

APAI Lab02: PULP Embedded Programming

Davide Nadalini, Lorenzo Lamberti, Alberto Dequino, Luca Bompani, Francesco Conti.

(University of Bologna)

d.nadalini@unibo.it lorenzo.lamberti@unibo.it
alberto.dequino@unibo.it

Objective of the Class

Intro: PULP platform and the PULP-SDK

Tasks: some basics of C programming on PULP:

- Hello world
- vector sum,
- Matrix-vector multiplication
- Measuring execution performance

Programming Language: C

Lab duration: 3h

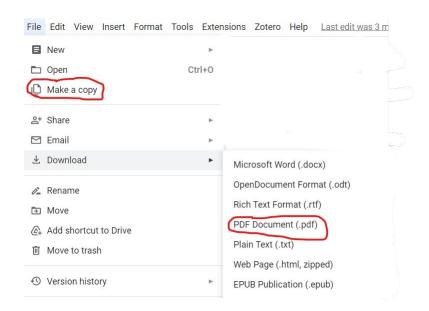
The class is meant to be interactive: coding together and on your own!



How to deliver the Assignment

You will deliver ONLY the GDOC assignment, no code

- Copy the google doc to your drive, so that you can edit. (File -> make a copy)
- Fill the tasks on this google doc.
- Export to pdf format.
- Rename the file to: LAB<number_of_the_lesson>_APAI_<your_name>.pdf
- Use Virtuale platform to load ONLY your .pdf file

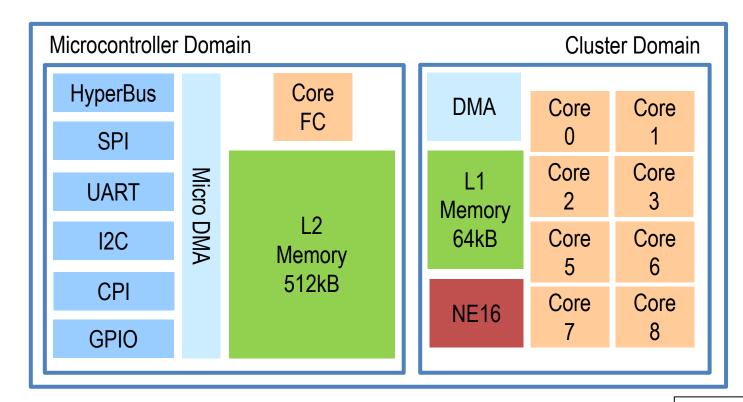






INTRO

Get confident with the PULP Platform



- Cores: 1 + 8
- On-chip Memories
 - A level 2 Memory, shared among all cores
 - A level 1 Memory, shared by the 8-cores cluster
- cluster-DMA: A multi-channel 1D/2D DMA, controlling the transactions between the L2 and L1 memories
- micro-DMA: A smart, lightweight and completely autonomous DMA () capable of handling complex I/O scheme
 - **Bus+Peripherals:** HyperBus, I2S, CPI, timers, SPI, GPIOs, etc...

GitHub HW Project: https://github.com/pulp-platform/pulp **HW Documentation**:

https://raw.githubusercontent.com/pulp-platform/pulp/master/doc/datasheet.pdf APAI2023 - Lab03

NB: this is the architecture you find on our nano-drones and GAP boards!





We will work in groups of 2 people

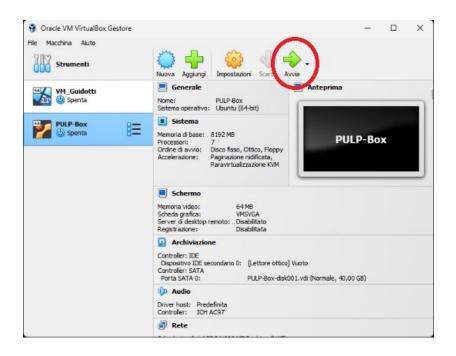
Lab PCs are limited! Find a colleague to work with!

Remember to send the assignment with both names in the file name!!



