Course guide
270649 - SA - Supercomputers Architecture

Last modified: 29/07/2025

Unit in charge: Barcelona School of Informatics
Teaching unit: 701 - DAC - Department of Computer Architecture.

Degree: MASTER'S DEGREE IN INNOVATION AND RESEARCH IN INFORMATICS (Syllabus 2012). (Optional subject).

Academic year: 2025 **ECTS Credits:** 6.0 **Languages:** English

LECTURER

Coordinating lecturer: JORDI TORRES VIÑALS

Others: Primer quadrimestre:
JORDI TORRES VIÑALS - 10

PRIOR SKILLS

Programming in C and Linux basics will be expected in the course. In addition, prior exposure to parallel programming constructions, Python language, experience with linear algebra/matrices, or machine learning knowledge will be helpful.

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
CEE4.1. Capability to analyze, evaluate and design computers and to propose new techniques for improvement in its architecture.
CEE4.2. Capability to analyze, evaluate, design and optimize software considering the architecture and to propose new optimization techniques.
CEE4.3. Capability to analyze, evaluate, design and manage system software in supercomputing environments.

Generical:
CG1. Capability to apply the scientific method to study and analyse of phenomena and systems in any area of Computer Science, and in the conception, design and implementation of innovative and original solutions.

Transversal:
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.

Basic:
CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.
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.
CB9. Possession of the learning skills that enable the students to continue studying in a way that will be mainly self-directed or autonomous.

TEACHING METHODOLOGY

Class attendance and participation: Regular attendance is expected, and is required to be able to discuss concepts that will be covered during class.

Lab activities: Some exercises will be conducted as hands-on sessions during the course using supercomputing facilities. The student's own laptop will be required to access these resources during the theory class. Each hands-on session will involve writing a lab report with all the results. There are no days for theory classes and days for laboratory classes. Theoretical and practical activities will be interspersed during the same session to facilitate the learning process.

Reading/presentation assignments: Some exercise assignments will consist of reading documentation/papers that expand the concepts introduced during lectures. Some exercises will involve student presentations (randomly chosen).

Assessment: There will be one midterm exam in the middle of the course. The student is allowed to use any type of documentation (also digital via the student's laptop)

LEARNING OBJECTIVES OF THE SUBJECT

1.To train students to follow by themselves the continuous development of supercomputing systems that enable the convergence of advanced analytic algorithms as artificial intelligence.

STUDY LOAD

Type	Hours	Percentage
Hours large group	24,0	16.00
Hours small group	24,0	16.00
Self study	96,0	64.00
Guided activities	6,0	4.00

Total learning time: 150 h

CONTENTS

00. Welcome: Course content and motivation

01. Supercomputing basics

02. Heterogeneous supercomputers

03. Supercomputer management and storage systems

04. Benchmarking supercomputers

05. Data center infrastructures

06. Parallel programming models

07. Parallel performance models

08. Parallel programming languages for heterogeneous platforms

09. Artificial Intelligence is a computing problem

10. Deep Learning essential concepts

11. Using Supercomputers for DL training

12. Accelerate the learning with parallel training on multi-GPUs

13. Accelerate the learning with distributed training on multiple parallel servers

14. How to speed up the training of Transformers-based models

ACTIVITIES

00. Welcome

Specific objectives:
1

Related competencies :
CB9. Possession of the learning skills that enable the students to continue studying in a way that will be mainly self-directed or autonomous.
CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.
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.
CEE4.1. Capability to analyze, evaluate and design computers and to propose new techniques for improvement in its architecture.
CEE4.3. Capability to analyze, evaluate, design and manage system software in supercomputing environments.
CEE4.2. Capability to analyze, evaluate, design and optimize software considering the architecture and to propose new optimization techniques.
CG1. Capability to apply the scientific method to study and analyse of phenomena and systems in any area of Computer Science, and in the conception, design and implementation of innovative and original solutions.
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.

