



LESSON

MongoDB: A Developer Data Platform

Google slide deck available [here](#)

This work is licensed under the [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License](#)
(CC BY-NC-SA 3.0)

Overview



Learning Objectives

At the end of this lesson, learners will be able to:

- Explain the key features of MongoDB's developer data platform.
- Describe how data is stored, processed, and accessed in MongoDB.
- Describe the use cases for change streams when accessing data.

Suggested Uses

- Lecture for one hour class or a part of a longer lecture period
- Handouts / asynchronous learning
- Supplemental reading material - read on your own / not part of formal teaching
- Complement to University courses [M001: MongoDB Basics](#).

This lesson is a part of the courses [MongoDB: A Developer Data Platform](#) and [Introduction to Modern Databases with MongoDB](#).

At a Glance



Length:
45 minutes



Level:
Foundational



Prerequisites:
None

This work is licensed under the [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License](#)
(CC BY-NC-SA 3.0)

Share your feedback: We hope these curriculum materials will be a valuable resource for you and your learners. Let us know how the materials work for you, what we can improve on, and how MongoDB for Academia can support you via our brief [feedback form](#).

MongoDB for Academia: MongoDB for Academia offers resources for educators and students to support teaching and learning MongoDB. Check out our [educator resources](#) and join the Educator Community. Students can receive \$50 in Atlas credits and free certification through the [GitHub Student Developer Pack](#).

Last Update: July 2022

What is the MongoDB Data Platform?

As a complete ecosystem
not just a database, it
supports:

Storing the data

Accessing the data

Processing the data

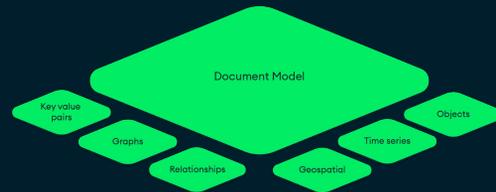


Built around the most intuitive way to model data — the document data model

Document data model maps to how modern developers think & code

Documents are inherently flexible while allowing data governance when required

Documents can address a wide variety of use cases and can be used to model both structured & “semi-structured” data

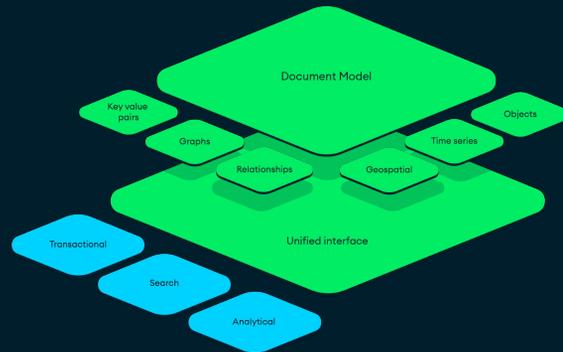




Unified query interface for a broad set of workload types

A single and consistent way to work with data across the transactional, search, and analytical workloads in your organization

Strongly consistent by default with support for multi-document ACID transactions





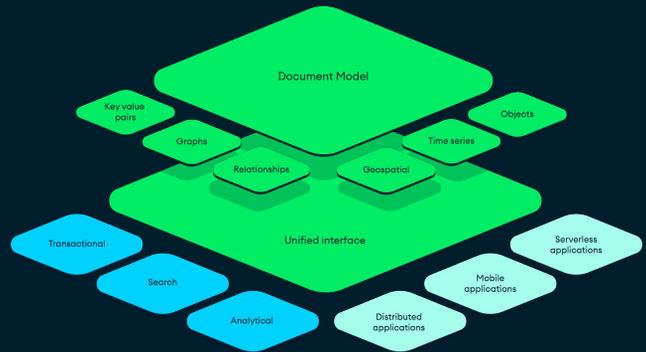
Support a wide range of modern application types as they grow and evolve

In addition to supporting common application types & requirements, the platform also supports —

Distributed applications; create a global data layer in a few clicks

Reactive mobile applications powered by local data that automatically syncs to the cloud

