

Apache Spark

Lugar: CITIC (Centro de Investigación en Tecnoloxías da Información e as Comunicacions), A Coruña

Data e Horario: 24 de Marzo de 2015, de 10 a 14h e de 15:30 a 19:30h

Resumo: Apache Spark é un sistema de propósito xeral para computación en clusters de alto rendemento. Trátase dunha evolución do sistema de dúas etapas baseado en disco do paradigma MapReduce a primitivas de memoria para incrementar o rendemento de certas aplicacións. O presente curso, de carácter práctico, prepara aos/ás asistentes naqueles coñecementos e capacidades necesarias para o uso de Apache Spark no desenvolvemento de aplicacións para o procesado de datos así como para o despregue da propia solución.

Contidos

- INTRODUCCIÓN AO APACHE SPARK
- INSTALACIÓN
 - o Cluster standalone
 - o Cluster Hadoop
- ADMINISTRACIÓN
 - o Configuración básica
 - o Guía de boas prácticas
- GUÍA DE PROGRAMACIÓN DE APLICACIÓNS
- MÓDULOS INTEGRADOS
 - o Spark Streaming
 - o Spark SQL
 - o MLlib
 - o Graphx
 - o Bagel

Organiza

Rede Galega de Tecnoloxías

Cloud e Big Data para



FINANCIADA POR:



PROGRAMA DE CONSOLIDACIÓN E ESTRUCTURACIÓN DE UNIDADES DE INVESTIGACIÓN COMPETITIVAS. MODALIDADE REDES. Nº EXPTE. R2014/041



Colabora:

