UAB
Universitat Autònoma de Barcelona

2023/2024

Distributed Systems

Code: 102740
ECTS Credits: 6

Degree	Type	Year	Semester
2502441 Computer Engineering	OB	3	1
2502441 Computer Engineering	OT	4	1

Contact

Name: Porfidio Hernandez Bude

Email: porfidio.hernandez@uab.cat

Teaching groups languages

You can check it through this link. To consult the language you will need to enter the CODE of the subject. Please note that this information is provisional until 30 November 2023.

Teachers

Sergio Villar Mesurado

Prerequisites

There are no prerequisites. However, students should be familiar with Computer Fundamentals, Computer Structure, Operating Systems and Computer Networks. It is also recommended for students to have notions of Linux System.

Objectives and Contextualisation

The subject of Distributed Systems aims to show the fundamental concepts, and most common problems of distributed systems, as well as the most common techniques of resource management in this type of systems. In the same line, the student will be introduced to the use of common tools in these environments, for the resolution of computational problems coming from the sciences and engineering.

Competences

Computer Engineering
- Acquire personal work habits.
- Capacity to design, develop, evaluate and ensure the accessibility, ergonomics, usability and security of computer systems, services and applications, as well as of the information that they manage.
- Communication.

- Conceive and develop centralised or distributed computer systems or architectures by integrating hardware, software and networks.
- Have the capacity to analyse and evaluate computer architectures, including parallel and distributed platforms, and develop and optimise software for the same.
- Have the capacity to analyse, evaluate, select and configure hardware platforms for the development and execution of computer applications and services.
- Have the capacity to conceive network technology based systems, applications and services, including Internet, Web, e-commerce, multimedia, interactive services and mobile computers.
- Have the capacity to design and implement system and of communications software.
- Work in teams.

Learning Outcomes

1. Analyse accessibility and security needs in accordance with the different types of users and different types of applications.
2. Analyse and evaluate the programming models and paradigms for distributed systems.
3. Analyse the characteristics of distributed systems.
4. Apply the concepts of operating systems and networks to develop the necessary software components to manage a high performance computer system and the communications involved in these systems.
5. Classify the different types of architecture for distributed systems, considering aspects relative to hardware and its interconnection, and to the software components of the system.
6. Communicate efficiently, orally or in writing, knowledge, results and skills, both in the professional environment and before non-expert audiences.
7. Conceive network technology based hardware systems, including Internet, web, e-commerce, multimedia, interactive services and mobile computation.
8. Describe the basic performance parameters of distributed computing systems.
9. Design computer solutions that integrate accessibility and security needs in a distributed system.
10. Design efficient base software for distributed systems.
11. Design network technology based systems, applications and services.
12. Evaluate and select distributed systems depending on performance index.
13. Evaluate the functionality and features of applications that run on distributed platforms.
14. Identify the architectures of distributed systems.
15. Identify the basic components that define the accessibility, ergonomics, usability and security of computer systems.
16. Identify the characteristics of distributed systems considering the available network technologies.
17. Identify the software levels that a distributed system should comprise.
18. Prevent and solve problems.
19. Propose IT solutions based on distributed systems that integrate in the architecture of the hardware components of the system, as well as the interconnection of the same, and the design of the necessary software.
20. Work cooperatively.

Content

Topic 1: Parallel and Distributed computing systems

- Parallel Infrastructures (Clusters, Grid and Clouds).
- Use cases: Hadoop/Spark.

Topic 2: Process synchronization and communication

- Processes and threads.
- Process communication (messages, RPC, communication by groups)
- Physical and logical clocks.
- Mutual exclusion and Deadlock.

- Selection Algoritms and fault tolerance.

Topic 3: Distributed Shared Memory (DVSM)

- Memory management in multiprocessors
- Consistency models (strict, sequential, causal, pram)
- DVSM implementations based on pages, variables and objects.

Topic 4: Job Scheduling

- Level of Schedulers.
- Taxonomies
- Scheduling: Tasks, jobs, workflows, virtual machines and containers.

