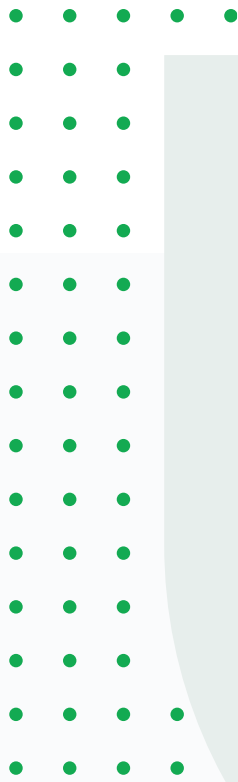




Data Modernization with MongoDB and Google Cloud

AUGUST 2021



As an IT leader or architect, you may discover that your software architecture is encountering performance issues. A traditional SQL database can limit you from easily scaling horizontally, and rigid table schemas lock you into restrictive data models. You may be considering moving your datastore from a mainframe or RDBMS to a more modern database to take advantage of advanced analytics, scale at a faster rate, and cut costs. Whether you are considering this move for a new project or while migrating an existing system, there is a clear path to success with MongoDB Atlas and Google Cloud.

Introduction

Since the 1980s, relational databases such as MySQL and Oracle have served as the go-to data management backbone for enterprise-grade applications. As a replacement for the mainframe hierarchical datastores and flat files of early computers, these relational databases greatly improved implementation and administration.

But software has come a long way since then. The requirements of modern software development have evolved, and rapid data growth coupled with new design paradigms are exposing the limitations of these older database systems. Operating with on-premises (“on-prem”) legacy relational database management systems (RDBMS) now carries with it downsides such as lack of flexibility and limited scalability—not to mention the difficulties of managing data synchronization across disparate platforms such as mobile and desktop.

Outdated data paradigms have become a major barrier to digital transformation. Moving to a modern database has become critical to the success of many organizations.

For many teams, the path forward is to switch to a **modern, document-based database that operates in the cloud**—e.g., MongoDB Atlas on Google Cloud.

At MongoDB, we’ve seen over 22,000 customers switch to a more modern, document-based model to take advantage of a more flexible and intuitive database designed for the developers who use it. This guide is designed for teams that want to understand why and how those customers made the switch from an RDBMS to a cloud-based modern database—whether they were migrating existing apps or building new ones from scratch.

We’ll look at *why* you need to modernize your database, the key differences between RDBMS and MongoDB Atlas, the steps you need to take to successfully migrate your data, and where to get more personalized help.

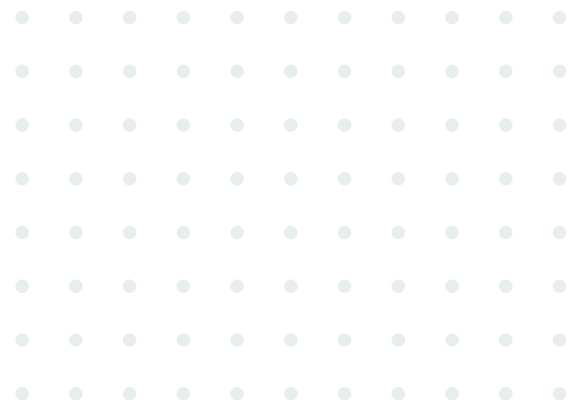
The Case for Migration

An on-prem RDBMS is a popular choice for many applications due to its long history and mainstream usage. Often, it's simply *the database most people know*. As a result, your organization has likely built processes—and maybe even entire teams—around maintaining and managing these legacy systems. This can make it complicated for leadership to face the need for change.

However, the issues of these legacy data systems will only continue to compound. As your organization continues its digital transformation, you may be finding that these outdated legacy layers are making adoption of newer architectures such as [microservices](#), **scale-out**, and **real-time streaming** difficult, if not impossible. On-prem legacy data layers are inflexible, cost-prohibitive, and often lack stability and scalability.

You need a system that can evolve alongside your needs, not an aging and inflexible infrastructure that creates a barrier to technology innovation. Rigid databases simply don't support digital transformation. Your data should map to how you want to interact with it in your application—not the other way around. As your application grows with new features, you will feel the constraints of data stored in an outdated format, particularly when you have a table that experiences dynamic growth.

The absence of agility, incurrence of high costs and potential compliance risks of running your own infrastructure, as well as barriers to scalability, are just a few of the reasons to turn to a more modern solution.



A Solution: MongoDB Atlas on Google Cloud

As a solution, an increasing number of companies are moving to a more modern database on the public cloud. This solution not only reduces the operational overhead of managing infrastructure, but also provides teams with access to on-demand services that give them the agility they need to meet faster application development cycles.

