controller

Since the CPU has only one interrupt pin (IRQ signal), we need a device that works as an interface between various interrupt sources and the CPU. The device has 8 input IRQ# pins, and outputs only one IRQ signal that is fed into the CPU. The 8 input IRQs come from various interrupt sources in the system, as shown in table 3.31.

IRQ Number	Device
0	Programmable timer
1	Keyboard controller
2	N/C
3	Reserved
4	N/C
5	N/C
6	N/C
7	N/C

Figure 3.31: IRQ lines and their connected devices.

When a device signals an IRQ on its IRQ line, the interrupt controller interrupts the CPU by activating the IRQ line between the controller and the CPU. The interrupt controller also enters a 'disabled state', it doesn't interrupt the CPU anymore until it is enabled again.

The interrupt service routine (ISR) will ask the interrupt controller for the number of the device which caused the interrupt. It then re-enables the interrupt controller so that when the CPU returns from the ISR, it can be interrupted again (notice that interrupts are disabled during ISR execution, unless re-enabled by the program itself).

If multiple devices interrupt the CPU at the same time, the controller gives priority to the device with lower interrupt number. This means that IRQ0 has the highest priority, and IRQ7 has the lowest priority. The device with the higher priority is served first, then next device, and so on, until no more devices have pending interrupts.

Interrupt controller appears to the programmer in memory region 0x1E802000 - 0x1E802FFF. We note that all memory addresses mentioned in this section are physical addresses. Refer to section 3.2.3 for more information on address translation in MIPS architecture. Table 3.32 shows memory-mapped registers of interrupt controller.

Memory Address	Access Type	Register Name	Function
0x1E802000	READ	IRQ_NO	IRQ number
0x1E802000	WRITE	PIC_STATE	Enable/disable the controller

Figure 3.32: Interrupt controller registers.

It is important to note that all registers are read and written using LW and SW instructions, respectively. We summarize the purpose of each register in the following points:

- **IRQ_NO**: A read-only register that is accessed by reading the word at phsyical location 0x1E802000. It returns the IRQ number of the interrupting device. It is usually read inside the ISR of the operating system. The OS reads this register (instead of polling each device) to know which device has interrupted the CPU, then executes the suitable interrupt handler.
- **PIC_STATE**: A write-only register that is accessed by writing the word at physical location 0x1E802000. If the OS writes 0 to this location, interrupt controller is disabled. If the OS writes 1, interrupt controller is enabled again. Any other value is invalid and shouldn't be used. The OS usually needs to re-enable the interrupt controller after handling an interrupt, in order to allow more interrupts to happen if there are any.

3.8 Programmable Timer

The timer module is analogous to programmable interval timer (PIT) of IBM PC. It contains a counter register that is incremented (by 1) on every clock cycle. The timer contains another register, called threshold, which specifies the maximum value of the counter. When the counter reaches threshold, it is reset to 0, and an IRQ signal is emitted.

Input clock must be 50MHz. If system clock has a frequency greater than 50MHz, it must be divided before it is fed into the timer. The timer is attached to IRQ line 0, the highest priority IRQ line. The timer can be used to keep track of the time of day, by programming it to signal an interrupt every 1sec, 1msec, or any other value.

A multi-tasking OS can use the timer interrupt to trigger the scheduler, in order to guarantee that no process will take more than a specific quantum of time. If a process times out, the timer will signal an interrupt, and the ISR will call the scheduler routine which will perform the task switching procedure.

The threshold register can be seen as a frequency divider for the input frequency (50MHz). For example, if the threshold register is programmed with a value of 2, the counter will be reset every 2 clock cycles, generating an interrupt every $2 * period$ units of time, where *period* is the period of the input frequency (period of 50MHz is 20ns). This means that the frequency of the interrupt signal is 25MHz.

For instance, assume that the OS wants the timer to generate an interrupt every 10ms. This means that the period of the interrupt signals is 10ms. The value of the threshold register should be 10ms/20ns, where 20ns is the period of the 50MHz input clock. In this case, threshold is 500000. This simply means that counter will reset every 500000 cycles.

Programmable timer appears to the programmer in memory region 0x1E801000 - 0x1E801FFF. We note that all memory addresses mentioned in this section are physical addresses. Refer to section 3.2.3 for more information on address translation in MIPS architecture. Table 3.33 shows memory-mapped registers of interrupt controller.

