270676 - CCBDA - Cloud Computing and Big Data Analytics

Coordinating unit: 270 - FIB - Barcelona School of Informatics
Teaching unit: 701 - DAC - Department of Computer Architecture
Academic year: 2019
Degree: MASTER’S DEGREE IN INNOVATION AND RESEARCH IN INFORMATICS (Syllabus 2012). (Teaching unit Optional)
ECTS credits: 6
Teaching languages: English

Prior skills

General knowledge of:
- TCP/IP networking
- Operating Systems basic administration and use of the OS from the programs
- Software development

Basic knowledge of:
- Unix command line.
- Python programming language.
- Git version control system.

Warning. Students are supposed to have the above background before starting the laboratory sessions. Complimentary fast-paced materials will be provided before class to help students meet the above requirements.

Degree competences to which the subject contributes

Basic:
CB7. Ability to integrate knowledges and handle the complexity of making judgments based on information which, being incomplete or limited, includes considerations on social and ethical responsibilities linked to the application of their knowledge and judgments.
CB8. Capability to communicate their conclusions, and the knowledge and rationale underpinning these, to both skilled and unskilled public in a clear and unambiguous way.

Specific:
CEE2.1. Capability to understand models, problems and algorithms related to distributed systems, and to design and evaluate algorithms and systems that process the distribution problems and provide distributed services.
CEE2.3. Capability to understand models, problems and mathematical tools to analyze, design and evaluate computer networks and distributed systems.

General:
CG4. Capacity for general and technical management of research, development and innovation projects, in companies and technology centers in the field of Informatics Engineering.
CG5. Capability to apply innovative solutions and make progress in the knowledge to exploit the new paradigms of computing, particularly in distributed environments.

Transversal:
CTR1. ENTREPRENEURSHIP AND INNOVATION: Capacity for knowing and understanding a business organization and the science that rules its activity, capability to understand the labour rules and the relationships between planning, industrial and commercial strategies, quality and profit. Capacity for developing creativity, entrepreneurship and innovation trend.
CTR3. TEAMWORK: Capacity of being able to work as a team member, either as a regular member or performing directive activities, in order to help the development of projects in a pragmatic manner and with sense of responsibility; capability to take into account the available resources.
CTR6. REASONING: Capacity for critical, logical and mathematical reasoning. Capability to solve problems in their area of study. Capacity for abstraction: the capability to create and use models that reflect real situations. Capability to
design and implement simple experiments, and analyze and interpret their results. Capacity for analysis, synthesis and evaluation.

**Teaching methodology**

Lectures, reading and discussion of technical and research papers, Presentation of topics (and papers) by students. Laboratory sessions and a practical class project.

Students are required to bring their laptop to carry out the laboratory sessions and practical class project.

This subject is taught only in English language.

**Learning objectives of the subject**

1. Present the student with new execution environments required to manage the computing resources and simplify the development and integration of the different types of applications and services at nowadays Internet-scale systems.
2. Collaborate in the design, implementation and presentation of a cloud computing environment that is required for a class project.
3. Find and understand useful information to create innovative solutions.

**Study load**

<table>
<thead>
<tr>
<th>Total learning time: 150h</th>
<th>Hours large group: 27h</th>
<th>18.00%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hours medium group: 0h</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>Hours small group: 16h 30m</td>
<td>11.00%</td>
</tr>
<tr>
<td></td>
<td>Guided activities: 10h 30m</td>
<td>7.00%</td>
</tr>
<tr>
<td></td>
<td>Self study: 96h</td>
<td>64.00%</td>
</tr>
</tbody>
</table>

Lectures, reading and discussion of technical and research papers, Presentation of topics (and papers) by students. Laboratory sessions and a practical class project.

Students are required to bring their laptop to carry out the laboratory sessions and practical class project.

This subject is taught only in English language.
## Lectures: Cloud Computing fundamentals

**Degree competences to which the content contributes:**

**Description:**  
Fundamental concepts: The effect of scale on system properties.  
- Issues in large-scale systems: virtualization, service orientation and composition, availability, locality, performance and adaptation.  
- Models for large-scale systems: system models for analysis, architectural models and service/deployment models.  
- Scaling techniques: basic techniques, scalable computing techniques for architectural models.  
- Middleware and Applications: computing, storage, web, content distribution, Internet-scale systems or services.  
- Environment and applications requirements.

## Laboratory sessions: Practical view of Cloud Computing

**Degree competences to which the content contributes:**

**Description:**  
Big Data Analytics in the Cloud  
- APIs: The Doors in the Cloud  
- Current required layers in Big Data Software Stack  
- New Software requirements for Advanced Analytics  
- New Hardware requirements for Advanced Analytics

## Assignment: Experimental part

**Degree competences to which the content contributes:**

**Description:**  
Development of a prototype application using Cloud service offerings (such as AWS, Google AppEngine, Open Stack, OpenNebula)  
- Development of a prototype application using advanced analytics services either provided regarding APIs or Software as a Service.
## Planning of activities

| Introduction | Hours: 6h  
Theory classes: 2h  
Practical classes: 0h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 4h |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description:</strong></td>
<td>Cloud Computing Definition. Service Oriented Architectures. Web Services. Business considerations</td>
</tr>
<tr>
<td><strong>Specific objectives:</strong></td>
<td>1</td>
</tr>
</tbody>
</table>

| Cloud Computing Architecture | Hours: 6h  
Theory classes: 2h  
Practical classes: 0h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 4h |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description:</strong></td>
<td>Technology. Architecture. Modelling and Design.</td>
</tr>
<tr>
<td><strong>Specific objectives:</strong></td>
<td>1</td>
</tr>
</tbody>
</table>

| Virtualization | Hours: 6h  
Theory classes: 2h  
Practical classes: 0h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 4h |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description:</strong></td>
<td>Foundations. Grid, cloud and virtualization.</td>
</tr>
</tbody>
</table>

| Data Storage | Hours: 12h  
Theory classes: 4h  
Practical classes: 0h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 8h |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Specific objectives:</strong></td>
<td>1</td>
</tr>
</tbody>
</table>
## Cloud Services

