

Big Data desde cero

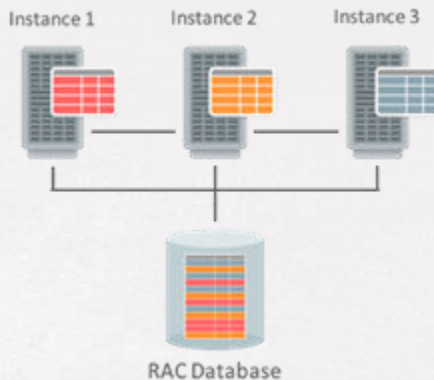
Contexto, tecnologías y aplicación

PLATAFORMA BIG DATA

- Qué es un Cluster
- Tipos de servidores de un cluster
- Qué se considera una Plataforma Big Data
- Capas de una plataforma Big Data

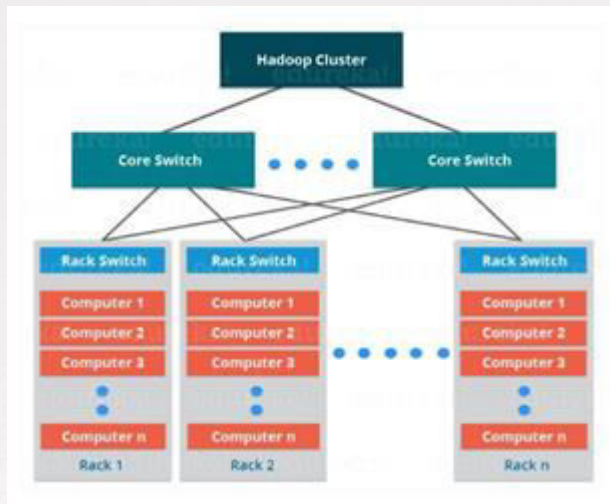
QUÉ ES UN CLUSTER

- Cluster es un conjunto de **servidores que trabajan coordinados**
- Existen **muchos tipos de cluster**, según el producto que lo conforme



TIPOS DE SERVIDORES DE UN CLUSTER

- Según el producto, el cluster tendrá unos tipos de servidores u otros.
- En Big Data tendremos **Cluster Hadoop**



- **Maestros (Masters)** → coordinan el cluster
- **Trabajadores (Workers)** → alojan y procesan los datos
- **Ingesta (Gateway o EdeNodes)** → conexión del cluster con el exterior
- **Utilidades (UtilityNodes)** → servicios adicionales

- Ampliar el número de Workers amplían el procesamiento y almacenamiento del cluster.

QUÉ SE CONSIDERA PLATAFORMA EN UN PROYECTO BIG DATA

- **Cluster Hadoop** es el **core** de una Plataforma Big Data. Pero Big Data no sólo es Hadoop.
- Una Plataforma Big Data engloba más tecnologías y se puede considerar un **conjunto de capas:**



Las capas agrupan componentes y tecnologías **por finalidad**

CAPAS DE UNA PLATAFORMA BIG DATA

- Importante comprender conceptualmente qué agrupa cada capa:

