Architetture dei Sistemi di Elaborazione 02GOLOV	Computer Architectures 02LSEYG
Laboratory 0x02	**Expected delivery of lab_02.zip including:** **- program_1.s** **- lab_02.pdf (fill and export this file to pdf)**
Delivery date 23/10/2025	

All the previous steps for compiling the source code, simulating with gem5 and visualizing the pipeline have been collected into a workspace available at:

https://github.com/cad-polito-it/ase_riscv_gem5_sim

To create your simulation workspace, you can run the following git commands:

- For HTTPS clone:

```
~/my_gem5Dir$ git clone https://github.com/cad-polito-it/ase_riscv_gem5_sim.git
```

- For SSH:

```
~/my_gem5Dir$ git clone git@github.com:cad-polito-it/ase_riscv_gem5_sim.git
```

If you wish to install on your machine the tools (without using the VM or LABINF's PCs) the repository contains information (and scripts) on installing the necessary tools out of the box (and generate a *setup_default* file accordingly). See README

IMPORTANT: The repository contains three *setup_default* files:

- *setup_default*: an example of out of box installation and different tool paths.
- *setup_default.vm*: tool paths used in the virtual machines.
- *setup_default.labinf*: tool paths used in the LABINF machines.

Follow the HOWTO instructions available on the GitHub Repository for simulating a program. You simply run the following command:

```
~/ase_riscv_gem5_sim$ ./simulate.sh
```

And select the program you want to simulate and see on the pipeline visualizer.

A useful list of RISC-V instructions can be found at:

https://msyksphinz-self.github.io/riscv-isadoc/#_rv32i_rv64i_instructions

- Exercise 1

The `riscv_in_order_hen_patt.py` (in the gem5 folder inside the workspace) file contains the pipeline configuration. Here you can modify the operation latency and the issue latency of integer and floating-point functional units (ALU, Multiplier, Divider).

You must configure the Gem5 simulator with the ***Initial Configuration*** as described below:

```
INTEGER_ALU_LATENCY = 1
INTEGER_MUL_LATENCY = 1
INTEGER_DIV_LATENCY = 1
FLOAT_ALU_LATENCY = 8
FLOAT_MUL_LATENCY = 24
FLOAT_DIV_LATENCY = 42
```

Write an assembly program (**`program_1.s`**) for the *RISCV* architecture described before being able to implement the following high-level code:

```
for (i = 31; i >= 0; i--) {
        v4[i] = v1[i]*v1[i] – v2[i];
        v5[i] = v4[i]/v3[i] – v2[i];
        v6[i] = (v4[i]-v1[i])*v5[i];
}
```

Assume that the vectors v1[], v2[], and v3[] have been previously allocated in memory and contain 32 single-precision **floating-point values;** also assume that v3[] does not contain any '0'. Additionally, the vectors v4[], v5[], v6[] are empty vectors allocated in memory.

Calculate the data memory footprint of your program:

Data	Number of bytes
V1	*128 byte*
V2	*128 byte*
V3	*128 byte*
V4	*128 byte*
V5	*128 byte*
V6	*128 byte*
Total	*768 byte*

- Calculate the CPU performance equation (CPU time) of the above program by assuming a clock frequency of 15 MHz:

$$\text{CPU time} = \left(\sum_{i=1}^{n} \text{CPI}_i \times \text{IC}_i\right) \times \text{Clock cycle period}$$

By definition:

- CPI_i is equal to the number of clock cycles required by the related functional unit to execute the instruction (EX stage);

- IC_i is the number of times an instruction is repeated in the referenced source code.

- Recalculate the CPU performance equation assuming that you can triple the speed of just one unit at a time:
  b) FP multiplier (MUL) unit latency: 24 → 8 clock cycles
  c) FP divider (DIV) unit latency: 42 → 14 clock cycles

Table 1: Calculate the CPU time ***by hand***

	Initial CPU time (a)	CPU time (b – MUL speeded up)	CPU time (c – DIV speeded up)
program_1.s	$2.684*10^{-4}$	$2.002*10^{-4}$	$2.087*10^{-4}$

- Using the simulator, calculate the CPU time again and fill in the following table:

Table 2: Collect the CPU time ***using the simulator***

	Initial CPU time (a)	CPU time (b – MUL speeded up)	CPU time (c – DIV speeded up)
program_1.s	$2.6*10^{-4}$	$1.91*10^{-4}$	$2*10^{-4}$

Are there any differences? If so, where and why? If not, please provide some comments in the box below:

