海量分布式存储系统**Doris** 的高可用架构设计分析

Doris（https://github.com/itisaid/Doris）是一个海量分布式 KV 存储系统，其设计目标是支持中等规模高可用可伸缩的 KV 存储集群。跟主流的 NoSQL 系统 HBase 相比较（Doris0.1 VS HBase0.90），Doris 具有相似的性能和线性伸缩能力，并具有更好的可用性以及更友好的图形用户管理界面。

对于一个数据存储系统而言，高可用意味着两个意思：

高可用的服务：任何时候，包括宕机、硬盘损坏、系统升级、停机维护、集群扩容等各种情况下，都可以对系统进行读写访问操作。

高可靠的数据：任何情况下，数据可靠存储，不丢失。

那么高可用的架构设计也就主要是在各种软硬件故障情况下，系统如何保障数据可靠存储，服务可用。

1 分布式存储系统的高可用架构

对于一个大规模集群的存储系统而言，服务器宕机、交换机失效是常态，架构师必须为这些故障发生时，保证系统依然可用而进行系统设计。系统架构层面，保证高可用的主要手段是——冗余：服务器热备，数据多份存储。使整个集群在部分机器故障的情况下可以进行灵活的失效转移（Failover），保证系统整体依然可用，数据持久可靠。Doris 系统架构如图 11.1 所示。