Topic 5: Distributed file system

- Architecture of parallel I / O.
- Design of Distributed File Systems.
- Taxonomy

Topic 6: Use Cases: Large Scale Video on Demand (LVoD)

- Types of VoD
- Design of a local VoD system
- Scalability
- LVoD

Methodology

Theoretical classes

Will be taught through lectures, although students will be encouraged to actively participate in the resolution of examples. Campus Virtual will be used for communication between lecturers and students (material, updates, announcements, etc.). The content of each of the classes is detailed in the planning sheet of the subject (schedule), which will be published on the first day of class, in the Virtual Campus. A reference book and teacher's notes will be used for those subjects that require it; in the same line, the transparencies of the subject will be available for each topic, and the resources used in class: test questions, exams from previous courses, papers on specific topics, use of chats and forums for discussion of topics from the subject etc. A weekly tutoring schedule will be enabled during the semester, which will be shown in the course schedule.

Exercise-based classes

- a.- Presentation of works. The following section aims to motivate the student and develop strategies for oral presentation, writing and public defense of work on issues related to distributed systems. At the beginning of the course, students will be grouped into work teams following the teacher's instructions, which will assign work topics to each team. During the semester, that will be detailed in the Virtual Campus, the teams will have tutorial meetings with the exercise teacher to monitor the evolution of the work.

    - 4 groups, 100,130 students, 2/4 students / work team, 10 work teams / group.

    - Defense of work (Presentation to class for 20 minutes, questions 10 minutes per team). (5 hours / group).

    - "Tutored sessions" with the teacher, dedicated to dealing with methodological and planning aspects. (2 sessions 15m / team)

- The hours dedicated in presentation class of works are indicated for each course in the schedule of the Degree. Registration to problem groups will be done through the corresponding Virtual Campus tool. The deadline to do so will be indicated in the schedule of the subject.

- b.- During exercise sessions, a list of problems will be resolved. Students are encouraged to solve the problems (7 hours remaining) on their own in advance. Students will also be encouraged to present their own solutions in class. The problem sessions will be indicated in the schedule of the subject.

Laboratory Practices

The laboratory practices will be carried out in sessions distributed during the course in the Departament laboratory; according to the corresponding schedule published in the Virtual Campus. The teachers of laboratory practices will generate the lists of dates and times. The groups of laboratory practices should be of two people. It is important to respect the scheduled associated with each group, in order to keep the practice server load balanced, trying to avoid spikes for proper system operation.

Transversal Competences: Transversal competencies assigned (T02.04 - Prevent and solve problems, T03.01 - Work cooperatively and T04.01 - Communicate efficiently, orally or in writing, knowledge, results and skills, both in professional environments and before public not experts), will work and evaluate mainly in the activities of exercises and laboratory practices. It is about fostering work team, as well as accurately presenting the results or analysis to specific problems; through the exhibition and discussion in class of works related to the topics of the subject. The registration to the groups of laboratoty practices will be done through the corresponding tool of the Virtual Campus. The deadline to do so will be indicated in the schedule of the subject.

Annotation: Within the schedule set by the centre or degree programme, 15 minutes of one class will be reserved for students to evaluate their lecturers and their courses or modules through questionnaires.

Activities

Title	Hours	ECTS	Learning Outcomes
Type: Supervised			
Exercises	12.5	0.5	2, 3, 1, 12, 13, 6, 11, 18, 19, 20
Laboratory practices	12	0.48	2, 3, 12, 13, 6, 7, 8, 9, 11, 10, 18, 19, 20
Theory	22	0.88	2, 3, 4, 12, 5, 7, 8, 11, 17, 14, 16, 18, 19
Type: Autonomous			
Personal work	93	3.72	2, 3, 1, 5, 7, 15, 17, 14, 16

Assessment

The course consists of three parts: Theory, Problems and Practices. The Theory part represents 50% of the final grade for the course, the Problems part 20%, and the Practices part the remaining 30%. It is essential that the weighted average of the three tests is at least 5.0 to pass the course. In addition, a minimum grade of 5.0 is required for each part of the subject to pass it. This subject does not provide for the single assessment system.

A student will be considered non-evaluable (NA) if he has not taken part in a set of activities whose weight is equivalent to a minimum of two-thirds of the subject's total grade.

