



Taller Big Data:  
**Introducción a Hadoop 3**  
11 de Xuño

# Javier Cacheiro

Técnico Superior Sistemas – CESGA  
Responsable Área Big Data

 [jlopez@cesga.es](mailto:jlopez@cesga.es)

 [@javicacheiro](https://twitter.com/javicacheiro)

 <http://github.com/javicacheiro>

 <https://es.linkedin.com/in/javier-lopez-cacheiro-9485737>

# ¿Qué aprenderé durante el taller?

- Cómo conectarte a la plataforma Hadoop 3
- Cómo transferir datos de forma eficiente
- Las herramientas que están disponibles
- Cómo lanzar estas herramientas

# ¿Qué **NO** se enseñará durante el taller?

Dado el carácter introductorio del taller y la gran variedad de herramientas disponibles, en este taller **no se enseñará a utilizar cada una de las herramientas**, sino que simplemente se mostrará como se puede acceder a ellas y cómo se lanzan.

**Posteriormente se realizarán talleres específicos** que se centrarán en herramientas concretas como Spark donde sí se enseñará su uso en detalle.

# Contenido

## 1. Introducción al servicio Big Data

- 1.1. Conceptos básicos
- 1.2. Descripción del hardware
- 1.3. Descripción del software
- 1.4. Portal BD|CESGA

## 2. Conexión al servicio Hadoop 3

- 2.1. VPN
- 2.2. Acceso por línea de comandos: SSH
- 2.3. Acceso a través del interfaz web: WebUI > HUE
- 2.4. Acceso por escritorio remoto

# Contenido

## 3. Transferencia de datos

- 3.1. Cómo transferir datos de forma eficiente usando el servicio DTN
- 3.2. Cómo transferir datos usando SCP
- 3.3. Cuotas en los sistemas de ficheros
- 3.4. Migración de datos de la antigua plataforma

## 4. Elementos básicos

- 4.1. HDFS: Almacenamiento distribuido
- 4.2. YARN: Ejecución y monitorización de trabajos

# Contenido

## 5. Herramientas disponibles

5.1. Spark

5.2. Jupyter

5.3. Hive

5.4. Impala

5.5. Sqoop

5.6. Modules

## 6. Dónde obtener información adicional

6.1. Tutoriales

6.2. Guía de usuario

6.3. Documentación oficial

# Logística

Duración del taller: 3 horas

Descansos: 10:50-11:00 / 11:50-12:00

Control de asistencia:

- Recordar firmar al terminar

Certificados de asistencia:

- Se entregarán al final

Encuesta de satisfacción:

- Os llegará por email



# Transparencias

Transparencias que utilizaremos:

<http://bigdata.cesga.gal/workshop>

Al final del taller os facilitaremos además un **pack de documentación** que os sirva de referencia



# Estamos aquí para ayudar

[bigdata@cesga.gal](mailto:bigdata@cesga.gal)

981 569810