Opening the VM and VSCode

- 1. On the lab's PCs, open the file explorer and go to This PC, C:/VM_Nadalini/
- 2. Double click on PULP-box.ova
- 3. VirtualBox opens, just click on "Fine"
- 4. Wait for the VM to be imported
- 5. Open the VM with "Avvia"





Password is 'pulp'



Opening the VM and VSCode

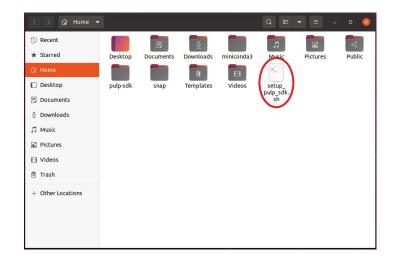
1. Open a terminal (right click – open a new terminal)

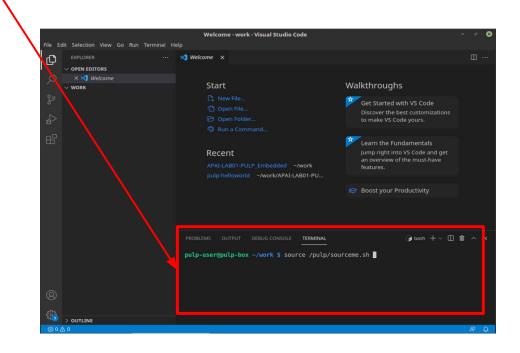
2. Open a text editor (For example "VSCode"):

Now you can use the integrated terminal (open with CTRL+J) to run your applications!

IMPORTANT: every time you open a new terminal to work on PULP, launch

\$ source setup_pulp_sdk.sh





\$ code .

Getting Started: *Helloworld*

IMPORTANT: activate the pulp-sdk module file <u>every</u> time a new shell is open.

```
$ source setup_pulp_sdk.sh
```

HOW TO RUN THE CODE:

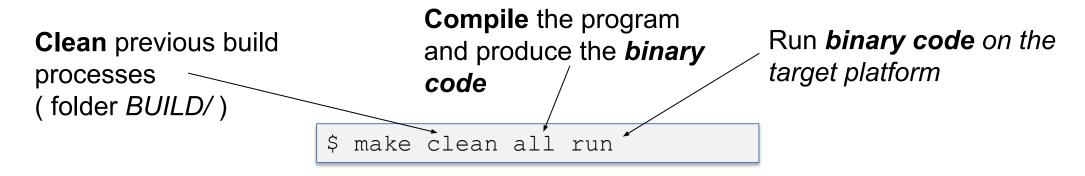
```
$ git clone https://github.com/EEESlab/APAI24-LAB02-PULP-Embedded-Programming
$ cd APAI24-LAB02-PULP-Embedded-Programming
$ cd pulp-helloworld/
$ make clean all run
test.c
```

Can you see the Helloworld from PULP! ?

```
int main()
{
    printf("Helloworld from PULP!\n");
}
```



Behind the box: Build Automation with Makefiles



Build automation is the process of automating the main steps required to create a software, including *compiling*, assembling, linking and (possibly) testing

Make is one of the most widespread utilities

configuration files are called Makefiles

A Makefile contains rules in the form:

target: prerequisites
<TAB> command

Makefile

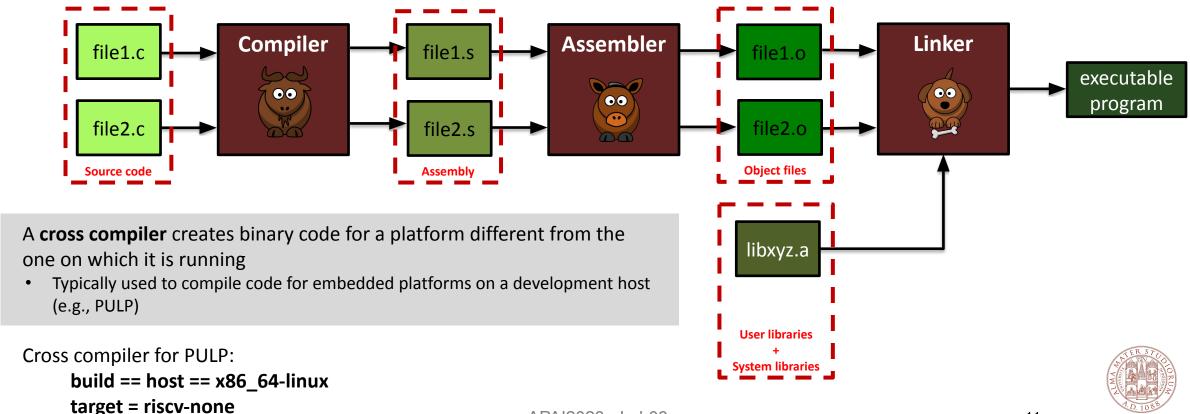
```
# This is a comment
APP_SRCS += test.c
APP_CFLAGS += -03 -g
APP_LDFLAGS +=
include $(RULES_DIR)/pmsis_rules.mk
```

