



# Bases de datos NoSQL: Familias de columnas

XS-3210 Sistemas de Información, II-2023

Docente: Lic. Michael Sánchez Soto

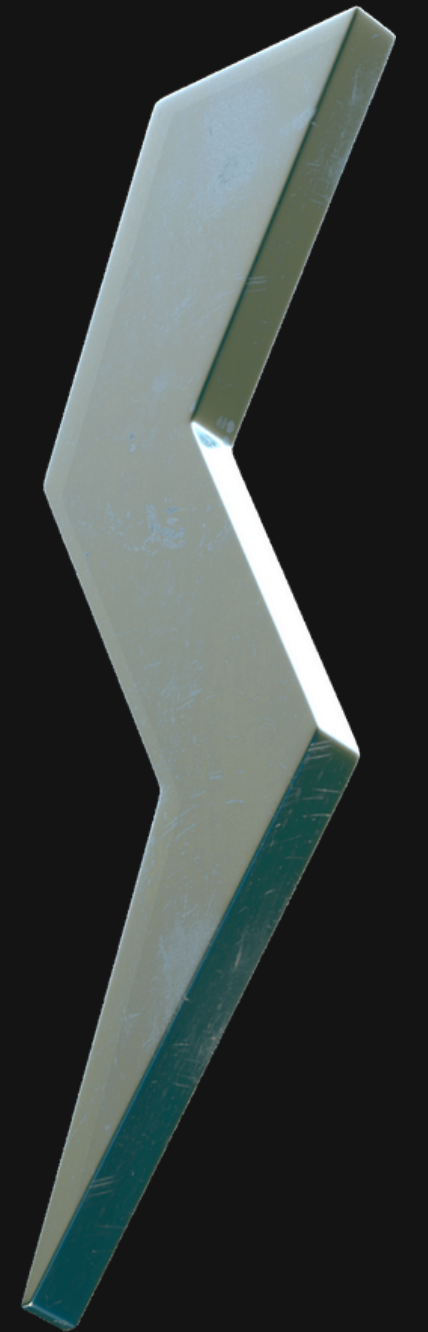
Integrantes:

Iván Daniel Rodríguez Cruz, Cesar Peñaranda Chaves, Joseph Elí Rivera Noguera, Luis Carlos González y Andrea Sánchez Corella.

Miércoles 18 de octubre de 2023

# Bases de datos NoSQL: Familias columnares

Las bases de datos NoSQL "columnar" almacenan datos en columnas en lugar de filas, lo que agiliza el acceso a la información, especialmente en grandes conjuntos de datos. Son una excelente elección para Big Data debido a su escalabilidad horizontal, lo que significa que pueden manejar volúmenes masivos de datos al agregar nodos adicionales. Además, su flexibilidad les permite adaptarse a las necesidades cambiantes de las empresas y a la complejidad creciente de los objetivos de análisis, garantizando la integridad del almacenamiento de datos (Amazon, 2023).



# Problema de estudio

Un comercio de electrónicos requiere un seguimiento detallado de ventas para análisis de mercado y control financiero. Para lograrlo, se pretende usar una base de datos columnar en NoSQL en lugar de herramientas como Excel debido al gran volumen de productos vendidos, lo que permitirá consultas más eficientes y análisis más sólidos. Además, de tener una mayor flexibilidad al momento de que los datos recobrados aumenten o se necesite mayor capacidad de almacenaje.

