DATABASE DESIGN FOR RETAIL INVENTORY AND PRODUCT CATALOGS

Couchbase

How companies are building their product and pricing catalog with Couchbase

Table of Contents

Architecting high-performance retail and e-commerce application that scale with demand

Today's leading retailers and e-commerce companies – Walmart, Tesco, eBay, Fanatics, StubHub, Staples, Cars.com, and hundreds more – are embracing NoSQL database technology to build and run modern web and mobile applications. Why? Because NoSQL is better than relational technology in meeting the performance, scalability, availability, agility, and affordability requirements of these applications.

Rising customer expectations and competitive pressures are driving the need for NoSQL. Companies face increasing demands to deliver great customer experiences – i.e., fast, personalized, context- and location-aware. At the same time, they have to manage growing volumes of users and data, while reducing costs and time to market. These pressures create a new set of requirements for the operational database.

The operational database must also be versatile to support many use cases in the space, including:

- Product and pricing catalog
- Shopping cart
- Customer profile management
- Session store
- 360-degree customer view
- Loyalty program management

To meet the demands of today's digital consumer, retailers are leveraging the Couchbase NoSQL database as they move away from monolithic solutions to microservices-based architectures. In this paper, we'll explore how these companies are building the product and pricing catalog with Couchbase, including data modeling and querying.

The following table shows the requirements needed to engage today's digital customer, how Couchbase addresses those requirements, and examples of customer use cases.

Requirements needed	How Couchbase meets these requirements
Service data requests with sub-millisecond latency	**Built-in data cache** that can serve documents with sub-millisecond latency Criteo is deploying massive amount of data to handle real-time queries within 30 ms response time.
Scale to meet peak demand (e.g. Black Friday, Christmas), easily and affordably	Easy scalability with a **single node type**. New node can be added to a cluster with 2 clicks. Scaling out when peak demand or down after peak. PMU is using elastic scalability to handle Grand Prix. They add a node for a day and remove it after Grand Prix.
Provide 24x7x365 availability	**Intra-cluster replication and auto-failover** within a cluster provides high availability. **Enterprise support** ensures high availability and diagnostics whenever a failure occurs. **Professional services** help tune the architecture to meet the availability needed.
Easily accommodate evolving data types and queries	Couchbase is also a **document database**, with a **JSON document** model, **SQL queries**, and **fast in-memory indexing** (Memory Optimized Index) Large Sport Retailer in France is using semi-structured data model to handle product categorization and inheritance by denormalizing the catalog.
Replicate data across data centers globally	Cross Data Center Replication (XDCR) allows **active/active replication** between data centers to meet the PRA requirements. Most Couchbase retail customers are deploying Couchbase in two data centers.
Integrate with big data tools like Hadoop, Spark, Kafka and others	Couchbase has connectors to Hadoop, Spark, Kafka, ElasticSearch and other big data technologies. PayPal leverages the Kafka connector to Couchbase in their architecture.
Accelerate and simplify development	Couchbase is a perfect fit to deploy microservices and accelerate development. Tesco completely redesigned their platform using Couchbase as the main database behind microservices for their product catalog, profile, and more.

Building a product catalog with Couchbase

The product catalog is at the heart of your e-commerce application. And products can be very complex. In addition to a name and base price, a product can have significantly more data associated with it. Products can belong to multiple categories. Product data can include numerous features, specifications, descriptions, and images. Data can include reviews, ratings, and other third-party generated content.

To provide some context, Couchbase supports Tesco's 10M+ product catalog and scales to support 35K requests per second. Tommy Hilfiger relies on Couchbase to serve product catalog content in its digital showrooms, while Cox Automotive processes up to 100M vehicles a day on Couchbase.

There are several ways to model product catalog data with a relational database. However, they're simply workarounds to get past limitations of relational modeling. No one approach is ideal. They all introduce new problems and challenges for developers.