MongoDB Atlas is designed to meet the demands of modern apps with a reliable and powerful foundation that:

1. Stores your data as documents rather than rows and columns, **allowing for greater flexibility.**
2. Offers a distributed system design that **scales horizontally without additional management.**
3. Provides a data synchronization mechanism that **automatically handles serialization, network recovery, and conflict resolution.**

Together, MongoDB Atlas and Google Cloud give you a fully managed, modern, cloud-native database platform that is ideal for building powerful, data-driven applications, deploying advanced data analytics, and maximizing the value of business data. Some benefits include:

- **Comprehensive monitoring and managed backup capabilities.** You don't need to hire administrators or operations engineers to take care of your database maintenance
- **Cutting-edge operational automation** and serverless platform services. Your database is managed by experts, freeing your team to concentrate on the features that matter for your business
- **Data securely stored and managed end-to-end** on Google Cloud. Its Zero Trust architecture ensures secure access to data while protecting against threats, with built-in regulatory compliance to abide by HIPAA, GDPR, and other government ordinances
- Data that can be retrieved by developers in a **simple document-based format**
- Runs on Google's secure private network backbone providing **lower latency for global scale applications**
- **Multi-region replication and global clusters** that ensure your customers see quick responses
- **Patches, upgrades, scaling, and provisioning are handled automatically**
- **Lower total cost of ownership (TCO)** and an accelerated time to value as compared to running your own hardware. Google Cloud's compute infrastructure offers higher IOPS at a lower cost

- **Redundancy and fail-over sequences.** Google Cloud offers at least three zone coverage in each available region for greater redundancy
- **MongoDB Atlas' [performance advisor](#)** provides fine-tuned metrics and specific guidance
- Access to **smart analytics and AI solutions** that democratize access to data across your organization and accelerate time-to-insights

For example, [by using MongoDB Atlas and Google Cloud, Forbes](#) was able to gain 58% faster build time, accelerate release cycles by 4x, support a 28% increase in subscriptions, and reduce TCO by 25%. Auto Trader [switched to MongoDB Atlas on Google Cloud](#) to “reduce complexity and stay nimble.”

This shift to [MongoDB Atlas on Google Cloud](#) gives developers best-in-class automation, workload intelligence, and proven practices that guarantee availability, scalability, security, and compliance. It also helps users prioritize developer agility, productivity, and time to market.



Advantages of MongoDB Atlas on Google Cloud over RDBMS

Let's now look in detail at some of the advantages of MongoDB Atlas on Google Cloud over RDBMS, specifically: scalability, productive developers, data that is shared across devices, and performance.

Intuitive architecture empowers developers and maximizes productivity

By representing documents in BSON, developers can write code that accesses data in a more natural way, one that integrates more easily with the rest of a backend stack. There's no need to learn a completely separate text-based language like SQL in order to run queries or update records across relationships. Data is accessed using the familiar JSON dot notation on object properties.

This simplification lets developers focus on writing application code quickly, without worrying about crafting performant database queries.

Furthermore, running MongoDB Atlas on Google Cloud provides developers with a clean, modern experience that allows them to build and innovate quickly.



Seamless scalability enables expansion and helps you meet key workload requirements

One of the key strengths of a document-based database system like MongoDB Atlas is that the documents exist as entirely self-contained objects. This means that data can be partitioned into manageable chunks and distributed across multiple nodes—a process called *sharding*. MongoDB's native support for sharding makes it possible to achieve horizontal scalability without any additional configuration. Data stored across nodes doesn't suffer from additional performance penalties.

MongoDB is optimized for frequently accessed (or “hot”) data. Long term data storage (“cold data”) and analytical workloads can reside in Google's Big Query infrastructure, which is optimized for infrequent access.

It's not just MongoDB that's built to scale. Google Cloud is designed for easily deploying cloud services at hyper scale — which means that no matter how your business evolves and what kind of data you have, MongoDB Atlas on Google Cloud has the right solution for you. In addition, MongoDB Atlas on Google Cloud offers easier access to Google's numerous other cloud service offerings to meet any additional application or workload requirements, such as streaming, IoT, ETL, analytics and AI/ML.

Shared data across devices

MongoDB Atlas can be configured with Realm Sync, which synchronizes data updates across devices. Whether your users are interacting with your app on mobile, desktop, or in the cloud, Realm Sync manages the receipt and delivery of any data changes in real-time. If internet connectivity is lost, Realm Sync is intelligent enough to batch pending updates, and it even handles merging between disparate timestamped data. All of this is available as part of MongoDB's native feature set.

Realm Sync, in conjunction with the speed and reliability of Google Cloud, automates conflict resolution and network recovery, ensuring that every user is accessing the most up-to-date version of their data.

