

MONGO DB

CLASS 1: INTRODUCTION

INTRODUCTION:

MongoDB is a popular open-source NoSQL (non-relational) database system designed for storing and retrieving data in a flexible and scalable manner. It's a powerful choice for modern applications that deal with:

- Large and complex data structures: MongoDB stores data in flexible documents, similar to JSON objects. Documents can contain various data types, including nested structures and arrays, making them ideal for representing real-world entities and their relationships.
- Evolving data models: Unlike relational databases with rigid schemas, MongoDB allows documents within a collection (analogous to a table) to have different structures. This flexibility is crucial for applications where data models are likely to change over time.
- High performance and scalability: MongoDB is built for horizontal scaling, meaning you can easily add more servers to handle growing data volumes or increased query load. It offers features like sharding to distribute data efficiently across a cluster.
- Data Model and Schema Analysis: This report delves into the design of your MongoDB data model. It explains how collections and documents are structured, how effectively they represent your data, and potential areas for optimization.
- Migration Planning: If you're considering migrating data to MongoDB, this report would assess the feasibility and potential challenges. It might compare data models, analyze query performance implications, and outline the migration process.
- Security Assessment: This report focuses on the security posture of your MongoDB deployment. It would evaluate access controls, encryption strategies, and any potential security vulnerabilities.
- Evaluating Performance: This type of report focuses on measuring the query performance, throughput, and scalability of MongoDB for a specific application. It might analyze factors like average query time, latency, and how well MongoDB handles increasing data volumes or workloads.
- Comparing MongoDB to Relational Databases: This report pits MongoDB against a relational database for a particular use case. It would highlight the strengths and weaknesses of each approach, considering factors like data model complexity, query patterns, and scalability requirements.

HOW MONGO DB WORKS:

Now, we will see how actually thing happens behind the scene. As we know that MongoDB is a database server and the data is stored in these databases. Or in other words, MongoDB environment gives you a server that you can start and then create multiple databases on it using MongoDB.

Because of its NoSQL database, the data is stored in the collections and documents. Hence the database, collection, and documents are related to each other.

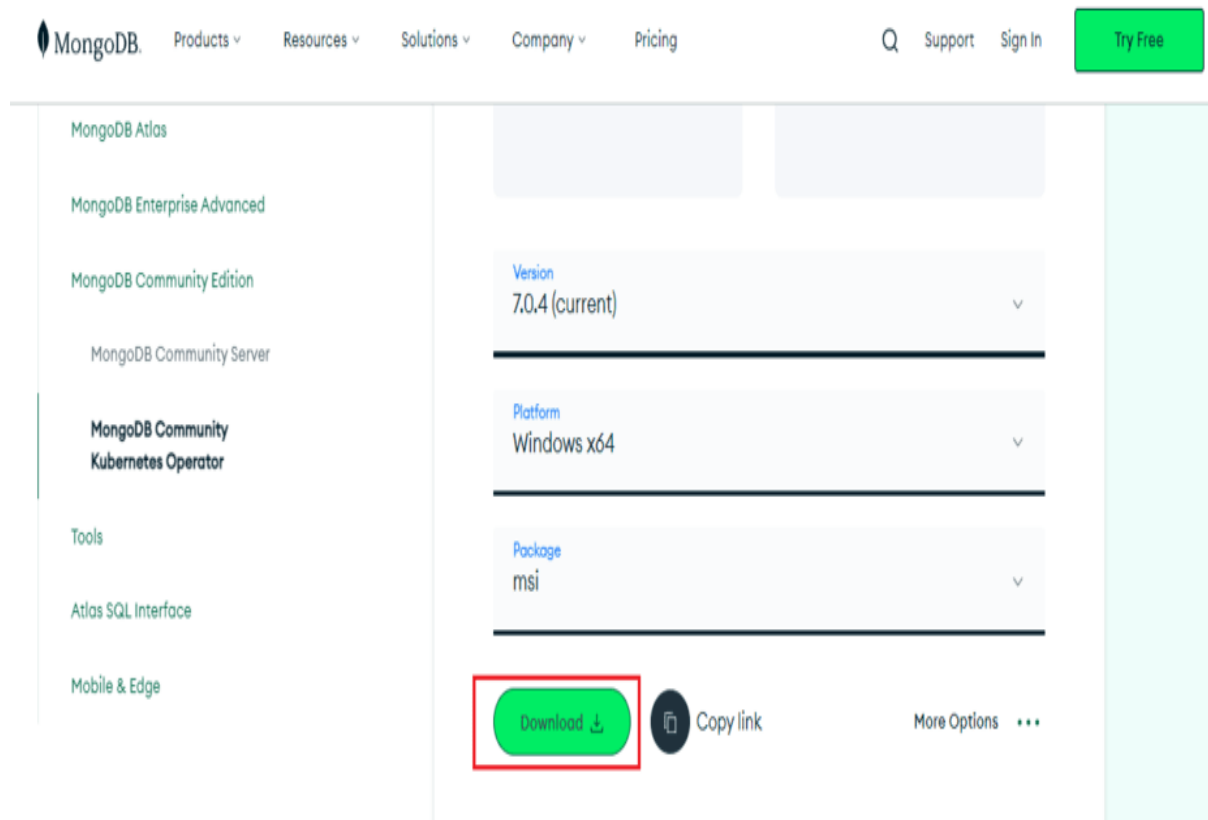
- The MongoDB database contains collections just like the MYSQL database contains tables. You are allowed to create multiple databases and multiple collections.
- Now inside of the collection we have documents. These documents contain the data we want to store in the MongoDB database and a single collection can contain multiple documents and you are schema-less means it is not necessary that one document is similar to another.
- The documents are created using the fields. Fields are key-value pairs in the documents, it is just like columns in the relation database. The value of the fields can be of any BSON data types like double, string, boolean, etc.
- The data stored in the MongoDB is in the format of BSON documents. Here, BSON stands for Binary representation of JSON documents. Or in other words, in the backend, the MongoDB server converts the JSON data into a binary form that is known as BSON and this BSON is stored and queried more efficiently.
- In MongoDB documents, you are allowed to store nested data. This nesting of data allows you to create complex relations between data and store them in the same document which makes the working and fetching of data extremely efficient as compared to SQL. In SQL, you need to write complex joins to get the data from table 1 and table 2. The maximum size of the BSON document is 16MB.