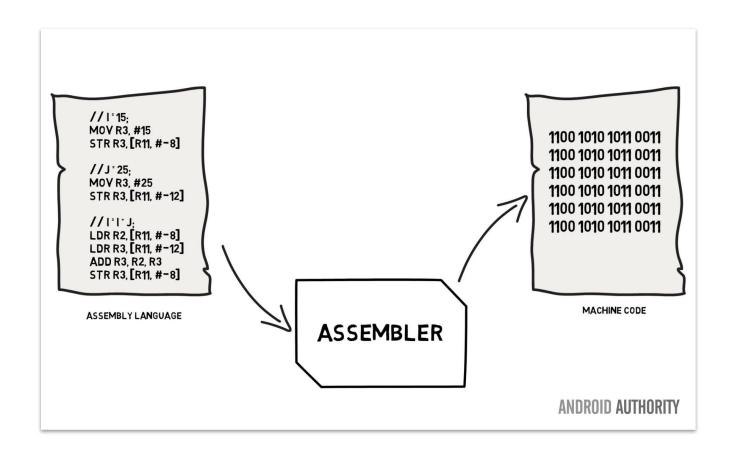


Compilation toolchain

A compilation toolchain includes several tools to achieve its final goal

- The **COMPILER** translates high-level source code (e.g., C) to a lower-level representation (e.g., assembly)
- The ASSEMBLER is program that translates assembly language to machine language
- The LINKER combines multiple object files into a single executable





Assembly vs. machine code

Machine code bytes	Assembly language statements	
B8 22 11 00 FF	foo: movl \$0xFF001122, %eax	
01 CA	addl %ecx, %edx	
31 F6	xorl %esi, %esi	
53	pushl %ebx	
8B 5C 24 04	movl 4(%esp), %ebx	
8D 34 48	<pre>leal (%eax,%ecx,2), %esi</pre>	
39 C3	cmpl %eax, %ebx	
72 EB	jnae foo	
C3	retl	

Instruction stream

B8 22 11 00 FF 01 CA 31 F6 53 8B 5C 24 04 8D 34 48 39 C3 72 EB C3

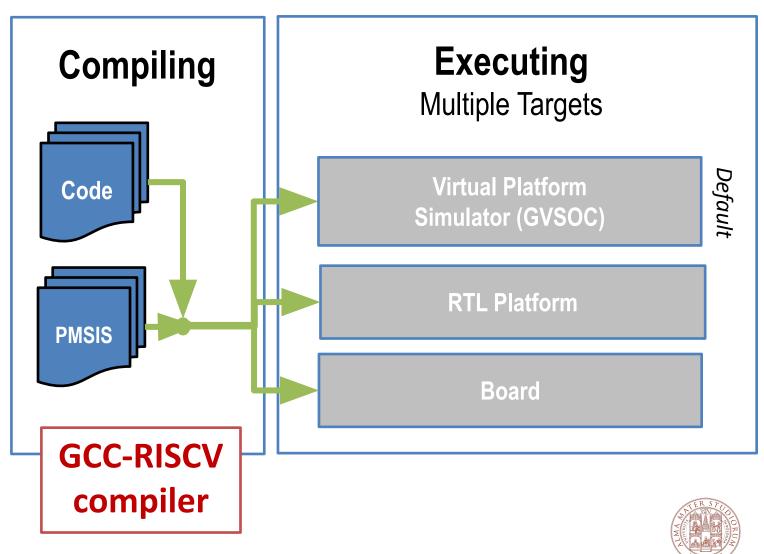


PULP Software Environment Workflow

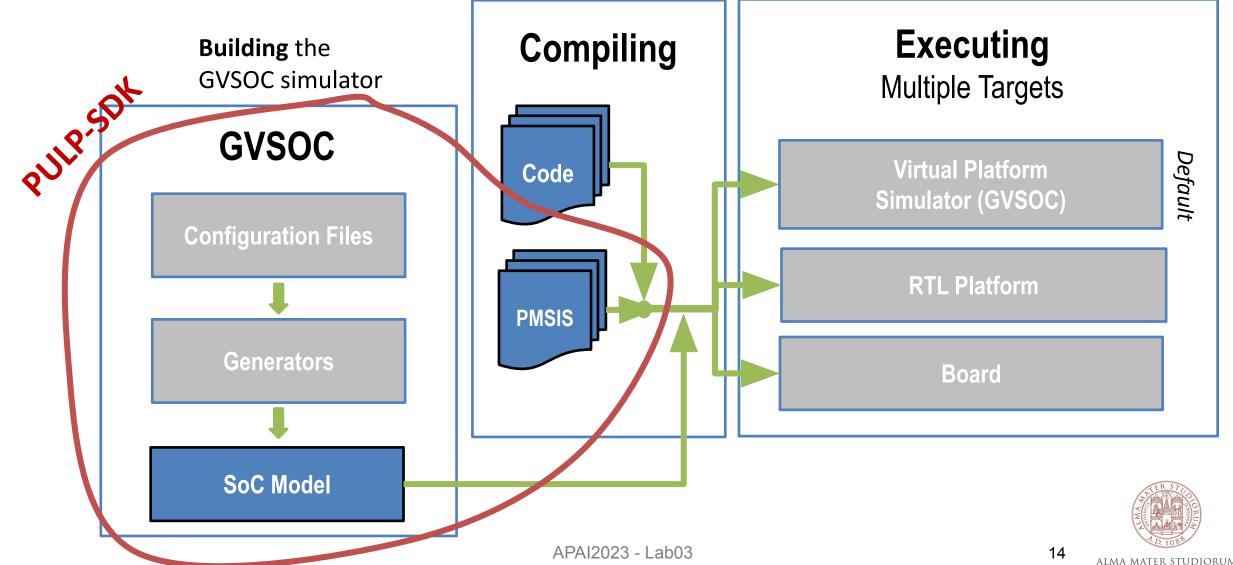
Application code: *main* function, application libraries, e.g. CNN inference function code...

