

Curso

# Administrator & Programming Training for Apache Hadoop

**Lugar:** CITIC (Centro de Investigación en Tecnoloxías da Información e as Comunicaci3ns), A Coruña  
**Data e Horario:** 20-21 de Xaneiro de 2015, de 10 a 14h e de 15:30 a 18:30h

**Resumo:** Apache Hadoop 3 un framework para o almacenamento e procesado a gran escala de conxuntos de datos en clusters. O presente curso, de car3cter pr3ctico, prepara aos/3s asistentes naqueles coñecementos e capacidades necesarias para o uso de Apache Hadoop no desenvolvemento de aplicaci3ns para o procesado de datos.

Contidos

## \* Administration

- Introduction to Apache Hadoop
  - The Case for Apache Hadoop
  - HDFS
  - Getting Data into HDFS
  - MapReduce
- Planning, Installing, and Configuring a Hadoop Cluster
  - Planning Your Hadoop Cluster
  - Hadoop Installation and Initial Configuration
  - Hadoop Clients
- Cluster Operations and Maintenance
  - Managing and Scheduling Jobs
  - Cluster Maintenance

## \* Programming

- Basic Programming with the Hadoop Core API
  - Writing a MapReduce Program
  - Unit Testing MapReduce Programs
  - Delving Deeper into the Hadoop API
  - Practical Development Tips and Techniques
  - Data Input and Output
- Problem Solving with MapReduce
  - Common MapReduce Algorithms
  - Joining Data Sets in MapReduce Jobs

Organiza

Rede Galega de Tecnoloxías

Cloud e Big Data para



FINANCIADA POR:



PROGRAMA DE CONSOLIDACI3N E ESTRUCTURACI3N DE UNIDADES DE INVESTIGACI3N COMPETITIVAS. MODALIDADE REDES. N3 EXPTE. R2014/041



UNI3N EUROPEA  
"Una maneira de facer Europa"

Colabora:

