Curso Big Data y Supercomputación. Transformando datos en conocimiento. (2da edición)

Thursday 08 de June de 2017, 09:00 horas.

[1]

El curso "Big Data y Supercomputación: Transformando datos en conocimiento", tendrá lugar del 8 al 14 de junio, en horario de tarde, en el Centro Universitario de Mérida [2]. Este curso tiene una duración de 23 horas, y se encuentra dirigido principalmente a egresados universitarios en Ingeniería Informática e Ingeniería de Telecomunicación, estudiantes de 3º y 4º de las mismas ingenierías y estudiantes de doctorado o máster de titulaciones afines.

El principal objetivo es que los asistentes puedan comprobar qué aporta la supercomputación cuando se necesita procesar conjuntos de datos con características que los convierten en datos de tipo big data (volumen, tipos de datos, etc.) para obtener conocimiento. Para ello, el curso se centrará, por una parte, en el flujo de datos de procesamiento de big data mediante distintos servicios del framework Hadoop (como Flume, HDFS o Spark), ya desplegados en un clúster. Además, cobrará especial importancia el desarrollo de software para paralelizar el procesamiento de esos datos, para así obtener conocimiento después de su ejecución en el supercomputador LUSITANIA. Por otra parte, también se abordará la administración de los servicios utilizados.

Se expedirán certificados de asistencia y aprovechamiento. El acceso será libre, aunque limitado a un máximo de 25 asistentes, por lo que se requiere la realización de una inscripción on-line previa [3].

[3]
Horario:

- Mañanas de 9 a 14 horas.

A continuación se muestra el programa completo [4] del curso:

Bloque 1: Presentación (1 hora)

- Presentación de LUSITANIA I y II.
- Entorno de trabajo.
- Proyectos de la Fundación.
- Introducción a big data y ejemplos para mostrar la problemática del big data.
- Retos al procesar datos big data.

Bloque 2: Supercomputación y framework de Hadoop (11 horas)

- MV de Cloudera a través de OpenNebula.
- Hadoop: Justificación y características.
- HDFS: Características, arquitectura, rendimiento, configuración, lectura y escritura.
- Práctica de Hadoop: clúster HDFS y uso y administración del mismo.
- Módulos básicos del ecosistema de Apache para Hadoop.
- Casos prácticos de recolección de datos para su almacenamiento en el clúster.
- Práctica de consultas a datos del clúster HDFS.
- Procesamiento de datos en tiempo real.
- Spark: motivación y servicios.
- Transformaciones y Acciones en Spark.
- Planificador de DAGs (Directed Acyclic Graph).
- Administración del HistoryServer y servicios de Spark.

Bloque 3: Data Science (8 horas)

- Metodología y buenas prácticas para el procesamiento de datos científicos.
- Herramientas para la exploración de datos en R.
- Detección de patrones en conjuntos de datos científicos en R.
- Práctica con Spark y estrategias de definición de transformaciones.
- Machine learning: conceptos básicos.
- Práctica: ML en Spark con MLib (Machine Learning Library) y bootstraping.

Bloque 4: Casos de éxito con Big Data (3 horas)

- CénitS: ¿Pig Data, cómo que Pig Data?, ¿pero no era Big Pata?
- CénitS: Proyecto ConSumar.
- Canalyticals: Optimización de inversión de Capex para implantar 4G en función de patrones de movilidad de los clientes de telefonía de la ciudad.

Attachment	Size
Programa del curso Big Data y Supercomputación (2ª edición) [4]	589.55 KB

Cursos [5]

Source URL:https://www.cenits.es/en/agenda/09052017-curso-big-data-supercomputacion-transformando-datos-conocimiento-2da-edicion

Links
[1] https://www.cenits.es/sites/cenits.es/files/eventos/bigdata2017_cartel_a4_rgb.jpg [2] http://www.unex.es/conoce-la-uex/centros/cum [3] http://www.cenits.es/formularios/Curso-Big-Data-2da-edicion [4] https://www.cenits.es/sites/cenits.es/files/eventos/programa_curso_big_data_-_2a_edicion.pdf [5] https://www.cenits.es/en/categoria/cursos