

Infraestructura de almacenamiento y procesamiento

BIG DATA ACADEMY



Separación de la infraestructura

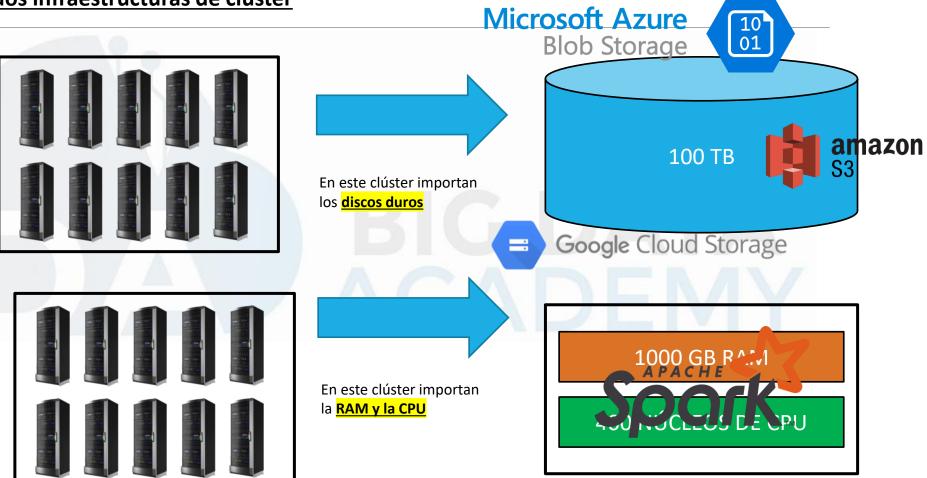


Separación de infraestructura

Como mínimo se necesitan dos infraestructuras de clúster

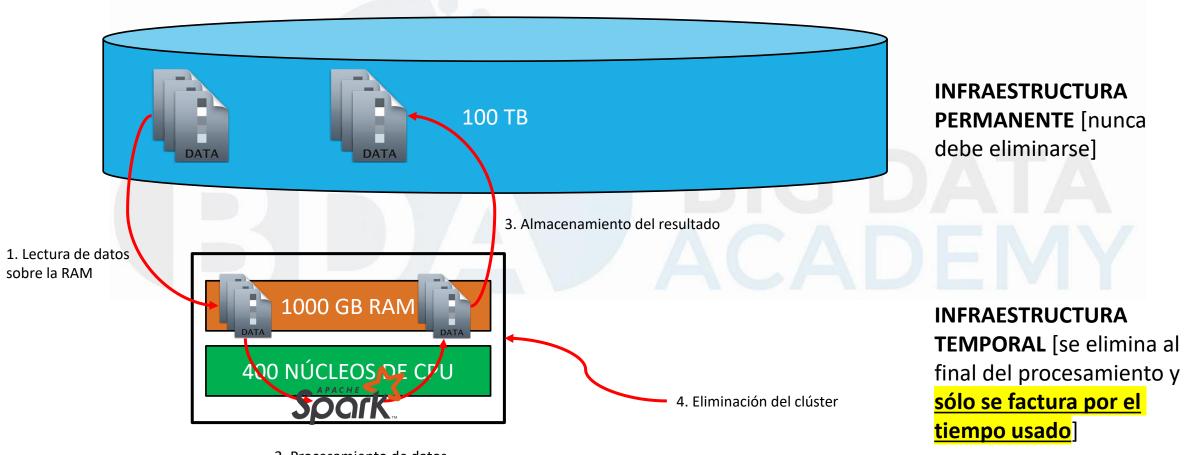
INFRAESTRUCTURA DE ALMACENAMIENTO

INFRAESTRUCTURA DE PROCESAMIENTO





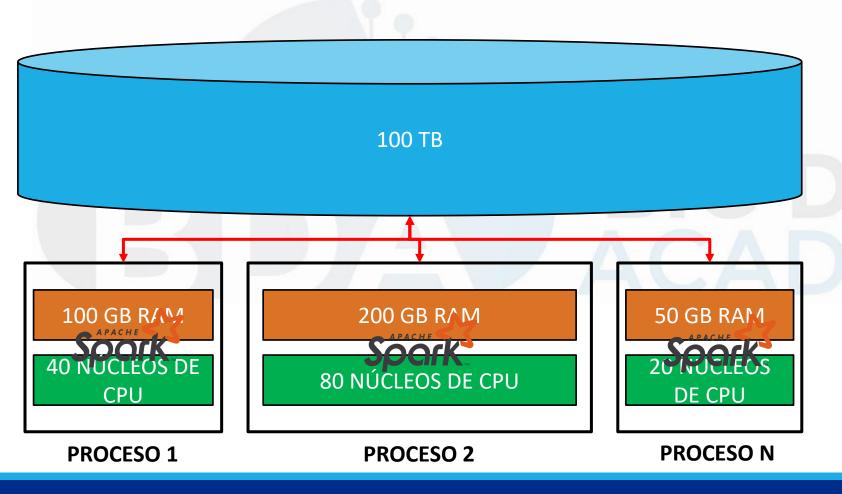
Patrón de procesamiento



2. Procesamiento de datos



Procesos sobre clúster



Cada proceso de negocio tiene su propio clúster que sólo vive por el tiempo de procesamiento, luego se destruye y sólo se factura por el tiempo de procesamiento



Infraestructura gratuita para practicar



Infraestructura gratuita para practicar

<u>Usaremos el servicio de la nube de DATABRICKS</u>, la cual ofrece un clúster de almacenamiento permanente y un clúster de procesamiento temporal (vive por 2 horas desde que lo dejes de usar)

