[ONLINE] PATC: Managing distributed data with Hecuba and dataClay

Objectives

The objective of this course is to give an overview of BSC storage solutions, Hecuba and dataClay. These two platforms allow to easily store and manipulate distributed data from object-oriented applications, enabling programmers to handle object persistence using the same classes they use in their programs, thus avoiding time consuming transformations between persistent and non-persistent data models. Also, Hecuba and dataClay enable programmers to transparently manage distributed data, without worrying about its location. This is achieved by adding a minimal set of annotations in the classes.

Both Hecuba and dataClay can work independently or integrated with the COMPSs programming model and runtime to facilitate parallelization of applications that handle persistent data, thus providing a comprehensive mechanism that enables the efficient usage of persistent storage solutions from distributed programming environments.

Both platforms offer a common interface to the application developer that facilitates using one solution or the other depending on the needs, without changing the application code. Also, both of them have additional features that allow the programmer to take advantage of their particularities.

Requirements

Basic programming skills in Python and Java.

Previous attendance to PATC course on programming distributed systems with COMPSs is recommended.

Please download and carefully read the following **instructions** regarding the logistics participants enrolling online PATC at BSC are expected to follow.

Learning Outcomes

In the course, the Hecuba and dataClay syntax, programming methodology and an overview of their internals will be given. Also, an overview of COMPSs at user level will be provided in order to take advantage of the distribution of data with both platforms. The attendees will get a first lesson about programming with the common storage interface that will enable them to start programming with both frameworks.

A hands-on with simple introductory exercises will be also performed for each platform, with and without COMPSs to distribute the computation. The students who finish this course will be able to develop simple Hecuba and dataClay applications and to run them both in a local resource and in a distributed platform (initially in a private cloud)

Academic Staff

Image not found or type unknown

Course Conveners

Department and Research group: Computer Science - Workflows and Distributed Computing

Yolanda Becerra, Data-driven Scientific Computing research line, Senior researcher
Anna Queralt, Distributed Object Management research line, Senior researcher

Course Lecturers

Department and Research group: Computer Sciences - Workflows and Distributed Computing

Alex Barceló, Distributed object Management research line, Researcher
Yolanda Becerra, Data-driven Scientific Computing research line, Senior researcher
Daniel Gasull, Distributed object Management research line, Research engineer
Juanjo Costa, Computer Architecture Department at UPC
Anna Queralt, Distributed object Management research line, Senior researcher

Materials

Image not found or type unknown

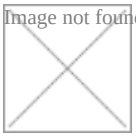
INTELLECTUAL PROPERTY RIGHTS NOTICE:

• The User may only download, make and retain a copy of the materials for his/her use for non?commercial and research purposes.

• The User may not commercially use the material, unless has been granted prior written consent by the Licensor to
do so; and cannot remove, obscure or modify copyright notices, text acknowledging or other means of identification or disclaimers as they appear.

• For further details, please contact BSC?CNS patc@bsc.es

Further information



NOTE: PATC courses do not charge fee

Barcelona Supercomputing Center - Centro Nacional de Supercomputación

---

Source URL (retrieved on *28 Mar 2024 - 11:55*): https://www.bsc.es/ca/education/training/patc-courses/online-patc-managing-distributed-data-hecuba-and-dataclay-1