ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
EIDGENÖSSISCHE TECHNISCHE HOCHSCHULE – LAUSANNE
POLITECNICO FEDERALE – LOSANNA
SWISS FEDERAL INSTITUTE OF TECHNOLOGY – LAUSANNE

Faculté Informatique et Communications
CS–202 Computer Systems
Argyraki K., Bugnion E. & Chappelier J.-C.

NOM : Hanon Ymous
(000000) #0000
Seat #: 0

CS–202 COMPUTER SYSTEMS

Final Exam

June 17th, 2024

INSTRUCTIONS (please read carefully)

IMPORTANT! Please strictly follow these instructions, otherwise your exam may be canceled.

1. You have three hours to complete this examination (3:15–6:15pm).
2. You must **use black or dark blue ink**, neither pencil nor any other color.
3. This is closed book exam.
   Personal notes, two times dual-sided A4 sheets (4 sides in total), allowed.

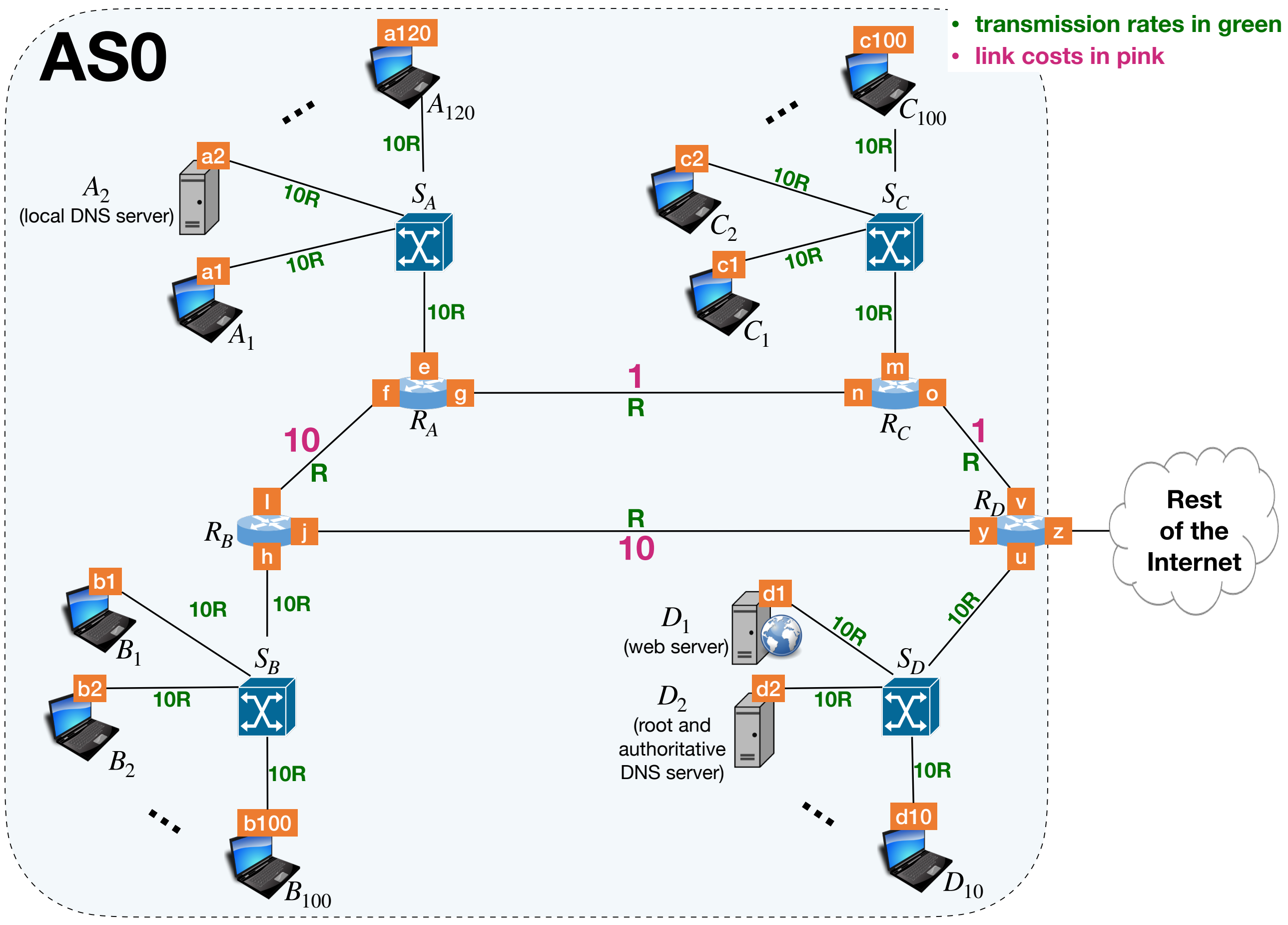
   On the other hand, you may not use any personal computer, mobile phone or any other electronic equipment.
4. Answer the questions directly on the exam sheet; do not attach any any additional sheets; only this document will be graded.
5. Carefully and *completely* read each question so as to do only what we actually ask for. If the statement seems unclear, or if you are in any doubt, ask one of the assistants for clarification.
6. The exam consists of six exercises, which can be addressed in any order, except Question 2 which is a continuation of Question 1. These exercises do not score the same; points are indicated; the total is 125 points. All exercises count for the final grade.