**Hours:** 12h  
- Theory classes: 4h  
- Practical classes: 0h  
- Laboratory classes: 0h  
- Guided activities: 0h  
- Self study: 8h

## Cloud Security

**Hours:** 6h  
- Theory classes: 2h  
- Practical classes: 0h  
- Laboratory classes: 0h  
- Guided activities: 0h  
- Self study: 4h

## Service Oriented Architectures

**Hours:** 6h  
- Theory classes: 2h  
- Practical classes: 0h  
- Laboratory classes: 0h  
- Guided activities: 0h  
- Self study: 4h

## Cloud Tools

**Hours:** 12h  
- Theory classes: 4h  
- Practical classes: 0h  
- Laboratory classes: 0h  
- Guided activities: 0h  
- Self study: 8h

## Cloud Applications

**Hours:** 9h  
- Theory classes: 3h  
- Practical classes: 0h  
- Laboratory classes: 0h  
- Guided activities: 0h  
- Self study: 6h
## Future Trends

**Hours:** 4h  
Theory classes: 2h  
Practical classes: 0h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 2h

## Collaborative class project

**Hours:** 34h  
Theory classes: 0h  
Practical classes: 0h  
Laboratory classes: 0h  
Guided activities: 4h  
Self study: 30h

## Specific objectives:
2, 3

## Lab: Basic knowledge toolbox

**Hours:** 2h  
Theory classes: 0h  
Practical classes: 0h  
Laboratory classes: 0h 30m  
Guided activities: 1h 30m  
Self study: 0h

## Lab: Doors in the cloud

**Hours:** 2h  
Theory classes: 0h  
Practical classes: 0h  
Laboratory classes: 1h 30m  
Guided activities: 0h 30m  
Self study: 0h

## Lab: Content Delivery Network

**Hours:** 2h  
Theory classes: 0h  
Practical classes: 0h  
Laboratory classes: 1h 30m  
Guided activities: 0h 30m  
Self study: 0h
### Lab: Extract and analyze data

**Hours:** 2h  
- Theory classes: 0h  
- Practical classes: 0h  
- Laboratory classes: 1h 30m  
- Guided activities: 0h 30m  
- Self study: 0h

### Lab: Interact with users and services

**Hours:** 2h  
- Theory classes: 0h  
- Practical classes: 0h  
- Laboratory classes: 1h 30m  
- Guided activities: 0h 30m  
- Self study: 0h

### Lab: Monitoring and Security

**Hours:** 2h  
- Theory classes: 0h  
- Practical classes: 0h  
- Laboratory classes: 1h 30m  
- Guided activities: 0h 30m  
- Self study: 0h

### Lab: Data storage

**Hours:** 2h  
- Theory classes: 0h  
- Practical classes: 0h  
- Laboratory classes: 1h 30m  
- Guided activities: 0h 30m  
- Self study: 0h

### Lab: Web Services

**Hours:** 2h  
- Theory classes: 0h  
- Practical classes: 0h  
- Laboratory classes: 1h 30m  
- Guided activities: 0h 30m  
- Self study: 0h
<table>
<thead>
<tr>
<th>Discuss: Virtualization</th>
<th>Hours: 4h</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Theory classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Practical classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Laboratory classes: 1h</td>
</tr>
<tr>
<td></td>
<td>Guided activities: 1h</td>
</tr>
<tr>
<td></td>
<td>Self study: 2h</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discuss: Cloud providers comparison</th>
<th>Hours: 4h</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Theory classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Practical classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Laboratory classes: 1h</td>
</tr>
<tr>
<td></td>
<td>Guided activities: 1h</td>
</tr>
<tr>
<td></td>
<td>Self study: 2h</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discuss: Federated Cloud Computing</th>
<th>Hours: 4h</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Theory classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Practical classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Laboratory classes: 1h</td>
</tr>
<tr>
<td></td>
<td>Guided activities: 1h</td>
</tr>
<tr>
<td></td>
<td>Self study: 2h</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discuss: Cloud governance</th>
<th>Hours: 4h</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Theory classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Practical classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Laboratory classes: 1h</td>
</tr>
<tr>
<td></td>
<td>Guided activities: 1h</td>
</tr>
<tr>
<td></td>
<td>Self study: 2h</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discuss: Future trends</th>
<th>Hours: 4h 42m</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Theory classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Practical classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Laboratory classes: 1h 12m</td>
</tr>
<tr>
<td></td>
<td>Guided activities: 1h 30m</td>
</tr>
<tr>
<td></td>
<td>Self study: 2h</td>
</tr>
</tbody>
</table>
Students will be evaluated on their participation in class, laboratory sessions, class attendance, reading and presenting reports and papers and assignments on specific topics.

The final grade for the course is the weighted average of the grades for the following components obtained en each part of the course:

- Lab sessions: 30%
- Papers Reading/Presentation and homework: 20%
- Course Projects: 30%
- Final exam: 20%

In order to be able to publicly defend the course project, students must have attended at least 70% of the classes and teams must have delivered on time the activities that have been planned during the course. The course project is the result of teamwork, which will be reflected in the grade given to the group as a whole. Each member of the group will be responsible for part of the project and might be graded individually on his or her contribution.

Bibliography

Basic:
