UNIVERSIDADE DA CORUÑA

| Teaching Guide | | | | |
|---|---|---|---|---|
| Identifying Data | | | | 2021/22 |
| Subject (*) | HPC on the Cloud | | Code | 614973106 |
| Study programme | Mestrado Universitario en Computación de Altas Prestacións / High Performance Computing (Mod. Virtual) | | | |
| Descriptors | | | | |
| Cycle | Period | Year | Type | Credits |
| Official Master's Degree | 1st four-month period | First | Optional | 6 |
| Language | SpanishGalicianEnglish | | | |
| Teaching method | Non-attendance | | | |
| Prerequisites | | | | |
| Department | Departamento profesorado másterEnxeñaría de Computadores | | | |
| Coordinador | Pardo Martínez, Xoán Carlos | E-mail | xoan.pardo@udc.es | |
| Lecturers | Pardo Martínez, Xoán Carlos | E-mail | xoan.pardo@udc.es | |
| Web | aula.cesga.es/courses/MASTERHPC7 | | | |
| General description | For several years, the use of parallel computing architectures was a fundamental aspect that allowed the development of important areas in multiple fields of basic and applied science. However, the high cost of traditional parallel systems limited its use practically to large industries and research centers. The use of low-cost computer networks, as well as computing using connected infrastructures through the Internet, has been a practical and cheap alternative to large systems for some time. Thus, Cloud computing has emerged as a paradigm of distributed computing that changes the way we use computers, allowing a transparent, safe and cheap access to huge computational resources from anywhere in the world.<br><br>The main objective of this subject is to introduce the Cloud Computing model, and how the field of High Performance Computing can use the cloud to deal with problems that, until now, were restricted to be solved in large supercomputers. You will see different examples of how it is possible to solve problems in the field of High Performance Computing using distributed services and resources accessible in the cloud. | | | |
| Contingency plan | Not applicable because it is a distance learning subject | | | |

| Study programme competences / results | |
|---|---|
| Code | Study programme competences / results |
| A1 | CE1 - Define, evaluate and select the most appropriate architecture and software to solve a problem |
| A6 | CE6 - Know the available tools for the distributed systems computing |
| B2 | CB7 - The students have to know how to apply the acquired knowledge and their capacity to solve problems in new or hardly explored environment inside wider contexts (or multidiscipinary) related to its area of development |
| B5 | CB10 - The students have to possess learning skills that allows them to continue to study in a mainly self-driven or autonomous manner |
| B6 | CG1 - Be able to search and select useful information to solve complex problems, using the bibliographic sources of the field |
| C1 | CT1 - Use the basic technologies of the information and computing technology field required for the professional development and the long-life learning |

| Learning outcomes | | | |
|---|---|---|---|
| Learning outcomes | Study programme competences / results | | |
| The student will know the basics of cloud computing and service virtualization. | AJ6 | | |
| The student will know and learn to use the basic services provided by the main Cloud public providers. | AJ1<br>AJ6 | | CJ1 |
| The student will know and know how to apply the main paradigms of distributed programming used in Cloud computing. | AJ1<br>AJ6 | BJ2 | CJ1 |
| The student will know and learn to use the services and resources available in the cloud to prepare and execute applications in the field of high performance computing. | AJ6 | | CJ1 |

| | | | |
|---|---|---|---|
| The student will acquire the necessary skills for the search, selection and management of resources (bibliography, software, etc.) related to Cloud computing in the field of high performance computing. | | BJ5<br>BJ6 | |

| Contents | |
|---|---|
| Topic | Sub-topic |
| Introduction to Cloud Computing | |
| Cloud Computing services: virtual clusters | |
| Distributed processing models and frameworks | |
| Services for distributed processing in the cloud | |

| Planning | | | | |
|---|---|---|---|---|
| Methodologies / tests | Competencies / Results | Teaching hours (in-person & virtual) | Student?s personal work hours | Total hours |
| Workbook | A1 A6 | 0 | 24 | 24 |
| ICT practicals | A1 A6 B2 B5 B6 C1 | 4 | 71 | 75 |
| Supervised projects | B2 B5 B6 | 0 | 40 | 40 |
| Objective test | A1 A6 B2 B6 | 2 | 0 | 2 |
| Personalized attention | | 9 | 0 | 9 |
| (*)The information in the planning table is for guidance only and does not take into account the heterogeneity of the students. | | | | |