Memory Address	Access Type	Register Name	Function
0x1E801000	READ	PIT_COUNTER	The counter
0x1E801000	WRITE	PIT_THRESHOLD	The threshold register

Figure 3.33: Programmable timer registers.

It is important to note that all registers are read and written using LW and SW instructions, respectively. We summarize the purpose of each register in the following points:

- **PIT_COUNTER**: A read-only register that is accessed by reading the word at phsyical location 0x1E801000. It returns the current value of the internal counter of the timer, which resets to zero whenever it reaches the value of the threshold.
- **PIT_THRESHOLD**: A write-only register that is accessed by writing the word at physical location 0x1E801000. Any write to this location will update the value of the threshold register which specifies the maximum limit of the counter.

3.9 Keyboard Controller

The keyboard controller provides an interface between our system and an external PS/2 keyboard that is used as an input device. When the user presses down (or up) one of the keys on the keyboard, a specific scan code is sent from the PS/2 keyboard to the keyboard controller, which immediately activates IRQ1 line to trigger an interrupt.

When the operating system is interrupted, the interrupt service routine must query the keyboard controller to retrieve the scan code of the key that has been pressed. Every key in the keyboard has a specific scan code, as shown by figure 3.34.

Keyboard controller module contains an internal 16-byte FIFO buffer to store up to 16 scan codes, with every scan code stored in one byte. So for example, if user presses on Shift key, then immediately presses down 'H' key, the two scan codes are stored in the buffer until the operating system reads them.

When a key is pressed up (released), the PS/2 keyboard shall send 0xF0, followed by the scan code of the released button. When the OS reads the buffer, it will find 0xF0 in the head of the buffer, and the scan code of the released button in the next byte of the buffer.

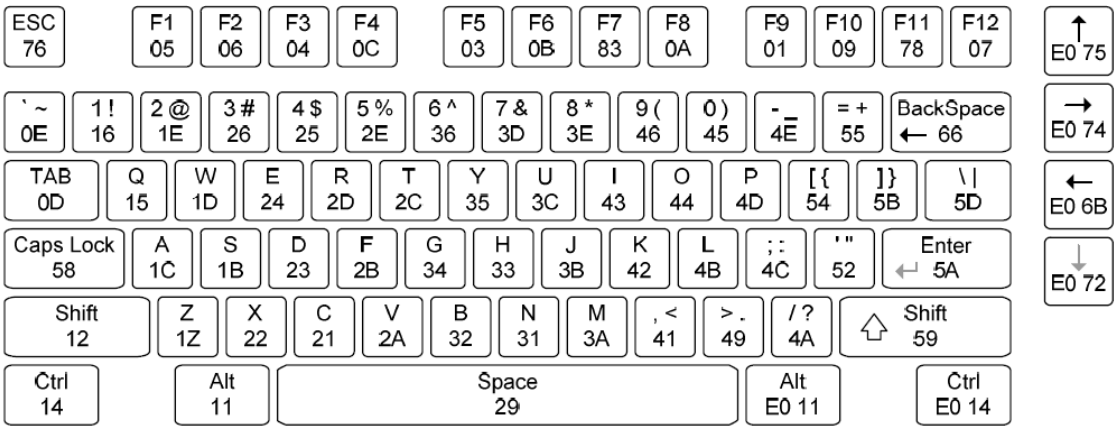


Figure 3.34: PS/2 keyboard scan codes

Keyboard controller appears to the programmer in memory region 0x1E800000 - 0x1E800FFF. We note that all memory addresses mentioned in this section are physical addresses. Refer to section 3.2.3 for more information on address translation in MIPS architecture. Table 3.35 shows memory-mapped registers of interrupt controller.

Memory Address	Access Type	Register Name	Function
0x1E800000	READ	KBD_BUFHEAD	Buffer head register

Figure 3.35: Keyboard controller registers

It is important to note that all registers are read using LBU instruction. We summarize the purpose of each register in the following points:

- **KBD_BUFHEAD**: A read-only register that is accessed by reading the byte at phsyical location 0x1E800000. It returns the value of the byte at the head of the internal buffer. We note that when the buffer head is read, it is flushed; next read will cause the keyboard controller to return the next byte in the buffer, and so on, until the buffer becomes empty. When the buffer is empty, a value of '0' is returned. When ISR executes, the OS should keep reading this register until it returns zero, to make sure that all values stored in the buffer are read and no scan code is lost.

Chapter 4

Implementation

Chapter 5

Software

Chapter 6

Results

Chapter 7

Conclusion

7.1 Future Work

Appendix A

Instruction Listing