

小菜鸟。

博客园

首页

新随笔

联系

订阅

管理

关系数据库与非关系数据库

1. 关系型数据库

关系型数据库，是指采用了关系模型来组织数据的数据库。

关系模型是在1970年由IBM的研究员E.F.Codd博士首先提出的，在之后的几十年中，关系模型的概念得到了充分的发展并逐渐成为主流数据库结构的主流模型。

简单来说，关系模型指的就是二维表格模型，而一个关系型数据库就是由二维表及其之间的联系所组成的一个数据组织。

关系模型中常用的概念：

- 关系：可以理解为一张二维表，每个关系都具有一个关系名，就是通常说的表名
- 元组：可以理解为二维表中的一行，在数据库中经常被称为记录
- 属性：可以理解为二维表中的一列，在数据库中经常被称为字段
- 域：属性的取值范围，也就是数据库中某一列的取值限制
- 关键字：一组可以唯一标识元组的属性，数据库中常称为主键，由一个或多个列组成
- 关系模式：指对关系的描述。其格式为：关系名(属性1，属性2， ，属性N)，在数据库中成为表结构

关系型数据库的优点：

- 容易理解：二维表结构是非常贴近逻辑世界的概念，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)大大减低了数据冗余和数据不一致的概率

2. 关系型数据库瓶颈

- 高并发读写需求

网站的用户并发性非常高，往往达到每秒上万次读写请求，对于传统关系型数据库来说，硬盘I/O是一个很大的瓶颈

- 海量数据的高效率读写

网站每天产生的数据量是巨大的，对于关系型数据库来说，在一张包含海量数据的表中查询，效率是非常低的

- 高扩展性和可用性

公告

昵称：小菜鸟。
园龄：2年5个月
粉丝：4
关注：2
[+加关注](#)

<	2018年1月						>
	日	一	二	三	四	五	六
31	1	2	3	4	5	6	
7	8	9	10	11	12	13	
14	15	16	17	18	19	20	
21	22	23	24	25	26	27	
28	29	30	31	1	2	3	
4	5	6	7	8	9	10	

搜索

找找看

谷歌搜索

随笔档案

- 2017年6月 (3)
- 2016年4月 (1)
- 2015年11月 (12)
- 2015年9月 (3)
- 2015年8月 (20)

阅读排行榜

1. 关系数据库与非关系数据库(5817)
2. 数据库中触发器的创建与注意事项(4186)

在基于web的结构当中，数据库是最难进行横向扩展的，当一个应用系统的用户量和访问量与日俱增的时候，数据库却没有办法像web server和app server那样简单的通过添加更多的硬件和服务节点来扩展性能和负载能力。对于很多需要提供24小时不间断服务的网站来说，对数据库系统进行升级和扩展是非常痛苦的事情，往往需要停机维护和数据迁移。

对网站来说，关系型数据库的很多特性不再需要了：

- 事务一致性

关系型数据库在对事物一致性的维护中有很大的开销，而现在很多web2.0系统对事物的读写一致性都不高

- 读写实时性

对关系数据库来说，插入一条数据之后立刻查询，是肯定可以读出这条数据的，但是对于很多web应用来说，并不要求这么高的实时性，比如发一条消息之后，过几秒乃至十几秒之后才看到这条动态是完全可以接受的

- 复杂SQL，特别是多表关联查询

任何大数据量的web系统，都非常忌讳多个大表的关联查询，以及复杂的数据分析类型的复杂SQL报表查询，特别是SNS类型的网站，从需求以及产品阶级角度，就避免了这种情况的产生。往往更多的只是单表的主键查询，以及单表的简单条件分页查询，SQL的功能极大的弱化了

在关系型数据库中，导致性能欠佳的最主要原因是多表的关联查询，以及复杂的数据分析类型的复杂SQL报表查询。为了保证数据库的ACID特性，我们必须尽量按照其要求的范式进行设计，关系型数据库中的表都是存储一个格式化的数据结构。每个元组字段的组成都是一样，即使不是每个元组都需要所有的字段，但数据库会为每个元组分配所有的字段，这样的结构可以便于标语表之间进行链接等操作，但从另一个角度来说它也是关系型数据库性能瓶颈的一个因素。

3. NoSQL

NoSQL一词首先是Carlo Strozzi在1998年提出来的，指的是他开发的一个没有SQL功能，轻量级的，开源的关系型数据库。这个定义跟我们现在对NoSQL的定义有很大的区别，它确实实字如其名，指的就是“没有SQL”的数据库。但是NoSQL的发展慢慢偏离了初衷，我们要的不是“no sql”，而是“no relational”，也就是我们现在常说的非关系型数据库了。

2009年初，Johan Oskarsson举办了一场关于开源分布式数据库的讨论，Eric Evans在这次讨论中再次提出了NoSQL一词，用于指代那些非关系型的，分布式的，且一般不保证遵循ACID原则的数据存储系统。Eric Evans使用NoSQL这个词，并不是因为字面上的“没有SQL”的意思，他只是觉得很多经典的关系型数据库名字都叫“**SQL”，所以为了表示跟这些关系型数据库在定位上的截然不同，就是用了“NoSQL”一词。