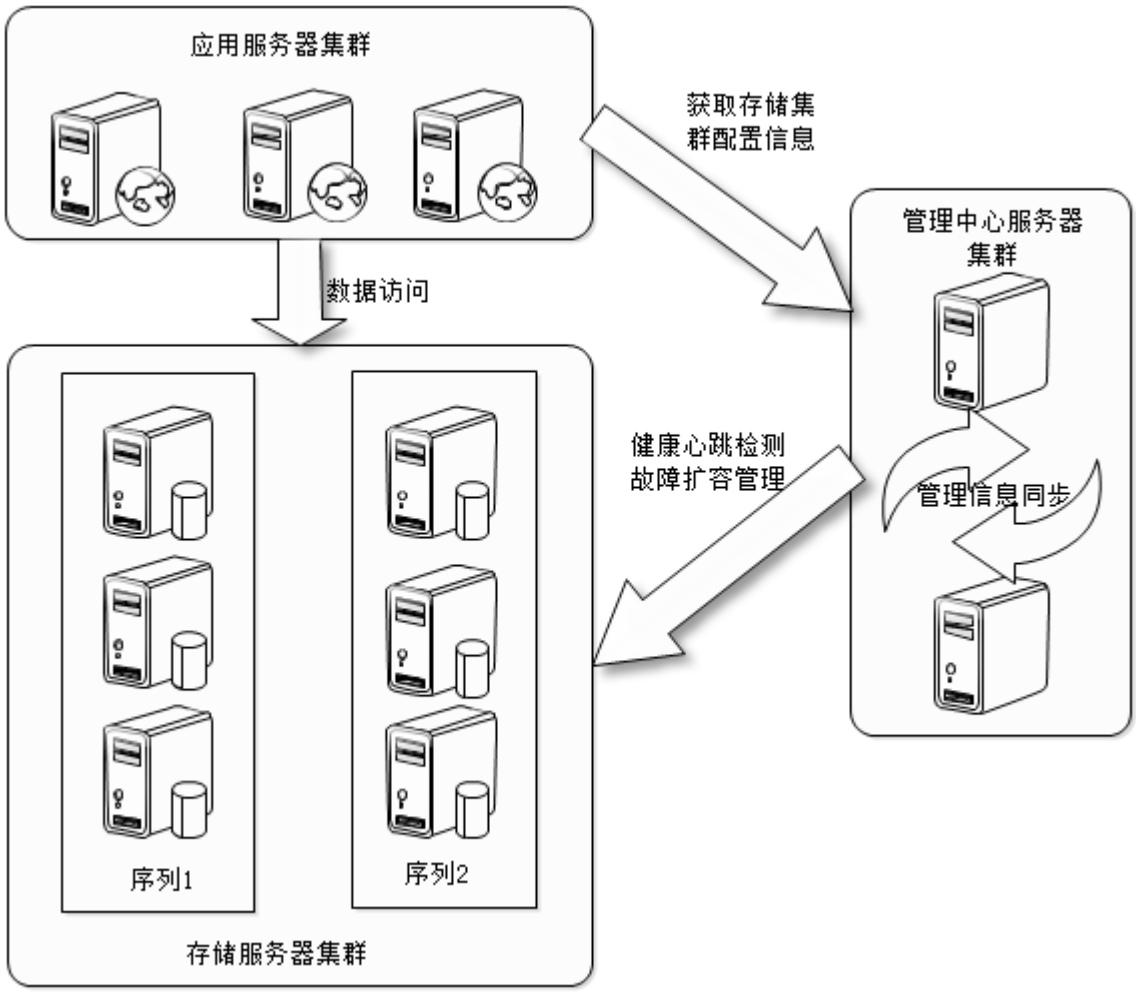

图 11.1 Doris 的整体架构

系统整体上可分为三个部分

- **应用程序服务器**：它们是存储系统的客户，对系统发起数据操作请求。
- **数据存储服务器**：存储系统的核心，负责存储数据、响应应用服务器的数据操作请求。
- **管理中心服务器**：这是一个由两台机器组成的主－主热备的小规模服务器集群，主要负责集群管理， 对数据存储集群进行健康心跳检测；集群扩容、故障恢复管理；对应用程序服务器提供集群地址配置信息服务等。

其中数据存储服务器又根据应用的可用性级别可以设置数据复制份数，即每个数据实际物理存储的拷贝数目，复制份数越多，可用性级别越高，当然需要的服务器也越多。为了便于管理和访问数据的多个拷贝，将存储服务器划分为多个序列，数据的多个拷贝存储在不同的序列中（序列可以理解为存储集群中的子集群）。

应用服务器写入数据的时候，根据集群配置和应用可用性级别使用路由算法在每个序列中计算得到一台服务器，然后同时并发写入这些服务器中；应用服务器读取数据的时候，只需要随机选择一个序列，根据相同路由算法计算得到服务器编号和地址，即可读取。通常情况下，系统最少写入的拷贝份数是两份。如图 11.2 所示。

