

## Trabajo práctico no. 2

### Base de datos NoSQL: Cassandra

1. Usando el manejador Cassandra define una tabla que permita guardar información sobre autos que vende una agencia a los cuales les da servicio. La agencia vende autos de diversas marcas. Hay que observar los siguientes criterios:

- Se deben poder guardar atributos típicos de un auto como: marca, modelo, número de serie, año y placa.
- Definir la clave primaria de tal manera que se aprovechen lo mejor posible las características de Cassandra para almacenar los datos.
- También debe ser posible guardar los distintos colores que un tipo de auto tiene y las distintas variedades de auto (austero, equipado, deportivo).
- Asimismo, se debe poder saber cuándo se le ha hecho servicio al auto, qué tipo de servicio y cuál fue el costo.

Ingresa algunas tuplas de datos para verificar la estructura de la tabla.

2. Usando el manejador Cassandra define una tabla que permita simular un *wiki*<sup>1</sup>.

El wiki inicialmente debe permitir manejar “páginas”, y su texto, de una manera similar a la indicada en la figura:

|               | row keys<br>(wiki page titles) | column family<br>"text"   |
|---------------|--------------------------------|---------------------------|
| row<br>(page) | "first page's title"           | "": "Text of first page"  |
| row<br>(page) | "second page's title"          | "": "Text of second page" |
|               |                                |                           |

Figure 14—The wiki table has one column family.

<sup>1</sup> Wiki: sitio web cuyas páginas pueden ser editadas por múltiples usuarios a través de un navegador web.

Lo que también se persigue es que para cada persona que modifique el texto se guarde también la información siguiente:

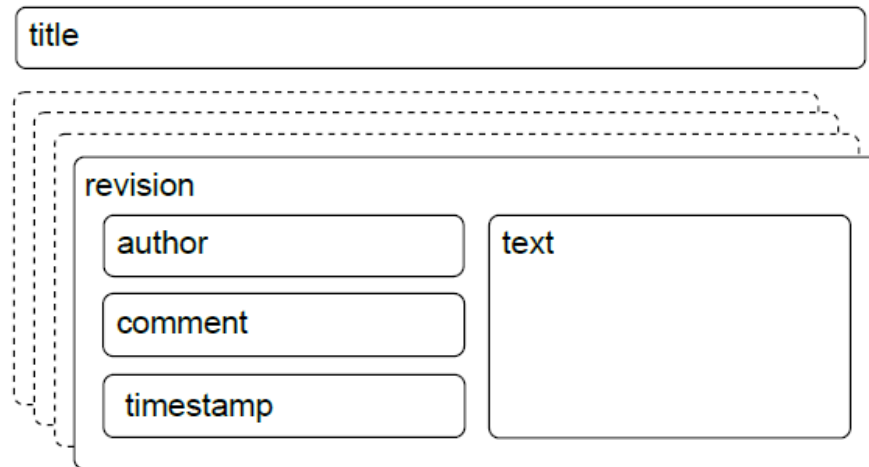


Figure 15—Requirements for a wiki page (including time dimension)

Esto es, que se guarde el nombre de la persona que modificó, comentarios sobre el trabajo que hizo y el tiempo en que esto ocurrió.

En un esquema de familias de columnas, la información guardada sería:

|               | keys<br>(title) | family<br>"text" | family<br>"revision"                |
|---------------|-----------------|------------------|-------------------------------------|
| row<br>(page) | "first page"    | "": "..."        | "author": "..."<br>"comment": "..." |
| row<br>(page) | "second page"   | "": "..."        | "author": "..."<br>"comment": "..." |
|               |                 |                  |                                     |

Figure 16—Updated wiki table architecture (time dimension not shown)

Inserta algunas tuplas para probar la estructura de la tabla.

## TAREA 2

### Fecha de entrega: clase 9.

3. Una cadena de supermercados desea saber cuál es la tendencia de sus clientes en cuanto a los productos que adquieren cuando realizan sus compras en sus diferentes sucursales.

Primeramente, la cadena quiere saber las fechas en que los clientes van a hacer sus compras, a qué sucursales acuden, cuáles productos compran y en qué cantidad. También quiere conocer el total de la compra realizada.

Adicionalmente la cadena quiere conocer por sucursal y por fecha, cuáles productos se vendieron en el día, el total de unidades de cada uno y el monto total de la venta de cada producto.

Define consultas para responder lo siguiente:

- a. La cantidad total de clientes que acudieron a una sucursal dada en una fecha específica.
- b. La cantidad total de clientes que acudieron a todas las sucursales en un día concreto.
- c. Los productos que adquirió un cliente dado en una fecha dada.
- d. El nombre de los clientes que acudieron a comprar el mes pasado.
- e. La cantidad total de productos (tipos, no unidades) que se vendieron en un día concreto.
- f. El nombre de los productos que se vendieron la semana pasada (día inicial: lunes).

### Observaciones:

Puedes utilizar más de una tabla e índices donde se requiera. Inserta una cantidad adecuada de tuplas para probar las estructuras y las consultas. Utiliza enteros para identificar a las sucursales y nombres muy cortos para los clientes y los productos.

Utiliza de manera adecuada los recursos que brinda Cassandra para definir las tablas, de tal forma que la especificación de la clave primaria ayude a almacenar lo mejor posible los datos, para que las consultas se respondan rápido y con el menor costo posible. También considera que los nulos se deben reducir al mínimo.

**La clave primaria debe estar definida como una clave compuesta.**

4. Un banco desea conocer información acerca de las empresas y clientes preferentes que utilizan los diversos productos que ofrece (cheques, tarjeta de crédito e inversiones).

El banco quiere conocer para cada cliente cuáles son los montos totales de los depósitos y los retiros que han efectuado en los diversos productos que éstos tienen y la fecha en que los realizaron.

Asimismo, desea saber por cada producto que maneja, y por cada día, los totales de depósitos y retiros que se realizaron en los mismos, por cada medio que ofrece el banco (ventanilla, internet, celular y terminales en comercios en el caso de tarjeta).

Define consultas para responder lo siguiente:

- a. El tipo de productos que maneja un cliente dado.
- b. La cantidad total de clientes que manejan inversiones.
- c. El nombre de los clientes que hicieron depósitos o retiros la semana pasada (día inicial: lunes).
- d. El nombre de los clientes, y el día, que hicieron retiros superiores a un monto dado.
- e. Los montos de los depósitos y retiros que se hicieron para un producto dado en un rango de fechas.
- f. El nombre de los productos, y del medio, que tuvieron depósitos superiores a un monto específico en un día concreto.

**Observaciones:**

Las mismas que para el ejercicio 3.

**Entregables:** para ambos ejercicios, archivos de texto con la definición de las tablas, las inserciones y las consultas. Enviarlos al correo del profesor: flopez@itam.mx.