

MongoDB 介绍

MongoDB 是目前在 IT 行业非常流行的一种非关系型数据库(NoSql)。

MongoDB 很好的实现了面向对象的思想，在 MongoDB 中，每一条记录都是一个 Document 对象。

MongoDB 数据库与传统的关系型数据库相比，它具有操作简单，性能优越的特点。

与关系型数据库相比，MongoDB 的优点：

(1) 弱一致性（最终一致），更能保证用户的访问速度

举例来说，在传统的关系型数据库中，一个 COUNT 类型的操作会锁定数据集，这样可以保证得到“当前”情况下的精确值。在某些情况下，需要这样作，例如通过 ATM 查看账户信息的时候很重要。但对于 Wordnik（Wordnik 是一项在线字典及百科全书服务）来说，数据是不断更新和增长的，这种“精确”的保证几乎没有任何意义，反而会产生很大的延迟。他们需要的是一个“大约”的数字以及更快的处理速度。Wordnik 开始使用 MySQL 数据库，后来逐渐从 MySQL 迁移到 MongoDB。据 Wordnik 技术团队描述，他们决定使用 MongoDB，是看中了它的弱一致性（最终一致）及文档结构的存储方式。

(2) 文档结构的存储方式，能够更快速的查询数据

对于一个层级式的数据结构来说，如果要将这样的数据使用扁平式的表状的结构来保存数据，这会使用查询效率变得极其低下。

举例来说，一个员工信息存在一个本子上，如果只存在一页纸上，查询速度会比较快，如果把这个员工信息查分成几个部分，分别存在不同的页上，而且这几页不是连续在一起的，东存一块，西存一块，查找这个员工信息时效率肯定低下。在关系型数据库中，把员工信息存放在员工表中，如果员工有所在部门，则把部门信息存放在部门表中，然后在员工表中增加一个字段“部门编号”外键关联到部门表。如果我们要查询员工信息，并且要查询出员工的所在部门名称，则需要通过关联查询实现。这种把一个完整的员工信息分割成几个部分分别存在不同的地方，然后再通过关联查询，才能查询出，这种效率肯定是低下的。

(3) 内置 GridFS，支持大容量的存储

GridFS 是一个出色的分布式文件系统，可以支持海量的数据存储。内置了 GridFS 的 MongoDB，能够满足对大数据集的快速查询。

（4） 内置 Sharding

MongoDB 提供了基于 Range 的 Auto Sharding 机制，它能够自动建立一个水平扩展的数据库集群系统，一个 Collection 可按照记录的范围，分成若干个段，切分到不同的 Shard 上。简单的说，在 MongoDB 中数据可以使用分布式的方式存在多台计算机节点上，在处理数据时，多个节点同时工作，降低处理时长，提高性能。

（5） 性能优越

在很多场合下，千万级别的文档对象，近 10G 的数据，MongoDB 对有索引的字段查询不会比 MySQL 等关系型数据库慢，而对非索引字段的查询操作，则 MongoDB 性能更突出。