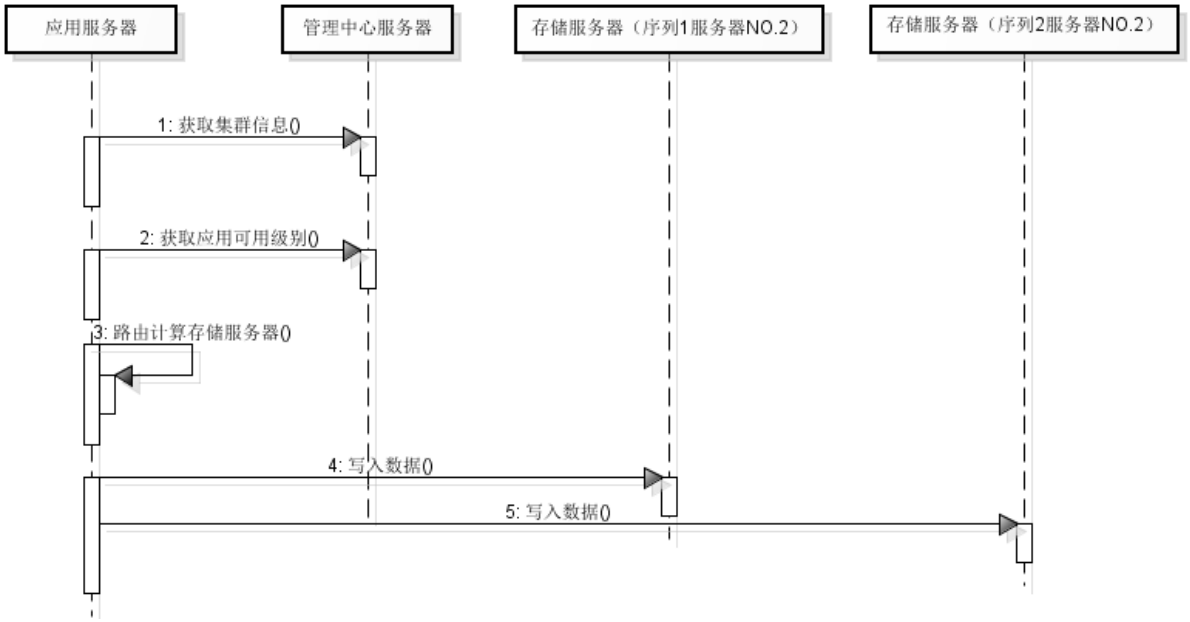

图 11.2 Doris 系统调用时序模型

在正常状态下，存储服务器集群中的服务器互不感知，不进行任何通讯；应用服务器也只在启动的时候从管理中心服务器获取存储服务器集群信息，除非集群信息发生变化（故障、扩容），否则应用服务器不会和管理中心服务器通讯。一般而言，服务器之间通讯越少，就越少依赖，发生故障时候互相影响就越少，集群的可用性就越高。

2 不同故障情况下的高可用解决方案

高可用的系统需要解决在不同故障情况下都保持较高的系统可用性，但是不同故障类型带来的问题复杂性不同，不可能使用一种解决方案处理所有情况，需要针对各种故障提供具体解决方案。

分布式存储系统的故障分类：

在讨论解决方案之前，我们先对故障进行分类，针对不同故障情况，分别处理对待。对于一个分布式存储系统而言，影响系统整体可用性的故障可以分成三类：

- **瞬时故障**：引起这类故障的主要原因是网络通讯瞬时中断；服务器内存垃圾回收或后台线程繁忙停止数据访问操作响应。其特点是故障时间短，在秒级甚至毫秒级系统即可自行恢复正常响应。
- **临时故障**：引起这类故障的主要原因是交换机宕机、网卡松动等导致的网络通讯中断；系统升级、停机维护等一般运维活动引起的服务关闭；内存损坏、CPU过热等硬件原因导致的服务器宕机；这类故障的主要特点是需要人工干预（更换硬件、重启机器等）才能恢复正常。通常持续时间需要几十分钟甚至几小时。故障时间可分为两个阶段：临时故障期间，临时故障恢复期间。
- **永久故障**：引起这类故障主要原因只有一个：硬盘损坏，数据丢失。虽然损坏硬盘和损坏内存一样，可以通过更换硬盘来重新启动机器，但是丢失的数据却永远找不回来，因此其处理策略也和前面两种故障完全不同，恢复系统到正常状态也需要更长的时间。故障时间可分为两个阶段：永久故障期间，永久故障恢复期间。

正常情况下系统访问结构

在只使用两份拷贝作为高可用策略的情况下，系统访问结构如图 11.3 所示。

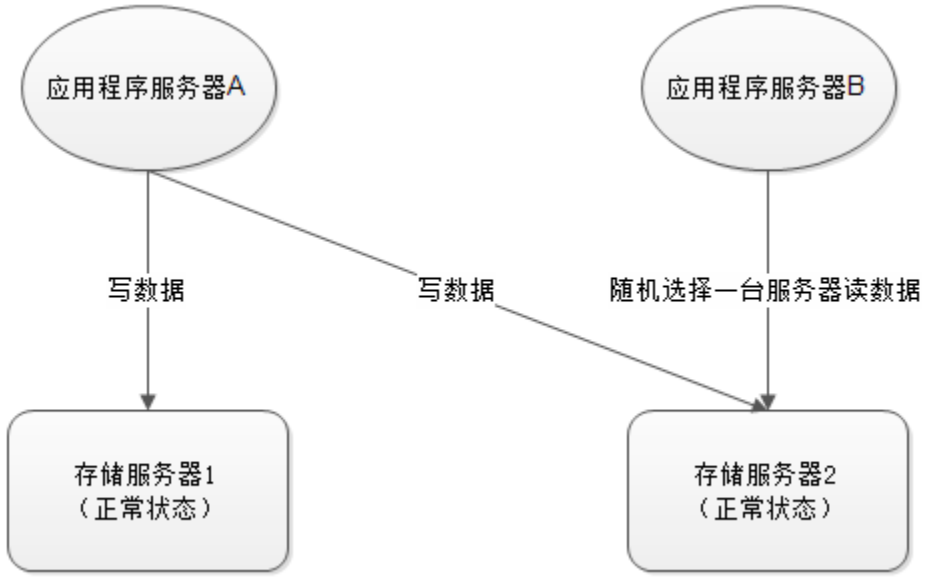

图 11.3 正常情况下 Doris 访问模型

应用程序在写数据的时候，需要路由计算获得两台不同的服务器，同时将数据写入两台服务器；而读数据的时候，只需要到这两台服务器上随机一台服务器读取即可。

瞬时故障的高可用解决方案

瞬时故障是一种严重性较低的故障，一般系统经过较短暂的时间即可自行恢复，遇到瞬时故障，只需要经过多次重试，就可以重新连接到服务器，正常访问。如图 11.4 所示。

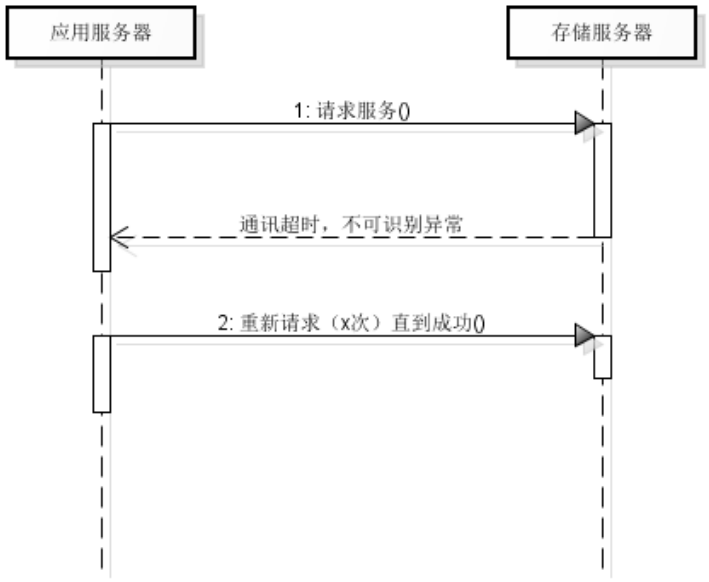

图 11.4 Doris 瞬时故障解决方案

如果应用多次重试后，仍然失败，那么有可能不是瞬时故障，而是更严重的临时故障，这时候需要执行临时故障处理策略。

当然也有可能是应用服务器自己的故障，比如系统文件句柄用光导致连接不能建立等，这时候需要请求管理中心服务器进行故障仲裁，以判定故障种类。