Serverless development frameworks



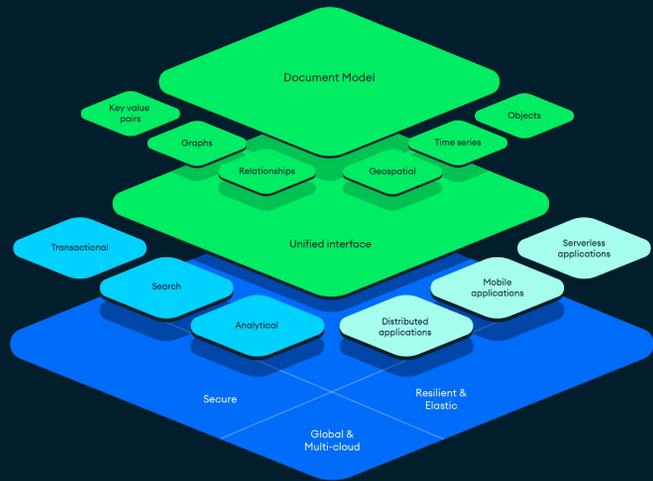


Built on a battle-tested platform that allows you to run anywhere

Cutting edge and comprehensive controls to ensure data security and privacy

Deploy in over 100+ regions across 3 cloud providers. Deploy across clouds to get the best from each provider with no lock-in.

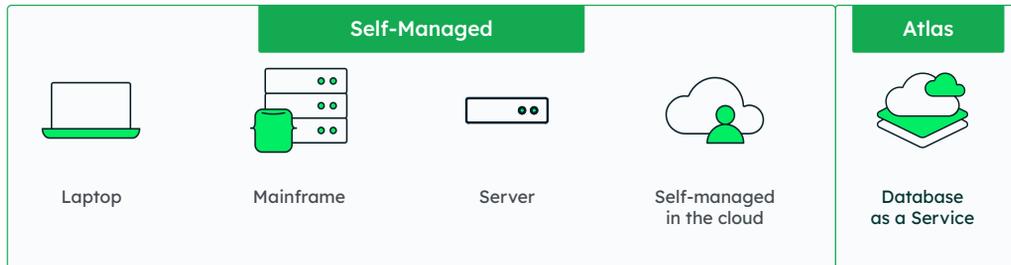
Continuous uptime with advanced automation to ensure performance, no matter the scale



The background of the slide is a dark green color. On the right side, there is a lighter green abstract shape that resembles a stylized leaf or a drop. In the top right corner of this shape, there is a small, light green leaf icon.

Storing the Data

Store the Data: Self-Managed or Atlas



Consistent developer experience

Same codebase, same APIs, same tools, wherever you run

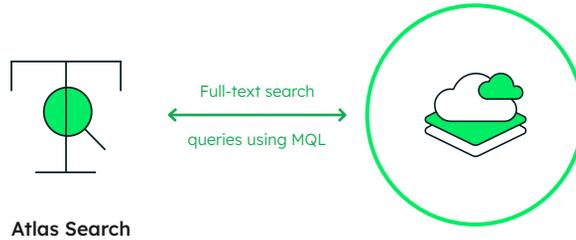
Here are the two approaches to storing your data, to clarify in either case you will need a database installation. This can either be self-managed or run in MongoDB's service, Atlas.

Self-managed means you can have the data on your laptop as part of your development environment, equally you can host it in your own server, mainframe, or yourself in a public cloud provider. The Ops Manager tool is recommended to manage, monitor, and backup these environments.

Atlas or MongoDB's "database as a service", is where MongoDB will manage and automate your deployment in one or many public cloud providers. These providers are currently Amazon's AWS, Microsoft's Azure, or Google's Cloud Platform. MongoDB and it's teams will manage and maintain your databases, in addition all of the same tooling and features available in Ops Manager are available in Atlas.



