

4. 第四次笔记

学生：张帅豪 18030100101

老师：李龙海

定义：

将分布式系统中多个节点的存储资源整合在一起，向用户应用程序呈现统一的存储空间和文件系统目录树。用户无需关心数据存储在哪个节点上。大文件被自动分块并分别存储到不同的节点上。

目标：

- 提高数据存储容量：系统的总存储容量是多个组成节点本地存储容量的聚合，系统总存储容量能够随着组成节点数的增加而增大。（水平可扩展）
- 提高数据吞吐量：系统的数据读写吞吐量能够随着组成节点数的增加而增大。（水平可扩展）
- 提高可靠性/可用性：部分存储节点发生故障时数据不丢失；部分存储节点失效时，用户仍然可以读写数据。（容错性）
- 降低数据访问延时：使用户可以从地理位置上最接近的节点读取数据。（例子：CDN内容分发网络）
- 提高分布式数据处理系统的运行效率。

基本手段：

复制

基本思想：

在多个不同的节点上保存相同数据的多个副本（replica）。

用途

- ①复制提供了冗余，如果一些节点不可用，剩余的节点仍然可以提供数据服务。
- ②多个节点上存储副本也可以提高数据吞吐率、改善访问性能。

带来的问题：

- ①硬件成本提高。
- ②在数据变更时保障多数据副本之间的一致性是个复杂问题。（需要根据具体应用场景进行权衡）

分区

基本思想：

将一个大型数据文件(或数据库)拆分成较小的子集（称为分区partition或切片shard），再将不同的分区指派给不同的节点。

用途

- ①提高了吞吐率：访问数据的负载被分散到多个节点上。
- ②提高了可靠性：鸡蛋被放到多个篮子里
- ③方便了实现数据的并行处理。

带来的问题：

- ①跨区查询（分布式索引问题）
- ②合理、动态分区问题（大数据如何拆分）
- ③负载均衡问题（各个切片如何合理地分配给不同节点）
- ④分布式事务处理

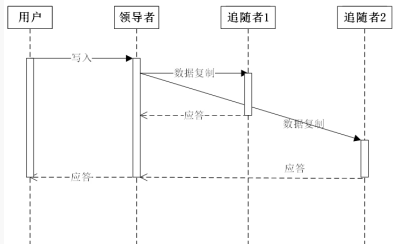
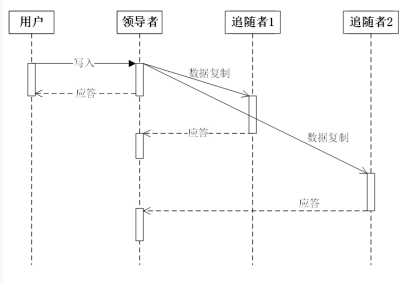
基于领导者的复制

基本思想：

- ①副本之一被指定为领导者(主库)；其他副本被指定为追随者（从库）
- ②客户端要向数据存储系统写入数据时，它必须将请求发送给领导者；领导者将新数据写入本地存储，同时也会将数据变更发送给所有的追随者。
- ③当客户想要从数据存储系统读取数据时，它可以向领导者或追随者查询。
- ④适合于读多写少的应用场景。

分类复制方法

| 单主复制： | 多主复制： | 无主复制： |
|---|--|---|
| ①一个领导者多个追随者。 | ①系统中有多个主库接受写入操作。每个主库都将该数据更改转发给所有其他节点。每个领导者同时扮演其他领导者的追随者。 | ①没有主库从库的概念。 |
| ②优点：实现相对简单。大多数应用场景都是读多写少。 | ②应用场景：多数据中心之间的复制；协同文档编辑。 | ②客户端直接将写入请求发送到各个副本；或者客户端将写入发送给某个节点，由该节点充当代理节点向其他节点转发写入请求。 |
| ③缺点：主库即是性能瓶颈，又是单点故障节点 (Single point of failure) | ③优点：写入性能高。 | |
| ④为提高容错性，自动将某个从库切换为主库时会面临脑裂问题。 | 缺点：写入冲突问题、多副本一致性问题解决方案太复杂。 | |

| 同步复制: | 异步复制: | 混合复制: |
|---|--|-------------------------------------|
| ①领导者写入自己的存储器, 并且接收到所有跟随者已经成功更新本地存储的应答后再向用户返回成功写入应答。 | ①领导者自己写入成功后立即向用户返回成功应答, 不等待其他跟随者的应答消息。 | 与部分跟随者节点之间采用同步复制, 另外的跟随者节点之间采用异步复制。 |
| ②优点: 副本之间的一致性好; | ②优点: 写入速度快; 容错性好; | |
| ③缺点: 写入速度慢; 一个节点失效会使整体写入功能失败 | ③缺点: 保持多副本一致性复杂; | |
|  |  | |

- 日志更新: 数据副本节点一般分两步处理数据写入(或复制)请求: 先将更新操作信息追加到更新日志
- 快照的补充