注：数据库事务必须具备ACID特性，ACID是Atomic原子性，Consistency一致性，Isolation隔离性，Durability持久性。

非关系型数据库提出另一种理念，例如，以键值对存储，且结构不固定，每一个元组可以有不一样的字段，每个元组可以根据需要增加一些自己的键值对，这样就不会局限于固定的结构，可以减少一些时间和空间的开销。使用这种方式，用户可以根据需要去添加自己需要的字段，这样，为了获取用户的不同信息，不需要像关系型数据库中，要对多表进行关联查询。仅需要根据id取出相应的value就可以完成查询。但非关系型数据库由于很少的约束，他也不能够提供像SQL所提供的where这种对于字段属性值情况的查询。并且难以体现设计的完整性。他只适合存储一些较为简单的数据，对于需要进行较复杂查询的数据，SQL数据库显的更为合适。

4. 关系型数据库 V.S. 非关系型数据库

关系型数据库的最大特点就是事务的一致性：传统的关系型数据库读写操作都是事务的，具有ACID的特点，这个特性使得关系型数据库可以用于几乎所有对一致性有要求的系统中，如典型的银行系统。

但是，在网页应用中，尤其是SNS应用中，一致性却不是显得那么重要，用户A看到的内容和用户B看到同一用户C内容更新不一致是可以容忍的，或者说，两个人看到同一好友的数据更新的时间差那么几秒是可以容忍的，因此，关系型数据库的最大特点在这里已经无用武之地，起码不是那么重要了。

相反地，关系型数据库为了维护一致性所付出的巨大代价就是其读写性能比较差，而像微博、facebook这类SNS的应用，对并发读写能力要求极高，关系型数据库已经无法应付(在读方面，传统上为了克服关系型数据库缺陷，提高性能，都是增加一级memcache来静态化网页，而在SNS中，变化太快，memcache已经无能为力了)，因此，必须用新的一种数据结构存储来代替关系数据库。

3. 数据库系统中事务的ACID原则(3255)

4. 索引建立的原则(2438)

5. HTML发展历史(2421)

推荐排行榜

1. 数据库系统中事务的ACID原则(1)

2. 索引建立的原则(1)

关系数据库的另一个特点就是其具有固定的表结构，因此，其扩展性极差，而在SNS中，系统的升级，功能的增加，往往意味着数据结构巨大变动，这一点关系型数据库也难以应付，需要新的结构化数据存储。

于是，非关系型数据库应运而生，由于不可能用一种数据结构化存储应付所有的新的需求，因此，非关系型数据库严格上不是一种数据库，应该是一种数据结构化存储方法的集合。

必须强调的是，数据的持久存储，尤其是海量数据的持久存储，还是需要一种关系数据库这员老将。

5. 非关系型数据库分类

由于非关系型数据库本身天然的多样性，以及出现的时间较短，因此，不想关系型数据库，有几种数据库能够一统江山，非关系型数据库非常多，并且大部分都是开源的。

这些数据库中，其实实现大部分都比较简单，除了一些共性外，很大一部分都是针对某些特定的应用需求出现的，因此，对于该类应用，具有极高的性能。依据结构化方法以及应用场合的不同，主要分为以下几类：

- 面向高性能并发读写的key-value数据库：

key-value数据库的主要特点即使具有极高的并发读写性能，Redis,Tokyo Cabinet,Flare就是这类的代表

- 面向海量数据访问的面向文档数据库：

这类数据库的特点是，可以在海量的数据中快速的查询数据，典型代表为MongoDB以及CouchDB

- 面向可扩展性的分布式数据库：

这类数据库想解决的问题就是传统数据库存在可扩展性上的缺陷，这类数据库可以适应数据量的增加以及数据结构的变化

好文要顶

关注我

收藏该文

小菜鸟。
关注 - 2
粉丝 - 4

+加关注

0

0

« 上一篇：数据库的比较

» 下一篇：varchar 保存英文中文区别。

posted @ 2015-08-16 22:09 小菜鸟。 阅读(5816) 评论(0) 编辑 收藏

刷新评论 刷新页面 返回顶部

注册用户登录后才能发表评论，请 [登录](#) 或 [注册](#)，[访问网站首页](#)。

- 【推荐】50万行VC++源码：大型组态工控、电力仿真CAD与GIS源码库
- 【推荐】加入腾讯云自媒体扶持计划，免费领取域名&服务器
- 【福利】限时领取，H3 BPM给你发年终奖



最新IT新闻：

- 程序员职业出路：工程师与管理岗位的“钟摆式轮回”
 - 微软再发对比：Edge速度比Chrome快48% 拦截钓鱼网站多18%
 - 12306已使用了两个公有云：日均可售票1500万张
 - Uber第三季度净营收首超20亿美元 环比增长21%
 - 华尔街讶异：中国人收入落后美国 手游消费竟反超
- » 更多新闻...



最新知识库文章：

- 步入云计算
 - 以操作系统的角度述说线程与进程
 - 软件测试转型之路
 - 门内门外看招聘
 - 大道至简，职场上做人做事做管理
- » 更多知识库文章...