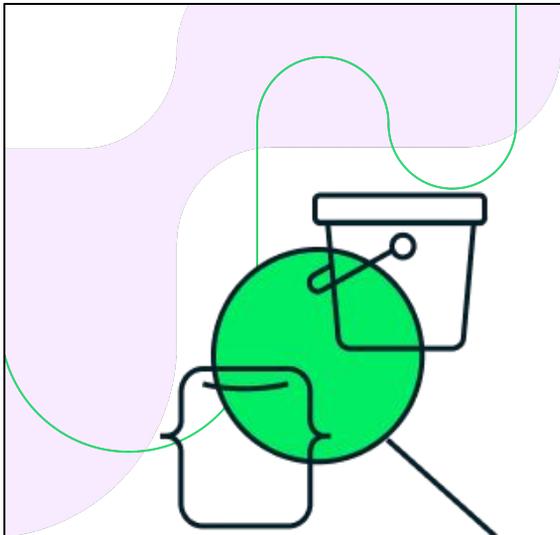
Store the Data: Atlas



Atlas Cloud Database
at the core, supporting the latest versions of MongoDB

Looking into the first aspect of how MongoDB is a data platform, specifically how the data is stored. At the core, Atlas runs the latest version of MongoDB. This ensure that it can avail any new features and the upgrades of versions are handled automatically by Atlas itself.

Atlas isn't just a database alone, it is a key element of the platform so it supports and runs other tooling such as Atlas Search. This utilises Apache's Lucene to provide fine grained full-text search queries using MongoDB query language (MQL).



The diagram features a central green circle. A white speech bubble is positioned to its left, and a white rectangular box with a small circle inside is positioned above it. A magnifying glass is shown over the bottom right of the green circle. A green line starts from the top of the green circle, loops around the top, and then connects to the top of the white box above it. The background consists of light purple and white abstract shapes. In the top right corner of the slide, there is a small black leaf icon.

Data Federation

- Cheaper storage (S3)
- Combine data across Atlas & S3 for queries
- Same query languages (MQL & AggFrmWrk)
- Federated queries
- Automatic data tiering / archiving

Data Lakes are increasingly more common, it is essentially a centralized storage point for your data, whether structured or unstructured.

You can query across data and combine the results from S3 and from Atlas databases.

It supports a variety of file formats including JSON , BSON , CSV, TSV, Avro, ORC, and Parquet. These can be queried using the same query language used with any MongoDB database, the MongoDB Query Language (MQL) and equally you can use the Aggregation Framework. The queries can be performed using the mongo shell, MongoDB Compass, or any MongoDB driver. These queries do not trigger any data movement or transformation of the data, the queries are performed against S3.

Federated queries allows you to query multiple databases and S3 in a single query

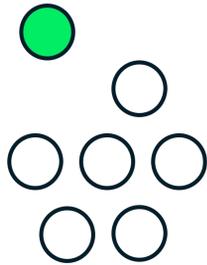
Accessing the Data





Designate read-only nodes

Separate operational and analytical queries to different nodes



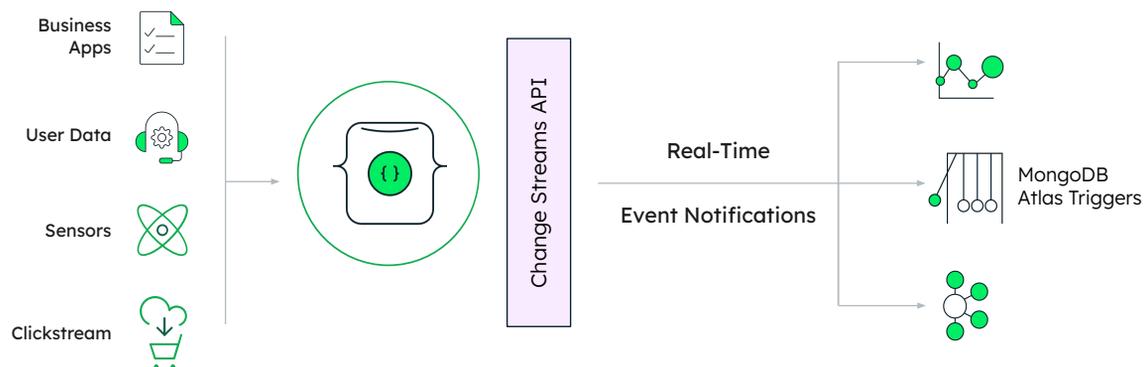
Workload Isolation

Moving to the second aspect of a data platform which is accessing the data. The first place to start with is workload isolation.

