**下载**

npm install gridfs-stream

**原理**

GridFS不是MongoDB自身特性，只是一种将大型文件存储在MongoDB的文件规范，所有官方支持的驱动均实现了GridFS规范。GridFS制定大文件在数据库中如何处理，通过开发语言驱动来完成、通过API接口来存储检索大文件。

**<使用场景>**

▲如果您的文件系统在一个目录中存储的文件的数量有限，你可以使用GridFS存储尽可能多的文件。

▲当你想访问大型文件的部分信息，却不想加载整个文件到内存时，您可以使用GridFS存储文件，并读取文件部分信息，而不需要加载整个文件到内存。

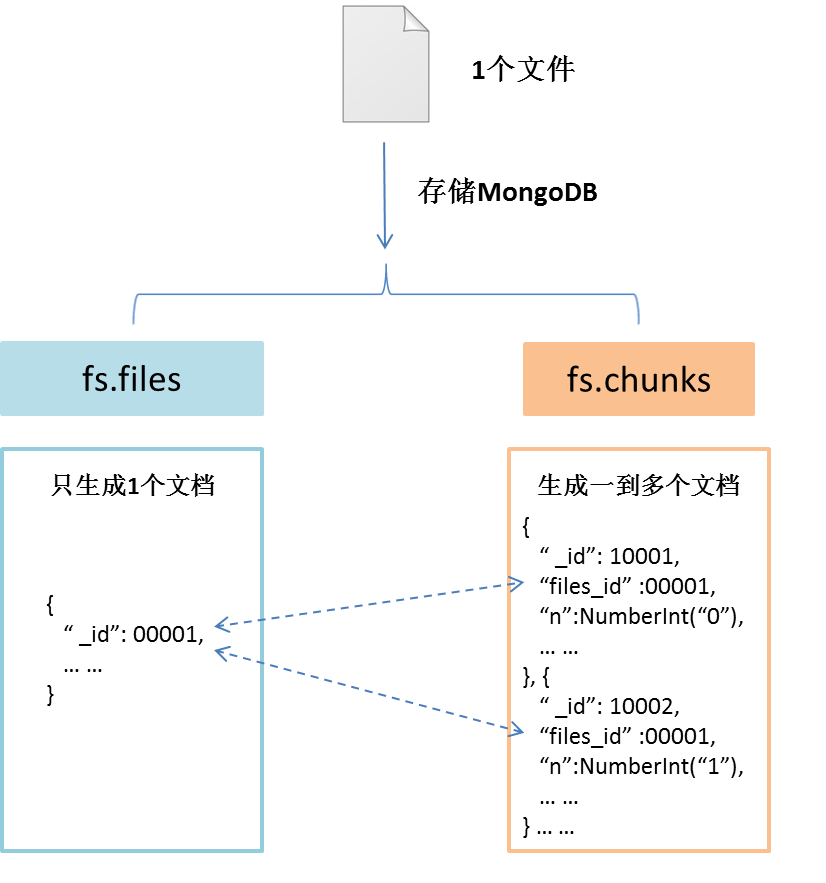
▲当你想让你的文件和元数据自动同步并部署在多个系统和设施，你可以使用GridFS实现分布式文件存储。

GridFS使用两个集合（collection）存储文件。一个集合是chunks, 用于存储文件内容的二进制数据；一个集合是files，用于存储文件的元数据。

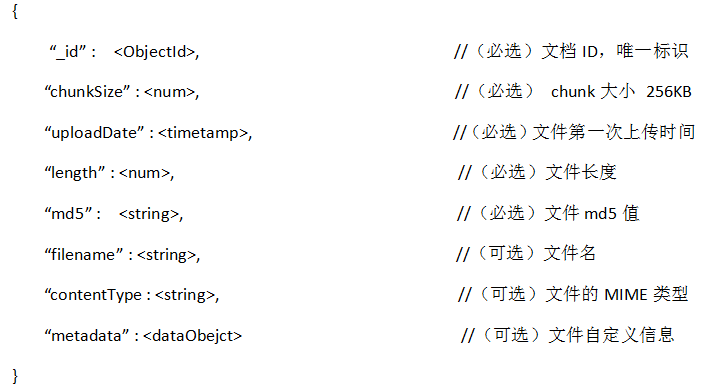
GridFS会将两个集合放在一个普通的buket中，并且这两个集合使用buket的名字作为前缀。MongoDB的GridFs默认使用fs命名的buket存放两个文件集合。因此存储文件的两个集合分别会命名为集合fs.files ,集合fs.chunks。

当把一个文件存储到GridFS时，如果文件大于chunksize （每个chunk块大小为256KB），会先将文件按照chunk的大小分割成多个chunk块，最终将chunk块的信息存储在fs.chunks集合的多个文档中。然后将文件信息存储在fs.files集合的唯一一份文档中。其中fs.chunks集合中多个文档中的file\_id字段对应fs.files集中文档”\_id”字段。

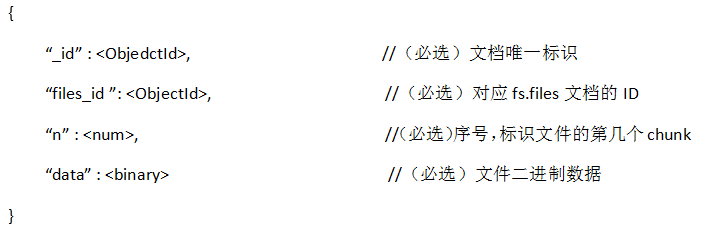
读文件时，先根据查询条件在files集合中找到对应的文档，同时得到“\_id”字段，再根据“\_id”在chunks集合中查询所有“files\_id”等于“\_id”的文档。最后根据“n”字段顺序读取chunk的“data”字段数据，还原文件。



fs.files 集合存储文件的元数据，以类json格式文档形式存储。每在GridFS存储一个文件，则会在fs.files集合中对应生成一个文档。  
▲fs.files集合中文档的存储内容如下：



fs.chunks 集合存储文件文件内容的二进制数据，以类json格式文档形式存储。每在GridFS存储一个文件，GridFS就会将文件内容按照chunksize大小（chunk容量为256k）分成多个文件块，然后将文件块按照类json格式存在.chunks集合中，每个文件块对应fs.chunk集合中一个文档。一个存储文件会对应一到多个chunk文档。  
▲ fs.chunks集合中文档的存储内容如下：



为了提高检索速度 MongoDB为GridFS的两个集合建立了索引。fs.files集合使用是“filename”与“uploadDate” 字段作为唯一、复合索引。fs.chunk集合使用的是“files\_id”与“n”字段作为唯一、复合索引。