OPERATING SYSTEMS & PARALLEL COMPUTING

Bachelor in Data and Business Analytics BDBA SEP-2023
OSPC-DBA.2.M.A

Area Information Systems and Technology
Number of sessions: 30
Academic year: 23-24
Degree course: SECOND
Number of credits: 6.0
Semester: 2º
Category: COMPULSORY
Language: English

Professor: **JOSE ANTONIO GARCÍA ESCANDÓN**

E-mail: joseantoniog@faculty.ie.edu

Short description

Entrepeneural engineer with a lot of experience creating custom made solutions with a variety of technologies. Have been around the industry raising capital, working at a unicorn, working in oen of the oldest companies in the world and even built my own.

Education

- Master in Big Data and Business Analytics - IE - Spain
- Engineering in Information Technology and Intelligent Systems - Universidad Panamericana - Mexico

Work Experience

- Galeo - IoT Engineer - Apr.2022 - Present
- Cover Genius (Insurance) - Integrations Engineer - MADRID, SPAIN - Feb. 2022 - PRESENT
- Mercedes-Benz Mexico - IT Tech and Innovation Lead - APR. 2019 – Feb. 2022
- Tresgarfin (FINTECH) Mexico City, Mexico - CEO and Founder - ABR. 2019 – PRESENT
- Mutuo Financiera (FINTECH) Tech Lead - AUG. 2017- APR. 2019

Office Hours

Office hours will be on request. Please contact at:

joseantoniog@faculty.ie.edu

SUBJECT DESCRIPTION

Today's problems (weather or earthquakes forecasting, crash test simulations, development of new medicaments or new materials, managing data in search engines, stocks problems, …) require a huge computing power and large storage capacities, which generally can only be provided by high performance computer systems or even networked computers. These systems can only be used efficiently for these kinds of problems with parallel programs. This course pinpoints a profound knowledge for some parallel programming techniques.

For Data management and Analysis, a huge volume of data or information must be analysed and processed quickly and it is the reason that the knowledge in operating systems and parallel computing is really necessary in order to take advantage of these techniques.

In this context, this course will cover a brief historical perspective of the evolution of operating systems over the last fifty years, will teach you its components and how to swiftly operate with it and then cover the major components of most modern operating systems until the most modern parallel computing systems. This course is about writing effective programs to harness the unprecedented power provided by modern parallel computers, so that the programs attain the highest possible levels of performance the machines are capable of. The parallel computers we focus on include multi-core processors as well as clusters and supercomputers made from them. This course also teaches an adequate analytical framework for understanding performance, including performance models, scalability analysis, and iso-efficiency.

Finally, we discuss programming approaches for executing applications on accelerators such as GPUs. This part introduces the CUDA (Compute Unified Device Architecture) programming environment and how the cloud has changed parallel computing.

LEARNING OBJECTIVES

The objective of this course is to provide students with basic principles of operating systems and parallel and distributed computing needed for data analysis and management. This course includes these topics: OS role, components and structure, process concepts, threads, CPU scheduling, memory management, concurrent processes, interprocess coordination and communication, parallelism versus concurrency, parallel architecture, algorithms, and decomposition.

At the end of the course; students should be able to:

- Write efficient parallel programs for multicore processors and distributed memory machines;
- List architectural elements of modern operating systems and processors and explain their impact on performance;
- Identify how to transform a problem solved without parallelization to one being solved with it.
- Optimise performance by applying parallelization and being able to detect the difference between.
- Explain the functionality of a parallel system and its component
- Create and debug programs to accomplish a computational task;
- Analyse scalability of parallel algorithms and derive their iso-efficiency function;
- Explain workings of covered parallel algorithms and reason about their efficacy and of variants;
- Answer questions about parallel application domains and their characteristics, as well as computational challenges in those domains;
- Develop a solid parallel algorithm; from theory to execution and implementation.
- Understand the cloud implications and new technologies from the cloud.