Being able to run operational workloads along with analytical workloads on the same platform is a big advantage as it means you do not need to provision a separate database or set of resources to service your analytical queries. This is done by designating a specific set of nodes to be read-only, which will service the analytical queries.

MongoDB Atlas provides one set of nodes which are focused on servicing your main traffic for your operational applications. These nodes continue to replicate any changes or updates to a second set of nodes. This second set of nodes serves your analytics queries. Your application can target through which of the two sets of nodes should service the request depending on the purpose of the request.

Access the Data: Change Streams



MongoDB as a data platform has another useful feature but that is not as well known. Change streams provide a mechanism for applications to subscribe to changes whether on a single collection, a database, or an entire deployment. These are streaming as they happen to the application so it can react immediately to them.

Real-time data changes occurring in a range of applications, such as business applications, user data, sensor data, or clickstream data.

Change Streams and specifically the Change Stream API allows for applications to access and consume these real-time event notifications. Change streams aren't just limited to your application as they underpin a number of other tools in the wider MongoDB ecosystem. Change Streams use the aggregation framework so it allows for additional filtering or transformations to be applied to the changes as they are sent out.

MongoDB Charts can use Change Streams as a data source which can additionally be used with the Charts Embedded SDK to include the functionality of both directly into your application.

MongoDB Atlas triggers, specifically the database triggers, are powered by Change Streams. It allows the database triggers to listen for changes in watched collections and map them to database events.

MongoDB's Kafka Connector also uses Change Streams. It passes these notifications

to Kafka so the details of the events can be stored in Kafka.

This functionality allows for reactivity to be built into your applications where any changes can be pushed from MongoDB to your application rather than your application having to regularly poll the database with additional queries to determine if there were any changes in terms of the data since the last time it checked.

Quiz



Quiz



Which of the following are true for MongoDB as a data platform?

Select all that apply.

- A. MongoDB does not support workload isolation
- B. Change Streams use the MongoDB Aggregation Framework

Quiz



Which of the following are true for MongoDB as a data platform?

Select all that apply.

- A. MongoDB does not support workload isolation
- B. Change Streams use the MongoDB Aggregation Framework

INCORRECT: MongoDB does not support workload isolation - Replica Set Tags are used in MongoDB within a replica set to support workload isolation.

CORRECT: Change Streams use the MongoDB Aggregation Framework - The Change Stream feature makes extensive use of the Aggregation Framework and it is what allows for changes to be made to the change stream event document prior to it being sent to a listening application.

Quiz



Which of the following are true for MongoDB as a data platform?

Select all that apply.

- A. MongoDB does not support workload isolation
- B. Change Streams use the MongoDB Aggregation Framework

This is incorrect. Replica Set Tags are used in MongoDB within a replica set to support workload isolation.

INCORRECT: MongoDB does not support workload isolation - Replica Set Tags are used in MongoDB within a replica set to support workload isolation.

Quiz



Which of the following are true for MongoDB as a data platform?

Select all that apply.

- A. MongoDB does not support workload isolation
- B. Change Streams use the MongoDB Aggregation Framework

This is correct. The Change Stream feature makes extensive use of the Aggregation Framework.

CORRECT: Change Streams use the MongoDB Aggregation Framework - The Change Stream feature makes extensive use of the Aggregation Framework and it is what allows for changes to be made to the change stream event document prior to it being sent to a listening application.

Processing the Data



Compass

Visual Data Explorer

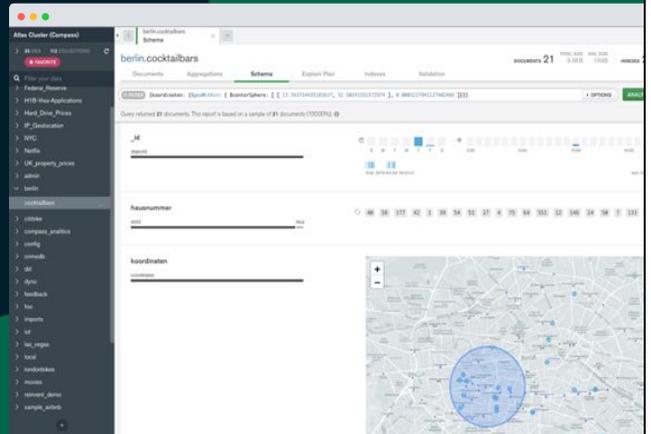
CRUD Operations

Visual Query Builder

Schema Analyzer

Schema Validation

Embedded Shell



Let's look at how the data can be processed.