瞬时故障，系统访问模型如图 11.5 所示。

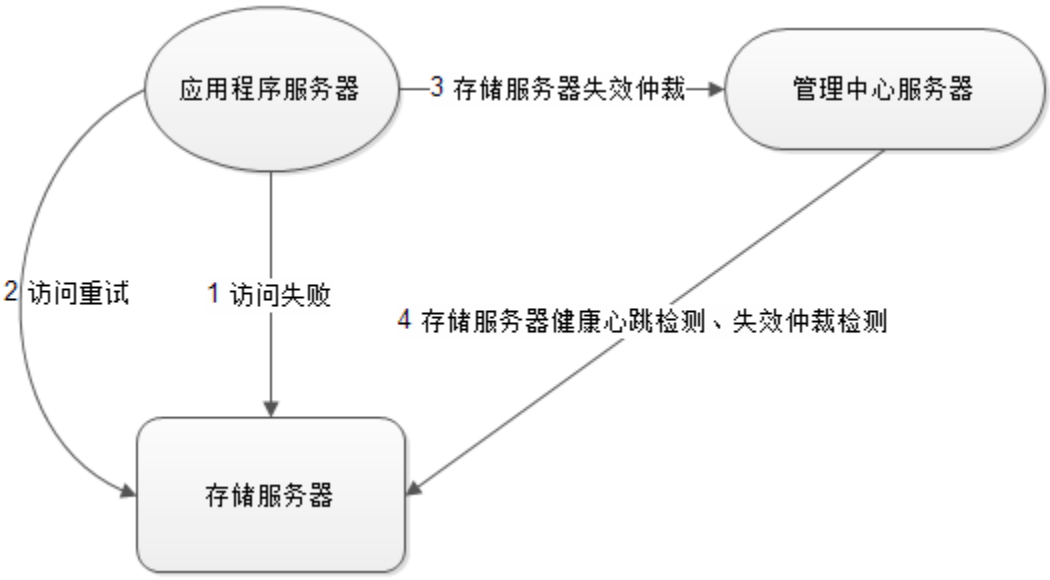

图 11.5 Doris 瞬时失效访问模型

临时故障的高可用解决方案

临时故障要比瞬时故障严重，系统需要人工干预才能恢复正常，在故障服务器未能恢复正常前，系统也必须保证高可用。由于数据有多份拷贝，因此读数据的时候只需要路由选择正常服务的机器即可；写数据的时候，正常服务的机器依然正常写入，发生故障的机器需要将数据写入到临时存储服务器，等待故障服务器恢复正常后再将临时服务器中的数据迁移到该机器，整个集群就恢复正常了。如图 11.6 所示。

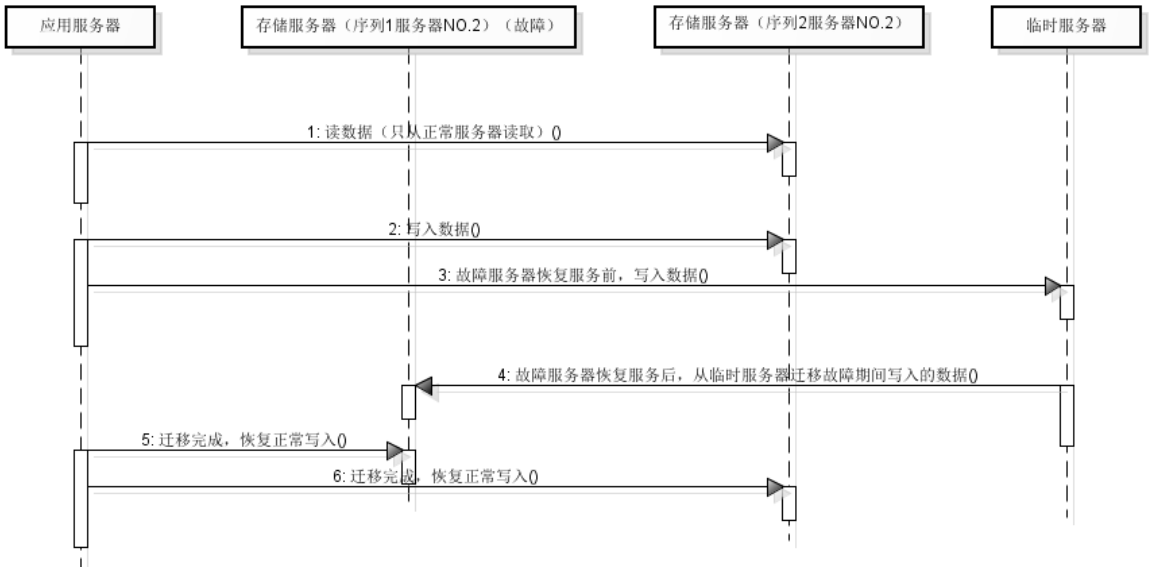

图 11.6 Doris 临时故障解决方案

其中临时服务器是集群中专门部署的服务器（根据可用性规划，临时服务器也可以部署为多台机器的集群），正常情况下，该服务器不会有数据写入，处于空闲状态，只有在临时失效的时候，才会写入数据。任何时候该服务器都不会提供读操作服务。

