

## 11-NoSQL：在高并发场景下，数据库和NoSQL如何做到互补？

你好，我是唐扬。

前几节课，我带你了解了在你的垂直电商项目中，如何将传统的关系型数据库改造成分布式存储服务，以抵抗高并发和大流量的冲击。

对于存储服务来说，我们一般会从两个方面对它做改造：

- 1.提升它的读写性能，尤其是读性能，因为我们面对的多是一些读多写少的产品。比方说，你离不开的微信朋友圈、微博和淘宝，都是查询QPS远远大于写入QPS。
- 2.增强它在存储上的扩展能力，从而应对大数据量的存储需求。

我之前带你学习的读写分离和分库分表就是从这两方面出发，改造传统的关系型数据库的，但仍有一些问题无法解决。

比如，在微博项目中关系的数据量达到了千亿，那么即使分隔成1024个库表，每张表的数据量也达到了亿级别，并且关系的数据量还在以极快的速度增加，即使你分隔成再多的库表，数据量也会很快增加到瓶颈。这个问题用传统数据库很难根本解决，因为它在扩展性方面是很弱的，这时，就可以利用NoSQL，因为它有着天生分布式的能力，能够提供优秀的读写性能，可以很好地补充传统关系型数据库的短板。那么它是如何做到的呢？

这节课，我就还是以你的垂直电商系统为例，带你掌握如何用NoSQL数据库和关系型数据库互补，共同承担高并发和大流量的冲击。

首先，我们先来了解一下NoSQL数据库。

### NoSQL，No SQL？

NoSQL想必你很熟悉，它指的是不同于传统的关系型数据库的其他数据库系统的统称，它不使用SQL作为查询语言，提供优秀的横向扩展能力和读写性能，非常契合互联网项目高并发大数据的特点。所以一些大厂，比如小米、微博、陌陌都很倾向使用它来作为高并发大容量的数据存储服务。

NoSQL数据库发展到现在，十几年间，出现了多种类型，我来给你举几个例子：

- Redis、LevelDB这样的KV存储。这类存储相比于传统的数据库的优势是极高的读写性能，一般对性能有比较高的要求的场景会使用。
- Hbase、Cassandra这样的列式存储数据库。这种数据库的特点是数据不像传统数据库以行为单位来存储，而是以列来存储，适用于一些离线数据统计的场景。
- 像MongoDB、CouchDB这样的文档型数据库。这种数据库的特点是Schema Free（模式自由），数据表中的字段可以任意扩展，比如说电商系统中的商品有非常多的字段，并且不同品类的商品的字段也都不尽相同，使用关系型数据库就需要不断增加字段支持，而用文档型数据库就简单很多了。

在NoSQL数据库刚刚被应用时，它被认为是可以替代关系型数据库的银弹，在我看来，也许因为以下几个方面的原因：

- 弥补了传统数据库在性能方面的不足；
- 数据库变更方便，不需要更改原先的数据结构；
- 适合互联网项目常见的大数据量的场景；

不过，这种看法是个误区，因为慢慢地我们发现在业务开发的场景下还是需要利用SQL语句的强大的查询功能以及传统数据库事务和灵活的索引等功能，NoSQL只能作为一些场景的补充。

那么接下来，我就带你了解**NoSQL数据库是如何做到与关系数据库互补的**。了解这部分内容，你可以在实际项目中更好地使用NoSQL数据库补充传统数据库的不足。

首先，我们来关注一下数据库的写入性能。

## 使用NoSQL提升写入性能

数据库系统大多使用的是传统的机械磁盘，对于机械磁盘的访问方式有两种：一种是随机IO；另一种是顺序IO。随机IO就需要花费时间做昂贵的磁盘寻道，一般来说，它的读写效率要比顺序IO小两到三个数量级，所以我们想要提升写入的性能就要尽量减少随机IO。

以MySQL的InnoDB存储引擎来说，更新binlog、redolog、undolog都是在做顺序IO，而更新datafile和索引文件则是在做随机IO，而为了减少随机IO的发生，关系数据库已经做了很多的优化，比如说写入时先写入内存，然后批量刷新到磁盘上，但是随机IO还是会发生。

索引在InnoDB引擎中是以B+树（[上一节课](#)提到了B+树，你可以回顾一下）方式来组织的，而MySQL主键是聚簇索引（一种索引类型，数据与索引数据放在一起），既然数据和索引数据放在一起，那么在数据插入或者更新的时候，我们需要找到要插入的位置，再把数据写到特定的位置上，这就产生了随机的IO。而且一旦发生了页分裂，就不可避免会做数据的移动，也会极大地损耗写入性能。

