

# Computing Infrastructures
## January 13, 2025

Course Section: ☐ Prof. Ardagna ☐ Prof. Palermo ☐ Prof. Roveri

Student ID (Codice Persona): .....

Last Name: .....
(LAST NAME IN CAPITAL LETTERS)

First Name: .....
(FIRST NAME IN CAPITAL LETTERS)

**Exam Duration: 1hour and 30min**

Students are not permitted to use mobile phones and similar connected devices. Course materials and programmable devices (e.g. programmable calculators) cannot be used as well. **Any violation of the rules is considered a cheating action.**

**Answers must be given on the Answer Sheets and in English**. Any box filled or answer provided on the other sheets will be ignored. Students must use a pen (black or blue) to mark the answers (no pencil).

Write the LAST and FIRST name in CAPITAL LETTER, and in this order, in all places where requested. **Where it is requested only the STUDENT ID (Codice Persona), do not write your name**.

**Check that the first number of the code for the Answer Sheet** is the same as for the other sheets. The code can be found in the top-right corner of each page in the form `+NN/KK/XX+`. The parts that should correspond is ONLY the first digit `NN`.

Mark clearly the box corresponding to your answers, without overlapping on other boxes. If you make a mistake on them, circle the word *Question* together with the related number, and write the correct letter to its side.

**Numerical exercises require writing the formulas and procedure used to solve the problem just after the question in the left space**. Exercises without the procedure used to reach the result will not be considered for the evaluation. Only the numeric answer and its unit should be reported on the corresponding dotted line in the Answer Sheet.

**The answers to the *Open Questions* should be written using ONLY the space available on in the boxes within the Answer Sheets**. The answers should be readable by the professor. Unreadable answers will not be considered for the evaluation.

Scores: correct answers take positive points, unanswered questions take 0 points, **wrong answers can have negative points**. An indication of the points is available at the beginning of each section. The final score can be re-modulated at the end of the evaluation.

## True false questions

Correct answer: +1, No answer: 0, Wrong Answer -0.5

*Answers must be given on the ANSWER SHEETS. Any box filled here will be ignored. Pay attention to the position (A or B) of the True/False answers, since they are not always in the same position.*

**Question 1** SANs are primarily used for block-level access to data, while NAS devices provide file-level access.

A False B True

**Question 2** RAID can be used with both hardware and software implementations.

A False B True

**Question 3** GPUs (Graphics Processing Units) are primarily used in datacenters for accelerating graphical applications and gaming.

A False B True

**Question 4** Cloud computing applications do not require any server to run.

A True B False

**Question 5** Availability zones within a Compute Region can have a roundtrip time larger than few milli seconds.

A False B True

**Question 6** Edge computing is a distributed computing paradigm that brings computation and data storage closer to the location where it is needed.

A False B True

**Question 7** The power consumption of an average datacenter is in the order of a few GigaWatts.

A False B True

**Question 8** Leaf-spine topologies are more suitable for small-scale datacenter deployments with limited server connections.

A False B True

**Question 9** Warehouse-scale computers are typically composed of a large set of homogeneous servers.

A False B True

**Question 10** Virtualization can help simplify IT management by reducing the number of physical machines that need to be managed.

A True B False

## Exercises

Correct answer: +2, No answer: 0.

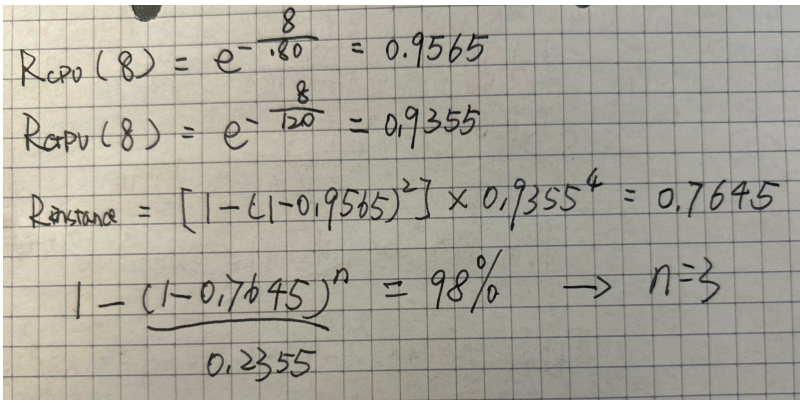
*The formulas and procedures used to solve the exercises should be included here close to the question. The numeric answer, and only that, must be given on the ANSWER SHEETS. Any number written only here will be ignored. The correct number is ONLY a necessary condition for a correct answer. If the formulas are not available after each exercise, they will be considered as not answered.*

