

浅谈非关系型数据库在未来移动通信中的应用

管红光¹ 周炜²

(1.华为技术有限公司南京研究所; 2.南京邮电大学 江苏南京 210000)

摘要: 移动通信的未来是朝着高速率的数据传输, 随时随地的移动接入, 高度集中的服务的4G方向发展, 该文将结合非关系型数据库的自身特点说明它在未来移动通信中的应用。

关键词: 非关系型数据库 移动通信 应用

中图分类号: TP3

文献标识码: A

文章编号: 1674-098X(2013)06(c)-0040-01

1 非关系数据库简介

非关系型数据库, 即 NoSQL (NoSQL=Not Only SQL), 是一项具有革命性的数据库技术。NoSQL的发展最早可以追溯到1991年Berkeley DB第一版的发布。Berkeley DB是一个Key/Value (键/值) 类型的Hush数据库。NoSQL发展至2009年趋势越发高涨, 这得益于web2.0技术在网络技术中的应用。传统的关系数据库中的表都是一些格式化的数据结构, 每个元组的属性字段的组成都一样, 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表与表之间进行连接等操作, 但另一方面, 它也是关系型数据库性能瓶颈的一个因素。而非关系数据库以键值对存储, 它的结构不固定, 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。总而言之, NoSQL有如下主要特点: 首先是NoSQL可以处理超大量的数据; 其次, NoSQL可以运行在便宜的PC服务器上, 并且易于集群扩充; 第三, 由于其灵活的结构特点, NoSQL可以突破传统数据库的性能瓶颈, 从而实现对数据高并发读写的需求。

2 未来移动通信中的数据存储

在传统移动通信中, 移动交换中心数据库 (例如GSM中的HLR和VLR中的数据库), 记录了用户的基本信息数据, 如: IMSI (在不支持IMSI的情况下, 可存储MIN)、MDN、ESN (电子序列号)、主叫用户类别 (CPC)、移动台等级标志等信息。在传统数据库中, 系统会根据这些属性建立一张用户表, 在实际运行时, 每新增一个用户, 则数据表新增一条记录, 分别记录新增用户的IMSI、MDN等数据。在传统通信业务中, 一般只有语音或短信等业务, 所以每个用户的信息格式基本都是相同的, 因此可以使用传统数据库技术 (如Oracle)。但是, 移动通信移动通信的未来趋势是朝着高质量、高速率的数据传输, 随时随地的移动接入, 高度集中的服务, 多兼容、多样化、高智能的方向发展的, 而它的最终目的是实现个性化的通信。所以, 在未来的移动通信中, 每个用户的信息都是个性化的, 例如用

户A定制了GPS业务, 系统必须随时记录用户所处的位置; 而用户B定制了亲情号码业务, 系统必须登记他/她的亲情号码的相关信息; 同时, 用户C定制了新闻套餐, 系统必须每天为用户C更新新闻信息; D用户则希望或得最新的时尚资讯……在移动交换中心中, 每个用户的个性定制都不一样, 因此在传统数据库中, 若想将所有的用户信息记录到一张表中, 那么这张表的属性字段的总数将是一个庞大的数字, 这显然是无法实现。在传统的数据库中, 解决的方法, 是进行分表操作, 即针对GPS用户建立一张表, 再针对亲情号码业务建立一张表……因此我们查询用户A的定制业务时, 系统需要对用户表和GPS用户表进行关联查询, 但在实际应用中, 每个用户可能定制了不同的业务, 那么传统数据库系统需要对几十张表进行关联查询, 同时一个移动通信系统中往往存储了几百万乃至上千万用户的数据信息, 对于大数据量的数据库系统, 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, 我们希望通过单表的主键查询, 以及单表的简单条件查询就能获取用户的相关信息。这里, NoSQL数据库MongoDB很好得解决了这一问题。

和传统数据库不同, MongoDB数据库存储的数据是键值对的集合, 如上文, 用户A的相关信息在数据库中存储的形式如下: {IMSI:0000000001, NAME:A, ADDRESS:"江苏省南京市新模范马路66号"}。其中, 这个元组中有三个键值对, 三个键名分别是IMSI、NAME和ADDRESS。0000000001, "A" 和 "江苏省南京市新模范马路66号" 分别是三个键的值, 这三个键值对组成了记录用户A信息的文档。而用户B的文档为: {IMSI:0000000002, NAME:"B", FRIEND:{IMSI:0000000011, NAME:G}}, 与用户A不同, 用户B的记录并没有ADDRESS这个键值, 而是FRIEND, 它记录了用户B亲情好友G的相关信息。同样, 用户C和用户D可以根据自己的定制业务存储不同的键值对的文档, 不同用户的文档最终移动通信业务的集合。最后, 在移动通信系统中, 每个用户都需要一些共同的基本键值对如IMSI、姓名等, 我们可以在集合中设置IMSI键为主键, 通过主键系统能够建立一个检索, 从而系统能

够快速地查询大数据的集合中各个用户的信息。

另外, 未来移动通信的高速率传输可以为用户提供更多的个性化媒体服务, 用户可以定制各种不同的图片、音乐等媒体文件; 而MongoDB不仅能够存储文字、数字等数据, 还能存储二进制大型文件 (如音乐、图片、视频等), 这位未来移动通信提供了更好、更快捷的数据存储方案。

3 MongoDB分片集群在未来移动通信中的应用

MongoDB另一个核心应用就是其分片, MongoDB分片的基本思想就是将集合切成小块。这些小块分散到若干片里面, 每个片只负责该集合数据的一部分。而应用程序不必知道哪片对应哪些数据, 甚至不需要知道该数据是否已经拆分, 所以在分片之前要运行一个路由进程, 该进程名为mongos。这个路由器知道所有数据的存放位置, 所以应用可以连接它来正常发送请求。在未来移动通信数据库系统中, 我们可以根据用户的IMSI号对所有用户数据的集合进行分片, 例如IMSI号为158开头的所有用户数据存至分片主机1, 139开头的所有用户数据存至分片主机2……最后, 系统通过路由进程访问数据库。这样, 我们既可以有效利用闲置的数据库主机, 在用户量增加时, 我们可以适当地增加分片主机, 从而达到各个主机的负载均衡, 提高运行效率。

4 结语

非关系型数据库目前在社交网站等web2.0网络应用中得到广泛应用, 对于未来的移动通信系统, 亟需解决的是数据库与移动通信协议、信令之间绑定问题。它对大数据高并发读写的优异性必将使之在未来移动通信数据库中占有一席之地。

参考文献

- [1] Kristina Chodorow, Micheal Dirolf 著. MongoDB权威指南[M]. 程显峰, 译. 北京: 人民邮电出版社, 2011.
- [2] 罗茜文. 技术创新: 新时期移动通信发展的主旋律[J]. 移动通信, 2010(1).
- [3] 范凯. NoSQL数据库综述[J]. 程序员, 2010(6): 76-78.