Data model	Problems and challenges
Entity attribute values	• Querying a product requires lots of joins • Querying a product requires hard-coded values • Adding a product requires multiple inserts • Indexes are inefficient due to low cardinality • Requires pivots, will not scale well
Class table inheritance	• Querying a product requires a join • Querying different products requires lots of joins • Adding a product requires two inserts • Adding a new product type requires schema changes
CLOB	• Can't create indexes to improve query performance • Have to read and write the entire value • Difficult to filter because the fields are not exposed

Data modeling

It's much easier to model product catalogs with a NoSQL document database because all data for a single product can be stored together in a single document instead of multiple rows, often in multiple tables. And not only is it easier to model the data, it's simpler and faster to access – there's no need to perform a query with multiple joins or to pivot the results.

Here's what a sample document might look like with Couchbase:

Product (Book)

```
key:
products::books::chasm_city

document:
{
  "type": "book",
  "title": "Chasm City",
  "author": "Alastair Reynolds",
  "skus": [
    {
      "sku": 123456,
      "isbn-13": "9780441010646",
      "format": "paperback",
      "pages": 704
      "published": 2003
      "price": 9.99},
    {
      "sku": 234567,
      "isbn-13": "9780575068773",
      "format": "hardcover",
      "pages": 528
      "published": 2001
      "price": 19.99}]
  "reviews": {
    "average": 4.4,
    "count": 26},
  "categories": [
    "Science Fiction and Fantasy",
    "Science Fiction",
    "Space Opera"]
}
```

Product (Movie)

```
key:
products::movies::2001_a_space_odyssey

document:
{
  "type": "movie",
  "title": "2001: A Space Odyssey",
  "director": "Stanley Kubrick",
  "released": 1968
  "skus": [
    {
      "sku": 789012,
      "format": "blu-ray",
      "duration": 149,
      "price": 9.99,
      "reviews": {
        "average": 4.7,
        "count": 150}},
    {
      "sku": 890123,
      "format": "laserdisc",
      "duration": 149,
      "price": 124.99,
      "reviews": {
        "average": 4.5,
        "count": 130}}]
  "categories": [
    "Science Fiction"]
}
```

Querying

Unique to Couchbase, product data can be queried via a SQL-based query language, N1QL (“nickel”), or via simple key-value operations. In the following examples, the results are displayed as tables for simplicity. However, within applications, the results can be JSON documents, providing much more flexibility – especially when querying multiple product types.

Example One: List all products by title and price

```
SELECT p.title, p.director, v.format, v.price FROM products p UNNEST p.versions v
WHERE p.type= "movie" AND v.price<10
```

Title	Type	Format	Price
2001: A Space Odyssey	Movie	Blu-Ray	9.99
2001: A Space Odyssey	Movie	Laserdisc	124.99
Chasm City	Book	Paperback	9.99
Chasm City	Book	Hardcover	19.99

Example Two: Find all movies under $50

```
SELECT p.title, p.director, v.format, v.price FROM products p UNNEST p.versions v
WHERE p.type=  "movie" AND v.price<10
```

Title	Director	Format	Price
2001: A Space Odyssey	Stanley Kubrick	Blu-Ray	9.99

Other types of retail and e-commerce applications built on Couchbase

Shopping cart

Facilitate purchases with the shopping cart service and avoid cart abandonment. Plus, with Couchbase's built-in cache, get the performance at scale needed to meet shopper demand. Tesco and Staples rely on Couchbase for shopping cart, checkout, search, and more.

Session store

Manage critical online session data in real time with Couchbase so you can monitor user stats, maintain security, track behavior, place ads, provide access to content, and more. eBay leverages Couchbase to store 13B writes per day as it tracks user sessions across its site.

Personalization

Capture omnichannel customer engagements to create a single view of your customer and provide real-time, personalized updates at every customer touchpoint. Comcast built a Customer 360 platform with Couchbase.

Mobile wallet

Whether it's accessing reservation details, loyalty club information, or user profile data, companies like United Airlines use Couchbase to power their mobile wallet.

Click & Collect

Customers are embracing online purchases and in-store pickup because it's cost-effective and flexible. Learn how companies like Doddle use Couchbase to power their Click & Collect service.

Loyalty program

Pull in data from various sources to build, track, and maintain customer data that fuels your loyalty program, including digital and paper coupon management and profile management. The Cincinnati Reds use Couchbase to power their fan loyalty program.

