Student ID (*Matricola*)

0	0	0	0	0	0
1	1	1	1	1	1
2	2	2	2	2	2
3	3	3	3	3	3
4	4	4	4	4	4
5	5	5	5	5	5
6	6	6	6	6	6
7	7	7	7	7	7
8	8	8	8	8	8
9	9	9	9	9	9

Computing Infrastructures

Course 095897

M. Gribaudo, R.Mirandola

05-09-2016

Last Name / Cognome:

.....

First Name / Nome:

.....

Answers must be given exclusively on the answer sheet (last sheet): DO NOT FILL ANY BOX IN THIS SHEET

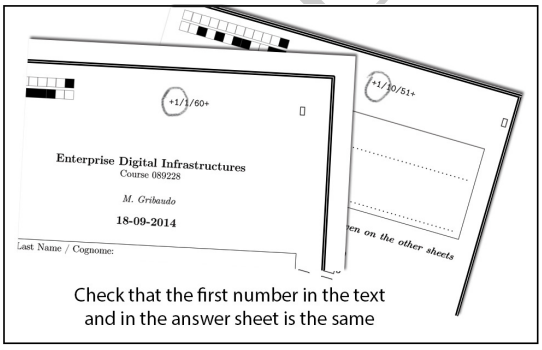
Students must use pen (black or blue) to mark answers (no pencil).
Students are permitted to use a non-programmable calculator.

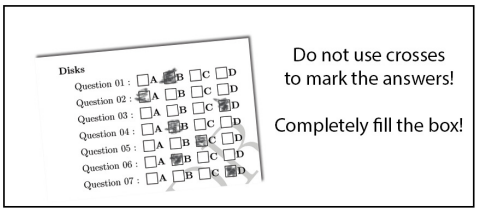
Students are NOT permitted to copy anyone else's answers, pass notes amongst themselves, or engage in other forms of misconduct at any time during the exam.

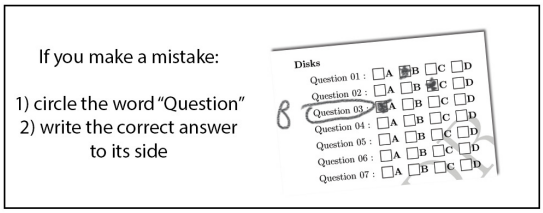
Students are NOT permitted to use mobile phones and similar connected devices.

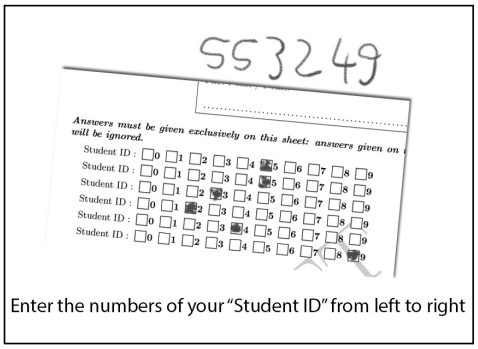
Scores for the multiple-choice part: correct answers +1 point, unanswered questions 0 points, wrong answers -0.333 points.

You cannot keep a copy of the exam when you leave the room.









Disks

Consider the following set of disks connected in RAID 6, with generator $g = 2$, $MTTF = 1461$ days, $MTTR = 15$ days, rotating at 8400 RPM, with an average seek of 3 milliseconds, trasfer time of 128 MB / second, and a controller overhead of 150 microseconds:

Disk 1	Disk 2	Disk 3	Disk 4	Disk 5	Disk 6
Block A1	Block A2	Block A3	Block A4	Block A5	Block A6
D0 1	D1 0	D2 0	D3 10	P 11	Q
Block B1	Block B2	Block B3	Block B4	Block B5	Block B6
D0 1	D1 0	D2	P 11	Q 81	D3
Block C1	Block C2	Block C3	Block C4	Block C5	Block C6
D0 1	D1 0	P	Q	D2	D3 10

Figure 1: A RAID 6 configuration.

