Written examination in Distributed Systems

November 25th 2014

Last name: _____

First name: _____

Student number: _____

I confirm with my signature that I will process the written examination alone and that I feel healthy and capable to participate this examination.
I am aware, that from the moment, when I receive the written examination, I am a participant of this examination and I will be graded.

Signature: _____

- Provide on all sheets (including the cover sheet) your *last name*, *first name* and *student number*.
- Use the provided sheets. Own paper must *not* be used.
- Place your *ID card* and your *student ID card* on your table.
- You are allowed to use a *self prepared*, *single sided DIN-A4 sheet* in the exam. Only *handwritten originals* are allowed, but no copies.
- You are allowed to use a non-programmable calculator.
- Answers, written with pencil or red pen are *not* accepted.
- Time limit: *90 minutes*
- Turn off your mobile phones!

Result:

Question:	1	2	3	4	5	6	7	8	**Σ**	Grade
Maximum points:	10	11	7	11	15	12	12	12	90	—
Achieved points:										

Last name: First name: Student number:

Question 1)

Points: .....

Maximum points: 2+4+4=10

a) Name four types of clients, which exist in the client-server model.

b) What is the height of a stack of storage media, if HDDs (capacity: 4 TB, thickness: 2.5 cm) are used for storing 20 PB of data?

c) Calculate the time which is required to fill a 3.5" HDD (Capacity: 6 TB, Transfer rate: 125 MB/s) completely with data.

Last name: First name: Student number:

Question 2)

Points: .....

Maximum points: 2+2+2+2+2+1=11

a) Name the two installation concepts for cluster systems.

b) What is an Active/Active-Cluster?

c) What is an Active/Passive-Cluster?

d) What is the meaning of Failover?

e) What is the meaning of Failback?

f) What is the objective of High Throughput Clustering?

Last name: First name: Student number:

Question 3)

Points: .....

Maximum points: 3+1+1+1+1=7

a) Give a short definition of Grid Computing.

b) What is an Intra-Grid?

c) What is an Extra-Grid?

d) What is an Inter-Grid?

e) For exercise sheet 8, you implemented with the infrastructure services of the Amazon Web Services, a highly available High Throughput Cluster of virtual web servers. Which web server software did you use?

Last name: First name: Student number:

Question 4)

Points: .....

Maximum points: 1+1+1+1+1+1+1+1+2+1=11

a) Physical resources are offered in...

☐ Grids ☐ Clouds

b) Virtualized resources are offered in...

☐ Grids ☐ Clouds

c) Full-automatization (*industrialized IT*) is an attribute of...

☐ Grids ☐ Clouds

d) Weak automatization (*traditional IT*) is an attribute of...

☐ Grids ☐ Clouds

e) Virtual Organizations are implemented in...

☐ Grids ☐ Clouds

f) Resources without central control are usually connected in...

☐ Grids ☐ Clouds

g) Resources under central control are usually connected in...

☐ Grids ☐ Clouds

h) Consumption-based billing (Pay-as-you-go principle) is an attribute of...

☐ Grids ☐ Clouds

i) Four HTTP methods are enough to work with resources inside storage services like S3 or Google Cloud Storage. Write the HTTP methods into the table.

HTTP method	**Description**
	Create or replace resource
	Request resource
	Append something to a resource
	Erase resource

j) Why is it recommendable, that storage services do not only implement support for the four HTTP methods of subtask i), but also for the HTTP method `HEAD`?

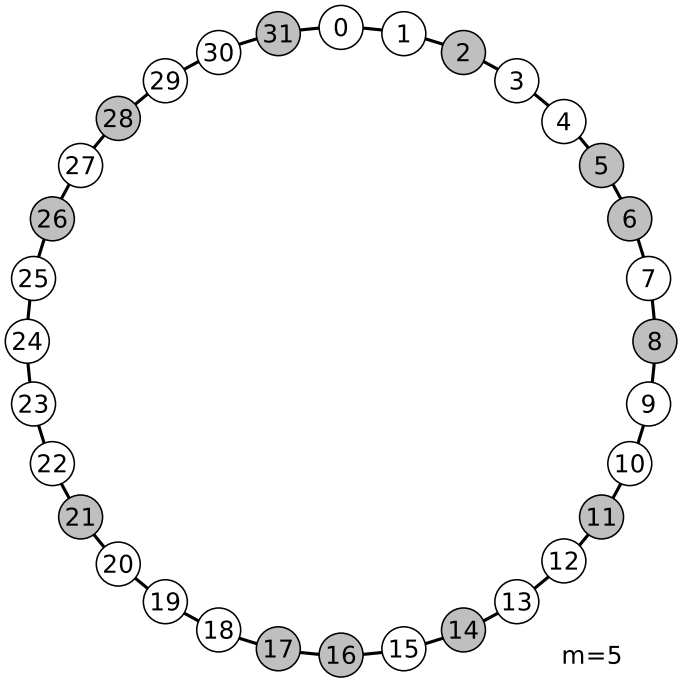
Last name: First name: Student number:

Question 5)

Points: .....

Maximum points: 10+1+1+1+1+1=15

a) Calculate the Finger Table values of node $n = 9$ and insert the correct values into the provided Finger Table.



Finger Table of node $n = 9$

Entry	**Start**	**Node**
1		
2		
3		
4		
5		

The table has 5 entries, because m contains the length of the ID in bits and $m = 5$

The Start value of entry i of the table on node n is $(n + 2^{i-1}) \bmod 2^m$

The Node value of entry i points to the first node, which follows to n at a distance of at least 2^{i-1}

b) Which node is responsible for the key (resource) with ID 22 ?

c) What is the maximum number of instances, an EBS volume can be connected with?

d) According to which principle works the storage service S3?

☐ block-based storage service ☐ object-based storage service

e) According to which principle works the storage service EBS?

☐ block-based storage service ☐ object-based storage service

f) Which storage services require the user/customer to choose and deploy a file system?

☐ block-based storage services ☐ object-based storage services

Last name: First name: Student number:

Question 6)

Points: .....

Maximum points: 12

Name four cloud services (only platform and infrastructure services are allowed!) you used for solving the exercise sheets. Also explain in a few words which functionality of theses services you used. It should become clear why you used each single service.

Name of service	Sort of service	Explain the functionality you used and also the reason for using the service
	☐ PaaS ☐ IaaS	
	☐ PaaS ☐ IaaS	
	☐ PaaS ☐ IaaS	
	☐ PaaS ☐ IaaS	

Last name: First name: Student number:

Question 7)

Points: .....

Maximum points: 3+7+2=12

Your local time in Frankfurt am Main is Monday 09:00 (UTC+1). You need to copy 3 TB of data into the storage service S3. You have two options:

- **Scenario 1:** You immediately start at 09:00 (UTC+1) to upload the 3 TB of data to S3 via the internet. Consider the data rate between your computer and S3 is 100 Mbit/s.
- **Scenario 2:** You use the AWS Import/Export service. Therefore you copy the data to a HDD, which is connected via USB 3.0. The transfer rate (for write) is 125 MB/s.

  After you copied the data, you pack the HDD into a parcel and send it via a package delivery company to Amazon. DHL, UPS and FedEx can deliver a parcel from Frankfurt am Main in less than 24 hours to most places in Europe.

  You need 15 Minutes to put the HDD into a parcel and another 15 Minutes to bring the parcel to the branch office of a package delivery company.

  The parcel must arrive at the branch office of the package delivery company no later than 16:30 (UTC+1) to arrive at Amazon in Ireland at 9:00 (UTC) the next working day.

  An Amazon employee needs to copy the data from the HDD to the S3 service. The transfer rate of the HDD (for read) is 150 MB/s.

  Consider 2 hours additional overhead for the in-house mail at Amazon to ship the HDD to the correct employee.

Calculate...

a) for the first scenario, how long it takes until the data is copied to S3.

b) for the second scenario, how long it takes until the data is copied to S3.

c) the data rate of the second scenario.

(For all subtasks, the calculation steps must be visible.)

Last name: First name: Student number:

Question 7 – Additional Page)

Maximum points: 3+7+2=12

Last name: First name: Student number:

Question 8)

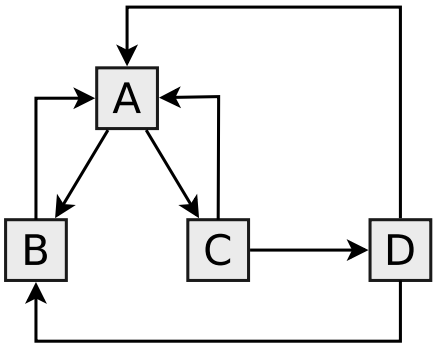
Points: .....

Maximum points: 12

- PR_p = PageRank of a web page p
- $L_{IN}(p)$ = Set of documents, which refer to $p \Longrightarrow$ incoming links
- $L_{OUT}(p)$ = Set of documents, to which p refers $\Longrightarrow$ outgoing links
- d = damping factor between 0 and 1

$$PR(p) = (1-d) + d * \sum_{p_i \in L_{IN}(p)} \frac{PR(p_i)}{\text{amount } L_{OUT}(p_i)}$$

Calculate the missing iterations of the PageRank algorithm for the given example scenario with $d = 0.8$.



	0	1	2	3	4	5	PR
A	1		1.48		1.5184		1.553216
B	1		1.16		1.096		1.071424
C	1		0.92		0.8688		0.852416
D	1		0.44		0.5168		0.522944