Runtime SW: Peripheral *Drivers,* RTOS, Board support packages (BSP)

PMSIS layer



PULP Software Environment Workflow



PULP SDK (software development kit)

The PULP-SDK (https://github.com/pulp-platform/pulp-sdk) includes a **PULP platform simulator (GVSOC)** and the SW libraries.

Check the /pulp/pulp-sdk folder:

- rtos/ runtime code and software stack
- tools/ configuration files, python generators, pulp runner and the gysoc components
- tests/ sample code to test the platform's features
- applications/ full-application codes

\$ cd /pulp/pulp-sdk

```
oulp-user@pulp-box /pulp/pulp-sdk $ ll
rwxrwxr-x 9 pulp-user pulp-user
                                  4096 Feb 3 2021 ./
rwxrwxr-x 6 pulp-user pulp-user
           6 pulp-user pulp-user
           2 pulp-user pulp-user
                                               2021 configs/
           8 pulp-user pulp-user
                                  4096 Feb 3
                                               2021 .git/
             pulp-user pulp-user
                                               2021 .gitignore
             pulp-user pulp-user
           1 pulp-user pulp-user 11357 Feb 3
             pulp-user pulp-user
           1 pulp-user pulp-user
          4 pulp-user pulp-user
rwxrwxr-x 2 pulp-user pulp-user
drwxrwxr-x 10 pulp-user pulp-user
                                  4096 Feb 3
pulp-user@pulp-box /pulp/pulp-sdk $
```



You could get the latest update from the open-source community and rebuild the **GVSOC** simulator

```
$ git pull origin main
 source configs/pulp-open.sh
$ make build
```

GVSoC – Features

Virtual platform features:

- C++ for fast native simulation
- Python for instantiation + configuration
- Complete set of traces to see what happen

Timing model:

- Fully-event based, instances can generate events at specific time
- Includes timing models for interconnects, DMACs, memories...
- Performance counters for information from the execution

Simulation performance:

- Around 1MIPS simulation speed
- Functionally aligned and calibrated with HW
- Timing accuracy is within 10-20% of target HW





TASK1

TASK 1: vector sum example

Preparation

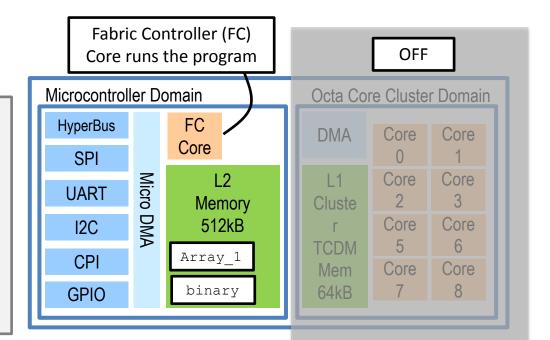
1. Clone repository for the lab, if you didn't already do it

git clone

https://github.com/EEESlab/APAI24-LAB02-PULP-Embedded-Programming

2. Go into folder for task1

cd APAI24-LAB02-PULP Embedded Programming/ cd vector sum/ make clean all run



Tasks

Read your assignment!

The vector sum() function returns the element-wise add of the values in array 1[N].

Note: the main include the function

testbench: $S = \sum_{i=1}^{N} i = \frac{N \cdot (N-1)}{2}$

Good practice for test!

- Array initialization (a[i] = i)
- **Function Call**
- **Check Result**

Take now a look to the code. Task What happen if you change: 1.2 (line 13) #define N 350 APAI2023 -

☐ How would you solve it?



