

Instrucciones

Seminario

Apache Hadoop

Curso:

**Ciencia de Datos: Un Enfoque Práctico en la
Era del Big Data**

Ponente:

Sara Del Río García (srio@decsai.ugr.es)

1. Introducción

Este documento contiene las instrucciones que se pueden llevar a cabo **de forma previa** al seminario. Además, se especifican los requerimientos mínimos con los que se debe contar.

2. Requerimientos

- Es necesario contar con una máquina con **procesador de 64 bits** y mínimo **4 GB de RAM**.
- **Descargar el contenido del curso**. Para más información, consultar apartado 3.1 "Material para el desarrollo de las prácticas".
- **Tener VirtualBox instalado en la máquina e importar la imagen preparada** para el desarrollo de las prácticas. Para más información, consultar apartado 3.2 "Instalar y configurar VirtualBox".

3. Instrucciones

3.1 Material para el desarrollo de las prácticas

El material para el desarrollo del curso de Hadoop consta de 3 directorios, donde cada directorio contiene el material asociado a cada una de las partes del seminario:

- **Directorio "Parte 1"**: contiene el material asociado para la instalación de Hadoop en una máquina.
- **Directorio "Parte 2"**: contiene el material para la ejecución de un ejemplo sencillo sobre la instalación realizada previamente.
- **Directorio "Parte 3"**: contiene el material para la ejecución de un ejemplo más complejo sobre la instalación realizada previamente.

3.2 Instalar y configurar VirtualBox

Para instalar VirtualBox, descargar del enlace:

<https://www.virtualbox.org/wiki/Downloads>

el paquete adecuado e instalar siguiendo el asistente o lo que proceda dependiendo de la plataforma.

By downloading, you agree to the terms and conditions of the respective license.

- **VirtualBox platform packages**. The binaries are released under the terms of the GPL version 2.
 - **VirtualBox 5.0.16 for Windows hosts** ➞ x86/amd64
 - **VirtualBox 5.0.16 for OS X hosts** ➞ amd64
 - **VirtualBox 5.0.16 for Linux hosts**
 - **VirtualBox 5.0.16 for Solaris hosts** ➞ amd64

Ilustración 1: Download VirtualBox

Una vez instalado VirtualBox, importar la imagen para VirtualBox (archivo "Hadoop-CM.ova" dentro del directorio "Part 1"). Para ello, hacer clic sobre el menú "Archivo" y seguidamente hacer clic sobre "Importar servicio virtualizado...". A continuación se mostrará un asistente que permitirá seleccionar el archivo a importar (archivo "Hadoop-CM.ova"), hacer clic sobre "Next" y en la siguiente pantalla, sobre el botón "Importar".