Question 1 What data is contained in block $A6$?

A $Q = 11$
B $Q = 81$
C Cannot be computed
D $Q = 162$

SOLUTION:
$Q = 1 + 2^3 \cdot 10 = 81$

Question 2 What data is contained in blocks $B3$ and $B6$?

A $D_2 = 0$, $D_3 = 10$
B $D_2 = 69$, $Q = 730$
C Cannot be computed
D $D_2 = 69$, $Q = 365$

SOLUTION:
$1 + D_2 + D_3 = 11$ and $1 + 4 \cdot D_2 + 8 \cdot D_3 = 81$.
$D_2 = 10 - D_3$ and $4(10 - D_3) + 8 \cdot D_3 = 80$
$D_3 = (80 - 40)/4 = 10$, $D_2 = 10 - 10 = 0$.

Question 3 What data is contained in blocks $C3$, $C4$ and $C5$?

A $P = 11$, $Q = 81$, $D_2 = 0$
B Cannot be computed
C $P = 11$, $Q = 162$, $D_2 = 0$
D $P = 12$, $Q = 85$, $D_2 = 1$

SOLUTION:
RAID 6 can recover up to 3 failures, so the solution cannot be computed

Question 4 The rotational latency of each disk is:

A 7.143 ms
B 14.286 ms
C 3 ms
D 3.571 ms

SOLUTION:
The rotational latency is half of the time required to perform one rotation:
$T_l = \frac{60000}{2 \cdot 8400} = 3.571$ ms.

Question 5 The average time required to read a 512 B block on one of the disk of the array:

A 10.147 ms
B 6.725 ms
C 10.296 ms
D 6.575 ms

SOLUTION:
The total average transfer time is:
$T_a = T_c + T_l + T_s + F/r_t = 0.15 + 3 + 3.571 + 512/(128 * 1024 * 1024) * 1000 = 6.725$ ms

Question 6 The mean time to data loss of the system is:

A 231003 days B 11858 days C 115501 days D 4743 days

SOLUTION:
$MTTDL = \frac{2MTTF^3}{6 \cdot 5 \cdot 4 \cdot MTTR^2} = \frac{2 \cdot 1461^3}{120 \cdot 15^2} = 231003$ days.

Virtualization and IaaS

Question 7 Which of the following is **NOT** a typical network configuration of a modern data-center ?

A D-Cell B NAS-Head C Fat tree D Three layers

SOLUTION:
The NAS-Head is a type of unit used for storage. It is thus the one that is not meaningful here.

Question 8 Which of the following could be an example of resource aggregation that can be achieved using virtualization?

A Partitioning a single disk to make it appear as several units

B Sharing a printer in a small network

C JBOD (*Just a Bunch Of Disks*) storage units

D Using 4G LTE and WiFi in a smartphone to increase network bandwidth

SOLUTION:
Using 4G LTE and WiFi in a smartphone to increase network bandwidth is an example of aggregation of two resources (4G LTE and WiFi) to create a single faster virtual connection

Question 9 Which of the following properties is **NOT** strictly required to allow virtualization in the Popek and Goldberg theorem?

A Equivalence B Efficiency C Encapsulation D Resource control

SOLUTION:
Encapsulation is a property of virtualization, but it is not required by the Popek and Goldberg theorem.

Question 10 A virtual machine has a network adapter in NAT mode. The virtual machine has IP address 10.0.1.15, while the physical machine has IP address 192.168.5.13. The virtual machines is running a web-server listening on port 80. In order to access the web-server running on the virtual machine using address `http://192.168.5.13:8080` from another PC in the same sub-network, the following port-forwarding rule must be added:

A From 8080 on the host to 80 of the guest

B From 8080 on the guest to 80 of the host

C From 80 on the guest to 8080 of the host

D From 80 on the host to 8080 of the guest

SOLUTION:
The port forwarding rule must allow to see port 8080 on the host, as if it was port 80 on the guest.