The first tool within the data platform to introduce for process is MongoDB Compass, this is a visual data explorer. It runs as an application on your desktop or laptop and can connect to a local or to a remote MongoDB deployment. It lets you view the data with many different visualizations.

It allows for creating/reading/updating/deleting data or CRUD operations, on the documents in your collections.

It also includes visual query builders to help you construct complex queries for MongoDB Query Language and for MongoDB Aggregations.

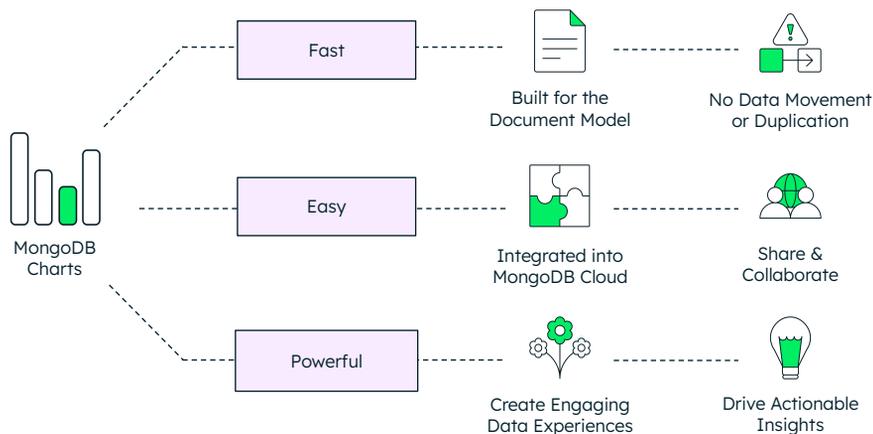
MongoDB Compass also provides a schema analyzer which is the best way to identify schemas for a given dataset/datasets.

Schema validation supports JSON Schema format which you can edit/configure in Compass and look at live document previews of passing and of failing documents.

An embedded shell is also included in MongoDB Compass. This is a complete and functional Javascript environment that allows for interacting with MongoDB deployments. The MongoDB embedded shell allows for testing queries and for operations to be performed directly against your database within MongoDB Compass.



Process the Data: MongoDB Charts



MongoDB Charts is the best way to visualize MongoDB data. With Charts...

- It's **fast** to visualize data
- It's **easy** to get started and share data
- And Charts is **powerful** to create engaging data experiences that drive actionable insights for internal and external users.

It is an analytical tool designed specifically for the document model, this provides it with additional flexibility when compared to using analytical tools designed for relational models with MongoDB data.

Charts versatility from Business Intelligence tool capabilities to embedded analytics with its quick and easy visualizations for the document model, help drastically **improve developer productivity**.

- **Data visualized in just minutes.** Charts is built specifically for the document model, no ETL, no time loss to data manipulation or duplication required to visualize rich JSON data.
- **Data shared quickly and easily.** Integrated with Atlas as part of the MongoDB Cloud Data Platform, there's no setup for Charts. Go straight from Atlas into Charts to start sharing and collaborating on your MongoDB data.
- **Powerful capabilities to create engaging data experiences for actionable insights for the right user in the right context.** Whether via dashboard or a rich embedding SDK with authentication features, create engaging data-driven

- experiences for everyone. From an internal dashboard to a customer-facing app, these experiences drive actionable insights across a wide spectrum of users.

Quiz



Quiz



Which of the following are true for MongoDB as a data platform?

Select all that apply.

- A. Compass provides a local tool to process MongoDB data
- B. Compass does not provide an embedded MongoDB Shell
- C. Charts does not allow for external embedding into other applications
- D. Business Intelligence Connector works with any ODBC compliant tool

Quiz



Which of the following are true for MongoDB as a data platform?

Select all that apply. More than 1 answer choice can be correct.