## NoSQL数据库是怎么解决这个问题的呢？

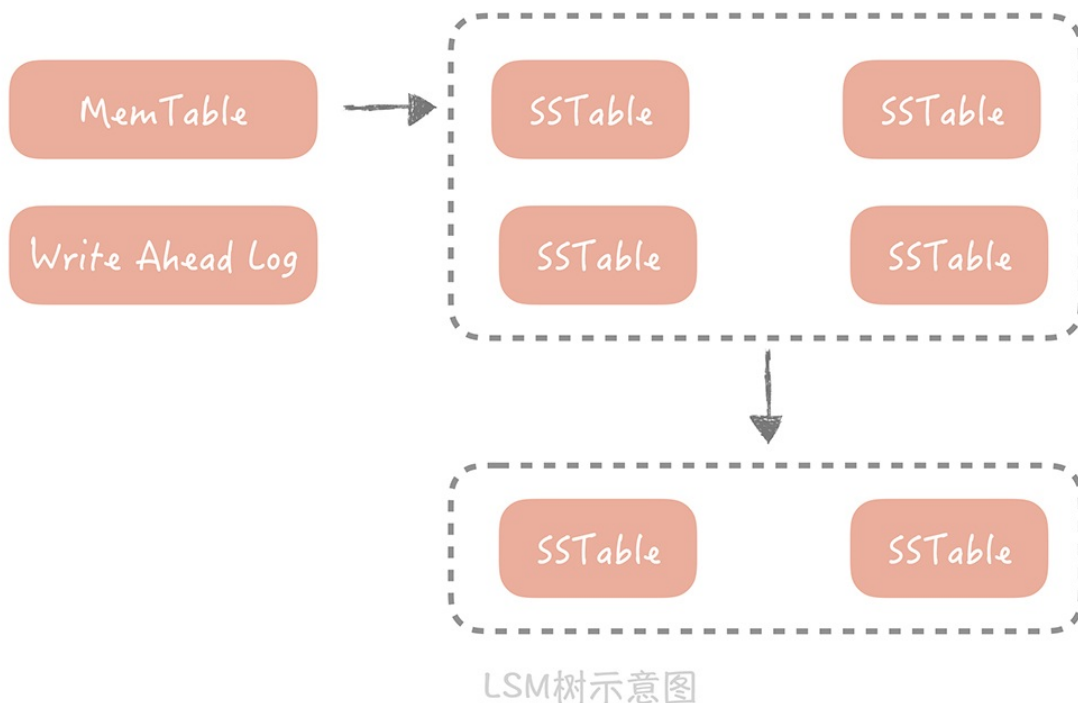
它们有多种的解决方式，这里我给你讲一种最常见的方案，就是很多NoSQL数据库都在使用的**基于LSM树的存储引擎**，这种算法使用最多，所以在这里着重剖析一下。

LSM树（Log-Structured Merge Tree）牺牲了一定的读性能来换取写入数据的高性能，Hbase、Cassandra、LevelDB都是用这种算法作为存储的引擎。

它的思想很简单，数据首先会写入到一个叫做MemTable的内存结构中，在MemTable中数据是按照写入的Key来排序的。为了防止MemTable里面的数据因为机器掉电或者重启而丢失，一般会通过写Write Ahead Log的方式将数据备份在磁盘上。

MemTable在累积到一定规模时，它会被刷新生成一个新的文件，我们把这个文件叫做SSTable（Sorted String Table）。当SSTable达到一定数量时，我们会将这些SSTable合并，减少文件的数量，因为SSTable都是有序的，所以合并的速度也很快。

当从LSM树里面读数据时，我们首先从MemTable中查找数据，如果数据没有找到，再从SSTable中查找数据。因为存储的数据都是有序的，所以查找的效率是很高的，只是因为数据被拆分成多个SSTable，所以读取的效率会低于B+树索引。



和LSM树类似的算法有很多，比如说TokuDB使用的名为Fractal tree的索引结构，它们的核心思想就是将随机IO变成顺序的IO，从而提升写入的性能。

在后面的缓存篇中，我也将给你着重介绍我们是如何使用KV型NoSQL存储来提升读性能的。所以你看，NoSQL数据库补充关系型数据库的第一种方式就是提升读写性能。

## 场景补充

除了可以提升性能之外，NoSQL数据库还可以在某些场景下作为传统关系型数据库的补充，来看一个具体的例子。

假设某一天，CEO找到你并且告诉你，他正在为你的垂直电商项目规划搜索的功能，需要支持按照商品的名称模糊搜索到对应的商品，希望你尽快调研出解决方案。

一开始，你认为这非常的简单，不就是在数据库里面执行一条类似：“select \* from product where name like ‘%\*\*\*%’”的语句吗？可是在实际执行的过程中，却发现了问题。