### Question 11

A scientific computation that needs to be carried out within the PoliMi data center uses a server composed of 2 CPUs and 4 GPUs. Knowing that:

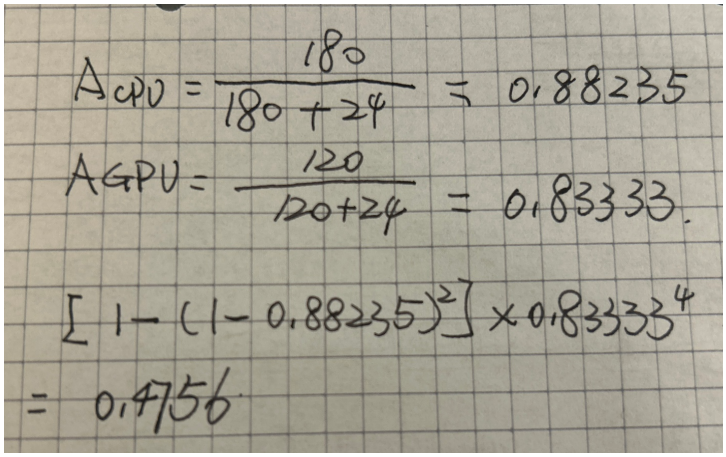
- The computation takes 8 days to complete,
- The computation requires at least one CPU and all GPUs within the server to be operational to complete successfully;
- $MTTF_{CPU} = 180$ days and $MTTF_{GPU} = 120$ days.

How many parallel instances of the computation must be launched to ensure a probability higher than 98% that at least one computation produces results successfully? Notes: (i) Use at least 4 decimal places for all intermediate calculations. (ii) All other components of the server can be considered ideal.



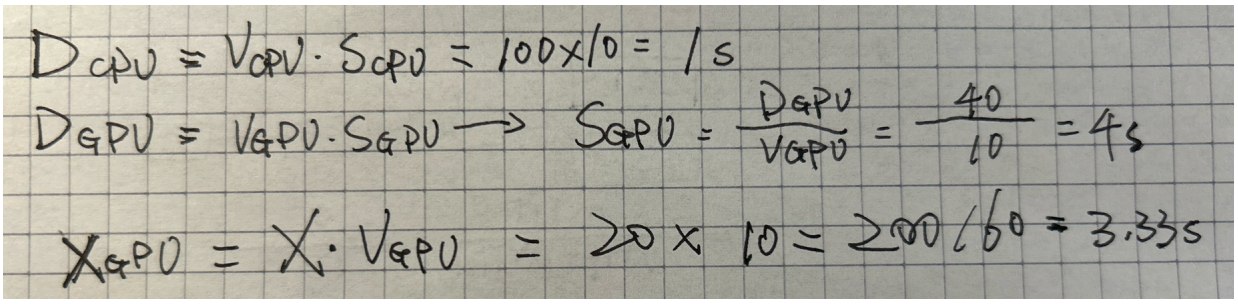

### Question 12

If the time required to recover and replace a component of the server described in the previous exercise (whether it is a CPU or a GPU) is equal to 24 days, what is the total availability of the server? Notes: (i) Use at least 5 decimal places for all intermediate calculations. (ii) All other components of the server can be considered ideal.

