
XXII Jornadas de Ingeniería del Software y Bases de Datos (JISBD 2017)

<http://fg.ull.es/sistedes2017/jisbd>

TUTORIAL

Tecnologías de bases de datos NoSQL

Responsable(s):

Diego Sevilla Ruiz, Facultad de Informática, Universidad de Murcia,
dsevilla@um.es



Duración:

3 horas.

Dirigido a:

Investigadores, doctorandos y estudiantes de posgrado en el ámbito de los sistemas distribuidos, de las bases de datos en general, y de las bases de datos NoSQL en particular.

Objetivos:

A la hora de implementar cualquier sistema de información actual, las tecnologías NoSQL aparecen como una alternativa más a considerar para el almacenamiento y recuperación de los datos. Su conocimiento, estudio de ventajas, inconvenientes, mecanismos de optimización, requisitos hardware y software de diseño y planificación son necesarios para la selección de la tecnología adecuada. Este tutorial introduce las distintas tecnologías NoSQL desde esta perspectiva, particularidades de modelado, idoneidad para ciertos problemas, opciones de optimización, nuevos lenguajes de consulta, etc. Se estudiarán sus diferentes tipos: clave/valor, documentales, columnares y de grafos. Durante su estudio aflorarán técnicas como *Map-Reduce* y procesamiento en *Streaming*. Finalmente, se estudiarán nuevas bases de datos basadas en el paradigma NewSQL.

Contenidos:

1. Introducción a las tecnologías NoSQL
 - a. Escalabilidad horizontal
 - b. Modelado NoSQL
 - c. Map-Reduce
 - d. Tipología de sistemas NoSQL

2. Bases de Datos Clave/Valor y Documentales
 - a. Datos semi-estructurados
 - b. JSON
 - c. Modelo de agregación
 - d. Map-Reduce y búsquedas

- e. Estudio de casos
 - f. MongoDB
 - g. CouchDB
3. Bases de datos Columnares
- a. El paradigma columnar
 - b. Familias de columnas
 - c. Organización de datos
 - d. Índices
 - e. Optimización de búsquedas
 - f. Estudio de casos
 - g. Hbase
 - h. Cassandra
4. Bases de datos de grafos
- a. Grafos
 - b. Algoritmos sobre grafos
 - c. Lenguajes de consulta
 - d. Estudio de casos
 - e. OrientDB
 - f. Neo4j
5. Tecnologías NewSQL
- a. Introducción a tecnologías SQL-on-Hadoop
 - b. SparkSQL
 - c. Hive
 - d. Pig

Material:

El material proporcionado a los asistentes vía web y utilizado en el tutorial consistirá en los siguientes elementos:

- Presentación utilizada en formato PDF.
- Máquina virtual VirtualBox con las bases de datos instaladas.
- Repositorio Github con los ejemplos.

Para la realización de la parte práctica del tutorial, los asistentes que lo deseen pueden traer un equipo portátil con el material anterior descargado (se usará VirtualBox).