你发现这类语句并不是都能使用到索引，只有后模糊匹配的语句才能使用索引。比如语句“select \* from product where name like ‘%电冰箱’”就没有使用到字段“name”上的索引，而“select \* from product where name like ‘索尼%’”就使用了“name”上的索引。而一旦没有使用索引就会扫描全表的数据，在性能上是无法接受的。

于是你在谷歌上搜索了一下解决方案，发现大家都在使用开源组件Elasticsearch来支持搜索的请求，它本身是基于“倒排索引”来实现的，**那么什么是倒排索引呢？**

倒排索引是指将记录中的某些列做分词，然后形成的分词与记录ID之间的映射关系。比如说，你的垂直电商项目里面有以下记录：

商品ID	商品名称
1	西门子电冰箱
2	小米电视机
3	美的电冰箱

那么，我们将商品名称做简单的分词，然后建立起分词和商品ID的对应关系，就像下面展示的这样：

关键词	商品ID
西门子	1
电冰箱	1, 3
小米	2
电视机	2
美的	3

这样，如果用户搜索电冰箱，就可以给他展示商品ID为1和3的两件商品了。

而Elasticsearch作为一种常见的NoSQL数据库，**就以倒排索引作为核心技术原理，为你提供了分布式的全文搜索服务，这在传统的关系型数据库中使用SQL语句是很难实现的。**所以你看，NoSQL可以在某些业务场景下代替传统数据库提供数据存储服务。

## 提升扩展性

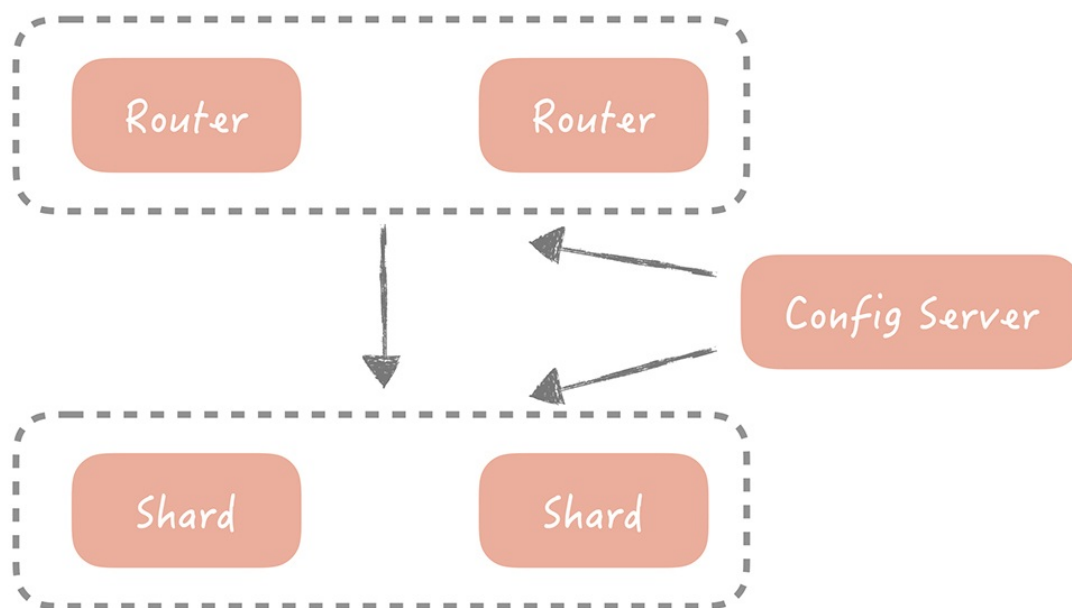
另外，在扩展性方面，很多NoSQL数据库也有着先天的优势。还是以你的垂直电商系统为例，你已经为你的电商系统增加了评论系统，开始你的评估比较乐观，觉得电商系统的评论量级不会增长很快，所以就为它分了8个库，每个库拆分成16张表。

但是评论系统上线之后，存储量级增长的异常迅猛，你不得不将数据库拆分成更多的库表，而数据也要重新迁移到新的库表中，过程非常痛苦，而且数据迁移的过程也非常容易出错。

这时，你考虑是否可以考虑使用NoSQL数据库来彻底解决扩展性的问题，经过调研你发现它们在设计之初就考虑到了分布式和大数据存储的场景，**比如像MongoDB就有三个扩展性方面的特性。**