Question 11 A virtual machine has a network adapter in Bridge mode. The the physical machine has IP address 192.168.5.13. Which of the following could be a valid IP address for the virtual machine?

A 192.168.5.13 B 192.168.5.131 C 10.0.1.15 D None of the other

SOLUTION:
In Bridge mode, a VM has an address on the same sub-network of the host, but different from it. Among the possibility, the one that fulfills these properties is 192.168.5.131.

Question 12 A program A in a physical machines is executed in 100 sec, and in a virtualized environment in 121 sec. A program B takes 130 sec on a physical machine, and requires 120 sec in a virtualized environment. This can occur because:

A It cannot occur: there must have been a measuring error somewhere

B Program B has a smaller virtualization overhead than program A

C Program B has a larger fraction of sensitive instructions than program A

D Program B has a smaller fraction of sensitive instructions than program A

SOLUTION:
Since both the virtualization overhead and the fraction of sensitive instructions must be non-negative numbers, the execution time in a virtualized environment cannot decrease. Since program B reduces its execution time, there must have been some measurement errors somewhere

Performance - (10 points)

Let consider a client server infrastructure where we are interested in modelling the Client (modelled with a single CPU) and a Server (modelled with a CPU and 2 disks) and without network delay.

During a period of observation T=25 minutes we have collected the following information:

- Number of system completions: 450
- Service time of Client CPU: 0.25 s
- Service time of Server CPU: 0.3 s
- Service time of Disk 1: 0.25 s
- Service time of Disk 2: 0.325 s
- Utilization of Client CPU: 0.15
- Utilization of Server CPU: 0.36
- Utilization of Disk 1: 0.45
- Utilization of Disk 2: 0.39

1. Provide the following definition
   (a) How can be defined the service time S_k of a service center?
   (b) How can be defined the throughput X_k of a service center?
2. Define the system model
3. Compute
   (a) the throughput X of the system
   (b) The service demands of all the servers and determine the bottleneck
   (c) The number of visits at all servers
   (d) The throughput of all the servers
4. Determine the maximum system throughput and the minimum average response time. In order to improve the system, two possible upgrade are considered:
   (a) The bottleneck server is substituted with one having double speed
   (b) A new disk is added equal to the original disk 1 and the load is balanced between the two disks

   Study and discuss the effects that each of the two different upgrades would have on the the maximum system throughput and the minimum average response time of the original system.

SOLUTION:

1) See slides

2) We can define an open model like the one in the Figure
3)

a) $X = \frac{C}{T} = \frac{450}{1500} = 0.3 job/s$

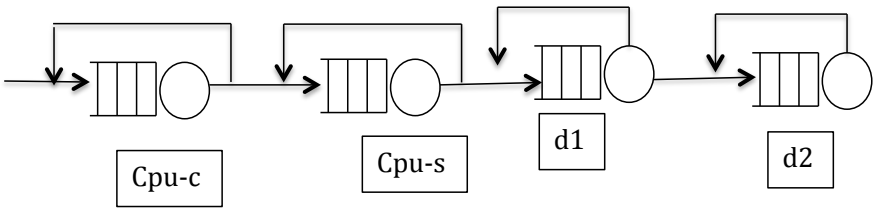


Figure 2: QN model

b) $D_k = \frac{U_k}{X}$, so we obtain $D_{CPU-C} = \frac{0.15}{0.3} = 0.5$, $D_{CPU-S} = \frac{0.36}{0.3} = 1.2$, $D_{D1} = \frac{0.45}{0.3} = 1.5$, $D_{D2} = \frac{0.39}{0.3} = 1.3$
The bottleneck is thus Disk 2.

c)$V_k = \frac{D_k}{S_k}$, so we obtain $V_{CPU-C} = \frac{0.5}{0.25} = 2$, $V_{CPU-S} = \frac{1.2}{0.3} = 4$, $V_{D1} = \frac{1.5}{0.25} = 6$, $V_{D2} = \frac{1.3}{0.325} = 4$.

