

Programa Curso de Posgrado Big Data (Procesamiento para Grandes Datos)

Contenidos:

Unidad 1: Conceptos básicos de paralelismo Procesamiento paralelo. Arquitecturas paralelas. Servidores, Clusters y Cloud. Modelos de programación. Métricas. Herramientas.

Unidad 2: Introducción a Big data. Relación con Paralelismo. Fundamentos. Objetivos. Modelos de datos y modelos de procesamiento. Paradigma Map-Reduce. Apache Hadoop. ¿Por qué paralelismo sobre Big Data?

Unidad 3: Sistemas de almacenamiento para Big Data. Sistemas de archivos distribuidos. Clasificación. Apache HDFS. Bases de datos relacionales. Bases de datos NoSQL. Hive, Shark, MongoDB, Cassandra.

Unidad 4: Procesamiento paralelo para Big Data en la Nube. Las/os alumnas/os deberán desarrollar un proyecto que involucre construir una infraestructura en la nube y combinar los conocimientos adquiridos.

Unidad 5: IoT y Big Data. Obtención de millones de datos provenientes de redes de objetos inteligentes. Gestión y análisis de datos provenientes de redes de sensores.