

(5) 常用数据库类型 (掌握)

数据库根据存储方式可以分为关系型数据库 (SQL) 和非关系型数据库 (Not Only SQL, NoSQL)。

(1) 关系型数据库

关系数据库是在一个给定的应用领域中, 所有实体及实体之间联系的集合。关系型数据库支持事务的 ACID 原则, 即原子性 (Atomicity)、一致性 (Consistency)、隔离性 (Isolation)、持久性 (Durability) 这四种原则保证在事务过程当中数据的正确性。

(2) 非关系型数据库

非关系型数据库是分布式的、非关系型的、不保证遵循 ACID 原则的数据存储系统。NoSQL 数据存储不需要固定的表结构, 通常也不存在连接操作。在大数据存取上具备关系型数据库无法比拟的性能优势。

常见的非关系数据库分为:

- **键值数据库:** 类似传统语言中使用的哈希表。可以通过 key 来添加、查询或者删除数据库, 因为使用 key 主键访问, 会获得很高的性能及扩展性。Key/Value 模型对于信息系统来说, 其优势在于简单、易部署、高并发。
- **列存储 (Column-oriented) 数据库:** 将数据存储于列族中, 一个列族存储经常被一起查询, 比如人们经常会查询某人的姓名和年龄, 而不是薪资。这种情况下姓名和年龄会被放到一个列族中, 薪资会被放到另一个列族中。这种数据库通常用来应对分布式存储海量数据。
- **面向文档 (Document-Oriented) 数据库:** 文档型数据库可以看作是键值数据库的升级版, 允许之间嵌套键值, 而且文档型数据库比键值数据库的查询效率更高。面向文档数据库会将数据以文档形式存储。
- **图形数据库:** 允许人们将数据以图的方式存储。实体会作为顶点, 而实体之间的关系则会作为边。比如有三个实体: Steve Jobs、Apple 和 Next, 则会有两个 Founded by 的边将 Apple 和 Next 连接到 Steve Jobs。

不同存储方式数据库的优缺点, 关系型数据库和非关系型数据库的优缺点, 如表。

数据库类型	特点类型	描述
关系型数据库	优点	<ul style="list-style-type: none"> ● 容易理解: 二维表结构是非常贴近逻辑世界的一个概念, 关系模型相对网状、层次等其他模型来说更容易理解 ● 使用方便: 通用的 SQL 语言使得操作关系型数据库非常方便 ● 易于维护: 丰富的完整性 (实体完整性、参照完整性和用户定义的完整性) 大大降低了数据冗余和数据不一致的概率
	缺点	<ul style="list-style-type: none"> ● 数据读写必须经过 SQL 解析, 大量数据、高并发下读写性能不足 (对于传统关系型数据库来说, 硬盘 I/O 是一个很大的瓶颈) ● 具有固定的表结构, 因此扩展困难 ● 多表的关联查询导致性能欠佳
非关系型数据库	优点	<ul style="list-style-type: none"> ● 高并发: 大数据下读写能力较强 (基于键值对的, 可以想象成表中的主键和值的对应关系, 且不需要经过 SQL 层的解析。所以性能非常高) ● 基本支持分布式: 易于扩展, 可伸缩 (因为基于键值对, 数据之间没有耦合性, 所以非常容易水平扩展) ● 简单: 弱结构化存储
	缺点	<ul style="list-style-type: none"> ● 事务支持较弱 ● 通用性差 ● 无完整约束, 复杂业务场景支持较差