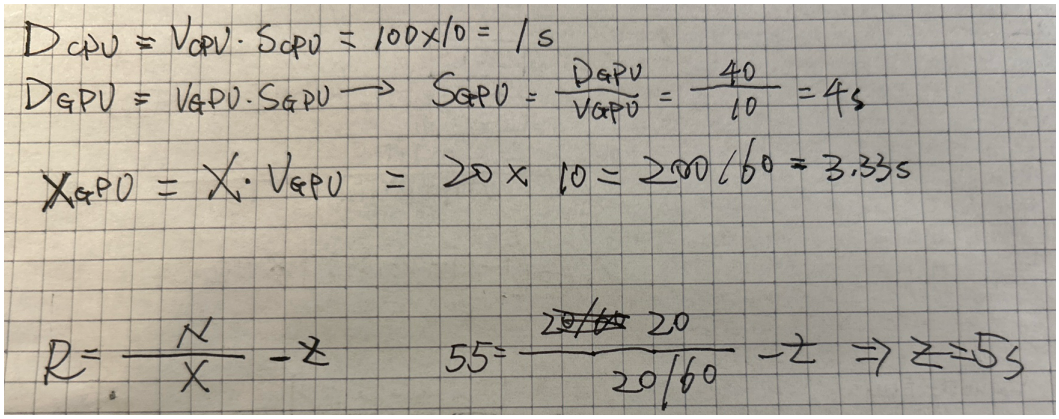


**Question 13**
A company is planning to use a RAID 5 array composed of 8 disks for critical data storage. The desired Mean Time to Failure for the entire RAID system ($MTTF_{RAID5}$) is 16 years. Each disk in the array has a $MTTF_{disk}$ equal to 500 days. What should be the MTTR required *in hours* to meet the target RAID $MTTF_{RAID5}$.

**Question 14**
Consider a system composed of three stations: the CPU that is characterized by $V_{CPU} = 100$ visits and an average service time of $S_{CPU}$ =10ms; the disk, characterized by a throughput of 12 IOPS, and a demand of $D_{DISK}$=150ms; and the GPU whose demand is $D_{GPU}$=40s and the number of visits $V_{GPU}$=10. Finally, the system throughput is $X$=20 jobs/min while the response time when there are $N = 20$ end-users in the system is $R = 55$s.
Compute the CPU demand and the GPU throughput.
Write in the answer sheet: $D_{CPU} = \ldots; X_{GPU} = \ldots$



**Question 15**
Considering the system described in Question 14, what is the users' think time Z?

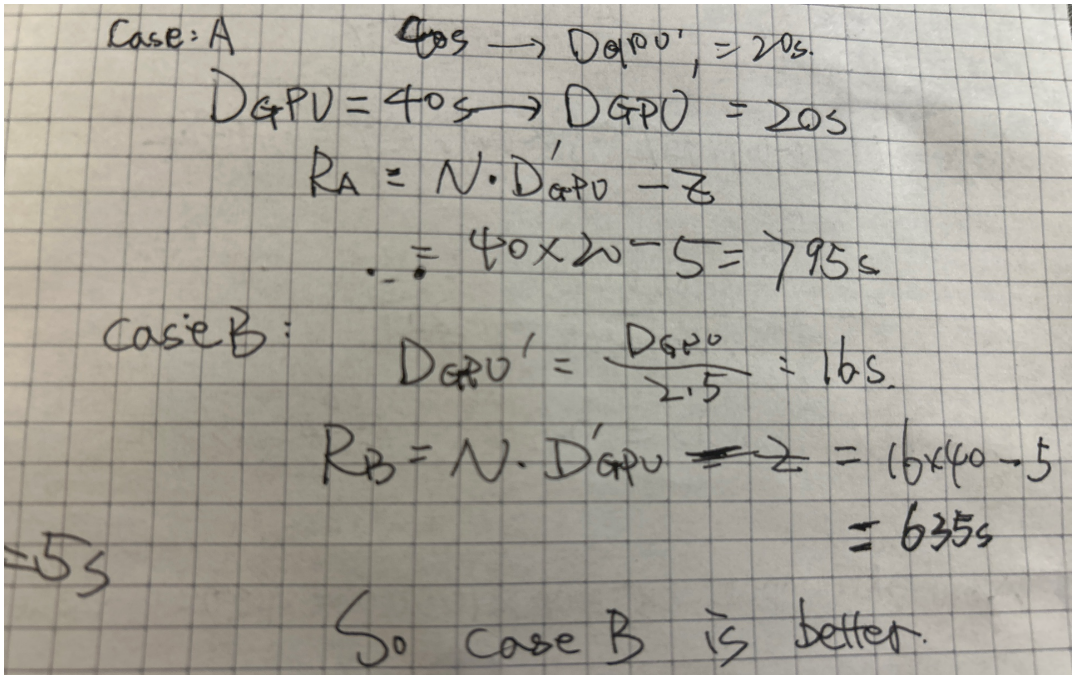



**Question 16**
Based on the system in Question 14, the number of end users is predicted to reach 40 in one month. Considering the response time lower bound, which option is better?

a. Upgrade the system by adding one more GPU (you can assume the new GPU is equal to the one initially available and to balance evenly the GPU processing).

b. Replace the GPU with one 2.5 times faster than the original one.