ADVANTAGES OF MONGO DB:

- It is a schema-less NoSQL database. You need not to design the schema of the database when you are working with MongoDB.
- It does not support join operation.
- It provides great flexibility to the fields in the documents.
- It contains heterogeneous data.
- It provides high performance, availability, scalability.
- It supports Geospatial efficiently.
- It is a document oriented database and the data is stored in BSON documents.
- It also supports multiple document ACID transition(string from MongoDB 4.0).
- It does not require any SQL injection.
- It is easily integrated with Big Data Hadoop

INSTALLING MONGO DB:

Step 1: Go to the [MongoDB Download Center](#) to download the MongoDB Community Server.



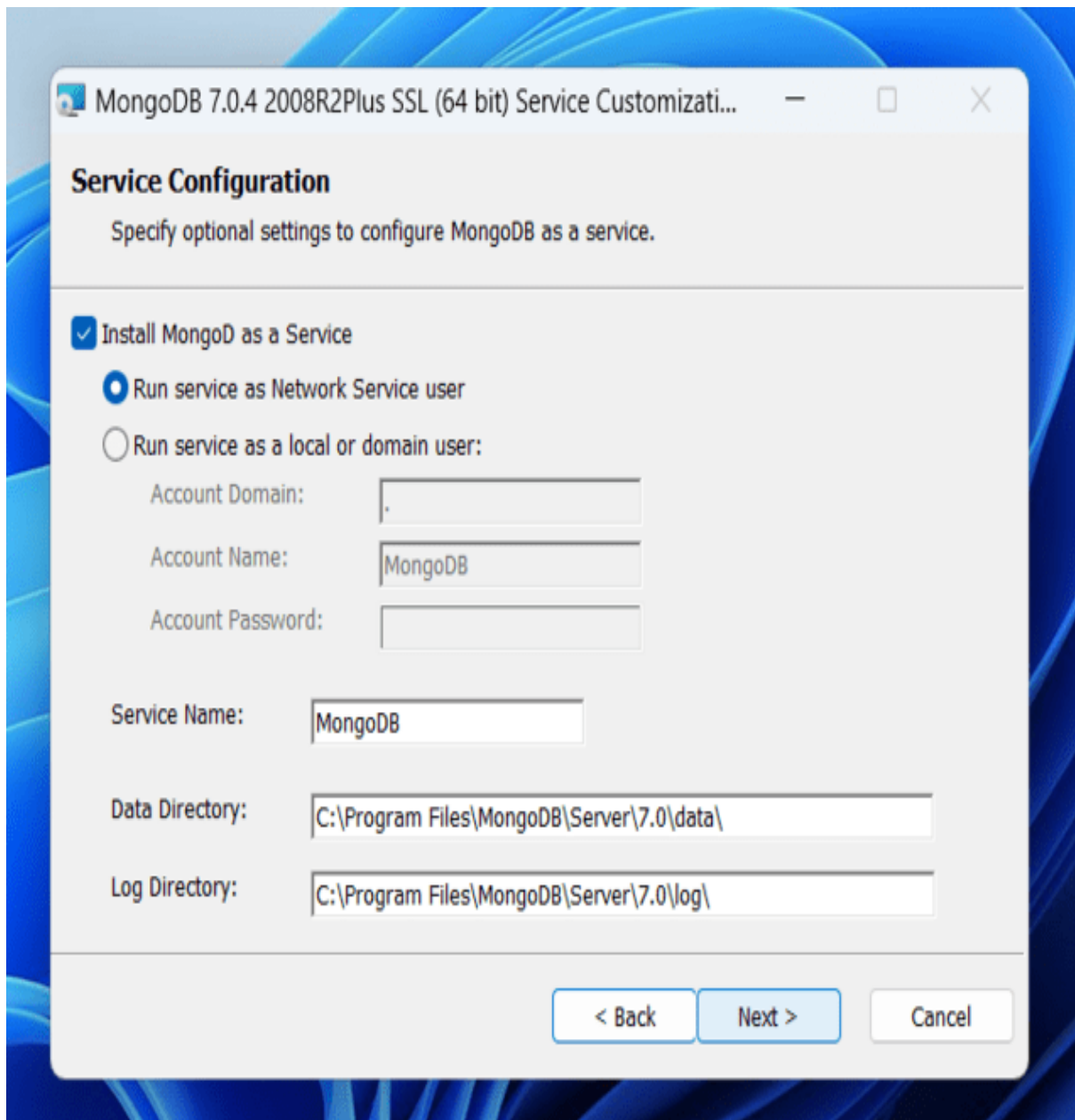
Here, You can select any version, Windows, and package according to your requirement. For Windows, we need to choose:

- Version: 7.0.4
- OS: Windows x64
- Package: msi

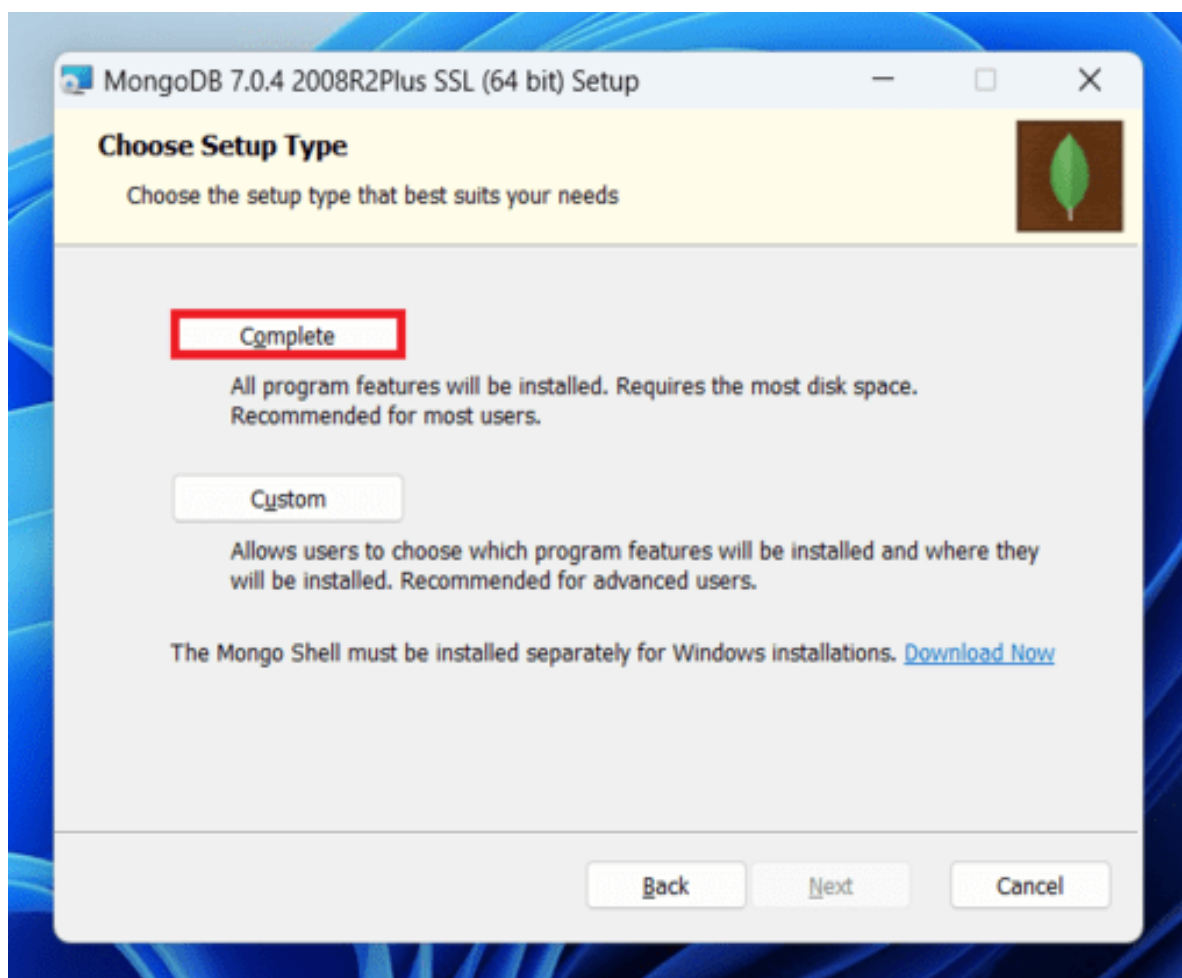
Step 2: When the download is complete open the msi file and click the *next button* in the startup screen:



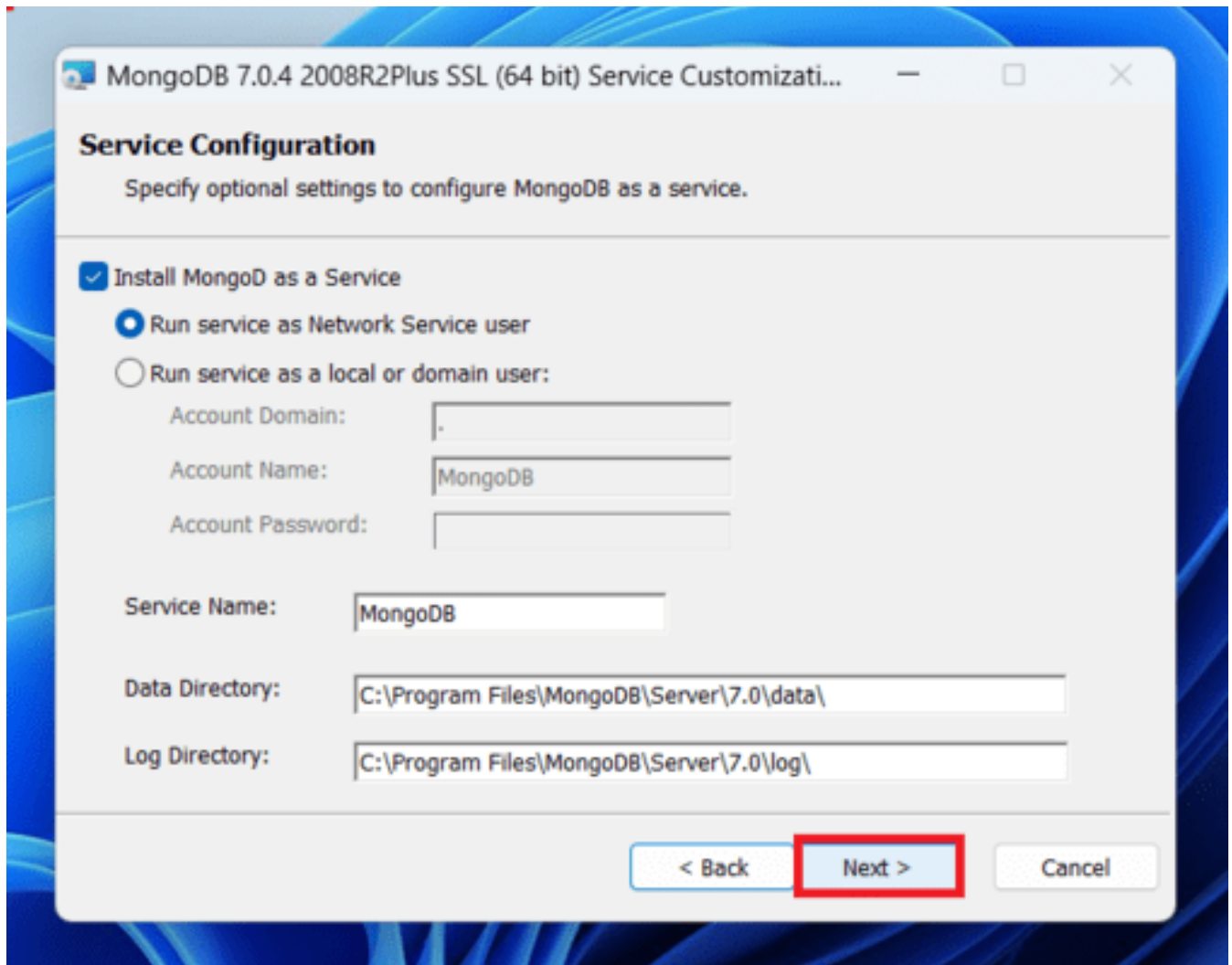
Step 3: Now accept the End-User License Agreement and click the next button:



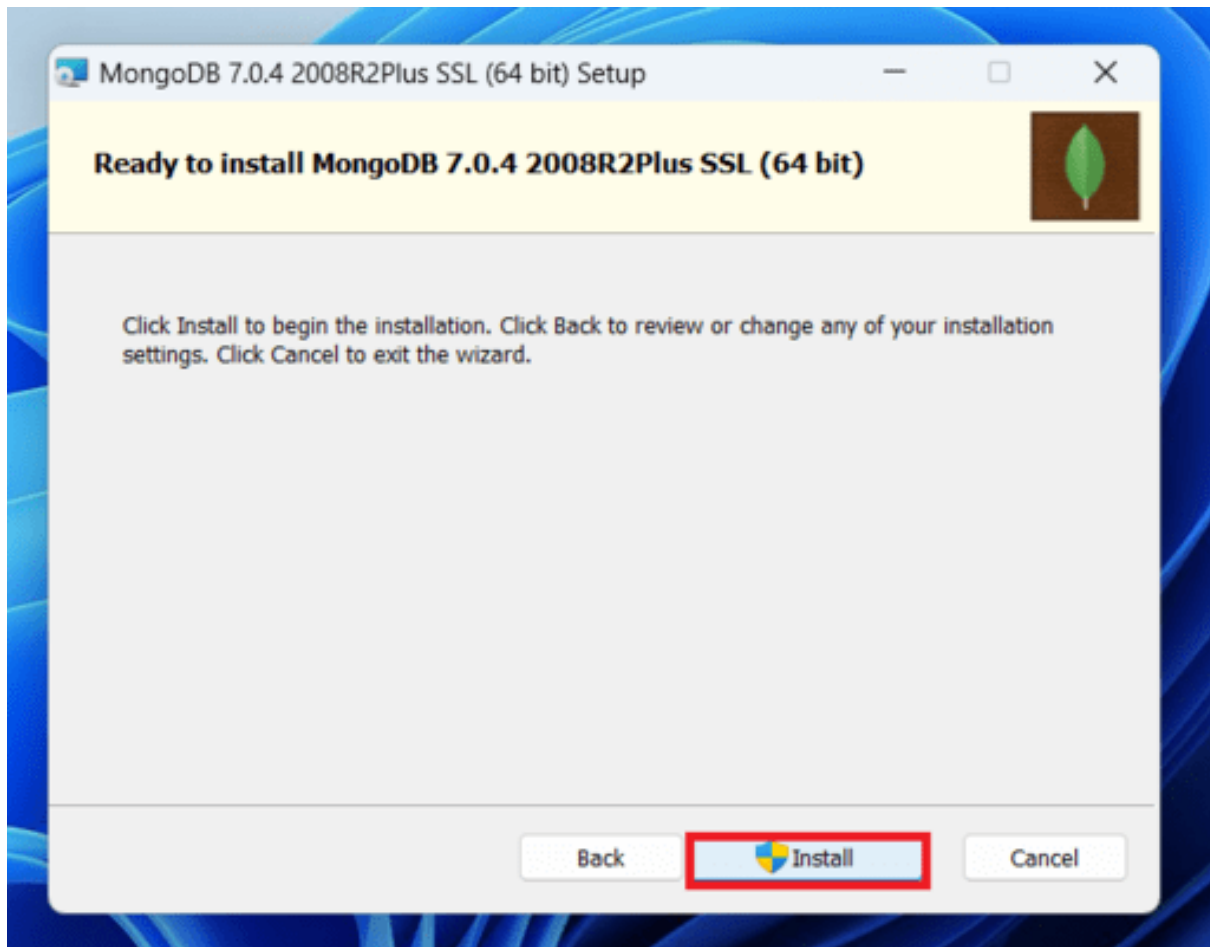
Step 4: Now select the *complete option* to install all the program features. Here, if you can want to install only selected program features and want to select the location of the installation, then use the *Custom option*:



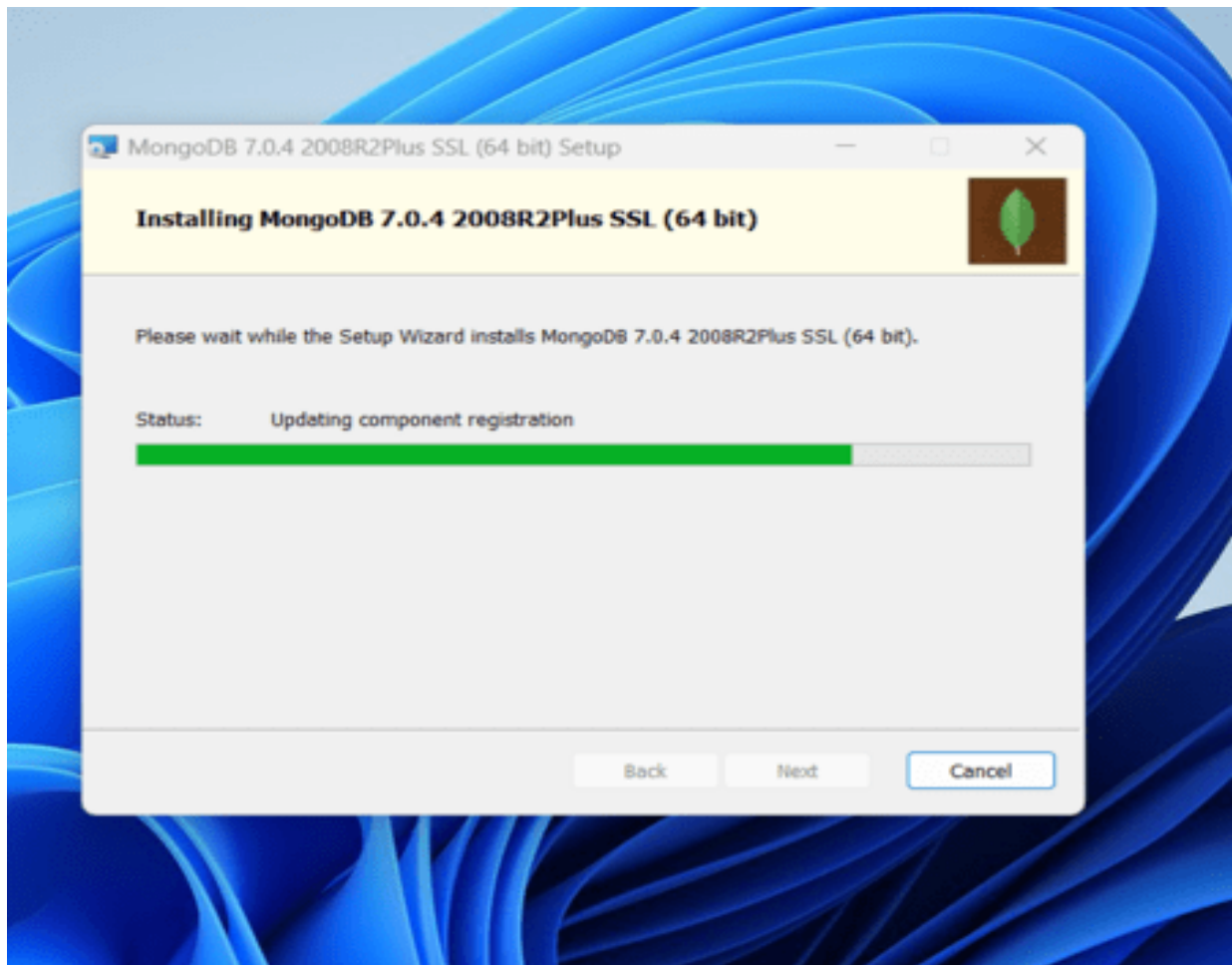
Step 5: Select “Run service as Network Service user” and copy the path of the data directory. Click Next:



Step 6: Click the *Install button* to start the MongoDB installation process:

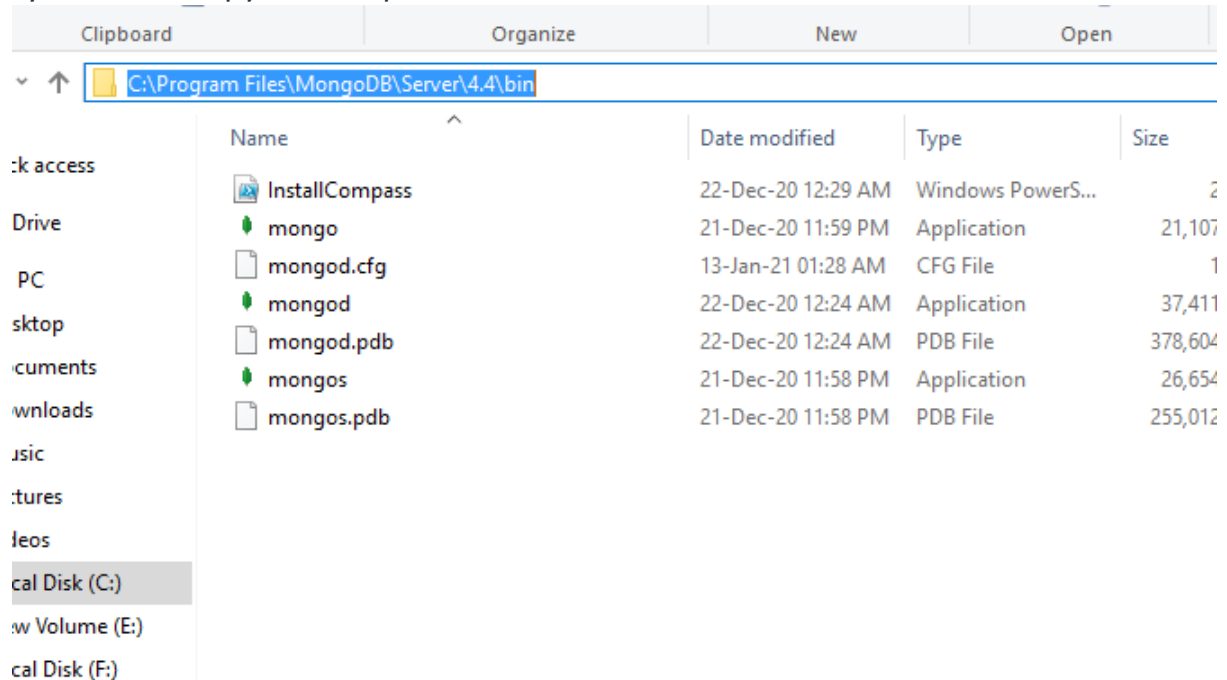


Step 7: After clicking on the install button installation of MongoDB begins:

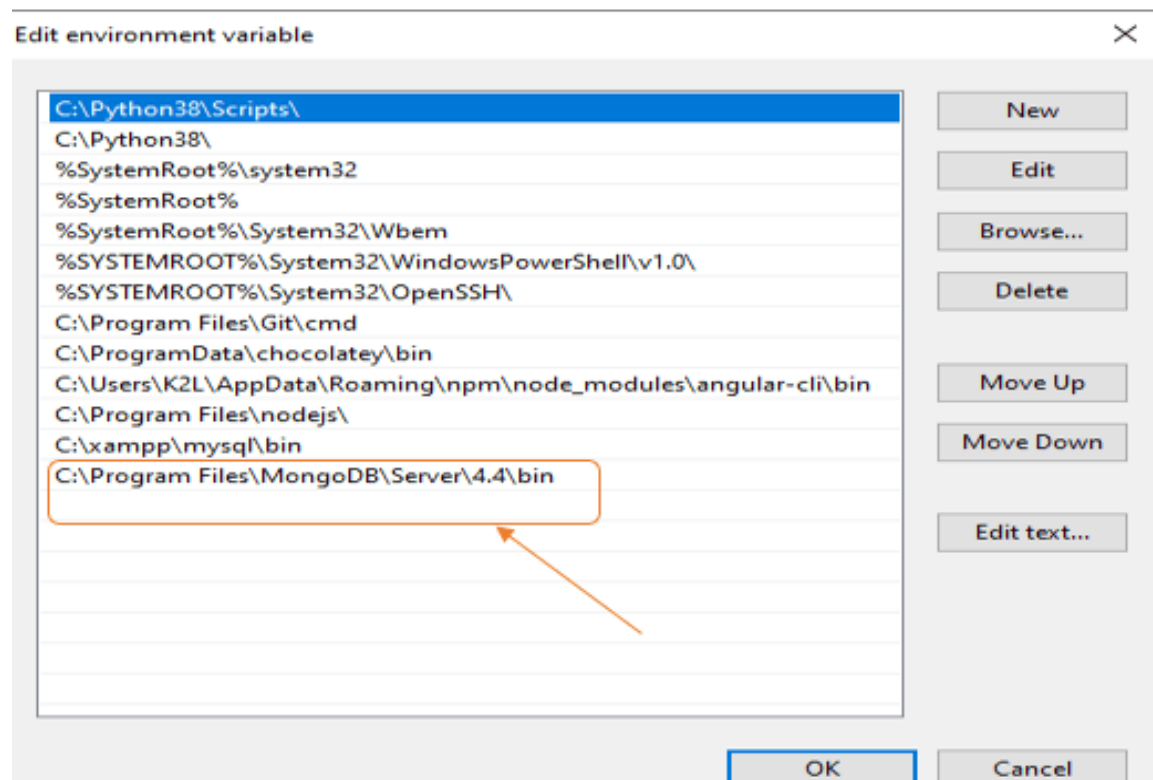


Step 8: Now click the ***Finish button*** to complete the MongoDB installation process:

Step 9: Now we go to the location where MongoDB installed in step 5 in your system and copy the bin path:



Step 10: Now, to create an environment variable open system properties >> Environment Variable >> System variable >> path >> Edit Environment variable and paste the copied link to your environment system and click Ok:



Step 11: After setting the environment variable, we will run the MongoDB server, i.e. mongod. So, open the command prompt and run the following command:

```
mongod
```

When you run this command you will get an error i.e. *C:/data/db/ not found*.

Step 12: Now, Open C drive and create a folder named “data” inside this folder create another folder named “db”. After creating these folders. Again open the command prompt and run the following command:

```
mongod
```

Now, this time the MongoDB server(i.e., mongod) will run successfully.

```
C:\Users\Nikhil Chhipa>mongod
{"t":{"$date":"2021-01-31T00:56:54.081+05:30"},"s":"I", "c":"CONTROL", "id":23285, "ctx"
ify --sslDisabledProtocols 'none'"}
{"t":{"$date":"2021-01-31T00:56:54.087+05:30"},"s":"W", "c":"ASIO", "id":22601, "ctx"
}
{"t":{"$date":"2021-01-31T00:56:54.088+05:30"},"s":"I", "c":"NETWORK", "id":4648602, "ctx"
{"t":{"$date":"2021-01-31T00:56:54.090+05:30"},"s":"I", "c":"STORAGE", "id":4615611, "ctx"
bPath":"C:/data/db/","architecture":"64-bit","host":"DESKTOP-L9MUQ7N"}}
{"t":{"$date":"2021-01-31T00:56:54.090+05:30"},"s":"I", "c":"CONTROL", "id":23398, "ctx"
rgetMinOS":"Windows 7/Windows Server 2008 R2"}}
{"t":{"$date":"2021-01-31T00:56:54.090+05:30"},"s":"I", "c":"CONTROL", "id":23403, "ctx"
gitVersion":"913d6b62acfb344dde1b116f4161360acd8fd13","modules":[],"allocator":"tcmalloc",
}}}}
{"t":{"$date":"2021-01-31T00:56:54.090+05:30"},"s":"I", "c":"CONTROL", "id":51765, "ctx"
ndows 10","version":"10.0 (build 14393)"}
{"t":{"$date":"2021-01-31T00:56:54.090+05:30"},"s":"I", "c":"CONTROL", "id":21951, "ctx"
{"t":{"$date":"2021-01-31T00:56:54.157+05:30"},"s":"I", "c":"STORAGE", "id":22270, "ctx"
:{"dbpath":"C:/data/db/","storageEngine":"wiredTiger"}}
{"t":{"$date":"2021-01-31T00:56:54.158+05:30"},"s":"I", "c":"STORAGE", "id":22315, "ctx"
ize=1491M,session_max=33000,eviction=(threads_min=4,threads_max=4),config_base=false,statist
le_manager=(close_idle_time=100000,close_scan_interval=10,close_handle_minimum=250),statisti
ess],"}
{"t":{"$date":"2021-01-31T00:56:54.395+05:30"},"s":"I", "c":"STORAGE", "id":22430, "ctx"
95788][3708:140713908197088], txn-recover: [WT_VERB_RECOVERY_PROGRESS] Recovering log 20 thr
{"t":{"$date":"2021-01-31T00:56:54.631+05:30"},"s":"I", "c":"STORAGE", "id":22430, "ctx"
```

RUN MONGO SHELL:

Step 13: Now we are going to connect our server (mongod) with the mongo shell. So, keep that mongod window and open a new command prompt window and write **mongo**. Now, our mongo shell will successfully connect to the mongod.

Important Point: Please do not close the mongod window if you close this window your server will stop working and it will not be able to connect with the mongo shell.

```
C:\Users\Nikhil Chhipa>mongo
MongoDB shell version v4.4.3
connecting to: mongodb://127.0.0.1:27017/?compressors=disabled&gssapiServiceName=mongodb
Implicit session: session { "id" : UUID("96cca5da-dc9f-4a40-aabb-732ee37600c0") }
MongoDB server version: 4.4.3
---
The server generated these startup warnings when booting:
  2021-01-28T20:56:52.570+05:30: Access control is not enabled for the database. Read and write access
configuration is unrestricted
---
---
  Enable MongoDB's free cloud-based monitoring service, which will then receive and display
metrics about your deployment (disk utilization, CPU, operation statistics, etc).

  The monitoring data will be available on a MongoDB website with a unique URL accessible to you
and anyone you share the URL with. MongoDB may use this information to make product
improvements and to suggest MongoDB products and deployment options to you.

  To enable free monitoring, run the following command: db.enableFreeMonitoring()
  To permanently disable this reminder, run the following command: db.disableFreeMonitoring()
---
>
```

Now, you are ready to write queries in the mongo Shell.

WHAT IS DATA?

DATA is statically raw and unprocessed information. For example – name, class, marks, etc. In computer language, a piece of information that can be translated into a form for efficient movement and processing is called data. Data is interchangeable information.

WHAT IS DATA BASE?

DATABASE is a collection of data that is organized, which is also called structured data. It can be accessed or stored in a computer system.

- Database is a container for collections.
- Each database gets its own set of files.
- A single MongoDB server has multiple databases.

DATA TYPE:

Databases can hold various kinds of information, including text, numbers, images, videos and more.

SQL and NO-SQL:

SQL: SQL stands for Structured Query Languages. To access a database, we use SQL statements.

NO-SQL: NO-SQL databases (also known as “not only SQL”) are non tabular databases that store data differently than traditional relational databases. SQL

SQL vs NoSQL IN MONGO DB:

Feature	SQL (Relational Databases)	MongoDB (NoSQL)
Data Model	Structured tables with rows and columns (fixed schema)	Flexible documents (JSON-like) with dynamic schema
Schema	Predefined schema enforced for all data	Flexible schema, documents within a collection can have different structures
Relationships	Defined through foreign keys and joins between tables	Embedded documents or references within documents
Scaling	Primarily vertical scaling (upgrades)	Horizontal scaling (adding more servers)
Queries	Uses SQL for querying data	Uses its own query language specific to document structure
ACID Properties	Typically enforces ACID properties (Atomicity, Consistency, Isolation, Durability)	Often prioritizes Availability over Consistency (following the CAP theorem)

Choosing Between SQL and MongoDB

The choice between SQL and MongoDB depends on your specific needs:

USE SQL IF:

You have a well-defined and stable data model.

Your application relies on complex queries involving joins across multiple tables.

Maintaining strict data consistency is critical.

USE MONGO DB IF:

You have a flexible or evolving data model.

You need high performance for large datasets or complex documents.

Horizontal scalability is a major requirement.