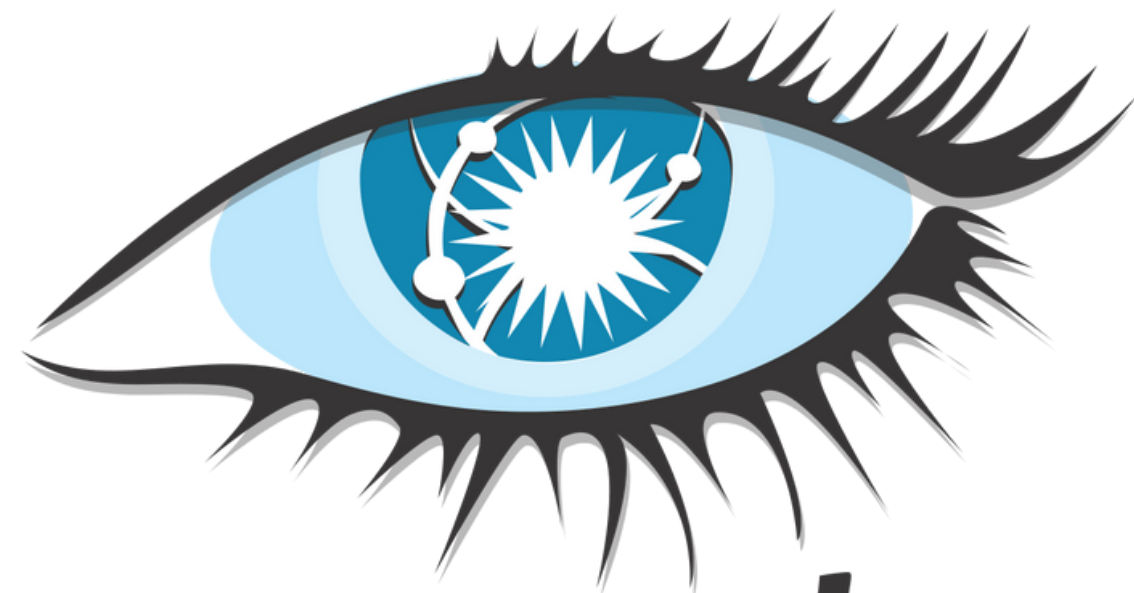


# Productos utilizados

Para este trabajo se utilizaron dos plataformas líderes en las bases de datos columnares, Amazon Redshift (servicio en la nube ligada a AWS) y Apache Cassandra (CMD de Windows).

