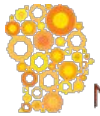
GridFS

Ramesh S

GridFS

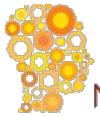
- GridFS is a specification for storing and retrieving files that exceed the BSON-document size limit of 16 MB.



Mongo Factory

GridFS

- Instead of storing a file in a single document, GridFS divides the file into parts, or chunks, and stores each chunk as a separate document.
- By default, GridFS uses a default chunk size of 255 kB
- The last chunk is only as large as necessary.
- Files that are no larger than the chunk size only have a final chunk, using only as much space as needed plus some additional metadata.

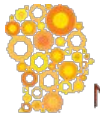


Mongo Factory

GridFS

GridFS stores files in two collections:

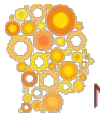
- **`chunks`** stores the binary chunks.
- **`files`** stores the file's metadata.



Mongo Factory

How to Store a File

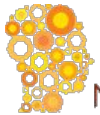
```
mongofiles -d database1 put abc.mpg
```



Mongo Factory

List all files in a db using list

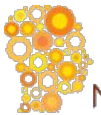
```
mongofiles -d database1 list
```



Mongo Factory

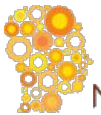
Get a file using get

```
mongofiles -d database1 get abc.mpg
```



Delete a file using delete

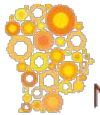
```
mongofiles -d database1 delete abc.mpg
```



Mongo Factory

Command list

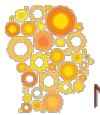
```
mongofiles -d database1 put abc.mpg

mongofiles -d database1 list

mongofiles -d database1 get abc.mpg

mongofiles -d database1 delete abc.mpg
```



Mongo Factory

When to Use GridFS

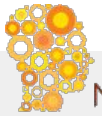
- Use GridFS for storing files larger than 16 MB.
- In some situations, storing large files may be more efficient in a MongoDB database than on a system-level filesystem.



Mongo Factory

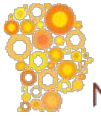
When to Use GridFS

- If your filesystem limits the number of files in a directory, you can use GridFS to store as many files as needed.
- When you want to access information from portions of large files without having to load whole files into memory, you can use GridFS to recall sections of files without reading the entire file into memory.
- When you want to keep your files and metadata automatically synced and deployed across a number of systems and facilities, you can use GridFS.
- When using geographically distributed replica sets, MongoDB can distribute files and their metadata automatically to a number of mongod instances and facilities.



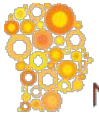
When Not To Use GridFS

- Do not use GridFS if you need to update the content of the entire file atomically.
- As an alternative you can store multiple versions of each file and specify the current version of the file in the metadata.
- You can update the metadata field that indicates “latest” status in an atomic update after uploading the new version of the file, and later remove previous versions if needed.



When Not To Use GridFS

- Furthermore, if your files are all smaller than the 16 MB BSON Document Size limit, consider storing each file in a single document instead of using GridFS.
- You may use the BinData data type to store the binary data.



Mongo Factory