SOLUTION: Casting Variables

(line 13) #define N 350

Output: Result is **not** correct. Got 37105 instead of 61425

If printing the array 1 values after initialization:

1 2 3 4 5 6 7 8 9 10 11 12 13 14 [....] 247 248 249 250 251 252 253 254 255 0 1 2 3 4 5 6 7 8 [....] 88 89 90 91 92 93 94

- unsigned char datatype can represent numbers from 0 to 255!! If >255, assigned values get truncated (value % 256)!
- Solution: cast array_1 to int or short int (and the function's arguments!)

(optional) Passing parameters via 'Makefile'

(OPTIONAL) After applying the fix We can pass parameters via Makefile: comment line 13, change the *Makefile* and launch:

```
$ make clean all run N=350
```

```
APP = test
N?=50

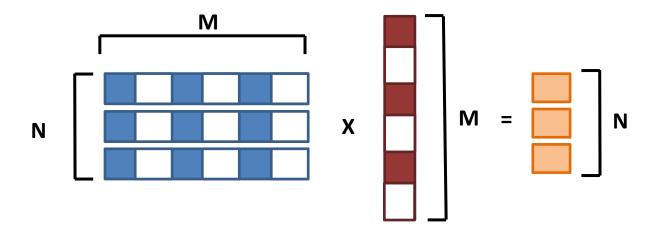
APP_SRCS += test.c
APP_CFLAGS += -03 -g
APP_CFLAGS += -DN=$(N)
APP_LDFLAGS +=
include $(RULES_DIR)/pmsis_rules.mk 19
```





Intro: TASK2

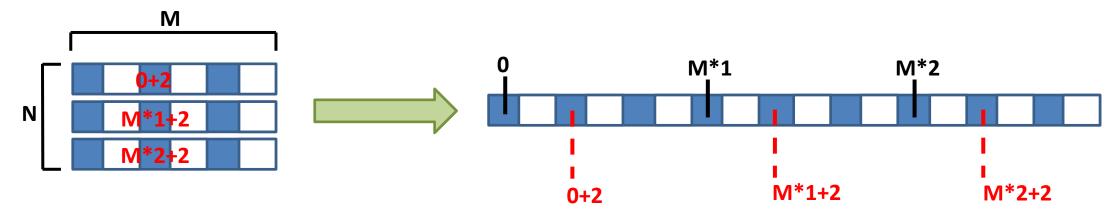
TASK2: Matrix-Vector Product



```
$ cd matrix_vector/
$ make clean all run
```

Example II: Matrix-Vector Product

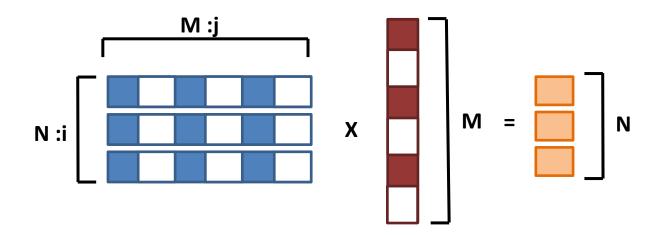
A multi-dimensional array, including a matrix, is efficiently represented as a memory-contiguous array (**NOT** Mat[M][N]).



In general: Mat[i][j] -> M[i*M+j]

```
$ cd matrix_vector/
$ make clean all run
```

Example II: Matrix-Vector Product



```
// generic matrix-vector multiplication
int gemv(int N, int M, float * mat, float *vec, float * output_vec) {
    for (int i=0; i<N; i++) {
        for (int j=0; j<M; j++) {
          vec_o[i] += mat_i[i*size_M+j] * vec_i[j];
        }
    }
}</pre>
```

Assembly Code

A disassembler is a computer program that translates machine language into assembly language \rightarrow the inverse operation to that of an assembler.

Disassembly, the output of a disassembler, is often formatted for human-readability rather than suitability for input to an assembler, making it principally a reverse-engineering tool. [Wikipedia]

To obtain the assembly code use the command:

\$ make dis > dis.txt

Check the RI5CY User manual:

https://riscv.org/wp-content/uploads/2017/05/riscv-spec-v2.2.pdf https://www.pulp-platform.org/docs/ri5cv_user_manual.pdf