临时故障发生期间，系统访问模型如图 11.7 所示。

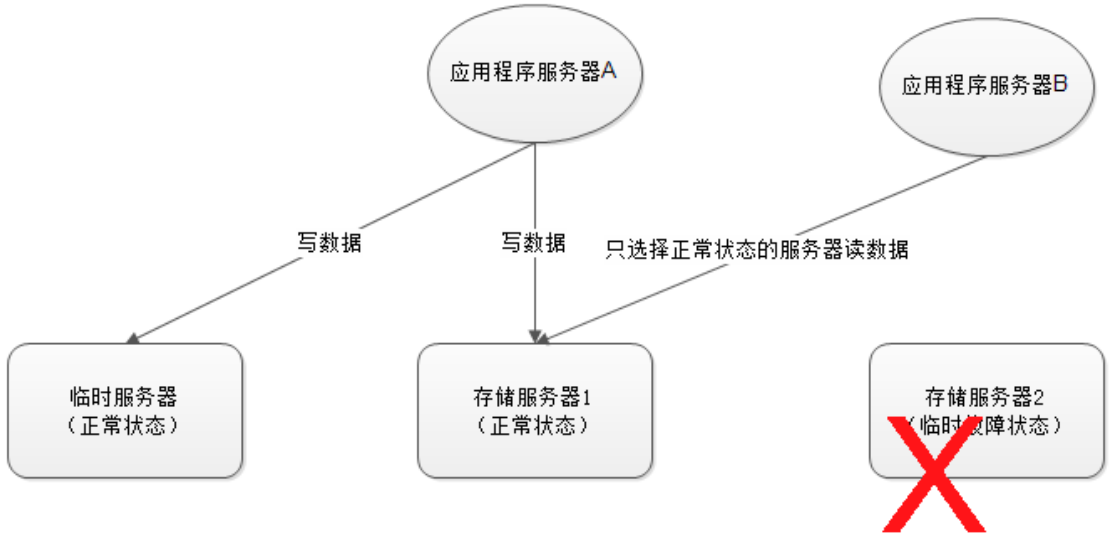

图 11.7 临时故障 Doris 访问模型

临时故障解决，系统恢复期间，访问模型如图 11.8 所示。

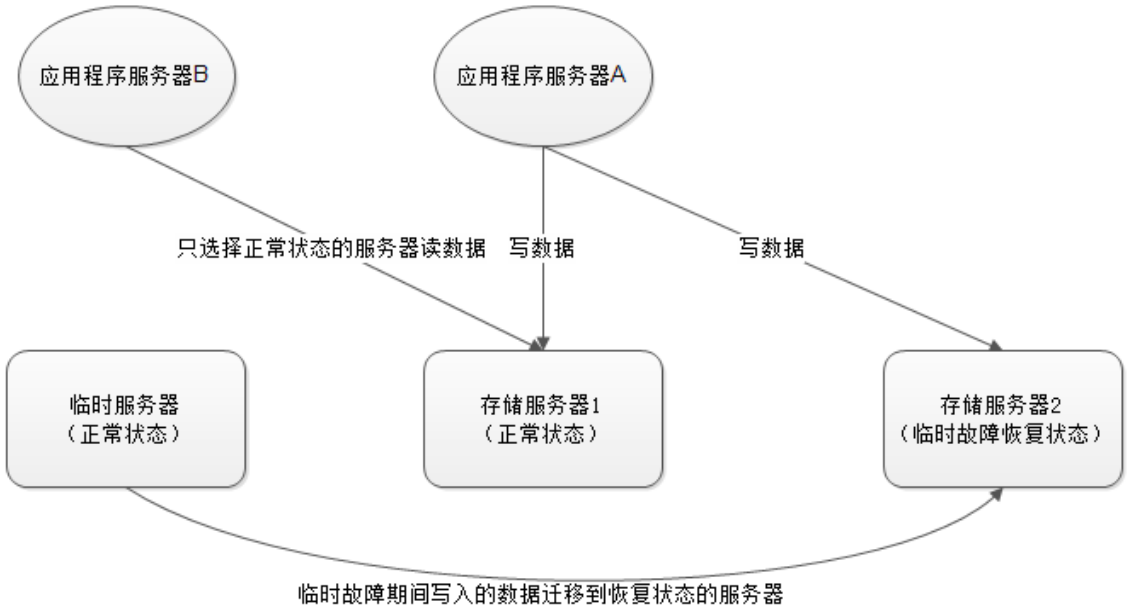

图 11.8 临时故障恢复期间 Doris 访问模型

临时故障期间写入临时服务器的数据全部迁移到存储服务器 2 后，故障全部恢复，存储服务器 2 恢复到正常状态，系统可按正常情况访问。

永久故障的高可用解决方案

永久故障是指服务器上的数据永久丢失，不能恢复。由于故障服务器上的数据永久丢失，从临时服务器迁移数据就没有意义，必须要从其他序列中正常的服务器中拷贝全部数据才能恢复正常状态。

永久故障发生期间，由于系统无法判断该故障时临时故障还是永久故障，因此系统访问结构和临时故障一样。当系统出现临时故障超时（超过设定时间临时故障服务器仍旧没有启动）或者人工确认为永久故障，系统启用备用服务器替代原来永久失效的服务器，进入永久故障恢复，访问模型如图 11.9 所示。

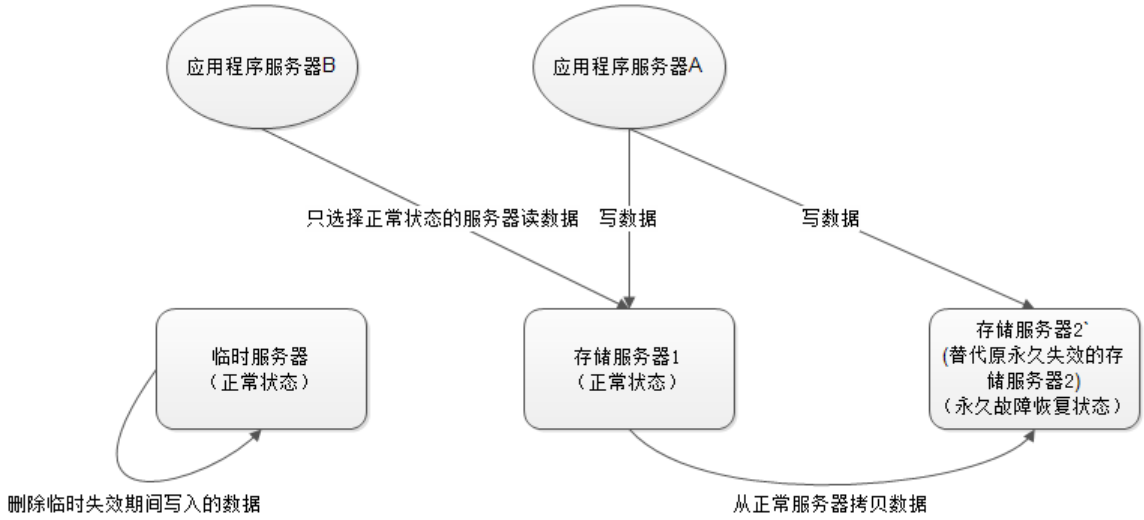

图 11.9 永久故障恢复期间 Doris 访问模型