

Modelo de agregación clave-valor

Máster en Business Analytics y Big Data
Bases de Datos No Convencionales



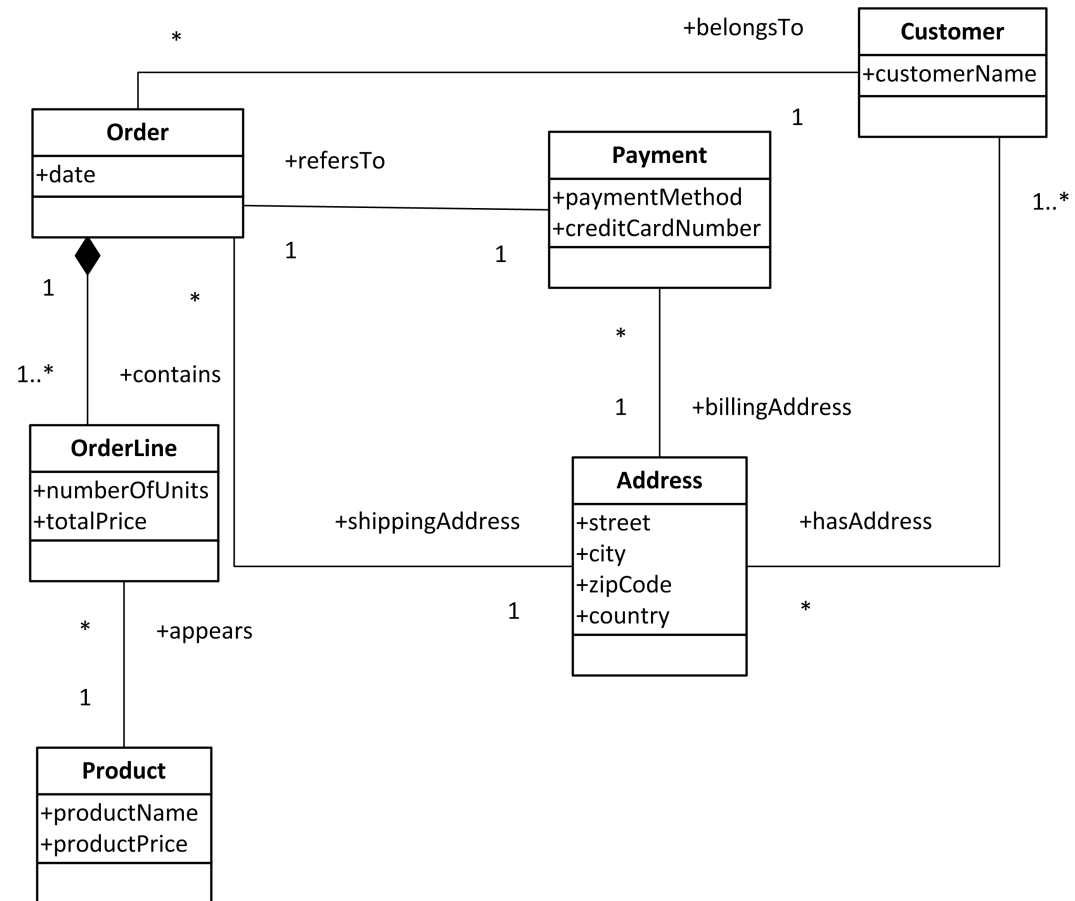
Basados en la documentación de la UOC. Autores: E. Rodríguez y J. Conesa

Contenidos de la sesión

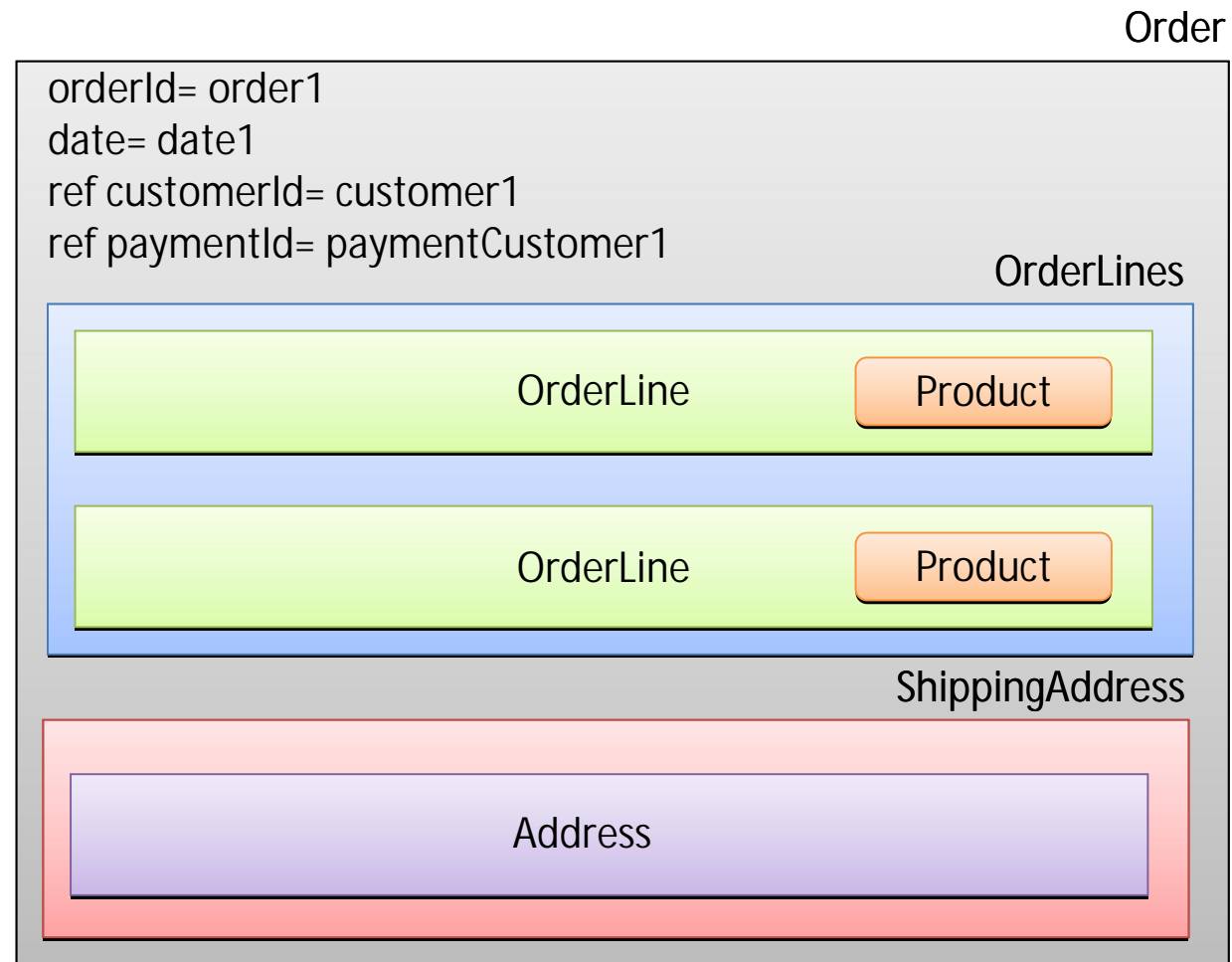
- Ejemplo motivador
- Modelo de datos
- Recomendaciones de uso
- Algunos ejemplos de uso
- Recursos y enlaces