LEAVE THIS EMPTY

Question 1	Question 2	Question 3	Question 4	Question 5	Question 6	**TOTAL**
25	30	19	12	14	25	125

Question 1 – Subnets, prefixes, and packets [25 points]

Consider the topology in Figure 1, which shows a single Autonomous System (AS), AS0.



(For convenience, a copy is also provided on draft paper.)

Figure 1: Network topology for Questions 4 and 5.

- There are 4 routers, R_A, R_B, R_C, and R_D. Router R_D is the only border router of AS0.
- There are 4 Ethernet switches, S_A, S_B, S_C, and S_D. These switches do not have IP addresses.
- There are 120 end-systems A_1 to A_{120}; 100 end-systems B_1 to B_{100}; 100 end-systems C_1 to C_{100}; and 10 end-systems D_1 to D_{10}.
- The pink numbers represent the costs of links between the routers.
- The green numbers represent the transmission rates of links.
- A_2 and D_2 are DNS servers. D_2 is both a root DNS server, and an authoritative DNS server for `epfl.ch`. A_2 is neither root, nor authoritative DNS server for any domain.
- A_1 uses A_2 as its local DNS server.
- Whenever A_2 needs to contact a root DNS server, it contacts D_2.
- D_1 is a web server with DNS name `d1.epfl.ch`.
- The small orange boxes are network interfaces. E.g., router R_A has network interfaces e, f, and g.

① **[1 point]** Identify all the IP subnets inside AS0 by **marking them on Figure 1 on the left** (not on the provided extra green sheet).

② **[10 points]** The administrator of this AS owns the following IP prefixes:

- 5.0.0.0/24
- 5.0.1.0/24

From these two IP prefixes, assign an IP prefix to each IP subnet that contains end-systems. Each IP prefix you assign must have the smallest possible size. **Justify your answer.**

Answer and justification:

continues on back ☞

③ **[2 points]** For each of the 4 routers, state how many routing protocol it participates in. **Justify your answer.**

Answer and justification:

④ **[2 points]** How many and which IP prefix(es) do you expect border router R_D to advertize to other ASes? **Justify your answer.**

Answer and justification:

⑤ **[10 points]** All end-systems and packet switches have been rebooted. All caches (of all kinds) are empty. The user of end-system A_1 types in their web browser `http://d1.epfl.ch/image.png`. This is a large image (it does not reference any other object).

List the sequence of packets that are received or forwarded (sent) by network interface e of router R_A as a result of the end-user's action, up to and including the first packet that carries D_1's HTTP response. You do not need to list the packets received or forwarded by other network interfaces of router R_A.

