

Sistemes Paral·lels i Distribuïts de Còmput**2012/2013**

Code: 42242

ECTS Credits: 6

Degree	Syllabus	Type	Year	Semester
4313133 Còmput d'Altes Prestacions, Teoria de la Informació i Seguretat / High Performance Computing, Information Theory and Security	1094 Còmput d'Altes Prestacions, Teoria de la Informació i Seguretat / High Performance Computing, Information Theory and Security	O	1	0

Contact

Name: Elisa Ruth Heymann Pignolo

Email: Elisa.Heymann@uab.cat

Use of languages

Principal working language: anglès (eng)

Prerequisites

There are no pre-requisites.

Objectives and Contextualisation

At the end of this subject, students should have enough knowledge, methods and technical skills to research on innovative solutions to distributed systems problems.

Skills

- Analyse and evaluate parallel and distributed computing architectures, as well as develop and optimise advanced software for them.
- Apply the functions and operations of Internet, new generation network technologies and protocols, component models, intermediate software and services to systems design.
- Apply the methodology of research, techniques and specific resources for investigating and producing innovative results in a certain specialised field.
- Possess and comprehend knowledge that offers the basis and opportunity to be original in the development and/or application of ideas, frequently in a research context.
- Research innovative solutions, or more efficient solutions than those used currently, to problems regarding operating systems, servers, applications and distributed computing based systems.
- Show responsibility in the handling of information and knowledge, and in the management of multidisciplinary groups and/or projects.
- Students must possess learning abilities to enable them to continue studying in a way that will to a large extent have to be self-managed and autonomous.

Learning outcomes

1. Apply a wide range of design techniques to middleware and development tools for tuning applications to their environment.
2. Apply the knowledge acquired to the design of distributed storage systems, in order to design data and computation intensive applications
3. Apply the methodology of specific research, techniques and resources for investigating and producing innovative results in a certain specialised field
4. Possess and comprehend knowledge that offers the basis and opportunity to be original in the

- development and/or application of ideas, frequently in a research context
5. Select both the most adequate distributed system and language for the generation of the proposed solution to a distributed computing problem
 6. Show responsibility in the handling of information and knowledge, and in the management of multidisciplinary groups and/or projects
 7. Students must possess learning abilities to enable them to continue studying in a way that will to a large extent have to be self-managed and autonomous
 8. Understand parallel computing environments, as well as their implications related to performance and costs.

Content

- **Distributed Computing Platforms (12 hours)**

- Parallel and Distributed Architectures
- Kernel core
- Job Scheduling
- Distributed File Systems
- Cluster computing

- **Distributed Application Architectures (9 hours)**

- Programming in Distributed Systems: Concepts, basic structures and programming paradigms.
- Grid / Cloud.
- Case study with Java RMI.
- Security in Distributed Systems

- **Massive Data Handling. Data Bases, from relational DBMS to noSQL recent technologies (9 hours)**

- DBMS structure
- Parallel databases and architectures
- Parallel query processing
- Parallel SQL dataflow
- Parallel DB Search
- Data partition strategies
- Scaling a DB system: partitioning, sharding, benchmarking, profiling
- noSQL: non-relational distributed databases. MapReduce, BigTable, Hadoop, HBASE, Pig

Methodology

The methodology will combine classroom work, problem solving in class, work in the computing lab, performing works from recommended readings and independent study student. It will use the virtual platform and asked for papers related to the thematic blocks.

Distribution of the tasks:

The Works will be done in groups, 20%

Evaluation of work done and presented by the student, 30%

Evaluation of activities in sessions tutored, 20%

Theoretical and practical tests individual, 30%

Activities

Title	Hours	ECTS	Learning outcomes
Type: Directed			
Lab exercises	4	0.16	2, 3, 4, 6, 7, 8
Lectures	26	1.04	1, 2, 3, 4, 5, 6, 7, 8
Type: Supervised			
Discussions	10	0.4	1, 2, 3, 4, 5, 6, 7, 8
Type: Autonomous			
Autonomous Work	98	3.92	1, 2, 3, 4, 5, 6, 7, 8

Evaluation

Final evaluation will come out from the combination of: (1) work developed on the areas in the module, (2) attendance to lectures and participation in class and labs, and (3) a final exam.

Evaluation activities

Title	Weighting	Hours	ECTS	Learning outcomes
Additional Work	40%	4	0.16	1, 2, 3, 4, 5, 6, 7, 8
Discussions	20%	6	0.24	1, 2, 3, 4, 5, 6, 7, 8
Presentation	40%	2	0.08	4, 5, 7, 8

Bibliography

BOOKS

- Andrew S. Tanenbaum, "Computer Networks", 3ª Edición P.H. 1996.
- Grama, A. Gupta, G. Karypis, and V. Kumar, "Introduction to Parallel Computing, 2nd Ed. Addison-Wesley, 2003.
- Rajkumar Buyya, "High Performance Cluster Computing: Programming and Applications", PH, 1999.
- G. Coulouris, J. Dollimore and T. Kinderg, "Sistemas Distribuidos: Conceptos y Diseño", Addison-Wesley, 3ª Ed. 2001.
- Bell, Charles; Kindahl, Mats; Thalmann, Lars. "MySQL High Availability". O'Reilly, 2010.
- Chang, Fay, et al. "Bigtable: A Distributed Storage System for Structured Data." OSDI, 2006

- Dewitt, David, and Jim Gray. "Parallel Database Systems: The Future of High Performance Database Processing." *Communications of the ACM* 35, no. 6 (1992): 85-98
- Schwartz, Baron; Zaitsev, Peter; Tkachenko, Vadim; Zawodny, Jeremy D.; Lentz, Arjen; Balling, Derek J. "High Performance MySQL", O'Reilly, 2008.
- Taniar, David; Leung, Clement H.C.; Rahayu, Wenny; Goel, Sushant. "High Performance Parallel Processing and Grid Databases". Wiley, 2008.
- White, Tom. "Hadoop, the definitive Guide", O'Reilly, 2011.
- Ian Foster, Carl Kesselman. *The grid: blueprint for a new computing infrastructure*. Morgan-Kaufmann 2004.
- Mark Dowd, John McDonald, Justin Schuh. *The Art of Software Security Assessment*. Addison-Wesley 2007.
- Rickard Oberg. *Mastering RMI: Developing Enterprise Applications in Java and EJB*. John Wiley & Sons. 2001.

Websites recommended:

<https://cv.uab.cat/>