The dates of continuous evaluation and delivery of works will be published in the virtual campus, and may be subject to possible programming changes for reasons of adaptation to possible incidents. Always be informed in the virtual campus about these changes since it is understood that the CV is the usual mechanism of exchange of information between teacher and students.

In the case of repeating students, it is not expected that repeat students can be evaluated in a different way compared to non-repeaters; they will request the teacher responsible for the subject, via institutional mail (porfidio.hernandez@uab.cat), and during the first two weeks of the course; the validation of theory, laboratory practices and / or presentation work in class and test of presentation works by specific group. The associated grade will be obtained for these sections during the 2022-2023 academic year.

In order to pass the course with honours, the final grade must be a 9.0 or higher. Because the number of students with this distinction cannot exceed 5% of the number of students enrolled in the course, this distinction will be awarded to whoever has the highest final grade. In case of a tie, partial-test results will be taken into consideration.

The way in which each of the different parts of the subject (Theory, Exercises and Laboratory Practices) will be evaluated is detailed below:

The evaluation of the theory part will be done with two individual partial controls/tests in regular class. As part of continuous assessment, the first test will take place during lectures; the second will take place on the date specified by coordination.They will be carried out in the manner established by the regulations of the center, by default, the modality of the exam will be face-to-face; which will be reported in the Virtual Campus news. The day that each of these tests will be carried out, as well as its content, is indicated in the course planning sheets. The theory grade will be obtained from the average of the two theory tests, provided that the score of each control is greater than or equal to 5.0. On the day of the second test, the first test will NOT be retrievable.

The evaluation of the exercises part consists of two parts:

    a) On the subject of presentation work, it will have to:
    a1) Perform a written work (10 pages- IEEE-formed standard sections: abstract, introduction, related work, specific content paper, results, conclusions, future workand references) and deliver it to the reponsible teacher using the virtual campus.
    a2) Defense of the work (oral presentation in regular class of 20 minutes, questions 10 minutes) on the day and time assigned for that purpose in the Virtual Campus.
    b) Presentation Works .

Laboratory Practices.

Practices Students' attendance and punctuality at all laboratory sessions is mandatory to pass the practicals; If one or more absences are not justified, themaximum practical grade will be 4.50. The student will connect to a server of the Department, in the hours planned for the purpose, where in general the correct operation of the system can be guaranteed; being able to connect during other periods under your responsibility. Practice teachers will enable tutoring sessions to monitor the practices, the timetables will be available on the Virtual Campus.

The internship qualification consists of two parts:
a) The reports written for each of the sessions. 50% of the practice grade.
b) Practice exam in the last session. 50% of the practice grade. Minimum score of 4.00 in this exam to pass the internship.

The memory of each session must detail:

Objective of the practice
Description and approach to the practice
Description of the procedures used at the functional level
Description of the problems that arose during the implementation of the practice and the solutions found
Conclusions drawn from the realization of the practice
Configuration files and tests

For each evaluation activity, a place, date and time of revision in which the student can review the activity with the teacher will be indicated. In this context, claims may be made on the activity grade, which will be evaluated by the faculty responsible for the subject. If the student does not show up for this revision, this activity will not be reviewed later.

Recovery Activities:

The student can present himself for recovery as long as he has presented himself to a set of activities that represent a minimum of two-thirds of the total grade of the subject.

- All those students who, having followed the continuous assessment, have not achieved the minimum necessary to pass the subject; They will have the option to do a final exam of recovery, for each one of the sections of which the subject consists: theory, laboratory practices and exercises. The date anddate of these recovery exams will be published to the Virtual Campus.
- Regarding therecovery of the laboratory practices, the responsible teacher will decide the recovery model to apply, by group orindividualized (extension or correction of any of the practices, individualized practice exam, etc). In relation to the problem recovery section, theresponsible professor will decide the recovery model to apply, as in the previous case.
- In relation to the exercises recovery section, the responsible professor will decide the recovery model to apply, as in the previous case.
- For the part corresponding to the theory recovery, the students will be able to take a final theory exam. This exam will be face-to-face and will consist of two individual partial tets, and the student will have to submit to the partial test that has not passed (minimum score of each control greater than or equal to 5.0, to overcome the theory part).