Your answer: I ***tempi risultano inferiori nella simulazione perché il mio calcolo manuale non tiene conto dell'esecuzione "in parallelo" di istruzioni su diverse unità logiche, nel caso specifico operazioni su interi possono terminare prima ancora che la precedente operazione su float termini poiche' non condividono la stessa unita' di calcolo. Sommando semplicemente le durate ho quindi ottenuto una stima a rialzo.***

- Exercise 2

Using the Gem5 simulator, validate experimentally the Amdahl's law, defined as follows:

$$\text{speedup}_{\text{overall}} = \frac{\text{execution time}_{\text{old}}}{\text{execution time}_{\text{new}}} = \frac{1}{(1 - \text{fraction}_{\text{enhanced}}) + \frac{\text{fraction}_{\text{enhanced}}}{\text{speedup}_{\text{enhanced}}}}$$

a. Using the program developed before **program_1.s**
b. Modify the processor architectural parameters related (in `riscv_in_order_hen_patt.py`) to multicycle instructions latency in the following way:

la	2 × 6	= 12
flw		= 3
addi		= 1
slli		= 1

} 17 + 5 = 22

blt	= 1
add	= 1
flw	= 1
add	= 1
flw	= 1
add	= 1
flw	= 1

} 7 · 32 = 224

fmul_s	= f_mul_lat
fsub_s	= f_alu_lat
add	= 1
fsw	= 1

} 32 · (f_mul_lat + f_alu_lat) + 32 · 2

fdiv_s	= f_div_lat
fsub_s	= f_alu_lat
add	= 1
fsw	= 1

} 32 · (f_div_lat + f_alu_lat) + 32 · 2

fsub_s	= f_alu_lat
fmul_s	= f_mul_lat
add	= 1
fsw	= 1

} 32 · (f_mul_lat + f_alu_lat) + 32 · 2

addi	= 1	
slli	= 1	} 4 · 32 = 128
j	= 2	
blt	= 2	} 2
l:	= 1	
l:	= 1	} 3
ecall	= 1	

$$CPU_t = 379 + 32 \cdot \left(2 \cdot f\_mul\_lat + 3 \cdot f\_alu\_lat + f\_div\_lat\right) \cdot \frac{1}{15 \cdot 10^6}$$

① $379 + 32 \cdot \left(2 \cdot 24 + 3 \cdot 8 + 42\right) \cdot \frac{1}{15 \cdot 10^6} = 4027 \cdot \frac{1}{15 \cdot 10^6} = 2.684 \cdot 10^{-4}$

② $379 + 32 \cdot \left(2 \cdot 8 + 3 \cdot 8 + 42\right) \cdot \frac{1}{15 \cdot 10^6} = 3003 \cdot \frac{1}{15 \cdot 10^6} = 2.002 \cdot 10^{-4}$

③ $379 + 32 \left(2 \cdot 24 + 3 \cdot 8 + 14\right) \cdot \frac{1}{15 \cdot 10^6} = 3131 \cdot \frac{1}{15 \cdot 10^6} = 2.087 \cdot 10^{-4}$

Configuration 1:

```
INTEGER_ALU_LATENCY = 1
INTEGER_MUL_LATENCY = 1
INTEGER_DIV_LATENCY = 1
FLOAT_ALU_LATENCY = 6
FLOAT_MUL_LATENCY = 20
FLOAT_DIV_LATENCY = 38
```

Configuration 2:

```
INTEGER_ALU_LATENCY = 1
INTEGER_MUL_LATENCY = 1
INTEGER_DIV_LATENCY = 1
FLOAT_ALU_LATENCY = 4
FLOAT_MUL_LATENCY = 16
FLOAT_DIV_LATENCY = 30
```

Configuration 3:

```
INTEGER_ALU_LATENCY = 1
INTEGER_MUL_LATENCY = 1
INTEGER_DIV_LATENCY = 1
FLOAT_ALU_LATENCY = 2
FLOAT_MUL_LATENCY = 4
FLOAT_DIV_LATENCY = 8
```

Compute both manually (using the Amdahl's Law) and with the simulator the speed-up for any one of the previous processor configurations. Compare the obtained results and complete the following table.

Table 5: **program_1.s speed-up computed *by hand and by simulation***

Proc. Config. **Speed-up comp.**	Initial config. [c.c.]	Config. 1 [c.c.]	Config. 2 [c.c.]	Config. 3 [c.c.]
By hand	*4027 cc*	*3451 cc 16.69%*	*2747 cc 46,59%*	*1083 cc 271%*
By simulation	*3903 cc*	*3391 cc 15.09%*	*2751 cc 41,87%*	*1247 cc 212%*