Answer by completing Table 1 on the next page. The first row shows an example (which is not part of the correct answer). You may not need to fill all the rows of the table. **If you need to make any assumptions, state them.**

Answer:

#	Source MAC	Dest MAC	Source IP	Dst IP	Transp. prot.	Src Port	Dst Port	Application & Purpose
Example	x	y	x	y	UDP	5000	6000	Request for file...
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								

Table 1: Packets received or sent by network interface e of router R_A.

Question 2 – TCP and delay computation [30 points]

Consider the same network topology as in Question 1 (Figure 1, page 2) and the events of Question 1⑤. Assume that:

- Transport-layer, network-layer, and link-layer headers have insignificant size.
- A_1's receiver window is always 5 000 bytes.
- The links between the 4 routers have transmission rate R, in both directions.
- All the other links have transmission rate 10R, in both directions.
- All links have propagation delay D, in both directions.
- All network devices are store-and-forward (as we saw in class) and have infinite queues.

Recall that, in Question 1⑤, A_1 makes an HTTP request to D_1 and receives an HTTP response. A_1 sends no data to D_1 other than this HTTP request. No segment is lost or reordered during the entire exchange.

① **[5 points]** Assume the following sizes:

- Maximum Segment Size (MSS): 1 byte.
- HTTP request: 1 byte.
- HTTP response, including HTTP header and `image.png`: 12 bytes.

The diagram in Figure 2 shows the beginning of the communication between A_1 and D_1. The next sequence number (SEQ) that A_1 is expecting after the 3-way handshake is SEQ 1. Complete the diagram by showing:

- All the segments (including the segments that carry only ACKs) exchanged until A_1 receives the entire image.
- The SEQ numbers of D_1's segments.
- The ACK numbers of A_1's segments.
- The status of D_1's congestion-control algorithm.
- The values of D_1's congestion window and ssthresh.

Answer here: (do NOT answer on the provided DRAFT copy)