Faster intelligence and speed to market

MongoDB's document model is optimized for how an application accesses data. These days, user-generated content, like reviews and comments, are best kept in a document-based model like MongoDB, instead of a table. Rather than fretting about designing the perfect table schema, MongoDB's BSON documents permit you to develop features quickly, without the restriction of fitting data into rigid columns. MongoDB's flexible document model means that your database can grow alongside your application and your users' needs without costly migrations and downtime.

Running on Google Cloud infrastructure offers MongoDB Atlas customers maximum flexibility to quickly adapt as their organizations evolve. Using Google's ecosystem of tools, customers can harness the power of data and AI through Google Cloud's open APIs, machine learning services, and analytics engines for fast and accurate intelligence, and for faster development without having to perform long-term infrastructure planning.

The Path to Migration

Next, let's look at the path to [migrating to MongoDB Atlas](#) on Google Cloud, and understand some of the key technical differences you will face along the way. While starting from scratch is the easiest way to introduce a more modern database, it's not always realistic. Most enterprise datastores are already in legacy formats, and your team will likely need to migrate existing data.

In fact, around 30% of all MongoDB projects are now migrations from relational databases. As illustrated below, enterprises from a variety of industries have migrated successfully from relational database management systems to MongoDB for many different kinds of applications.

Sample Organizations Migrating to MongoDB

Organization	Migrated From	Application
Cisco	Commercial RDBMS	eCommerce Platform
eHarmony	Oracle & Postgres	Customer Data Management & Analytics
Shutterfly	Oracle	Web and Mobile Services
Sega	MySQL	Gaming Platforms
Under Armour	Microsoft SQL Server	eCommerce
Baidu	MySQL	100+ Web & Mobile Service
MTV Networks	Multiple RDBMS	Centralized Content Management
Telefonica	Oracle	Customer Account Management
Verizon	Oracle	Single View, Employee Systems
The Weather Channel	Oracle & MySQL	Mobile Networking Platforms

Figure 1: Case Studies

Moving your data is never a trivial process. The good news is that the [MongoDB Professional Services team](#) has helped many clients migrate to MongoDB and understands the key steps to success. [Cloud Factory](#) is aimed at guiding organizations toward a cloud operating model on Google Cloud. Professional Services has helped businesses like Porsche, Vonage, and Sega with their schema design and performance optimization to build systems that scale.

Migrating to a document-based database like MongoDB can be broken down into a series of small steps:

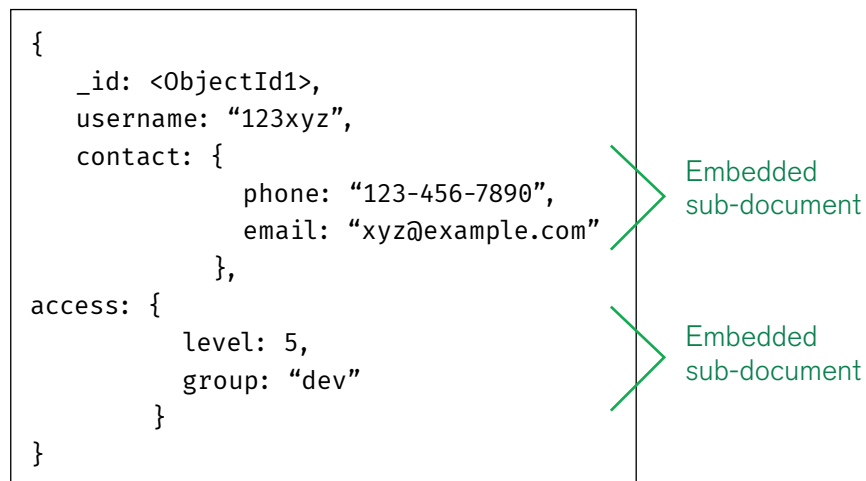
Define and Analyze Your Model

The first step in your migration is understanding whether a multi-cloud or hybrid-cloud system is best for you. This decision is ultimately based on identifying your application data, defining roles and responsibilities for access, and mapping out your infrastructure. In other words, your overall migration strategy and execution plan depends on whether or not you need to rethink your architecture and whether you already rely on any cloud-based services.

Design Your Application

Just as with any new software project, designing your data structure should be your first priority. For example, RDBMS schema design requires you to think about how the data in different tables relate to one another through an additional column called a foreign key. But MongoDB stores its data in a binary representation called [Binary JSON](#) (BSON). BSON supports sub-documents, which means that such relationships can be collapsed (embedded) into a single document in MongoDB.

Example BSON Document



Sub-documents translate into fewer database reads and fewer performance concerns across table JOINS. These techniques, not found in any RDBMS, require a change in thinking for data architects, developers, and DBAs.

The MongoDB Professional Services team can work closely with your data architects to develop a schema that encapsulates all of your existing business needs. If necessary, they can also help redesign your application for a cloud context.

Implement the Schema

After designing the schema, project teams have multiple options for importing data from existing relational databases to MongoDB. The tool of choice should depend on the stage of the project and the existing environment.