- A. Compass provides a local tool to process MongoDB data
- B. Compass does not provide an embedded MongoDB Shell
- C. Charts does not allow for external embedding into other applications
- D. Business Intelligence Connector works with any ODBC compliant tool

CORRECT: Compass provides a local tool to process MongoDB data - Compass provides a GUI which you can use to visualise and connect to MongoDB deployment whether remote or local.

INCORRECT: Compass does not provide an embedded MongoDB Shell - Compass in recent versions does include an embedded MongoDB Shell/Javascript environment.

INCORRECT: Charts does not allow for external embedding into other applications - Charts does allow for a variety of external embedding whether using a iFrame or via its Javascript SDK for embedding data externally.

CORRECT: Business Intelligence Connector works with any ODBC compliant tool - The Business Intelligence Connector is designed to be usable with a range of tools as it interprets SQL, translates it to MQL and queries the MongoDB database, it then reformats the results back to SQL and sends these back to whatever ODBC compliant tool is connected.

Quiz



Which of the following are true for MongoDB as a data platform?

Select all that apply. More than 1 answer choice can be correct.

- A. Compass provides a local tool to process MongoDB data
- B. Compass does not provide an embedded MongoDB Shell
- C. Charts does not allow for external embedding into other applications
- D. Business Intelligence Connector works with any ODBC compliant tool

This is correct. Compass provides a GUI which you can use to visualize and connect to MongoDB deployment whether remote or local.

CORRECT: Compass provides a local tool to process MongoDB data - Compass provides a GUI which you can use to visualise and connect to MongoDB deployment whether remote or local.

Quiz



Which of the following are true for MongoDB as a data platform?

Select all that apply. More than 1 answer choice can be correct.

- A. Compass provides a local tool to process MongoDB data
- B. Compass does not provide an embedded MongoDB Shell
- C. Charts does not allow for external embedding into other applications
- D. Business Intelligence Connector works with any ODBC compliant tool

*This is incorrect.
Compass in recent
versions does include an
embedded MongoDB
Shell/Javascript
environment.*

INCORRECT: Compass does not provide an embedded MongoDB Shell - Compass in recent versions does include an embedded MongoDB Shell/Javascript environment.

Quiz



Which of the following are true for MongoDB as a data platform?

Select all that apply. More than 1 answer choice can be correct.

- A. Compass provides a local tool to process MongoDB data
- B. Compass does not provide an embedded MongoDB Shell
- C. Charts does not allow for external embedding into other applications
- D. Business Intelligence Connector works with any ODBC compliant tool

This is incorrect. Charts does allow for a variety of external embedding whether using a iFrame or via its Javascript SDK for embedding data externally.

INCORRECT: Charts does not allow for external embedding into other applications - Charts does allow for a variety of external embedding whether using a iFrame or via its Javascript SDK for embedding data externally.

Quiz



Which of the following are true for MongoDB as a data platform?

Select all that apply. More than 1 answer choice can be correct.

- A. Compass provides a local tool to process MongoDB data
- B. Compass does not provide an embedded MongoDB Shell
- C. Charts does not allow for external embedding into other applications
- D. Business Intelligence Connector works with any ODBC compliant tool

This is correct. It is usable with a range of tools as it interprets SQL, translates it to MQL and queries the MongoDB database, it then reformats the results back to SQL and then back to the ODBC compliant tool.

CORRECT: BI Connector works with any ODBC compliant tool - It is usable with a range of tools as it interprets SQL, translates it to MQL and queries the MongoDB database, it then reformats the results back to SQL and then back to the ODBC compliant tool.

MongoDB is a data platform



Storing the data:
Self-Managed and Atlas

Accessing the data: Workload
isolation and Change Streams

Processing the data:
Compass, Charts, and the
Business Intelligence
Connector

We started this lesson highlighting that MongoDB is not just a database product, it is a complete ecosystem.

We pointed out that it is better viewed as a data platform with a number of tools that help with various aspects.

In storing the data, we looked at self-managed MongoDB, MongoDB Atlas which is the SaaS or DBaaS offering from MongoDB and includes features such as Atlas Search.

The slide features a dark teal background with a lighter teal abstract shape on the right side. A small green leaf icon is positioned in the top right corner. The text 'MongoDB Cloud' is centered in a white serif font.

MongoDB Cloud

Let's focus in the next section on how MongoDB Cloud is the foundation and underpinning of the MongoDB Data Platform. MongoDB Cloud encompasses the pieces of the data platform that are provided in a SaaS manner and where all the operational plus maintenance aspects are managed completely by MongoDB for the user.



Investing in a cloud-first, modern database

	2016	2017	2018	2019	2020	2021
MongoDB version	3.4	3.6	4.0	4.2	4.4	5.0
	Graph processing Faceted navigation Decimal data type Tunable consistency	Change streams Retryable writes Causal consistency JSON schema	Multi-doc transactions Type conversions Non-block sec. reads	Distributed transactions Client-side encryption Global PIT reads Materialized views Wildcard indexes	Refinable shard keys Hedged and mirrored reads Union aggregation stage Custom agg expressions Compound hash shard keys	Native time series support Live resharding Versioned API Multi-Cloud FLE Majority write concern default
Cloud releases	← CONTINUOUS DELIVERY →					
	MongoDB Atlas on AWS VPC Peering support	Cross-region replication Atlas on Google Cloud Atlas on Azure Storage auto-scaling Performance Advisor	mLab acquisition Global clusters Rich database auditing LDAP integration BYO encryption keys MongoDB Charts	Atlas Search Compute auto-scaling Analytics nodes Snapshot backups Query profiler AWS PrivateLink support	Multi-cloud clusters Online Archive Azure Private Link Multiple connection strings Cross-org billing Schema advice AWS IAM authentication GraphQL API	Edge-to-cloud sync Serverless instances Custom archiving rules Multi-cloud enhancements ... and more coming soon!

We've invested as much if not more in building out our cloud products and capabilities alongside continuous innovation on the core database software.

A major benefit of taking a cloud-first approach is the ability to change to a continuous delivery model with a cycle of three weeks rather than the prior annual release cycle for the software. This means users see new features far sooner and it allows for these to be regularly pushed out, further improving the developer experience.

A Cloud-First, Modern Database



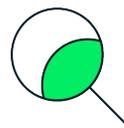
Global and
Multi-cloud
Reach



Performance
Advisory



Online Archive

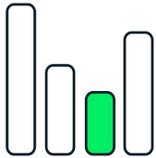


Atlas Search

MongoDB Cloud platform offers true multi-cloud capabilities as well as global capacities. It bakes in automatic instrumentation with performance advisory to dynamically monitor your workloads and make actionable index suggestions. In terms of easily archiving data, the online archive feature allows for colder storage to be used with defined rules to migrate data from your production clusters to cold storage/analytical storage (typically S3). Atlas Search allows for any application to easily add search functionality backed by Lucene via Atlas without having to stand up and maintain a separate search cluster, this means your data can be stored, used and searched in Atlas easily. Lucene allows for highly fine grained searches to be applied on your MongoDB database.



A Cloud-First, Modern Database



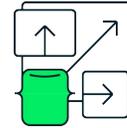
Atlas Charts



Global Clusters



Continuous Cloud Backup



Atlas Serverless

Atlas Charts allows for rapid data visualization within the environment without having to export to another separate data visualization tool. Global clusters allows for all users and applications regardless of their geographical location to have the optimal experience and latency, not to mention easily managing any kind of data compliance such as Europe's GDPR. In terms of backup, Atlas offers a continuous backup feature which allows you to define your own backup and retention policies, this can satisfy a recovery point objective as low as 60 seconds. Atlas Serverless offers a option for developers using the Serverless paradigm to use MongoDB's Atlas Serverless instances to easily scale up and down to their immediate needs of the application/traffic.

A Cloud-First, Modern Database



Data
Federation



Kubernetes
Operator



BI
Connector

Data Federation provides a query engine that allows for the federation of queries across a variety of different sources of data as well as the ability to move data between these sources. MongoDB's Atlas Operator allows you to simplify deployment, management and scaling of your Atlas clusters in Kubernetes. The BI Connector allows for a wide range of analytical tooling designed for relational/SQL databases to easily use MongoDB databases so you can use advanced visualization from the likes of Tableau to represent and interpret your data.

