



# UN SIS

## Universidad de la Sierra Sur

### Licenciatura en Informática

#### Base de Datos I

#### Apache Cassandra

#### Hernández Cruz Eduardo Gustavo



# Apache Cassandra

Es un almacén altamente escalable, eventualmente consistente y distribuido de estructuras clave-valor.

Iniciado por Facebook

- Código abierto en 2008
- Proyecto apache en 2009
- Licencia: Apache License 2.0
- Escrito en Java
- Multiplataforma
- Versión actual: 3.11.4 publicada el 11 de febrero del 2019.
- Web: <http://cassandra.apache.org/>



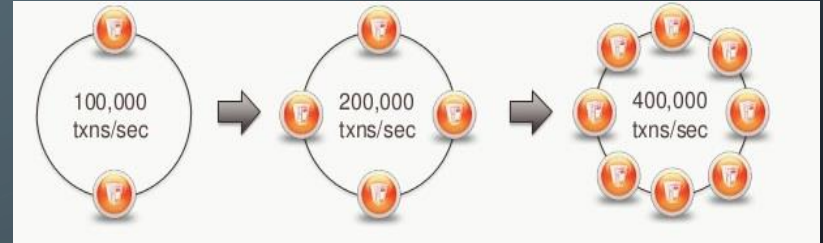
# Apache Cassandra

- Creado por Avanash Lakshaman y Prashany Malik para ser integrado en el motor de búsqueda de Facebook.
- Influido por Google BigTable (modelo de datos orientado a columnas) y Amazon Dynamo (arquitectura- distribuida peer-to-peer).



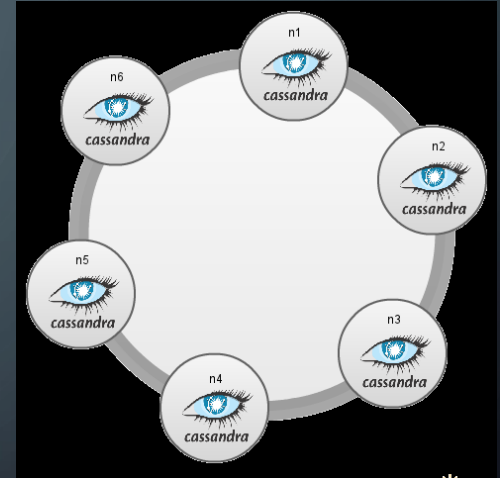
# Arquitectura

- Es distribuida.
- Escala linealmente.
- Escala de forma horizontal.
- Implementa una arquitectura Peer-to-Peer



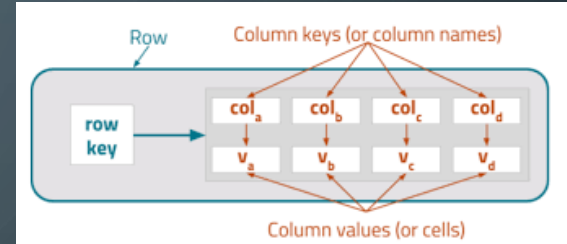
# Características principales

- Descentralizado.
- Soporta replicación y replicación de múltiples data center.
- Escalabilidad.
- Tolerancia a fallos.



# Modelo de datos

El modelo de datos de Cassandra consiste en particionar las filas, que son reorganizadas en tablas. Las claves primarias de cada tabla tiene un primer componente que es la clave de partición. Dentro de una partición, las filas son agrupadas por las columnas restantes de la clave. Las demás columnas pueden ser indexadas por separado de la clave primaria



# Lenguaje de consulta

```
CREATE KEYSPACE MiEspacioClaves
  WITH REPLICATION = { 'class' : 'SimpleStrategy', 'replication_factor' : 3 };

USE MiEspacioClaves;

CREATE COLUMNFAMILY MisColumnas (id text, Apellido text, Nombre text, PRIMARY KEY(id));

INSERT INTO MisColumnas (id, Apellido, Nombre) VALUES ('1', 'Perez', 'Juan');

SELECT * FROM MisColumnas;
```

Resultado:

id	nombre	apellido
1	Juan	Perez

(1 fila)

# Ventajas de Cassandra

- Escalabilidad horizontal: puede añadirse nuevo hardware si aumentan los requisitos de la base de datos.
- Respuesta rápida aunque la demanda crezca (si está bien configurado).
- Elevadas velocidades de escritura para gestionar volúmenes de datos incrementales (siempre que la implementación de la base de datos sea adecuada).





# Desventajas de Cassandra

- No hay JOINS (a cambio de más velocidad).
- No permite ordenar resultados en tiempo de consulta.
- Lenguaje CQL, con una sintaxis y semántica próxima a SQL, no está estandarizado.



# Referencias

<https://www.paradigmadigital.com/dev/cassandra-la-dama-de-las-bases-de-datos-nosql/>

[https://ocw.unican.es/pluginfile.php/2396/course/section/2473/NoSQL\\_Tema2\\_Cassandra.pdf](https://ocw.unican.es/pluginfile.php/2396/course/section/2473/NoSQL_Tema2_Cassandra.pdf)

[https://es.wikipedia.org/wiki/Apache\\_Cassandra](https://es.wikipedia.org/wiki/Apache_Cassandra)

<https://www.psafe.com/es/blog/que-es-apache-cassandra/>