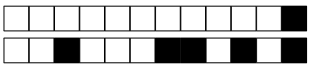
d) $X_k = X \cdot V_k$, so we obtain $X_{CPU-C} = 0.3 \cdot 2 = 0.6$, $X_{CPU-S} = 0.3 \cdot 4 = 1.2$, $X_{D1} = 0.3 \cdot 6 = 1.8$, $X_{D2} = 0.3 \cdot 4 = 1.2$

4)

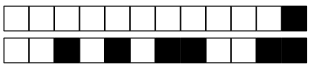
The original system has the following characteristics: $X_{\max} = \frac{1}{D_{\max}} = 0.6667 job/s$ and R > D, so R > 4.5 sec

a) Server D1 is substituted with one having double speed: so we have $D_{D1new} = 0.75$, the new $D_{\max}$ is now the $D_{D2} = 1.3$, so the first change leads to these new values: $X_{new1\,\max} = \frac{1}{D_{D2}} = 0.77 job/s$ and $R >$ D, so $R_{new1} > 3.75$ sec

b) A new server D1 is added equal to the original one and the load is balanced between the two servers: $D_{D1-1} = D_{D1-2} = 0.75$ the new $D_{\max}$ is now the $D_{D2} = 1.3$, so the change leads to these new values: $X_{new2\,\max} = \frac{1}{D_{D2}} = 0.77 job/s$ As concerns the response time, now D is unchanged with respect to the original system so $R_{new2} > 4.5$ sec



+1/8/53+



+1/10/51+

For your examination, preferably print documents compiled from auto-multiple-choice.

Dependability - (6 points)

In the following questions we will assume that both failure and repair events follow exponential distributions.

We have components of TypeA whose measured failure rate is $\lambda_A = 0.002$ days^{-1}.

Question 13 Calculate the number of days t during which a system of three components of TypeA in **series** has a probability of 0.8 to be continuously working without failure:

- [A] 37.19 days
- [B] 303.27 days
- [C] None of the other answers
- [D] 439.50 days

SOLUTION:
$R(t) = (e^{-0.002t})^3 = 0.8 \Longrightarrow \quad R(t) = e^{-0.006t} = 0.8 \quad \Longrightarrow \quad \frac{ln(0.8)}{-0.006} = t = 37.19$ days

Question 14 Calculate the reliability at time 50 of a system of two components of TypeA in **series**:

- [A] None of the other answers
- [B] 0,7408
- [C] 1,2214
- [D] 0,8187

SOLUTION:
$R(50) = (e^{-0.002 \cdot 50})^2 = 0.8187$

Question 15 Calculate the *maximum* number of components of TypeA we can put in **series** and the system still keeps a probability higher than 0.4 to continuously work for 10 days:

- [A] 45 components
- [B] 45.81 components
- [C] 46 components
- [D] None of the other answers

SOLUTION:
$R(10) = (e^{-0.002 \cdot 10})^n = 0.4 \Longrightarrow \quad R(t) = e^{-0.002tn} = 0.4 \quad \Longrightarrow \quad \frac{ln(0.4)}{-0.002 \cdot 10} = n = 45.81 \Longrightarrow n = 45$

Question 16 Calculate the reliability at time 5 of a system made of three components of TypeA in a configuration **2-out-of-3**:

- [A] None of the other answers
- [B] 0.0002
- [C] 0.9997
- [D] 0.97

SOLUTION:
$\mathrm{R}(5) = 1 - 1 \cdot (e^{-0.002 \cdot 5})^0 (1 - e^{-0.002 \cdot 5})^3 - 3 \cdot (e^{-0.002 \cdot 5})^1 (1 - e^{-0.002 \cdot 5})^2 = 0,9997049$

Question 17 Given that MTTR of a component of TypeA is 100 days, calculate the MTTR of a system made of three components of TypeA in **series**:

- [A] 100 days
- [B] 121.33 days
- [C] None of the other answers
- [D] 33.33 days

SOLUTION:
System Availability $= (\frac{500}{500+100})^3 = 0.5787$, System MTTF = Same as without repairs $= \frac{1}{3 \cdot 0.002} = 166,666$ –> System MTTR => $0.5787 = \frac{166,666}{166,666+MTTR}$ => MTTR=121.33

Question 18 Given that MTTR of a component of TypeA is 100 days, calculate the MTTF of a system made of three components of TypeA in **parallel**:

A 21500 days

B 7166.66 days

C None of the other answers

D 916.66 days

SOLUTION:
System Availability $= 1-(1-(\frac{500}{500+100}))^3 = 0,99537$, System MTTR = First component repaired of all the three broken $=\frac{100}{3} = 33.3333$ days. System MTTF $=> 0,99537 = \frac{MTTF}{MTTF+33.3333} =>$ MTTR= 7166,666

Last Name / Cognome:

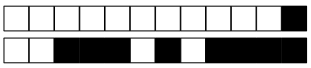
.....

First Name / Nome:

.....

Big Data and PaaS - (4 points)

Describe how the map, copy shuffle/sort and reduce phases are run within the MapReduce framework



Answer sheet:

Last Name / Cognome:

. . . .

First Name / Nome:

. . . .

Answers of the multiple-choice part of the exam must be given exclusively on this sheet

Student ID : ☐**0** ☐**1** ☐**2** ☐**3** ☐**4** ☐**5** ☐**6** ☐**7** ☐**8** ☐**9**

Student ID : ☐**0** ☐**1** ☐**2** ☐**3** ☐**4** ☐**5** ☐**6** ☐**7** ☐**8** ☐**9**

Student ID : ☐**0** ☐**1** ☐**2** ☐**3** ☐**4** ☐**5** ☐**6** ☐**7** ☐**8** ☐**9**

Student ID : ☐**0** ☐**1** ☐**2** ☐**3** ☐**4** ☐**5** ☐**6** ☐**7** ☐**8** ☐**9**

Student ID : ☐**0** ☐**1** ☐**2** ☐**3** ☐**4** ☐**5** ☐**6** ☐**7** ☐**8** ☐**9**

Student ID : ☐**0** ☐**1** ☐**2** ☐**3** ☐**4** ☐**5** ☐**6** ☐**7** ☐**8** ☐**9**

Disks

Question 01 : ☐**A** ☐**B** ☐**C** ☐**D**

Question 02 : ☐**A** ☐**B** ☐**C** ☐**D**

Question 03 : ☐**A** ☐**B** ☐**C** ☐**D**

Question 04 : ☐**A** ☐**B** ☐**C** ☐**D**

Question 05 : ☐**A** ☐**B** ☐**C** ☐**D**

Question 06 : ☐**A** ☐**B** ☐**C** ☐**D**

Virtualization and Iaas

Question 07 : ☐**A** ☐**B** ☐**C** ☐**D**

Question 08 : ☐**A** ☐**B** ☐**C** ☐**D**

Question 09 : ☐**A** ☐**B** ☐**C** ☐**D**

Question 10 : ☐**A** ☐**B** ☐**C** ☐**D**

Question 11 : ☐**A** ☐**B** ☐**C** ☐**D**

Question 12 : ☐**A** ☐**B** ☐**C** ☐**D**

Dependability

Question 13 : ☐**A** ☐**B** ☐**C** ☐**D**

Question 14 : ☐**A** ☐**B** ☐**C** ☐**D**

Question 15 : ☐**A** ☐**B** ☐**C** ☐**D**

Question 16 : ☐**A** ☐**B** ☐**C** ☐**D**

Question 17 : ☐**A** ☐**B** ☐**C** ☐**D**

Question 18 : ☐**A** ☐**B** ☐**C** ☐**D**