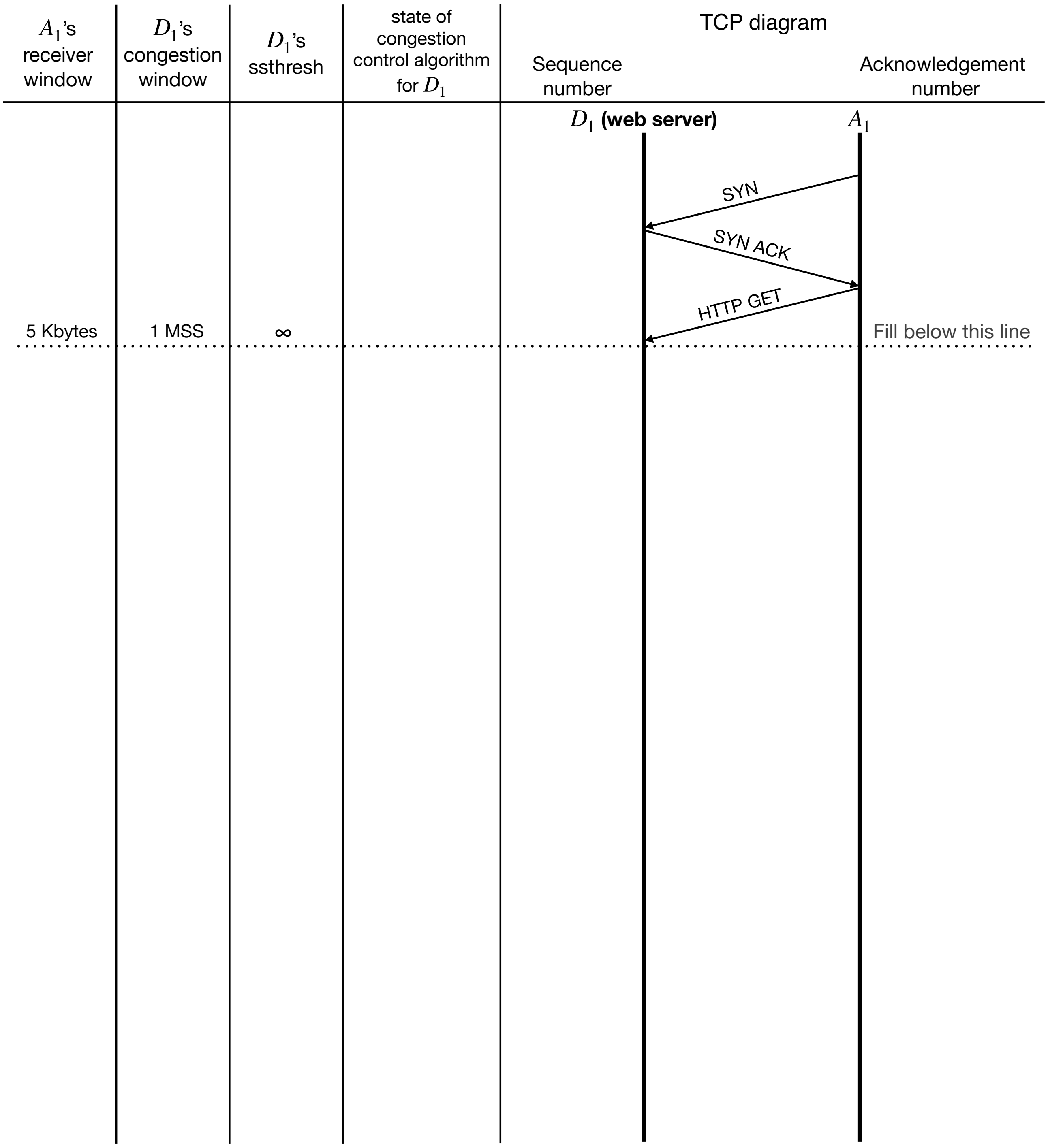


Figure 2: TCP diagram for Question 5.

continues on back ☞

② **[6 points]** Now assume the following sizes:

- MSS: 1000 bytes.
- HTTP request: 1000 bytes.
- HTTP response, including HTTP header and `image.png`: 12 000 bytes.

Does the TCP diagram change? If yes, in what way(s)? (You don't have to redraw the entire diagram from scratch, just state or sketch the changes, if any.) Is it flow control, or congestion control that determines the rate at which D_1 sends in this particular scenario? **Justify your answer.**

Answer and justification:

③ **[4 points]** Assume the same parameters as in sub-question ②.

Consider D_1's second segment, i.e., the first segment that carries (the first part of) D_1's HTTP response to A_1. How long does it take from the moment D_1 transmits the first bit of this segment until A_1 receives the last bit of this segment? **Justify your answer.**

Answer and justification:

④ **[10 points]** Assume the same parameters as in sub-question ②.

Now consider the entire sequence of D_1's segments that carry D_1's HTTP response to A_1. How long does it take from the moment D_1 transmits the first bit of the first segment until A_1 receives the last bit of the last segment? **Justify your answer.**

Be careful: D_1 and A_1 are not connected by a single link (as was the case in some of the practice exercises).

Answer and justification:

More room to answer on back ☞

(More room to answer Question 2④ if needed.)

⑤ **[5 points]** Which is the maximum number of segments that may be lost and not affect at all the answer to sub-question ④? If you need to assume a timeout value, assume that the timeout value is fixed at 4RTT. **Justify your answer.**

Answer and justification:

Question 3 – Forking and multi-threading [19 points]

Consider the program `forkloop.c` on the next page, compiled as follows:

```
gcc forkloop.c -o forkloop
```

In the program, `fork` and `wait` control the lifecycle of processes; `pthread_create` and `pthread_join` the lifecycles of threads, and `pthread_mutex_lock` of critical sections.

① **[4 points]** Draw a picture with the process parent-child tree for the execution of `./forkloop 3 0`. For each node in the tree, include the variables that determine the future execution of the process in the form $var = val$.

② **[1 point]** Provide one possible output of `./forkloop 3 0`.

③ **[2 points]** Is the output of `./forkloop 3 0` deterministic? **Justify completely**.

④ **[2 points]** Provide one possible output of `./forkloop 3 1`.

⑤ **[2 points]** Is the output of `./forkloop 3 1` deterministic. **Justify completely**.