Many users simply create their own scripts, which transform source data into a hierarchical JSON structure that can be imported into MongoDB using the `mongoimport` tool.

More cautious migrations involve running the existing RDBMS in parallel with the new MongoDB database, incrementally transferring production data. While retrieving records from the RDBMS, the application also writes them back out to MongoDB. A consistency checker, such as one with an MD5 checksum, can be used to validate the migrated data.

Shutterfly used this incremental approach to migrate the metadata of six billion images and 20TB of data from Oracle to MongoDB.

Incremental migration can be used when new application features are implemented with MongoDB, or where multiple applications are running against the legacy RDBMS. Migrating only those applications that are being modernized enables teams to divide projects into more manageable and agile development sprints.

Our [Business Value Consulting team](#) can estimate the cost of migration, helping you to plan the best strategy for your specific needs. We can also integrate with your development teams to provide guidance during sprints, code reviews, and any other work involved in making sure your data gets to the cloud safely.

Validate the Results

Achieving performance SLAs and expected data behaviors is critical in ensuring that your data has successfully migrated. The validation phase consists of performing real-time reads and writes to your new MongoDB Atlas datastore on Google Cloud, and provides assurances that you can finally retire your legacy systems.

What's more, you can set up [schema validation rules](#), with varying levels of strictness, to make certain that your new MongoDB database performs as expected in real production scenarios.

Optimize the Cloud

Code is always changing, and your data is no exception. MongoDB provides several solutions for your application developers to [optimize query performance](#).

Reliability and performance are critical, but so is the ability for your database to quickly scale up alongside your app's popularity. After migrating to Google Cloud, you'll be able to take advantage of its self-optimization capabilities so that your data flow uses hardware resources most effectively. You'll also have opportunities to optimize costs through more efficient platform management, as well as a pricing structure designed to provide the most value from your data and ROI from your cloud investment.

- Professional Services can engage with you on post-deployment performance reviews, determine where costs can be optimized, and set your apps up to ensure optimal use of native cloud features.

Where to Get More Help

Data modeling is an expansive topic. To help you make the right decisions, here's a summary of the key resources you should review:

- The MongoDB documentation provides an extensive section on [data modeling](#), starting from high-level concepts of the document data model before progressing to practical examples and design patterns, including more detail on referencing and embedding.
- You should also review [our "Building with Patterns" blog series](#) to learn more about specific schema design best practices for different use cases, including the following:
 - catalog and content management
 - IoT
 - mobile apps
 - analytics
 - single view, such as customer 360

It overlays these use cases with specific schema design patterns such as versioning, bucketing, referencing, and graphs.

- MongoDB University offers a no-cost, web-based [training course on data modeling](#). This is a great way to kick-start your learning on schema design with the document data model.
- [Instructor-led MongoDB training](#) helps you develop the expertise and confidence you need to be successful with MongoDB. Designed and led by MongoDB consultants, you and your team will be equipped to build solutions to complex development and operations problems.

Conclusion

A modern database is critical to success. MongoDB Atlas on Google Cloud is the best choice for the evolution of your application. By following the best practices outlined in this guide, you can feel comfortable and confident with the effort and path necessary for implementation. After your successful migration, teams can quickly start to realize a more flexible, scalable, and cost-effective infrastructure, innovating on applications that weren't possible before.

We Can Help

We are the company that builds and runs MongoDB. Over 15,000 organizations rely on our commercial products. We offer the following software and services to make your life easier:

[MongoDB Atlas](#) is the database as a service (DaaS) for MongoDB, available on AWS, Azure, and Google Cloud. It lets you focus on apps instead of ops. With MongoDB Atlas, you only pay for what you use with a convenient hourly billing model. Atlas auto-scales in response to application demand with no downtime, offering full security, resilience, and high performance.

[MongoDB Charts](#) is the best way to create visualizations of MongoDB data anywhere. Build visualizations quickly and easily to analyze complex, nested data. Embed individual charts into any web application or assemble them into live dashboards for sharing.

[MongoDB Realm](#) provides a solution that makes it easy for you to build powerful and engaging experiences on more devices.

[MongoDB Consulting](#) packages get you to production faster, help you scale and tune performance in production, and free you up to focus on your next release.

[MongoDB Training](#) helps you become a MongoDB expert, from design to operating mission-critical systems at scale. Whether you're a developer, DBA, or architect, we can make you better at MongoDB.

About Google Cloud

Google Cloud accelerates organizations' ability to digitally transform their business with the best infrastructure, platform, industry solutions and expertise. We deliver enterprise-grade solutions that leverage Google's cutting-edge technology – all on the cleanest cloud in the industry. Customers in more than 200 countries and territories turn to Google Cloud as their trusted partner to enable growth and solve their most critical business problems.