

Design a MongoDB database



[HTTPS://DOCS.MONGODB.ORG/MANUAL/](https://docs.mongodb.org/manual/)

Logical database design

In the NoSQL paradigm, you start by identifying the work or the queries that the application needs to do. These queries guide how you map your conceptual design to a MongoDB logical design.

What do you need to know about the app?

Usage pattern of data

How your application accesses the data

Which operations are time critical for your application

Ratio between read operations and write operations

You want to keep the frequently accessed data in RAM to ensure a quick retrieval time.

MongoDB refers to the working set of an application as the total data required for the application during “normal” operations

MongoDB logical database design

When converting your conceptual design to a Mongo database, you do not automatically create a collection for each entity.

In Mongo DB you can choose to **embed** related objects within one another or **reference** them by ID

You use your application functionality to determine whether to embed or reference. If you never allow the user to query an object by itself, you should consider embedding it.

MongoDB limits a document size to 16M

Structural constraints can also guide you

1:1 objects should be embedded or merged.

1:* Structural constraints

Mongo DB further separate 1:* into 1:few, 1:some and 1:many.

If you can put a limit on the many and it is small, consider an embedded array of objects for representation .

If the many really is medium sized (100s) then use an embedded array of references. Still embedding in the parent – just embedding the reference

If the many side is unlimited then, create 2 collections where each child references a parent object (_id values). This is known as a parent connection.

Many to many cardinality

A Many-to-Many relationship refers to the relationship between two entities A and B, where both sides can have one or more links to the other. In relational databases, these cases are modelled with a junction-table, however in MongoDB we can use **bidirectional embedding**, so we query A to search for the embedded references of the B objects, and then query B with a \$in operator to find these returned references

Musician Problem

Notown Records has decided to store information about musicians who perform on its albums (as well as other company data) in a database.

Each musician that records at Notown has an SSN, a name, an address, and a phone number. Poorly paid musicians do not have cell phones, often share the same address, and no address has more than one landline phone. Given their limited use, cell phones are not tracked.

Each instrument used in a song recorded at Notown has a unique identification number, a name (e.g., guitar, synthesizer, flute) and a musical key (e.g., C, B-flat, E-flat).

Each album recorded on the Notown label has a unique identification number, a title, a copyright date, a format (e.g., CD or MC), and an album identifier.

Each song recorded at Notown has a title and an author. The author of a song is a musician. There is 1 and only 1 author per song.

Each musician may play several instruments, and a given instrument may be played by several musicians.

Each album has a number of songs on it, but no song may appear on more than one album.

Each song is performed by one or more musicians, and a musician may perform a number of songs.

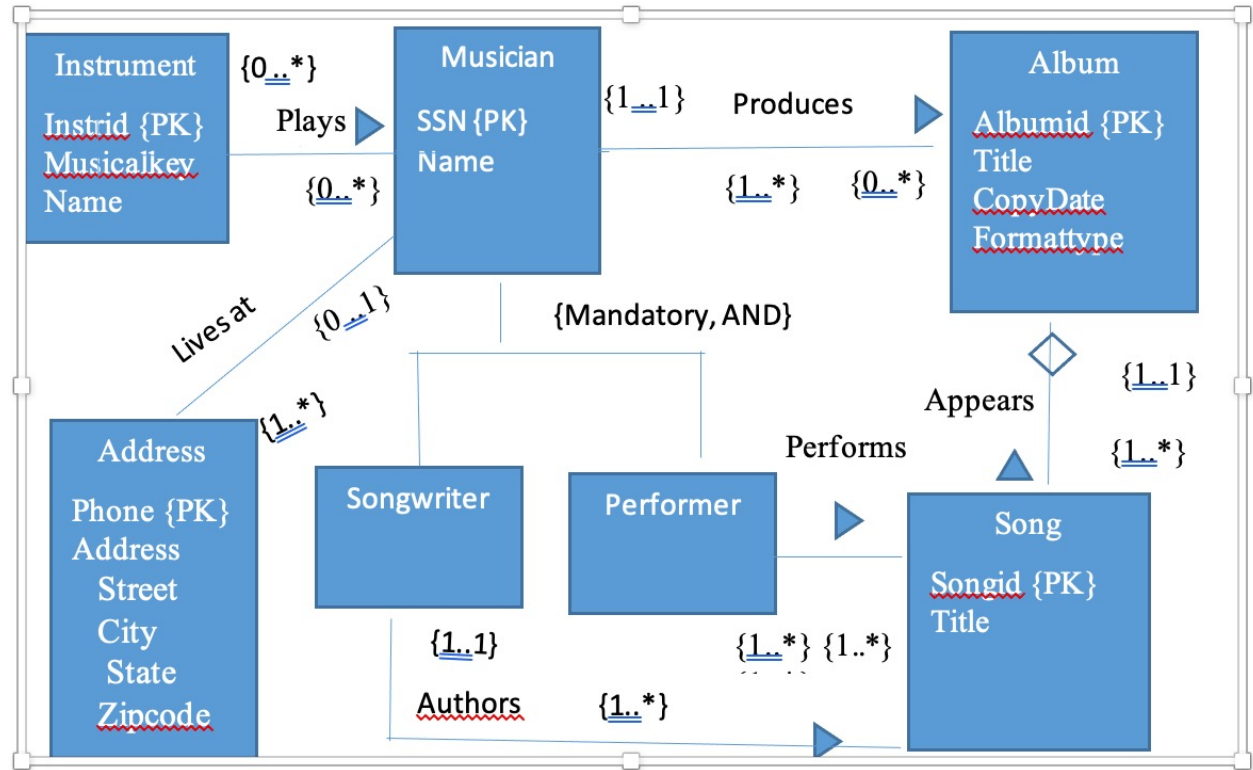
Each album has exactly one musician who acts as its producer. A musician may produce several albums, of course.

Design a conceptual schema for Notown and draw an UML diagram for your schema. Be sure to indicate all key and multiplicity constraints and any assumptions you make.

Conceptual Design to a Mongo Database

Application is
focused on
reporting
information
on musicians

Number of
Collections?



Collection for data

Summary

We start with the conceptual design and determine how many collections we need to represent the objects

We need to pay close attention to the application and the way it uses the data to determine the embedded documents versus references to other documents

References

<https://docs.mongodb.com/manual/tutorial/query-documents/>

<https://docs.mongodb.com/manual/tutorial/query-embedded-documents/>

<https://docs.mongodb.com/compass/current/>

<https://docs.mongodb.com/manual/tutorial/query-documents/>