⑥ **[1 point]** How many stacks are there (maximal value) for the process where $f(z = 2)$?

⑦ **[7 points]** Draw these stacks (for the process with $f(z = 2)$) with one thread in the critical section and the other thread blocked waiting on the mutex.
Stacks should be drawn from top to bottom (as on the hardware). Each call frame of the stack must be labeled with the name of the function (but the return IP address (RIP) does not need to be shown). The arguments and local variables of each call frame must be identified as follows:

- if the variable has a known integer value you must show the name along with the value (e.g. $foo = 4$);
- if the variable points to a null-terminated string, or an array of null-terminated strings with known values, show it as $foo = "bonjour"$ or $foo = ["hello", "world"]$, respectively;
- if the variable is a pointer, you must draw an arrow to the pointed address on a stack, on the heap, or on the global segment;
- if the value cannot be determined in this scenario, label it with a question mark (e.g. $foo = ?$).

Answers:

```c
// necessary #include <...>

#define MAX_N 12

pthread_mutex_t mut = PTHREAD_MUTEX_INITIALIZER;

void * g(void *arg) {
    int *p = (int *) arg;
    pthread_mutex_lock(&mut);
    *p = *p + 1000;
    pthread_mutex_unlock(&mut);
}

int f(int z) {
    if (z > 0 && z < MAX_N) {
        pthread_t thr[MAX_N];
        int acc = 0;
        for (int i = 0; i < z; i++) {
            pthread_create(&thr[i], NULL, g, &acc);
        }
        for (int i = 0; i < z; i++) {
            pthread_join(thr[i], NULL);
        }
        return acc;
    } else {
        return -2;
    }
}

int forkloop(int n, int m) {
    for ( ; n > 0; n--) {
        pid_t x = fork();
        if (x > 0) {
            wait(NULL);
        } else if (m > 0) {
            return f(n) + n ;
        } else {
            return n;
        }
    }
    return 0;
}

int main(int argc, char **argv)
{
    int a1 = atoi(argv[1]);
    int a2 = atoi(argv[2]);
    if (a1 < MAX_N)  {
        int y = forkloop(a1, a2);
        printf("done %d\n", y);
    }
}
```

Answers (continued):

(More room to answer Question 3 if needed.)

(More room to answer Question 3 if needed.)

Question 4 on back ☞

Question 4 – File System [12 points]

Assume a file system mounted at the root directory "/" that has only the following files and directories:

```
/usr/bin/gcc
/usr/bin/clang
/bin/gcc
```

`/usr/bin/gcc` corresponds to gcc version 41 and `/bin/gcc` corresponds to gcc version 42.

Assume also that the user has full access to the whole file system. Thus, the script containing the following commands is executed without errors:

```
mv /usr/bin/clang /bin
cp /usr/bin/gcc /bin/gcc
mv /bin/clang /usr/bin
```

`mv` is the standard POSIX utility to move a file using the rename system call
(`int rename(const char *old, const char *new);`);
`cp` is the standard utility to copy a file. If the destination file already exists, `cp` opens the file with the `open` system call, and modifies its content.

Finally, assume the following hypothesis about the file system and its content:

- the inode structure has only the following:
  - a length field indicating the length of the file;
  - a modification time field;
  - entries for direct and indirect blocks;
- for simplicity, all directories and files are small and fit within a single data block;
- the OS has a very large file system buffer cache, initially empty, which is used to cache inodes and data blocks;
- the OS file system buffer cache holds recently-accessed inodes and data blocks in memory;
- the OS writes back to disk all modified inodes and data blocks synchronously on all system calls;
- the inode numbers are:

`/`: 1	`/usr`: 11	`/usr/bin`: 22	`/usr/bin/gcc`: 33
`/usr/bin/clang`: 44		`/bin`: 55	`/bin/gcc`: 66

- the next available inode numbers are: 77, 88, 99. You may or may not have to use these.

Based on the above commands (script) and assumptions:

① **[2 points]** What is the directory structure of the file system mounted at "/" once the script finishes? For `gcc`, specify the version.

