DEPARTAMENTO
DE COMPUTACION
Facultad de Ciencias Exactas y Naturales - UBA

Trabajo Práctico 2

21 de Noviembre de 2015 — Bases de Datos

Integrante	LU	Correo electrónico
García, Diego	223/97	diego.garcia.mail@gmail.com
Morales, Marcelino	14/02	marcelino.morales@gmail.com
Schmit, Matías	714/11	matias.schmit@gmail.com
Tastzian, Juan Manuel	39/10	jm@tast.com.ar

Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires
Ciudad Universitaria - (Pabellón I/Planta Baja)
Intendente Güiraldes 2160 - C1428EGA
Ciudad Autónoma de Buenos Aires - Rep. Argentina
Tel/Fax: (54 11) 4576-3359
http://www.fcen.uba.ar

Índice

1. Introducción

2. Desnormalización

Para la resolución de este ejercicio entregamos el diseño físico de la DB realizado en formato JSON con el agregado de que en lugar de indicar un valor para cada atributo indicamos el tipo de datos. De esta forma hacemos mas legible el modelo. Además discutimos cada uno de los puntos de optimización pedidos comentando las distintas variantes que evaluamos para la resolución de los mismos. Para la definición de los tipos de datos a utilizar, al no estar especificados en el DER, los elegimos de la forma que nos pareció mas relevante al contexto.

El diseño de la DB propuesto es el siguiente:

```
db: {
Clientes: [{
DNI : int,
Nombre : string,
Edad : int,
CantCompras : int,
CantComprasMismaEdad : int
}],
Empleados: [{
NroLegajo,
Nombre,
ClientesAtendidos:[{
DNI : int,
Edad : int,
Nombre : string,
Fecha : string
}],
Sectores: [{
CodSector: string,
IdTarea: int
}]
}],
Articulos: [{
CodBarras : string,
Nombre : string,
CantVendidos: int,
CodSector : string
}],
Sectores: [{
CodSector: string,
Trabaja: [{
NroLegajo,
IdTarea
}]
}],
Tareas: [{
IdTarea: int,
Descripcion: string
}]
}
```

2.1. Resolución de Consultas

2.1.1. Empleados que atendieron a clientes de mayor edad

Para resolver esta consulta embebimos información de los clientes dentro del documento de Empleados, agregando una lista de clientes atendidos. De esta forma, al obtener el documento del empleado correspondiente podemos acceder directamente a los datos de todos los clientes que fueron atendidos

por él, y por lo tanto también podemos hacer un filtro sobre la entidad Empleados que devuelva los empleados que atendieron a los clientes mayores de edad.

2.1.2. Los articulos mas vendidos

Para resolver esta consulta se nos ocurren 3 opciones:

1. Embeber la información de compra dentro del artículo. El problema con este enfoque es que se replicaría la información del cliente en cada compra, y además la consulta pedida queda muy compleja de hacer.
2. Mantener una lista de artículos comprados dentro de cada cliente. Esto reduce la replicación de información, pero sigue sin solucionar la complejidad de la consulta pedida.
3. Finalmente, optamos por reutilizar la idea del punto 2 y agregar un contador de ventas dentro de cada artículo. Esto nos permite resolver la consulta de manera eficiente, con mínima redundancia, sin perder la información del modelo.

2.1.3. Los sectores donde trabajan exactamente 3 empleados

La consulta se resuelve buscando qué sectores tienen exactamente 3 elementos en el array "Trabaja". Esto vale porque la aridad de la ternaria no permite que haya un empleado de un sector con más de 1 tarea asignada

Para resolver esta consulta embebimos una parte de la información perteneciente a la entidad Empleado dentro del documento Sector, de esta forma resolvemos eficiéntemente la consulta (que se realiza contando la cantidad de elementos del array Trabajapor cada sector) y mantenemos replicada solo la información necesaria para resolverla.

2.1.4. El empleado que trabaja en más sectores

Para esta consulta decicimos guardar en cada empleado una lista de los sectores en los que trabaja. De esta forma podemos resolver esta consulta contando la longitud del array de sectores de cada empleado, y quedándonos con el de mayor longitud.

2.1.5. Ranking de los clientes con mayor cantidad de compras

Para optimizar al máximo esta consulta optamos por mantener un contador de compras dentro del documento de clientes. De esta forma resulta trivial el armado de un ranking de clientes con mayor cantidad de compras. Además esta opción solo requiere que al realizar una venta actualicemos el contador de compras del cliente.

2.1.6. Cantidad de compras realizadas por clientes de la misma edad

Para resolver esta consulta decidimos agregar un contador de compras extra a la entidad cliente que cuenta el número de compras realizadas por clientes de esa edad. De esta forma por cada compra realizada, debemos incrementar el contador de todos los clientes que tengan la misma edad que el contador. Ésto nos permite mantener un nivel óptimo de redundancia y evitar soluciones mas desprolijas como crear una tabla de edades por ejemplo.

3. Map Reduce

4. Sharding

5. Otras bases de datos NoSQL

6. Conclusiones