User profile

Deliver fast and seamless login, authentication, and profile management, including user preferences. As the number of users, complexity of user profile data, and user experience expectations accelerate, Couchbase handles the workload and scales to meet those demands. One of the world's largest consumer electronics manufacturers manages 1B user profiles with Couchbase.

Recommendations

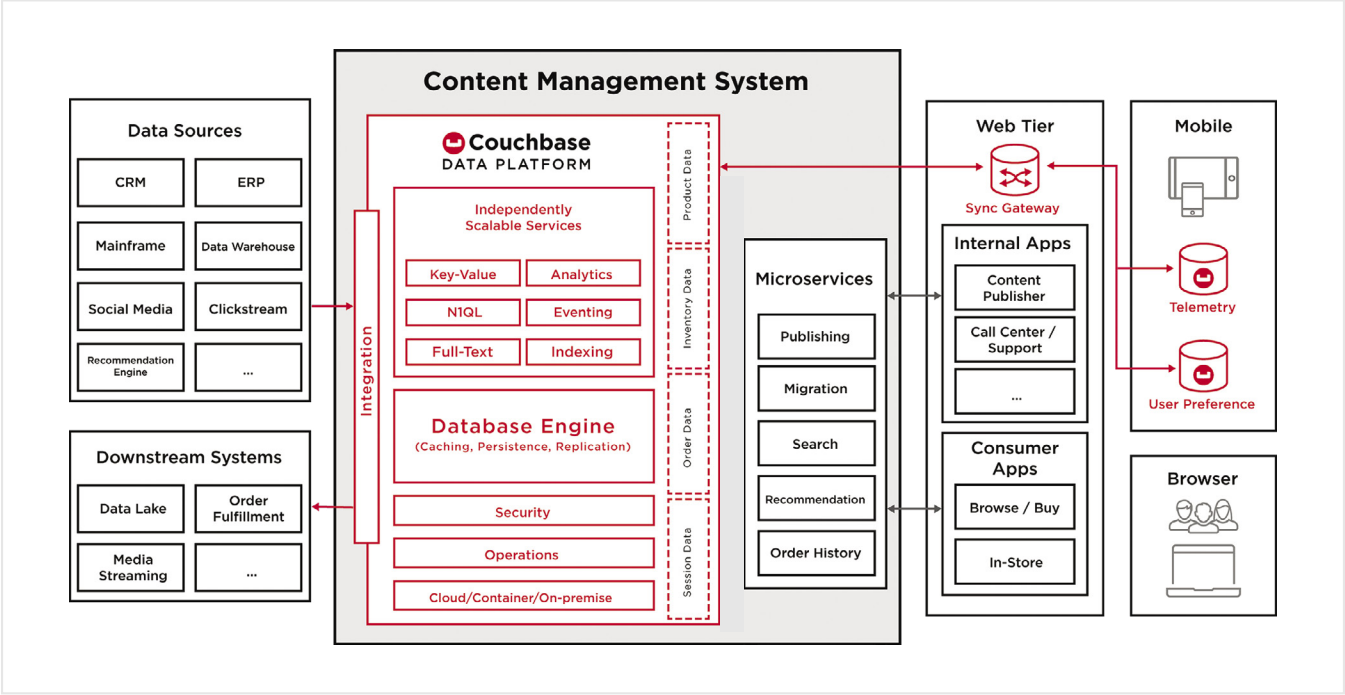
Real-time recommendations such as, "Customers who purchased this cheese also purchased this wine," or "You might like this hotel when you are in London," only work if they are delivered in real time. Couchbase is used by some of the world's largest retailers to serve their cross-sell and upsell programs, online and in their mobile applications.

Mainframe offload

By caching frequently used data, response times are improved and costs are substantially lowered through reduced workloads. If your app runs on a relational database like Oracle which requires costly hardware in order to scale, or a mainframe like Db2 which charges you by the MIP, a scale-out cache that runs on low-cost servers is more cost-effective.