Answer:

② **[10 points]** In the tables below, mark the inodes and data blocks that are read from and/or written to **disk** as a result of each command.

Enter the number of blocks associated with each inode (inode blocks, indirect blocks, data blocks) which are read from or written to **disk** for each syscall.

Mark read/write accesses to inode in table 1 and read/write accesses to data blocks in table 2 as follows:

No access: `X` One Read: `R` One Write: `W`

Examples:
One read and write: `RW` Two reads: `RR` Two reads and one write: `RRW`
and so on for different combinations of read (`R`) and writes (`W`).
The order of `RW` does not matter.

Creating/Moving a new inode/data block or updating an existing block is counted as a write.
In these tables, all blank answers will be interpreted as "*not answered*" rather than "`X`".

State your assumptions, if any.

Answers:

Table 1: **inode** blocks (commands refer to the former script):

	inodes									
	1	11	22	33	44	55	66	77	88	99
First `mv`										
`cp`										
Second `mv`										

Table 2: **data** blocks (commands refer to the former script):

	inodes									
	1	11	22	33	44	55	66	77	88	99
First `mv`										
`cp`										
Second `mv`										

(More room to answer Question 4 if needed.)

(More room to answer Question 4 if needed.)

Question 5 on back ☞

Question 5 – Scheduling [14 points]

Consider a system with a single CPU for computation and a single disk for IO, with the following characteristics:

- CPU:
  - Round Robin (RR) policy for scheduling computation requests (i.e., with a single queue);
  - scheduling quantum: 1 second;
  - when multiple tasks enter the tail of the pending queue at the same time, they enter in the following priority (i.e., the pending queue is a FIFO):
    - new task;
    - blocked task;
    - currently running task;
  - negligible time to send disk IO request.
- Disk:
  - elevator scan scheduling policy for scheduling IO requests;
  - has five cylinders;
  - head starts at cylinder zero and moves depending on the first request;
  - seek time to move from one cylinder to next: 1 second;
  - assume that there are no rotational or transfer latencies;
  - Scheduling policy is invoked when an IO request finishes.

The following table describes *five* tasks which do some CPU and Disk IO operations. Once the disk IO operation is completed, a task executes for the remainder of its CPU total time.
Here is the description of each column:

- Task Arrival Time: the time at which the task arrives;
- CPU Total Time: total time required for the computation of the task;
- IO Request Time: the time at which the task requests for IO; for example, if task 1 starts at time 1, it will request for IO at time 2 (=1 + 1);
- IO Access: the cylinder number accessed during the IO request.

Task ID	Task Arrival Time (in seconds)	CPU Total Time (in seconds)	IO Request Time (in seconds)	IO Access (Cylinder Number)
1	1	3	1	3
2	2	2	1	4
3	3	5	1	2
4	3	3	2	1
5	3	3	1	3

Given the list of five tasks above, fill in the time "diagram" on the next page (table) to show when each task computation and IO completes.
State your assumptions whenever necessary or appropriate.

Fill in the "time diagram as follows" from "Time=2" onwards:

- Running Task as the task that executes on the CPU during the period.
- Ready tasks is the ordered list of ready tasks during the period.
- Same for blocked tasks.
- "Current Disk Cylinder" expressed as "[Task ID:Cylinder Number]", for instance: [1:3], where Task ID is the task associated with the current disk seek and Cylinder Number should reflect the position of the disk at the *end* of period.
- "Pending Disk IO requests" as a list of [Task ID:Cylinder Number], for instance: [1:3], [1:4].

You may not need to fill all the rows of the table.

Time	Running Task (Task ID)	Ready Tasks (Tasks IDs)	Blocked Tasks (Tasks IDs)	Current Disk Cylinder	Pending Disk IO requests
0	–	–	–	[–:0]	–
1	1	–	–	[–:0]	–
2	2	–	1	[1:1]	–
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					

More room to answer on back ☞

(More room to answer Question 5 if needed.)

Question 6 – C Programming [25 points]

6.1 Wrong or right? [2 points]

Provided that the library `string` has been included (and there is a proper `main()` function), would the following portion of code compile?