- 其一是Replica，也叫做副本集，你可以理解为主从分离，也就是通过将数据拷贝成多份来保证当主挂掉后数据不会丢失。同时呢，Replica还可以分担读请求。Replica中有主节点来承担写请求，并且把对数据变动记录到oplog里（类似于binlog）；从节点接收到oplog后就会修改自身的数据以保持和主节点的一致。一旦主节点挂掉，MongoDB会从从节点中选取一个节点成为主节点，可以继续提供写数据服务。
- 其二是Shard，也叫做分片，你可以理解为分库分表，即将数据按照某种规则拆分成多份，存储在不同的机器上。MongoDB的Sharding特性一般需要三个角色来支持，一个是Shard Server，它是实际存储数据

的节点，是一个独立的Mongod进程；二是Config Server，也是一组Mongod进程，主要存储一些元信息，比如说哪些分片存储了哪些数据等；最后是Route Server，它不实际存储数据，仅仅作为路由使用，它从Config Server中获取元信息后，将请求路由到正确的Shard Server中。



MongoDB分片架构示意图

- 其三是负载均衡，就是当MongoDB发现Shard之间数据分布不均匀，会启动Balancer进程对数据做重新分配，最终让不同Shard Server的数据可以尽量均衡。当我们的Shard Server存储空间不足需要扩容时，数据会自动被移动到新的Shard Server上，减少了数据迁移和验证的成本。

你可以看到，NoSQL数据库中内置的扩展性方面的特性可以让我们不再需要对数据库做分库分表和主从分离，也是对传统数据库一个良好的补充。

你可能会觉得，NoSQL已经成熟到可以代替关系型数据库了，但是就目前来看，NoSQL只能作为传统关系型数据库的补充而存在，弥补关系型数据库在性能、扩展性和某些场景下的不足，所以你在使用或者选择时要结合自身的场景灵活地运用。

## 课程小结

本节课我带你了解了NoSQL数据库在性能、扩展性上的优势，以及它的一些特殊功能特性，主要有以下几点：

- 1.在性能方面，NoSQL数据库使用一些算法将对磁盘的随机写转换成顺序写，提升了写的性能；
- 2.在某些场景下，比如全文搜索功能，关系型数据库并不能高效地支持，需要NoSQL数据库的支持；
- 3.在扩展性方面，NoSQL数据库天生支持分布式，支持数据冗余和数据分片的特性。

这些都让它成为传统关系型数据库的良好补充，你需要了解的是，**NoSQL可供选型的种类很多，每一个组件都有各自的特点。你在做选型的时候需要对它的实现原理有比较深入的了解，最好在运维方面对它有一定的熟悉，这样在出现问题时才能及时找到解决方案。**否则，盲目跟从地上了一个新的NoSQL数据库，最终可能导致会出了故障无法解决，反而成为整体系统的拖累。



我在之前的项目中曾经使用Elasticsearch作为持久存储，支撑社区的feed流功能，初期开发的时候确实很爽，你可以针对feed中的任何字段做灵活高效地查询，业务功能迭代迅速，代码也简单易懂。可是到了后期流量上来之后，由于缺少对于Elasticsearch成熟的运维能力，造成故障频出，尤其到了高峰期就会出现节点不可用的问题，而由于业务上的巨大压力又无法分出人力和精力对Elasticsearch深入的学习和了解，最后不得不做大的改造切回熟悉的MySQL。所以，对于开源组件的使用，不能只停留在只会“hello world”的阶段，而应该对它有足够的运维上的把控能力。

## 思考时间

NoSQL数据库是可以与传统的关系型数据库配合，一起解决数据存储问题的，那么在日常工作中，你用到了哪些NoSQL数据库呢？在选型的时候是基于什么样的考虑呢？欢迎在留言区与我分享你的经验。

最后，感谢你的阅读，如果这篇文章让你有所收获，也欢迎你将它分享给更多的朋友。



# 高并发系统设计 40 问

## 攻克高并发系统演进中的业务难点

唐扬  
美图公司技术专家



新版升级：点击「 请朋友读」，20位好友免费读，邀请订阅更有**现金**奖励。

## 精选留言：

- Lion 2019-10-11 06:39:15  
我们正有这些痛点，思路不错，能更深入一些就更好了！ [4赞]
- gogo 2019-10-11 10:00:51  
老师您好，你文中微博关系的例子中提到传统关系型数据库扩展性的短板，又说明了nosql有很好的性能和扩展性，请问拿微博关系的例子来说，是可以单独用nosql存储呢，还是既要存在nosql中又要保存在mysql中呢？ [2赞]