  For the cases where some of the parts of the subject is suspended (note of theory, exercises, or laboratory practices; <5.0), after the recovery process; and that when making the calculation of the final grade according to the corresponding percentages is greater than 5.0, the note that will appear as the final grade of the subject will be a 4.5 (Following the recommendation of the coordination of the Section).

  Notwithstanding other disciplinary measures deemed appropriate, and in accordance with the academic regulations in force, assessment activities will receive a zero whenever a student commits academic irregularities that may alter such assessment. Assessment activities graded in this way and by this procedure will not be re-assessable. If passing the assessment activity or activities in question is required to pass the subject, the awarding of a zero for disciplinary measures will also entail a direct fail for the subject, with no opportunity to re-assess this in the same academic year. Irregularities contemplated in this procedure include, among others:

  - the total or partial copying of a practical exercise, report, or any other evaluation activity;
  - allowing others to copy; unauthorized use of AI (eg Copilot, ChatGPT or equivalent;
    - presenting group work that has not been done entirely by the members of the group;
    - presenting any materials prepared by a third party as one's own work, even if these materials are translations or adaptations, including work that is not original or exclusively that of the student;
    - having communication devices (such as mobile phones, smart watches, etc.) accessible during theoretical-practical assessment tests (individual exams).
    In future editions of this subject, the student who has committed irregularities in an evaluation act, will not be validated any of the evaluation activities carried out. In summary: copy, leave copy or plagiarize (or attempt) in any of the evaluation activities is equivalent to a SUSPENS, not compensable and without validation of parts of the subject in subsequent courses.
    An overall grade of 5 or higher is required to pass the subject. A "non-assessable" grade cannot be assigned to students who have participated in any of the individual partial tests or the final exam. The grade in the Transcript of Records (ToR) will be the lowest value between 3.0 and the weighted average grade, in the event of irregularities having been committed for any assessment activity (and therefore re-assessmentwill not be possible).It is important to bear in mind that no assessment activities will be permitted for any student at a different date or time to that established, unless for justified causes duly advised before the activity and with the lecturer's previous consent. In allother cases, if an activity has not been carried out, this cannot be

re-assessed. In the case of exercise resolution, a review may be requested afterthe date of the activity or the date of closure of the quiz. For all other assessment activities, a place, date and time of review will be indicated allowing students to review the activity with the lecturer. In this context, students may discuss the activity grade awarded by the lecturers responsible for the subject. If students do not take part in this review, no further opportunity will be made available. To consult the academic regulations approved by the Governing Council of the UAB, please follow this link: https://www.uab.cat/doc/TR_Normativa_Academica_Plans_Nous.

Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Exercises	20%	1.5	0.06	2, 3, 12, 13, 5, 6, 7, 9, 16, 18, 19, 20
Laboratory practices	30%	2	0.08	2, 4, 12, 13, 6, 11, 10, 18, 20
Parcial tests	50%	7	0.28	2, 3, 1, 4, 8, 15, 17, 14, 16, 19

Bibliography

General references:

- Andrew S. Tanenbaum, Maarten Van Steen "Distributed Systems: Principles and Paradigms", 2nd Ed. Addision-Wesley, Pearson 2007.
- G. Coulouris, J. Dollimore, T. Kinderg and Gordon Blair "Distributed Systems: Concepts and Design", Addison-Wesley, Pearson, 5ª Ed. 2012.
- Andrew S. Tanenbaum "Sistemas Operativos Distribuidos", Prentice Hall 1996.

Specific references

- Tom White, "Hadoop: The definitive guide", O´Reilly 2009
- Bill Chambers & M. Zaharia, "Spark The Definitive Guide, big data processing made simple", O´Reilly, 2018.

Software

Scala

https://downloads.lightbend.com/scala/2.12.12/scala-2.12.12.rpm

Spark

http://www-eu.apache.org/dist/spark/spark-3.0.1/spark-3.0.1-binhadoop2.7.tgz