Full-or-part-time: 0h 30m
Theory classes: 0h 30m

01. Supercomputing basics

Full-or-part-time: 4h 30m
Self study: 3h 30m
Theory classes: 1h

Exercise 01: Supercomputing impact

Full-or-part-time: 3h
Self study: 2h
Laboratory classes: 1h

02. Heterogeneous supercomputers

Full-or-part-time: 5h
Self study: 4h
Theory classes: 1h

Exercise 02: Getting started with storage and management systems

Full-or-part-time: 3h
Self study: 2h
Laboratory classes: 1h

03. Supercomputer management and storage systems

Full-or-part-time: 4h
Self study: 2h
Theory classes: 2h

Exercise 03: Exascale computers challenge

Full-or-part-time: 4h
Self study: 2h
Laboratory classes: 2h

04. Benchmarking supercomputers

Full-or-part-time: 6h
Self study: 4h
Theory classes: 2h

Exercise 04: Getting started with parallel programming models

Full-or-part-time: 5h
Self study: 4h
Laboratory classes: 1h

05. Data centers infrastructures

Full-or-part-time: 5h
Self study: 4h
Theory classes: 1h

Exercise 05: Getting started with parallel performance metrics

Full-or-part-time: 4h
Self study: 3h
Laboratory classes: 1h

06. Parallel programming models

Full-or-part-time: 9h
Self study: 3h
Theory classes: 6h

Exercise 06: Getting started with parallel performance models

Full-or-part-time: 4h
Self study: 3h
Laboratory classes: 1h

07. Parallel performance models

Full-or-part-time: 3h
Self study: 2h
Theory classes: 1h

Exercise 07: Emerging trends in supercomputing

Full-or-part-time: 6h
Self study: 5h
Laboratory classes: 1h

08. Parallel programming languages for heterogeneous platforms

Full-or-part-time: 2h
Self study: 1h
Theory classes: 1h

Exercise 08: Getting started with CUDA

Full-or-part-time: 6h 30m
Self study: 3h
Theory classes: 0h 30m
Laboratory classes: 3h

Midterm

Full-or-part-time: 12h 30m
Self study: 10h 30m
Theory classes: 2h

09. Artificial Intelligence is a Supercomputing problem

Full-or-part-time: 5h
Self study: 3h
Theory classes: 2h

Exercise 09: First contact with Deep Learning and Supercomputing

Full-or-part-time: 6h
Self study: 4h
Laboratory classes: 2h

10. Deep Learning essential concepts

Full-or-part-time: 2h
Self study: 1h
Theory classes: 1h

Exercise 10: The new edition of the TOP500

Full-or-part-time: 5h
Self study: 4h
Laboratory classes: 1h

11. Using Supercomputers for DL training

Full-or-part-time: 3h 30m
Self study: 2h
Theory classes: 1h 30m

Exercise 11: Using a supercomputer for Deep Learning training

Full-or-part-time: 7h
Self study: 4h
Laboratory classes: 3h

12. Accelerate the learning with parallel training using a multi-GPU parallel server

Full-or-part-time: 4h
Self study: 3h
Theory classes: 1h

Exercise 12: Accelerate the learning with parallel training using a multi-GPU parallel server

Full-or-part-time: 7h
Self study: 4h
Laboratory classes: 3h

13. Accelerate the learning with distributed training using multiple parallel servers

Full-or-part-time: 2h
Self study: 1h
Theory classes: 1h

Exercise 13: Accelerate the learning with distributed training using multiple parallel server

Full-or-part-time: 11h
Self study: 8h
Laboratory classes: 3h

14. How to speed up the training of Transformers-based models

Full-or-part-time: 1h
Theory classes: 1h

Exercise 14: How to speed up the training of Transformers-based models

Full-or-part-time: 7h
Self study: 4h
Laboratory classes: 3h

Final remarks

Full-or-part-time: 2h 30m
Self study: 2h
Theory classes: 0h 30m

GRADING SYSTEM

The evaluation of this course can be obtained by continuous assessment. This assessment will take into account the following:

20% Attendance + participation
10% Midterm exam
70% Exercises (+ exercise presentations) and Lab exercises (+ Lab reports)

Students who have not benefited from continuous assessment have the opportunity to take a final Course Exam. This exam includes evaluating the knowledge of the entire course (practical part, theoretical part, and self-learning part). During this course exam, the student is not allowed to use any documentation (neither on paper nor digital).

BIBLIOGRAPHY

Basic:
- Torres, Jordi. Supercomputing for Artificial Intelligence: Foundations, Architectures, and Scaling Deep Learning. Barcelona: WATCH THIS SPACE Book Series - Barcelona. Amazon KDP, 2025. ISBN 9798319328359.