

Big Data desde cero

Contexto, tecnologías y aplicación

PROVEEDORES PRODUCTOS BIG DATA

- Plataformas Apache Hadoop: Cloudera y Hortonworks
- Bases de Datos NoSQL
- Cloud: soluciones Hadoop, productos PaaS

PLATAFORMAS APACHE HADOOP: CLOUDERA Y HORTONWORKS

- Al instalar Hadoop se instala una Plataforma Big Data.
- Instalar “a mano” Hadoop es casi inviable → hay que recurrir a Plataformas Big Data.
- En Plataformas Hadoop líderes dos empresas:



PLATAFORMAS APACHE HADOOP: CLOUDERA Y HORTONWORKS

- Similar en propósito y funcionamiento - Diferentes en cuanto a cómo llevarlo a cabo.



- Tiene componentes propios.
- No soporta cualquier componente de Hadoop.
- Cloudera Manager como administrador
- (software propio).
- Obligatorio licencia para su uso empresarial
 - Incluye soporte



- Componentes 100% Apache Hadoop.
- Soporta cualquier componente de Hadoop.
- Ambari Server como administrador
- (software Open Source).
- No requiere licencia para su uso empresarial.
 - Se puede contratar soporte.

PLATAFORMAS APACHE HADOOP: CLOUDERA Y HORTONWORKS

- Ambas empresas se fusionaron

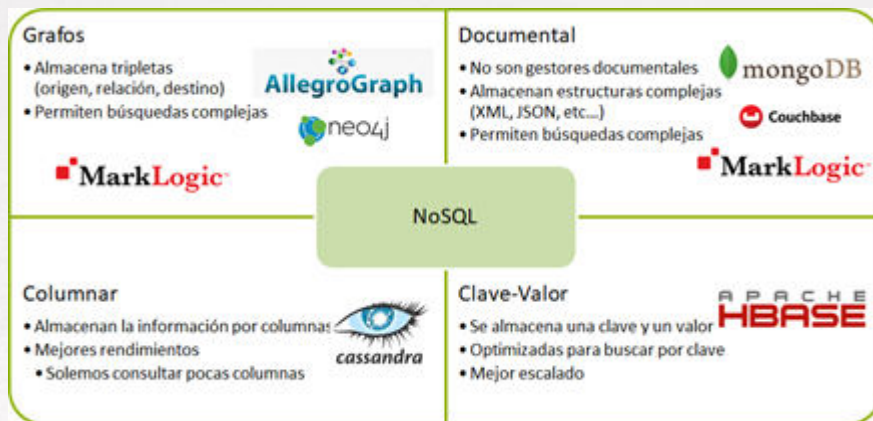


- El nombre de la “nueva” empresa es Cloudera

CLOUDERA

BASES DE DATOS NOSQL

- Bases de Datos NoSQL se engloban en las tecnologías Big Data
- En Big Data, hay muchas bases de datos NoSQL que no pertenecen al ecosistema Hadoop
- Se pueden agrupar según su funcionamiento, su estructura y su complejidad:



CLOUD: SOLUCIONES HADOOP, PRODUCTOS PAAS

- En Cloud se ofrecen Plataformas Hadoop propias.
- Además de Plataformas, hay productos PaaS concretos.
- Existe un amplio port-folio de productos Big Data en Cloud.



Amazon AWS



Microsoft Azure

CLOUD: SOLUCIONES HADOOP, PRODUCTOS PAAS

- Puede agobiar la gran cantidad de posibilidades.
- Pero siempre se podrán encuadrar las soluciones en las Capas de una Plataforma.

