

Apache HBase

Diana C. Gómez B.

CONTENIDO

01

¿QUE ES?

...

02

BENEFICIOS DE SU
USO

...

03

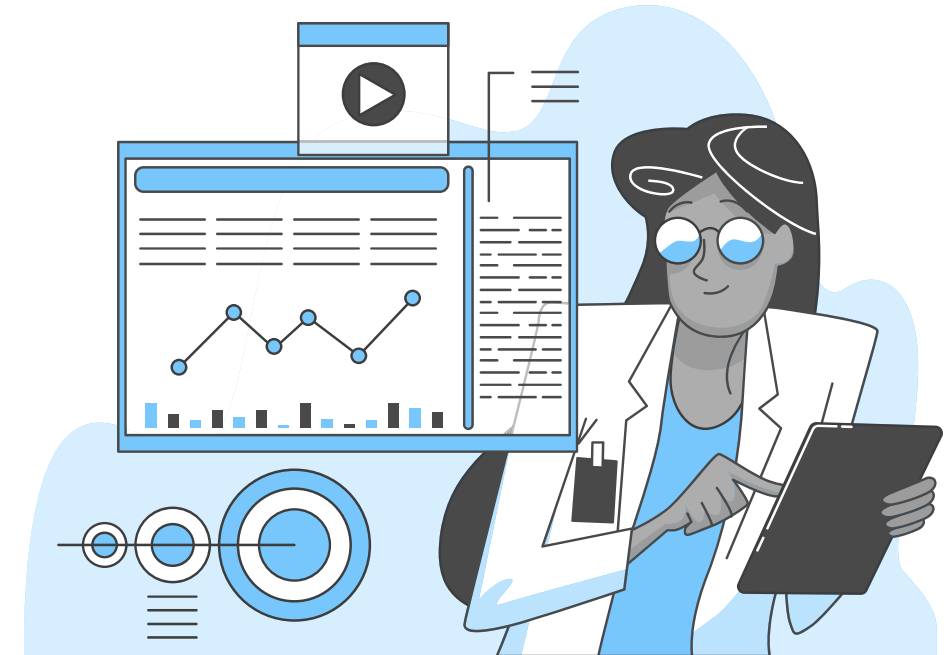
CASOS DE USO

...

04

CONCLUSIONES

...





¿QUE ES?

Apache HBase es un almacén de Big data distribuido, **No SQL** y de código abierto. Permite el acceso aleatorio, estrictamente coherente y en tiempo real a petabytes de datos. HBase es muy eficaz para gestionar conjuntos de datos grandes y dispersos.

BENEFICIOS DE HBASE



Escalabilidad

HBase está diseñado para gestionar el escalado en miles de servidores y el acceso a petabytes de datos.

...



Rapidez

Cada host tiene acceso a los datos en HDFS y S3, y atiende las solicitudes de lectura y escritura en milisegundos.

...



Tolerante a fallas

- HBase divide los datos almacenados en tablas en varios hosts del clúster
- Está diseñado para soportar errores de host individuales.

...

CASOS DE USO

01



Utiliza HBase para almacenar su Graph API, que es una API que permite a los desarrolladores acceder a los datos sociales de Facebook.

02



Utiliza HBase para almacenar sus tweets, menciones, seguidores y otra información relacionada con la plataforma

03



Utiliza HBase para almacenar y procesar datos de sus usuarios, como las películas y programas de televisión que han visto y las calificaciones que han dado.

04



Utiliza HBase para almacenar y procesar datos relacionados con la música, como las canciones que han sido reproducidas por los usuarios,

CONCLUSIONES

- Es altamente escalable, lo que lo convierte en una excelente opción para aplicaciones que necesitan manejar grandes volúmenes de datos de manera distribuida
- A diferencia de las bases de datos relacionales tradicionales que utilizan un modelo de filas, HBase adopta un modelo de columnas, lo que lo hace más eficiente para ciertas cargas de trabajo.
- HBase se integra profundamente con otras herramientas de Hadoop, como Hive, Pig, y MapReduce, lo que facilita el procesamiento de grandes volúmenes de datos distribuidos.
- La implementación de HBase requiere una infraestructura robusta y un clúster de nodos, lo que puede resultar costoso en términos de hardware y mantenimiento.

A decorative network diagram with blue nodes and lines. The nodes are represented by concentric circles, with the inner circle being a solid blue and the outer circle being a white ring with a blue border. The lines are thin and grey, connecting the nodes in a network pattern. There are three main clusters of nodes: one in the top right, one in the bottom left, and one in the bottom center. Each cluster is connected to a central node, and there are also some isolated nodes and lines. Ellipses (...) are used to indicate that the network continues beyond the visible nodes.

GRACIAS!