Assembly Code

```
1c008706 <gemv>:
                           blez a0,1c008740 <gemv+0x3a>
   1c008706: 02a05d63
   1c00870a: 00259f13
                           slli t5,a1,0x2
   1c00870e: 8fb6
                          mv t6,a3
   1c008710: 01e68e33
                           add t3,a3,t5
                          li t4,0
   1c008714: 4e81
   1c008716: a005
                            1c008736 <gemv+0x30>
  1c008718: 0048230b
                           p.lw t1,4(a6!)
  1c00871c: 0046a88b
                           p.lw a7,4(a3!)
  1c008720: 431c
                          lw a5,0(a4)
  1c008722: 431307b3
                           p.mac a5,t1,a7
  1c008726: c31c
                          sw a5,0(a4)
  1c008728: ffc698e3
                          bne a3,t3,1c008718 <gemv+0x12>
  1c00872c: 0e85
                          addi t4,t4,1
   1c00872e: 967a
                          add a2,a2,t5
                          addi a4,a4,4
   1c008730: 0711
   1c008732: 01d50763
                           beg a0,t4,1c008740 <gemv+0x3a>
   1c008736: 86fe
                          mv a3 t6
  1c008738: 8832
                              a6.a2
  1c00873a: fcb04fe3
                          bgtz a1,1c008718 < gemv+0x12>
   1c00873e b7fd
                          i 1c00872c < gemv + 0v26>
1c008738:
                8832
                                            a6,a2
  PC
                  opcode
                                        Instruction
```

In the Makefile:

- 1. Change from -01 to -03. What has changed in the assembly?
- 2. Remove –mnohwloop. What has changed in the assembly?
- NB. After changing the compiler flags, compile the code again and generate and disassembly using make dis

Program counter



Assembly Code: O3 optimization

```
APP = matrix-vector

APP_SRCS += test.c

APP_CFLAGS += -O3 -g -mnohwloops

APP_LDFLAGS +=

include $(RULES_DIR)/pmsis_rules.mk
```

```
1c008706 <gemv>:
                                                        The -03 optimized code get rid of a lw instruction
1c008706: 02a05c63
                      blez a0,1c00873e <gemv+0x38>
                      blez a1,1c00873e <gemv+0x38>
1c00870a: 02b05a63
                                                         because the accumulator is kept in the register file!
                      slli t4,a1,0x2
1c00870e: 00259e93
1c008712: 050a
                     slli a0,a0,0x2
1c008714: 00a70e33
                      add t3,a4,a0
                                                                  // generic matrix-vector multiplication
1c008718: 01d68333
                      add t1,a3,t4
                                                                  void gemv(int size N, int size M,
                      p.lw a1,4(a4!) # 10004 < l1 heap size+0x20>
1c00871c: 0047258b
                                                                  float * mat I, float *vec i, float * vec o)
1c008720: 87b6
                     mv a5,a3
                     mv a0,a2
1c008722: 8532
                                                                       for (int i=0; i < size N; i++) {
1c008724: 0045288b
                      p.lw a7,4(a0!)
                                                                         for (int j=0; j<size M; j++) {
1c008728: 0047a80b
                      p.lw a6,4(a5!)
                                                                 Inner
                                                                              // multitply accumulate operation
1c00872c: 430885b3
                      p.mac a1,a7,a6
                                                                 loop
                                                                              vec o[i] += mat i[i*size M+j] * vec i[j];
1c008730: feb72e23
                      sw a1,-4(a4)
1c008734: fef318e3
                      bne t1,a5,1c008724 < gemv+0x1e>
1c008738: 9676
                     add a2,a2,t4
1c00873a: feee11e3
                      bne t3,a4,1c00871c <gemv+0x16>
1c00873e: 8082
                     ret
```

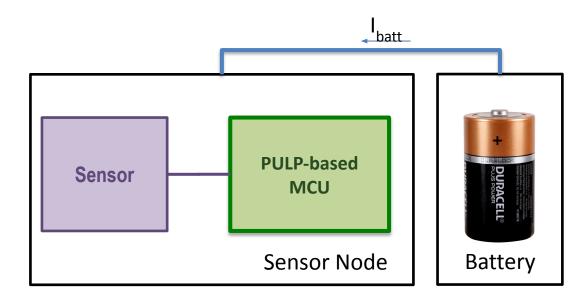
Al2023 - Lab03

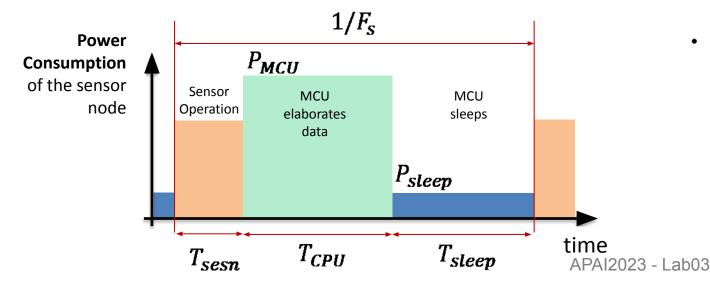
Assembly Code: HW loops

```
APP = matrix-vector
                                                   APP SRCS += test.c
                                                   APP CFLAGS += -O3 -g -mnohwloops
                                                   APP LDFLAGS +=
1c008706 < gemv >:
            04a05463
1c008706:
                              a0,1c00874e <gemv+0x48
                                                   include $(RULES_DIR)/pmsis_rules.mk
1c00870a:
            04b05263
                        blez
                              a1,1c00874e <gemv+0x48
1c00870e:
                              t3,a0,0x2
            00251e13
1c008712:
                              t3,t3,-4
            1e71
1c008714:
            00259e93
                              t4,a1,0x2
            002e5e13
                              t3,t3,0x2
1c008718:
                                                                   Hardware Loops Instructions have been placed!
1c00871c:
            01d68f33
                              t5,a3,t4
                        add
1c008720:
                             t3,t3,1
            0e05
                        addi
1c008722:
            015e407b
                        lp.setup
                                     x0,t3,1c00874c <gemv+0x46>
                                                                            // generic matrix-vector multiplication
1c008726:
            40df07b3
                              a5,t5,a3
                                                                            void gemv(int size N, int size M,
                              a5,a5,-4
1c00872a:
            17f1
                                                                            float * mat I, float *vec i, float * vec o)
                              a1,4(a4!) # 10004 < l1 heap size+0x20 >
1c00872c:
            0047258b
1c008730:
            8389
                              a5,a5,0x2
                        srli
1c008732:
                              a6.a3
                                                                                 for (int i=0; i < size N; i++) {
            8836
                        mν
                              a0,a2
1c008734:
            8532
                                                                                   for (int j=0; j<size M; j++) {
                        mν
                                                                          Inner
1c008736:
            0785
                              a5.a5.1
                        addi
                                                                                        // multiply accumulate operation
                                     x1,a5,1c008748 <gemv+0x42>
1c008738:
            0087c0fb
                        lp.setup
                                                                           loop
                                                                                        vec o[i] += mat i[i*size M+j] * vec i[j];
                              t1,4(a0!)
1c00873c:
            0045230b
1c008740:
            0048288b
                        p.lw
                              a7,4(a6!)
1c008744:
            431305b3
                        p.mac a1,t1,a7
1c008748:
            feb72e23
                              a1,-4(a4)
1c00874c:
            9676
                        add
                              a2,a2,t4
                                                                         2023 - Lab03
1c00874e:
            8082
                        ret
```

UNIVERSITÀ DI BOLOGNA

Assessing the MCU performance





Goal

- Extending the battery lifetime
 - Minimize the sensor node energy consumption
- Real time processing of sensor data

$$\min E_s + E_{MCU}$$
 s.t. $T_{CPU} < 1/F_s$

Assuming:

- a fixed sample rate F_s a negligible sensor energy cost $E_s \ll E_{MCU}(\text{e.g.}, T_{sesn} \ll 1/F_s)$
- a constant power envelope of the MCU for active (P_{MCU}) and sleep $(P_{Sleep} \ll P_{MCU})$ modes

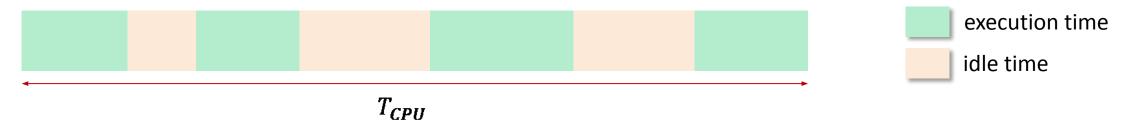
$$\min T_{CPU} P_{MCU} + T_{sleep} P_{sleep}$$
$$T_{CPU} + T_{sleep} = 1/F_s$$

Aim: minimize the processing time T_{CPU}

What can we optimize? -> the execution time

The processing time T_{CPU} composes of:

- the execution time, where the CPU
- the idle time, where the CPU waits for events or IRQ (the core may be clock gated to save power)



Given a CPU clock frequency, the <u>performance of a task (e.g. a C function) can be measured by</u> accounting:

- the number of elapsed clock cycles (N_{clk})
- the **number of instructions** (N_{instr}) executed within the task depends on the code optimization!
- **CPI (Clock Cycles Per Instruction)**: different instructions may take a different number of **clock cycles** (depending on the CPU microarchitecture)

$$T_{CPU} = N_{clk} \times T_{clk} = CPI_{avg} \times N_{instr} \times T_{clk}$$
 $CPI_{avg} = \frac{N_{clk}}{N_{inst}}$