TEACHING METHODOLOGY

IE University teaching method is defined by its collaborative, active, and applied nature. Students actively participate in the whole process to build their knowledge and sharpen their skills. Professor's main role is to lead and guide students to achieve the learning objectives of the course. This is done by engaging in a diverse range of teaching techniques and different types of learning activities such as the following:

Learning Activity	Weighting	Estimated time a student should dedicate to prepare for and participate in
Lectures	28.0 %	42.0 hours
Discussions	9.33 %	14.0 hours
Exercises in class, Asynchronous sessions, Field Work	9.33 %	14.0 hours
Group work	26.67 %	40.0 hours
Individual studying	26.67 %	40.0 hours
TOTAL	100.0 %	150.0 hours

PROGRAM

SESSION 1 (LIVE IN-PERSON)

Topics:

Operating system principles and computer architecture.

REQUIRED READING:

Book Chapters: Modern Operating Systems (3rd Edition): Chapter 1-2 (See Bibliography)

SESSION 2 (LIVE IN-PERSON)

Topics

- Processes and threads

REQUIRED READING:

Book Chapters: Modern Operating Systems (3rd Edition): Chapter 1-2 (See Bibliography)

SESSION 3 (LIVE IN-PERSON)

TOPICS:

- Operating Systems Kernels.
- Assignment: Exercises in Python

REQUIRED READING:

Book Chapters: Modern Operating Systems (3rd Edition): Chapter 1-2 (See Bibliography)

SESSION 4 (LIVE IN-PERSON)

TOPICS:

- Deadlocks

REQUIRED READING:

Book Chapters: Modern Operating Systems (3rd Edition): Chapter 3-9 (See Bibliography)

SESSION 5 (LIVE IN-PERSON)

TOPICS:

- Linux operating system
- Linux directory structure
- Linux commands
- Exercises

REQUIRED READING

Book Chapters: The Linux Commands Handbook (Pages 5 - 8) (techjunction.co)

SESSION 6 (LIVE IN-PERSON)

TOPICS:

- Concatenating commands
- Daemons
- Managed System Services

SESSION 7 (LIVE IN-PERSON)

TOPICS:

- Memory Management
- Algorithms:
    - First Come First Serve
    - Shortest Job First
    - Priority SchedulingRound Robin

SESSION 8 (LIVE IN-PERSON)

TOPICS:

- Storage
    - Files
        - Attributes
        - Operations
        - File Types
        - Structure
    - File Permissions

SESSION 9 (LIVE IN-PERSON)

TOPICS:

- Input/Output.
- Drivers

SESSION 10 (LIVE IN-PERSON)

TOPICS:

- Multiple Processor Systems.

REQUIRED READING:

Book Chapters: Modern Operating Systems (3rd Edition): Chapter 10-12 (See Bibliography)

SESSION 11 (LIVE IN-PERSON)

TOPICS:

- OS Labs (Windows, MacOSX, Linux).
- Detailed Labs PDF to read.
- Practice on your own PC.

SESSION 12 (LIVE IN-PERSON)

TOPICS:

- Security.
- User Permissions
- Group Permissions

REQUIRED READING:

Book Chapters: Modern Operating Systems (3rd Edition): Chapter 10-12 (See Bibliography)

SESSION 13 (LIVE IN-PERSON)

TOPICS:

- OS Labs (Linux).
- Detailed Labs PDF to read.
- Practice on your own PC

SESSION 14 (LIVE IN-PERSON)

REVIEW ALL SESSIONS

FORUM - Space to discuss doubts

REVIEW ALL SESSIONS

Kahoot in class simulating exam questions. Doubts and discussions for the exam.

SESSION 15 (LIVE IN-PERSON)

PARTIAL EXAM

SESSION 16 (LIVE IN-PERSON)

TOPICS:

- Parallel Computer Architecture.

REQUIRED READING:

Book Chapters: Parallel Programming for Multicore and Cluster Systems: Chapter 2 (See Bibliography)

SESSION 17 (LIVE IN-PERSON)

TOPICS:

- distributed-memory systems

REQUIRED READING:

Book Chapters: Parallel Programming for Multicore and Cluster Systems: Chapter 2 (See Bibliography)

SESSION 18 (LIVE IN-PERSON)

TOPICS:

- Interconnection networks and routing.

REQUIRED READING:

Book Chapters: Parallel Programming for Multicore and Cluster Systems: Chapter 2 (See Bibliography)

SESSION 19 (LIVE IN-PERSON)

TOPICS:

- Programming shared-address space.