作者回复2019-10-11 10:25:21

大部分放在mysql，像是粉丝数据是在nosql

- 饭团 2019-10-11 09:26:40  
老师，再问个问题！如果一类数据需要在nosql和关系形数据库都存储，需要存2次吗？现实开发中这样的情况多吗？能举个例子吗？ [1赞]

作者回复2019-10-11 10:25:49

是的呀 比如你的业务数据放在mysql, 索引数据放在es

● 饭团 2019-10-11 08:41:19

老师, 是不是这个意思:

首先不管是sql数据库还是nosql数据库, 提升写性能都是靠的将随机写转化为顺序写! 在这方面mysql使用WAL机制已经做的很好了! 但是关系型数据库主要是在扩展性方面有缺陷! 相对于除了kv型的nosql数据库, 其他类型的和mysql性能都差不多!

感觉mysql主要是在扩展性上有所欠缺! [1赞]

作者回复2019-10-11 09:13:29

benchmark结果来看, nosql的写入性能要好一些

● jc9090kkk 2019-10-11 15:41:56

感谢老师分享, 对于文中的知识点有一些疑问, 希望老师能解答下:

1.采用LSM树的nosql, 跟mysql的WAL机制类似, mysql通过redo log和bin log的两阶段提交来保证数据的crash-safe, 那么nosql的日志是用来做什么呢? 也是为了做数据恢复的吗?

2.像Hbase这样的nosql, 涉及的初表应该是为了做统计而设计的, 它支持事务吗? 如果支持的话, 使用场景是什么? 事务的隔离级别都有哪些? 因为以我的理解, 统计业务, 基本都是读请求, 很少有写请求, 用事务的场景应该极少吧?

3.因为nosql相比传统的关系型数据库来说, 拓展性更好, 那么就更适合做分布式系统? 后期老师有想法讲一讲比如类似redis或者mongoDB这样的系统, 分布式系统的一些实际经验或者说踩过什么坑么? 很想了解下。。。

思考问题:

公司现在用的nosql包括Redis, ES, Kafka, ClickHouse, 采用Redis是用来做cache和秒杀活动, 也会有一些异步化的处理, 采用Kafka+ES是一是为了采集业务端用户行为, 二是ES+Kibana可以提供可视化的数据分析平台供运营部门使用, 而且ES在后期的水平拓展的方案上来讲配置和维护比较简单(相比Sphinx而言), 为什么采用是因为公司之前接的是第三方的数据统计平台, 后期发现统计需要单独定制而且会污染业务代码不好维护, 而且有时候会发现数据日志还有丢失的情况导致数据统计异常, 所以采用Kafka+ES一起来支撑, 从而将业务端的一些耦合较重的埋点逻辑分离出去便于维护, ClickHouse主要是为了生成一些离线的数据报表。

● 吃饭饭 2019-10-11 15:41:54

读完这篇文章, 有些感悟: 技术方案层出不穷, 正是因为有了不同的需求, 才会衍生出众多的技术和框架, 只有真正明白技术为何产生和如何使用才会整体把控自身的需要, 有可能正是因为当前市场中没有能解决自身需求的框架, 我们才会真正考虑`写一个`: )

● 哇哦 2019-10-11 13:21:34

最近正好碰到分库分表的问题, 我们有一个记录表几千w,基本上都是插入和读取, 没有连表查询。这种情况是选择mysql不分库先分表(就一张表暂时先不分库), 还是直接用mogodb之类的(直接放进入建个索引就可以不管了?)。如果是分库分表有什么中间件和框架推荐吗(sharding-jdbc,mycat?),谢谢老师。

● 正在减肥的胖籽。 2019-10-11 11:42:25

老师可以举几个实战的例子吗?

比如: 像微博这种帖子列表, 某个主题下面的帖子列表.发帖量和用户查询量都比较大,这种使用Nosql, 我的理解是: 异步保存到数据库中(在app端先给一个占位的地方显示用户发的帖子)。然后同步到REDIS中. 我把握不好是: 同步到redis中, 1.过期时间需要设置多久? 2.如果按照不同维度排序展示(比如点赞量, 比如评论量)这种redis中应该如何存取? 老师可以针对这种场景讲解一篇吗?

作者回复2019-10-11 12:12:45

可以的 在实战篇中会讲到

- yaomon 2019-10-11 11:39:10

目前的项目就是 MySQL 配合 Elasticsearch 使用。Elasticsearch主要是做搜索用，写还是 MySQL，同时发 Kafka消息，消费端写ES。ES存在丢数据的问题，所以会定时全量/增量从 MySQL 中捞数据，覆写 ES，保证数据的一直性。

作者回复2019-10-11 12:12:00

是的 应该是这样