Reference architecture: catalog and inventory in retail

This reference architecture shows how Couchbase fits into a catalog and inventory service storing product, price, inventory, and session information under a unified services layer.



Customer case studies

Rue du Commerce is a leading e-commerce company in France with business solely done through the web. They replaced Elasticsearch and MariaDB with Couchbase to power their product catalog indexing application.

Main business challenges with Elasticsearch and MariaDB:	What they achieve with Couchbase:
• Needed scalable, high-performance indexing for superior shopping experience, even at peak periods • Required real-time price indexing to provide accurate omnichannel pricing	• Accurate product indexing by Google • Web traffic increased from 25% to 40% (15% traffic increase) • 500K products indexed in 5 minutes instead of 50 minutes (10x faster) • Product seen by Google in 50 ms vs. 1.4 sec previously (28x faster)

Facing issues around storing product catalog data in their relational database, Tesco, the third-largest retailer in the world, evaluated MongoDB™ but ultimately chose Couchbase as a consolidated product catalog database with microservices. Data is ingested via REST API from multiple MDM feeds (CSV, XML) and JSON documents capture multiple data structures: SKUs, product and accounting hierarchies, GTINs (barcodes, ISBNs, etc.).

Main business challenges with a relational database:	What they achieve with Couchbase:
• Product data stored in multiple relational databases, driving up maintenance costs • Deliver fast, easy access/sharing for product data throughout the company and ability to store and update product data for 10M items • Support frequently changing data and multiple data structures while improving customer experience across multiple purchase channels • Improve performance for peak periods like Black Friday, avoiding lengthy delays that resulted in lost revenue	• Scales easily and inexpensively to support 10 million products and 35 thousand requests per second • Low-latency access to millions of documents for great customer experiences • JSON improves developer agility through flexible schema for changing SKUs and support for SQL and text-based queries • First Black Friday success in 2016, earning excellent press reviews

Staples needed to better manage custom B2B product catalogs using 1.6 billion rules applied in real time. While it explored MongoDB™, an inability to scale easily and affordably led to canceled projects. Couchbase enabled Staples to simplify its catalog management using N1QL and JSON while also improving database performance and reliability.

Main business challenges with MongoDB™:	What they achieve with Couchbase:
• Needed to better manage personalized catalogs for B2B shoppers online, a practice that involved 1.6 billion rules applied in real time at time of purchase • Cancelled digital transformation projects due to inability to scale MongoDB™ and Db2 easily and affordably • Needed simplified replication, easy scalability to continue to grow infrastructure	• Simplified management of B2B product catalogs customized for each buyer with N1QL and JSON • Enhanced user experience by adding performance, boosting scalability, and enhancing reliability • Improved responsiveness and availability of e-commerce pricing engine

Louis Vuitton used Couchbase to replace manual tablet inventory updating, bringing its full product catalog directly to the mobile devices of every store employee. By synchronizing catalog data in real time, sales reps can instantly access up-to-date information, including availability in store and at other locations.

Main business challenges with a traditional database:	What they achieve with Couchbase:
• Overcome long, manual updates for each in-store tablet every morning to sync product catalog • Deliver up-to-date product catalog and real-time pricing despite poor network connectivity • Push data from cloud to 10,000 devices • Support 3 datacenters and Microsoft Azure cloud infrastructure • Deliver high availability to keep catalog online 24/7	• Always-on access to product catalog and store inventory • Sync time to mobile devices reduced by 92% • From less than 40% of employee users to 80%+ • Quick and easy deployment with Microsoft Azure

Learn more

To learn more, contact your Couchbase sales representative today or visit:
couchbase.com | couchbase.com/downloads

About Couchbase

Couchbase's mission is to be the data platform that revolutionizes digital innovation. To make this possible, Couchbase created the world's first Engagement Database. Built on the most powerful NoSQL technology, the Couchbase Data Platform offering includes Couchbase Server and Couchbase Mobile and is open source. The platform provides unmatched agility and manageability – as well as unparalleled performance at any scale – to deliver ever-richer and ever-more-personalized customer experiences.

3250 Olcott Street | Santa Clara, California 95054 | 1-650-417-7500 | www.couchbase.com