| Methodologies | |
|---|---|
| Methodologies | Description |
| Workbook | Instruction programmed through teaching materials, specially designed for autonomous and asynchronous learning, with an important weight of the references to the documentary sources used in the different contents. |
| ICT practicals | The students will resolve autonomously diverse problems which allow them to practice the theoretical topics of the subject. |
| Supervised projects | The subject of an individual assignment will be agreed with the teacher and the student will elaborate it more deeply in an autonomous way. |
| Objective test | At the end of the term there will be an exam on the contents of the subject. In this exam the theoretical and practical topics will be evaluated. |

| Personalized attention | |
|---|---|
| Methodologies | Description |
| Supervised projects<br>ICT practicals | The personalized attention during the ICT practices will serve to guide and check the students' work following the indications they were given.<br><br>To carry out the supervised assignments, students will be given the necessary initial indications and bibliographic references for consultation. During the elaboration, their progress will be monitored to offer additional guidelines to ensure the quality of the result according to predefined criteria.<br><br>Every teacher will provide a tutorial schedule to resolve students' questions related to the topics of the subject. Students will be encouraged to take advantage of the tutorial sessions as a fundamental part of their learning process.<br><br>To make personal attention easier an intensive use of online communication tools will be encouraged: videoconference, e-mail, chat, etc. |

| Assessment | | | |
|---|---|---|---|
| Methodologies | Competencies / Results | Description | **Qualification** |

| | | | |
|---|---|---|---|
| Objective test | A1 A6 B2 B6 | The test may contain multiple-choice questions, short answers or problems related to the contents covered in the subject | 40 |
| Supervised projects | B2 B5 B6 | The supervised projects will be about some topic agreed between the student and the teacher. It will be evaluated the compliance with specifications, originality, personal contribution, methodology, rigour and presentation of the results. | 20 |
| ICT practicals | A1 A6 B2 B5 B6 C1 | It will be evaluated the degree of compliance with the specifications, methodology, rigour and presentation of the results. | 40 |

| Assessment comments |
|---|
| In order to pass the subject, the student has to get a total score of 5 or higher.<br>Students that fail the subject can keep the marks of the labs and the supervised project in which they scored 5 or higher for the following year.<br>Second opportunity (July) and extraordinary<br>The evaluation will be the same as in the first opportunity. Students will have a second deadline before the final exam to submit failed practical assignments.<br>Condition to be considered "Absent"<br>Not handing in any assignments and not taking the exam.<br>Fraud<br>The fraud regulation of the UDC will be applied in case fraud was detected in any assignment or in the exam. |

| Sources of information | |
|---|---|
| Basic | - Erl T., Puttini R. and Mahmood Z. Cloud Computing, Concepts, Technology & Architecture (2013). Ed. Prentice-Hall.- White, T. Hadoop: The Definitive Guide, Storage and Analysis at Internet Scale, 4ª edición (2015). O'Reilly Media.- B. Chambers, M. Zaharia, "Spark: The Definitive Guide", O'Reilly, 2018 |
| Complementary | - Foster, I. and Gannon, D.B. Cloud Computing for Science and Engineering (2017). The MIT Press.<br />- Zaharia, M., Karau, H., Konwinski, A. y Patrick Wendell. Learning Spark: Lightning-Fast Big Data Analysis (2015), O'Reilly Media.<br />- Karau, H., Warren, R,. High Performance Spark: Best Practices for Scaling and Optimizing Apache Spark, (2017). O'Reilly Media.- Foster, I. and Gannon, D.B. Cloud Computing for Science and Engineering (2017). The MIT Press.- Zaharia, M., Karau, H., Konwinski, A. y Patrick Wendell. Learning Spark: Lightning-Fast Big Data Analysis (2015), O'Reilly Media.- Karau, H., Warren, R,. High Performance Spark: Best Practices for Scaling and Optimizing Apache Spark, (2017). O'Reilly Media. |

| Recommendations |
|---|
| Subjects that it is recommended to have taken before |
| Parallel Programming/614473102 |
| Subjects that are recommended to be taken simultaneously |
| High Performance Infrastructures/614473104 |
| Subjects that continue the syllabus |
| Data Analytics with HPC/614473108 |
| Other comments |
| Considering the strong interrelation between the theoretical and practical contents of the subject and the progressive introduction of new concepts closely related to each other, it is advisable a weekly review to make the most of the subject. |