Performance Counters: Measuring the CPU Time

Each RI5CY cores of the PULP platform provide a **performance counter**. These 32-bit counters can be <u>configured</u> to count the:

- Total number of cycles (also includes the cycles where the core is sleeping)
- Number of cycles the core was active (not sleeping)
- · Number of instructions executed
- Number of load data hazards
- Number of jump register data hazards
- Number of cycles waiting for instruction fetches, i.e. number of instructions wasted due to non-ideal caching
- · Number of data memory loads executed. Misaligned accesses are counted twice
- · Number of data memory stores executed. Misaligned accesses are counted twice
- Number of unconditional jumps (j, jal, jr, jalr)
- Number of both taken and not taken branches
- Number of taken branches
- · Number of compressed instructions executed

The performance counters of the cluster cores can also account:

- Number of memory loads to EXT executed. Misaligned accesses are counted twice. Every non-L1 access is considered
 external
- Number of memory stores to EXT executed. Misaligned accesses are counted twice. Every non-L1 access is considered
 external
- Number of cycles used for memory loads to EXT. Every non-L1 access is considered external
- Number of cycles used for memory stores to EXT. Every non-L1 access is considered external
- Number of cycles wasted due to L1/log-interconnect contention

Important: while on the PULP Virtual Platform multiple perf counters can be enabled concurrently, only one counters is awailable on a real HW device!

Performance Counters

Core RISC-V



Using the performance Counters on PULP

Using the PMSIS library APIs:

```
// enable the perf counters of interest
                 1 << PI PERF CYCLES
pi perf conf(
        1 << PI PERF INSTR );
// reset the performance counters
pi perf reset();
// start the performance counters
pi perf start();
// task to profile
foo();
// stop the performance counters
pi perf stop();
// collect and print statistics
uint32 t instr cnt = pi perf read(PI PERF INSTR);
uint32 t cycles cnt = pi perf read(PI PERF CYCLES);
```

Configure which performance counters to enable!

```
typedef enum {
 PI PERF CYCLES
 PI PERF ACTIVE CYCLES = 0, /*!< Counts the number of cycles the core was
 PI PERF INSTR
 PI PERF LD STALL
 PI PERF JR STALL
 PI PERF IMISS
 PI PERF LD
 PI PERF JUMP
                      = 8, /*! < Number of branches. Counts both taken and
 PI PERF BRANCH
                      = 9, /*!< Number of taken branches. */
 PI PERF RVC
                      = 10, /*!< Number of compressed instructions
 PI PERF LD EXT
                      = 12, /*! < Number of memory loads to EXT executed.
 PI PERF ST EXT
 PI PERF LD EXT CYC = 14, /*!< Cycles used for memory loads to EXT.
 PI PERF ST EXT CYC = 15, /*!< Cycles used for memory stores to EXT.
 PI PERF TCDM CONT = 16, /*!< Cycles wasted due to TCDM/log-interconnect
} pi perf event e;
```

/rtos/pmsis/pmsis api/include/pmsis/chips/default.h



TASK2

Profile the gemv

Make a new project by copying the *matrix-vector gemv* example into a new folder

For any of the previous compiler optimization, **measure the performance** <u>using the PMSIS performance counters</u> and report:

- number of clock cycles.
- number of instructions.

$$a \leftarrow a + (b \times c)$$

- Number of Multiply-Add operations
- the CPI.
- compute the number of clock cycles and instructions per elementary operation.
 - Define 1 elementary operation == 1Multiply-and-Accumulate

$$\#MAC = N \times M = 50 \times 50 = 2500$$

	-01	-03	-03 HWLoops
Clock Cycles			
Instr.			
MAC			
СРІ			
Intr/Cycles			
Instr / MAC			



SOLUTION: Profile the *gemv*

Make a new project by copying the *matrix-vector gemv* example into a new folder

```
$ cd matrix_vector/
make clean all run
```

Solution:

```
pi_perf_conf(1<<PI_PERF_CYCLES | 1<<PI_PERF_INSTR);
pi_perf_reset();

pi_perf_start();
// call the matrix-vector fucntion
gemv(N, M, matrix, vector, output_vec);
pi_perf_stop();

uint32_t instr_cnt = pi_perf_read(PI_PERF_INSTR);
uint32_t cycles_cnt = pi_perf_read(PI_PERF_CYCLES);
printf("Num.Istr: %d Num.Cycles: %d \n", instr_cnt,
cycles_cnt);</pre>
```

$$\#MAC = N \times M = 50 \times 50 = 2500$$

	-01	-03	-03 HWLoops
Clock Cycles			
Instr.			
MAC			
СРІ			
Intr/Cycles			
Instr / MAC			