Fully justify your answer.

```
#define NAME_SIZE 127
#define NB_VALUES 4

struct Foo {
  char name[NAME_SIZE+1];
  int whatever;
  double values[NB_VALUES];
};

void init_phys(struct Foo* s)
{
  double values[] = { 299792458, 9.80665, 6.02214076e23, 1.602176634e-19 };
  strcpy(s->name, "Some physics");
  s->whatever = 42;
  s->values = values;
}
```

Answer and justification:

continues on back ☞

6.2 Pointers [8 points]

On a 64-bit architecture where:

- `sizeof(short int)` is 2, such that $256 \times a + b$ is represented in memory with b first then a;
- the integer value of `(char)'A'` is 65, the one of `'B'` is 66, etc.;

what does the following code print?

Fully justify your answer **and provide** a drawing of the memory state of the variables `tab`, `ptr`, `p1`, `p2`, `p3` and `q` at line 6 just after the call line 25.

```
 1  #include <stdio.h>
 2  #include <string.h>
 3
 4  void f(int nb, const short int* q, size_t sz)
 5  {
 6    printf("%d: ", nb);
 7    for (size_t i = 0; i < sz; ++i) printf("%d, ", *(q + i));
 8    putchar('\n');
 9  }
10
11  void g(void* ptr)
12  {
13    const char* const p1 = ptr;
14    printf("1: \"%s\"\n", p1);
15
16    const short int* const p2 = ptr;
17    printf("2: %d\n", *p2);
18
19    f( 3, p2, 4          );
20    f( 4, p2, sizeof(ptr) );
21    f( 5, p2, strlen(ptr) );
22
23    const short int** p3 = &p1;
24    ++(*p3);
25    f( 6, *p3, 1);
26  }
27
28  int main(void)
29  {
30    short int tab[10] = { 65 * 256 + 66,  // 16'706
31                                     67,
32                          68 * 256 + 69,  // 17'477
33                                     70,
34                          71 * 256 + 72,  // 18'248
35                          73, 74, 75, 76, 77 };
36    g(tab);
37    return 0;
38  }
```

Answer and justification:

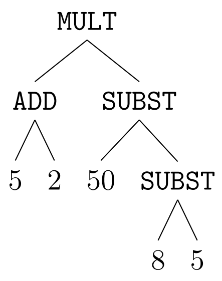
(More room to answer 6.2)

continues on back ☞

6.3 A bit of arithmetics [15 points]

We here consider writing *a few pieces* of a C code, the aim of which is to do arithmetic processing using (binary) tree representation.

Each node of the tree will have an operation (which can be represented as an `int`) and two operands (which are themselves (sub-)trees).
The leaves are also tree nodes which simply have a numerical value as their "operation" and two empty operands.
For instance, the arithmetic expression $(5+2)\times(50-(8-5))$ will be represented by the binary tree drawn on the right:



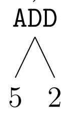
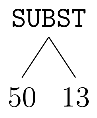
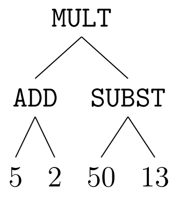
① **[2 points]** Assuming that the operations are for instance represented as

```
enum Operation { ADD, SUBST, MULT, DIV };
```

propose a type (data structure) to represent the arithmetic binary trees:

② **[1 point]** Declare a variable named `five`, stored on the stack, that would represent the leaf node 5:

③ **[5 points]** Define a `merge()` function that takes an operation and two trees and merge them into a higher level tree (to be returned).

For instance the merge of the tree  with the tree  using the operation `MULT` will return the tree: 

④ **[1 point]** Define a `leaf()` function which takes an integer value and returns a leaf node (or a pointer to it), similar to the variable `five` from subquestion ①, but allocated on the *heap*.

⑤ **[1 point]** Use the `merge()` and `leaf()` functions to declare a variable named `example` representing the expression $(5+2) \times (50-(8-5))$ (the tree of which is drawn above).

⑥ **[5 points]** Finally, define a `release()` function that *completely* deallocates a tree, assuming all its nodes have been allocated on the heap.