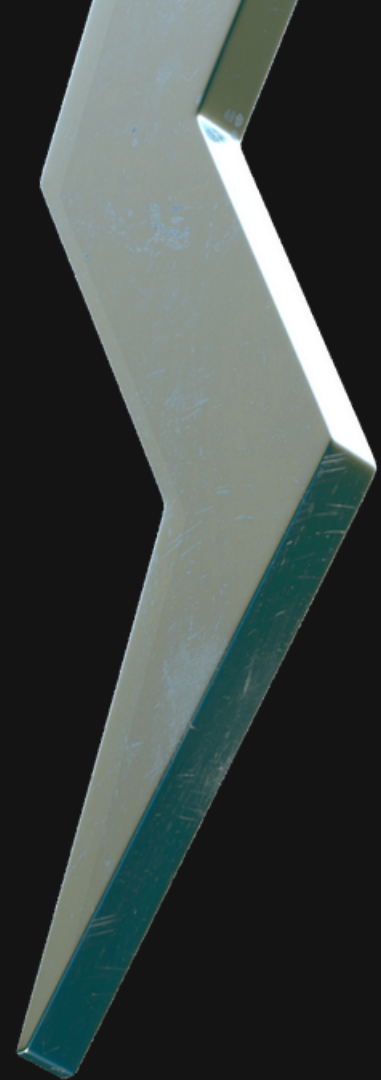


**amazon**  
**REDSHIFT**



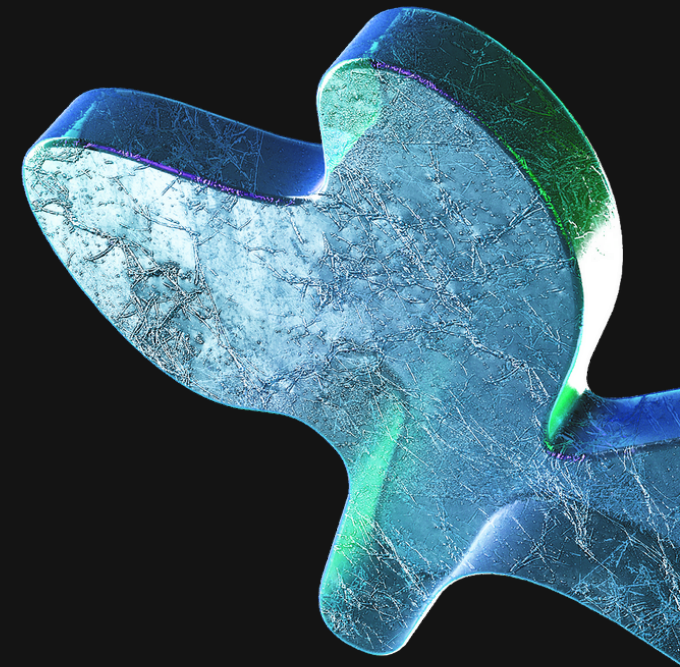
***cassandra***





¿Qué se hizo?

# Comparaciones entre softwares



- Consistencia
- Desempeño en la inserción de datos
- Desempeño en la recuperación de datos
- Facilidad de uso
- Herramienta de gestión
- Método utilizado para hacer consultas
- Almacenamiento

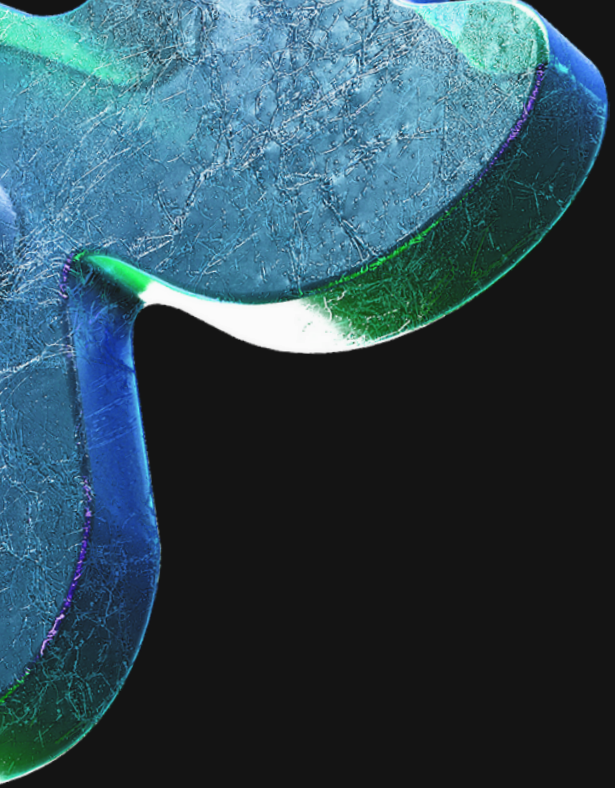
# Dificultades en Amazon Redshift

- Costos elevados del servicio de Amazon Redshift (facturación de \$1500).
- Problemas con los permisos y roles en los clústeres de los servicios en AWS para cargar la base de datos completa.
- Curva de aprendizaje más puntual por el hecho de tener que saber utilizar, no solo Amazon Redshift, sino los demás servicios de AWS para alimentar y dar seguimientos a las bases de datos creadas en esta plataforma.

# Dificultades en Amazon Redshift

Cargos por servicio		Cargos por cuenta	Facturas	Ahorros	Impuestos por servicio
Amazon Web Services, Inc. cargos por servicio <a href="#">Información</a>					<div>Expandir todo</div>
Total de servicios activos		Cargos totales por servicios antes de impuestos en USD			
7		1584,89 USD			
<div><div></div>Filtrar por nombre de servicio o región</div>					<div>&lt; 1 &gt;</div>





# Dificultades en Apache Cassandra

- Empezar

# Conclusiones