Ejemplo motivador

Imaginemos que queremos recuperar toda la información relativa a un pedido (Order) a efectos de gestión de envío.



Usando un modelo documental



¿Y si usáramos
un modelo
clave-valor?

OrderID



Modelo de Datos

Modelo de datos

- El modelo clave-valor tiene su origen en Berkeley DB (1986).
- Algunas bases de datos NoSQL que se basan en el modelo clave-valor son DynamoDB, Redis, Riak, Aerospike, Oracle Berkeley DB, entre otras.
- El agregado consiste en un par (clave, valor).
- La clave se puede extraer del dominio de aplicación (p.e. nombres de usuario, direcciones de correo electrónico, números de la seguridad social, coordenadas cartesianas etc.)
- La base de datos ignora la estructura asociada al contenido del agregado (valor), es decir, el esquema está implícito y su interpretación pasa a ser responsabilidad de la aplicación.
- A pesar de ello, algunas bases de datos clave-valor (p.e. Riak) permiten estructurar un mínimo el contenido del agregado.



Recomendaciones de uso

¿Cuándo usar un modelo clave-valor?

- En entornos que requieran escalabilidad
 - Sencillez en la configuración de la replicación y la distribución.
- En entornos que requieran una gran tolerancia a fallos
- Cuando la clave del agregado esté clara
- Cuando la mayoría de las consultas a realizar sean por clave
- Cuando existan pocas interrelaciones entre los datos
- Cuando no se requiera almacenar datos estructurados ni semiestructurados.
 - El esquema reside en la aplicación.
 - El esquema lo definen las consultas que se vayan a realizar con más frecuencia.

¿Cuándo no usar un modelo clave- valor?

- Cuando se necesiten transacciones.
- Cuando se requiera realizar consultas sobre los datos del agregado
- Cuando se necesitan frecuentemente datos relacionados en dos o más agregados.
 - No existe la posibilidad de hacer joins.
- Cuando se requiera un alto nivel de consistencia
 - La alta disponibilidad hace que sea difícil garantizar consistencia fuerte
- Cuando se necesitan explotar informes complejos, que requieran agregaciones complejas.



Ejemplos de uso

Ejemplos reales de bases de datos clave- valor

- El servicio público de sanidad del Reino Unido usa Riak para gestionar los datos de historiales clínicos de pacientes.
- Inagist utiliza Riak como sistema gestor de bases de datos.
- Mozilla utilizó Riak para analizar los resultados de su beta test en la versión 4.0.
- Wikia utiliza Riak para gestionar los datos de las sesiones de usuario.
- Pinterest utilizó Redis para poder adaptarse a su curva exponencial de crecimiento en usuarios y número de visitas.
- LinkedIn utiliza Voldemort para almacenar la información sobre las posibles relaciones entre sus usuarios.



Recursos y enlaces

Lecturas

- E. Redmond, J. Wilson (2012). Seven Databases in Seven Weeks: A Guide to Modern Databases and the NoSQL Movement. Pragmatic Bookshelf.
- P.J. Sadalage & M. Fowler. (2013). NoSQL Distilled. A brief Guide to the Emerging World of Polyglot Persistence, Pearson Education.
- C. Coronel, S. Morris & P. Rob (2013). Database Systems: Design, Implementation and Management 10e. Course Technology, Cengage Learning.
- R. Catell (2010). "Scalable SQL and NoSQL Data Stores". SIGMOD Record 39(4), pp 12-27. (<http://dl.acm.org/citation.cfm?id=1978919>).