Provide the estimated bounds in the two scenarios to motivate your answer.
Write in the answer sheet: *A or B* ; $R^A_{LOW} = \ldots$ ; $R^B_{LOW} = \ldots$



## Open Questions

Correct answer: +5, No answer: 0. Points are modulated considering the written text

*Write the answer using ONLY the space available in the boxes on the ANSWER SHEETS. The answers should be readable by the professor. Unreadable answers will be considered wrong.*

**Question 17**

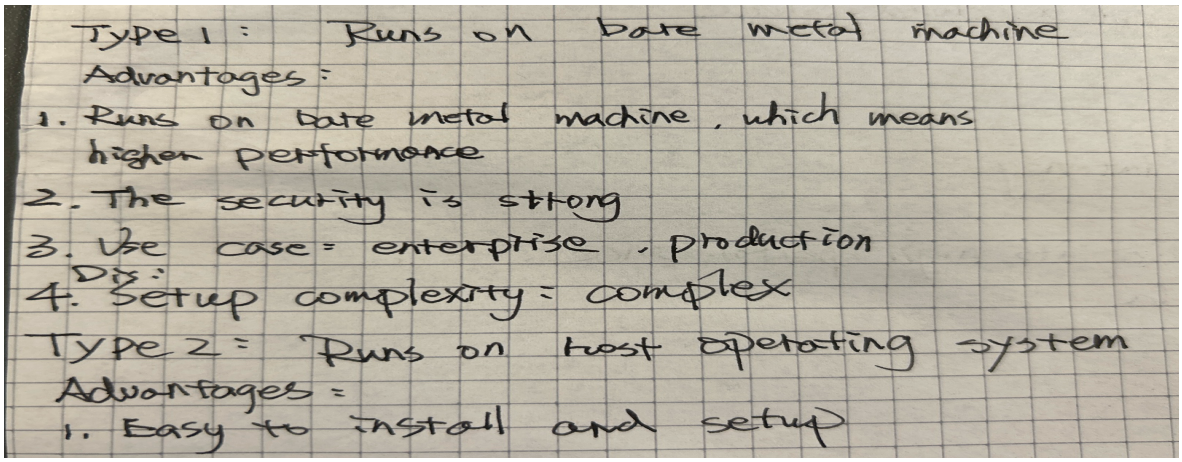
⇒ What are the advantages and disadvantages of using Type 1 hypervisors versus Type 2 hypervisors in a virtualized environment?

**Question 18**

⇒ What are the implications of hardware heterogeneity on software stack development in datacenters?

**!!!ANY ANSWER PROVIDED ON THIS PAGE WILL BE IGNORED!!!**

If needed, you can use the space hereafter to organize your answer.



+1/9/52+

# Computing Infrastructures - January 13, 2025

**Answer Sheets (Page 3)**

Student ID (Codice Persona): .....

## True/False Questions

Question 01 : ☐A ☒B

Question 02 : ☐A ☒B

Question 03 : ☒A ☐B

Question 04 : ☐A ☒B

Question 05 : ☒A ☐B

Question 06 : ☐A ☒B

Question 07 : ☒A ☐B

Question 08 : ☒A ☐B

Question 09 : ☐A ☒B

Question 10 : ☒A ☐B

## Exercises

Question 11 : 3

Question 12 : 0.4756 o 0.8558

Question 13 : 18,35 h (o 0.7644 giorni)

Question 14 : $D_{GPU} = 1\,SEC$ $X_{GPU} = 200 \frac{J}{MIN}$ o $3.33 \frac{J}{SEC}$

Question 15 : $E = 5\,SEC$

Question 16 : B ; $R^A_{MIN} = 795\,SEC$ ; $R^B_{MIN} = 635\,SEC$