Quiz



Quiz



Which of the following are true for MongoDB Cloud within the data platform? *Select all that apply.*

- A. MongoDB Atlas Search provides granular text searches
- B. Federated Data cannot be queried, data can only be archived to it
- C. MongoDB Charts only uses data from MongoDB Atlas collections

Quiz



Which of the following are true for MongoDB Cloud within the data platform? *Select all that apply.*

- A. MongoDB Atlas Search provides granular text searches
- B. Federated Data cannot be queried, data can only be archived to it
- C. MongoDB Charts only uses data from MongoDB Atlas collections

CORRECT: MongoDB Atlas Search provides granular text searches - MongoDB Atlas Search utilizes Apache Lucene to provide granular text searches. This service also manages automatically keeping the indexes updated as data changes.

INCORRECT: Federated Data cannot be queried, data can only be archived to it - Federated Data can be queried, it's purpose is to be both an archive as well providing the ability to query the data present there. It does not support indexes.

CORRECT: MongoDB Charts only uses data from MongoDB Atlas collections - Charts are designed to work directly from MongoDB Atlas databases and collections, it is not a standalone product but sits within MongoDB Cloud as a feature (visualization).

Quiz



Which of the following are true for MongoDB Cloud within the data platform? *Select all that apply.*

- A. MongoDB Atlas Search provides granular text searches *This is correct. MongoDB Atlas Search utilizes Apache Lucene to provide granular text searches. This service also manages automatically keeping the indexes updated as data changes.*
- B. Federated Data cannot be queried, data can only be archived to it
- C. MongoDB Charts only uses data from MongoDB Atlas collections

CORRECT: MongoDB Atlas Search provides granular text searches - MongoDB Atlas Search utilizes Apache Lucene to provide granular text searches. This service also manages automatically keeping the indexes updated as data changes.

Quiz



Which of the following are true for MongoDB Cloud within the data platform? *Select all that apply.*

- A. MongoDB Atlas Search provides granular text searches
- B. Federated Data cannot be queried, data can only be archived to it
- C. MongoDB Charts only uses data from MongoDB Atlas collections

*This is incorrect.
Federated Data can be queried, it's purpose is to be both an archive as well providing the ability to query the data present there. It does not support indexes.*

INCORRECT: Federated Data cannot be queried, data can only be archived to it - Federated Data can be queried, it's purpose is to be both an archive as well providing the ability to query the data present there. It does not support indexes.

Quiz



Which of the following are true for MongoDB Cloud within the data platform? *Select all that apply.*

- A. MongoDB Atlas Search provides granular text searches
- B. Federated Data cannot be queried, data can only be archived to it
- C. MongoDB Charts only uses data from MongoDB Atlas collections

This is correct. Charts are designed to work directly from MongoDB Atlas databases and collections, it is not a standalone product but sits within MongoDB Cloud as a feature (visualization).

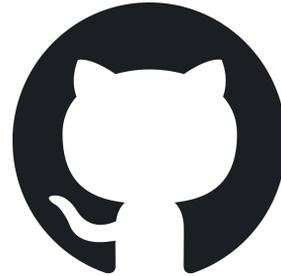
CORRECT: MongoDB Charts only uses data from MongoDB Atlas collections - Charts are designed to work directly from MongoDB Atlas databases and collections, it is not a standalone product but sits within MongoDB Cloud as a feature (visualization).

Continue Learning!



[MongoDB University](#) has free self-paced courses and labs ranging from beginner to advanced levels.

GitHub Student Developer Pack



Sign up for the [MongoDB Student Pack](#) to receive \$50 in Atlas credits and free certification!

This concludes the material for this lesson. However, there are many more ways to learn about MongoDB and non-relational databases, and they are all free! Check out [MongoDB's University](#) page to find free courses that go into more depth about everything MongoDB and non-relational. For students and educators alike, MongoDB for Academia is here to offer support in many forms. Check out our [educator resources](#) and join the Educator Community. Students can receive \$50 in Atlas credits and free certification through the [GitHub Student Developer Pack](#).