REQUIRED READING:

Book Chapters: Parallel Programming for Multicore and Cluster Systems: Chapter 6 (See Bibliography)

SESSION 20 (LIVE IN-PERSON)

TOPICS:

- Programming exercise - Shared address space using Python
- Detailed Labs PDF
- Practice on own computer / Raspberry PIs

SESSION 21 (LIVE IN-PERSON)

TOPICS:

- Programming scalable systems.

REQUIRED READING:

Book Chapters: Parallel Programming for Multicore and Cluster Systems: Chapter 5 (See Bibliography)

SESSION 22 (LIVE IN-PERSON)

TOPICS:

- Big Data.
- Hadoop
- NoSQL
- Cloud

SESSION 23 (LIVE IN-PERSON)

TOPICS:

- Parallel Computing Group Project with Python (1).

Forum

- Discussion on the project and doubts

SESSION 24 (LIVE IN-PERSON)

TOPICS:

- GPU Programming
- NVDIA: Pascal, Volta.

REQUIRED READING

- Textbook: CUDA Examples and Programming.

SESSION 25 (LIVE IN-PERSON)

TOPICS:

- Parallel Computing Group Project with Python (2).

Forum

- Discussion on the project and doubts

SESSION 26 (LIVE IN-PERSON)

TOPICS:

- Map Reduce
- YARN

SESSION 27 (LIVE IN-PERSON)

TOPICS:

- Parallel Computing Group Project with Python (3).
- Implementation on AWS, GCP or Oracle Cloud.

SESSION 28 (LIVE IN-PERSON)

TOPICS:

- Group Project Presentation (Parallel Computing on the Cloud)

SESSION 29 (LIVE IN-PERSON)

FORUM

Access to discuss all doubts.

REVIEW ALL SESSIONS

Kahoot in class simulating exam questions. Doubts and discussions for the exam.

SESSION 30 (LIVE IN-PERSON)

FINAL EXAM

EVALUATION CRITERIA

Criteria Percentage Comments

- Class Participation 10 % Class Participation
- Quizzes 20 % Quizzes- tests
- Group Presentation 20 % Group Project + Presentation
- Partial Exam 20 % Partial Exam on Operating Systems
- Final Exam 30 % Final Exam on Parallel Computing

criteria	percentage	Learning Objectives	Comments
Class Participation	10 %		
Individual Work	20 %		
Group Presentation	20 %		
Intermediate Tests	20 %		
Final Exam	30 %		

RE-SIT / RE-TAKE POLICY

A. Class participation and discussion

Class participation will be evaluated based on the following criteria:

Quality (not quantity) of your participation in class discussion. Discussion among peers is encouraged and questions can arise from the discussions which will be addressed as long as they are on the current topic handled.

B. Individual Work

Throughout the course the professor will indicate exercises/quizzes which should be delivered in a certain timeframe. The timeframe is strict and any submission not in the correct time and format will result in a 0.

C. Group project

The group project is an integral part of this course. Each group (randomly composed of 5-6 students) will be asked to choose an advanced topic in Parallel Computing in the cloud using modern technologies, to write a paper on that topic and to prepare a presentation. These presentations will be delivered in session 28.

D. Exams

There will be one partial exam on Operating Systems Section and one final exam on Operating Systems and Parallel Computing Sections. For these exams, you must bring your own computer. NO QUESTIONS ARE ALLOWED DURING THE EXAMS.

BIBLIOGRAPHY

Compulsory

- Andrew Tanenbaum. *Modern Operating Systems.* 5th. Pearson. ISBN 9781292459660 (Printed)

- Thomas Rauber, Gudula Rünger. (2023). *Parallel Programming for Multicore and Cluster Systems.* 3rd edition. Springer. ISBN 9783031289231 (Printed)

BEHAVIOR RULES

Please, check the University's Code of Conduct here. The Program Director may provide further indications.

ATTENDANCE POLICY

Please, check the University's Attendance Policy here. The Program Director may provide further indications.

ETHICAL POLICY

Please, check the University's Ethics Code here. The Program Director may provide further indications.