

基于云计算分布式存储架构与容错技术的研究

申高

(安阳职业技术学院, 河南 安阳 455000)

摘要 随着信息技术的飞速发展,传统的存储架构已无法满足大数据存储的要求,因而分布式存储得到了广泛应用。特别是随着计算机技术和云计算技术的发展,基于云计算的分布式存储概念被提出来,并在某些领域得到了应用和发展。我们应该看到的是,该存储架构在有利于实现资源共享、增大存储容量、提升访问便捷性的同时,也带来了错误率高、较为容易被攻击等问题。因此,系统研究基于云计算分布式存储架构与容错技术,具有很高的理论和实践意义。本文在简要论述相关技术概念的基础上,对云计算分布式存储架构进行了详细论述,并提出了具有创新性的容错技术。

关键词 云计算;分布式存储;容错技术;纠错码

中图分类号: TP3

文献标识码: A

文章编号: 1007-0745(2023)04-0016-03

在数据大爆炸的今天,人们对于数据的安全存储有着越来越高的要求。基于云计算的分布式存储技术相对于传统存储技术,大大提升了存储容量,提升了数据处理速度,提升了数据访问便捷性,但同时也对数据安全带来了越来越严重的挑战。因此,必须系统深入研究基于云计算分布式存储架构,并提出切实可行的容错技术,不断增强数据存储的安全性。

本文按照从理论到实际的思路,在简要介绍分布式存储、云计算和容错技术的基础上,对基于云计算分布式存储架构做了深入分析,并设计实现了基于汉明纠错码的容错技术。

1 关键技术简介

1.1 云计算技术

云计算(cloud computing)是基于互联网的相关服务的增加、使用和交付模式,通常涉及通过互联网来提供动态易扩展且经常是虚拟化的资源^[1]。云计算具有大规模、分布式、虚拟化、高可用性和扩展性、经济型和安全性等特点。

1.2 分布式存储技术

存储、计算和网络是构成现代化信息系统的三大部件。在2021年3月中央发布的“十四五”规划中,“大规模分布式存储”被列为明确的发展重点^[2]。所谓分布式存储,是相对于传统集中式共享存储的一种新技术概念。在过去30年里,中国企业普遍采用IBM、EMC、NetApp等国外知名厂商的磁盘阵列,其架构体

系主要是以RAID控制器为技术核心的集中式共享存储架构。分布式存储是通过软件定义存储技术,把多台x86服务器的CPU、内存、固态硬盘(SSD)、机械硬盘(HDD)等存储资源,进行统一资源管理,通过高速网络将数据分散存储在多台服务器中^[3]。

自云计算诞生以来,井喷式的数据处理业务量对IT的系统计算与存储能力都提出了非常高的要求。传统集中式存储面临数据孤岛、扩展困难、成本高昂等挑战,与此同时,分布式存储由于其易于扩展和管理、成本低、硬件灵活度高、可快速集成等明显优势,逐渐成为全球的技术发展趋势,也将作为中国数字化发展的重要基石。

1.3 容错技术

现代存储系统的本质是一个具备高度容错能力、能够在线更换故障或老旧部件的软硬件系统,需要同时满足几个方面的要求:高可靠性的数据存储基础、对数据错误具有自愈能力、正常运行和自愈过程均对前端应用无性能影响、以及较高的存储效率。存储系统的数据安全性就好比飞机的安全性,业务中断就好比重大飞行事故,数据丢失对于一个存储系统而言,就好比飞机失事,带来灾难性的破坏和难以估量的重大影响。

2 基于云计算分布式存储架构研究

2.1 主备机制

主备机制又称为主从机制,主节点(master)是提供数据的查询和更新服务,备节点(slave)在主节点

提供服务期间不提供对外服务。备节点从主节点获取更新消息,或主节点主动广播消息,一旦主节点挂掉,客户端可以通过负载均衡设备切换到备节点继续服务^[4]。像 redis、mysql 等做分布式方案都属于这种设计,通过 binlog/oplog 同步到备节点。

2.2 对等机制

集群中每个节点都是一样的对等结构,每一个节点都保存了数据的元信息以及数据,每个节点都提供数据的查询和更新服务。

在数据容灾方面,集群会对数据进行分片,且分片有多个副本的存在这些节点中,一旦某个节点挂掉,客户端通过某种哈希算法找其它副本即可继续服务。为了更好地提升云计算之下的分布式存储性能,使用图论设计环状的对等架构,成为对基本模式的改进与创新方式。在针对不同场景的读取和更新的优化上,一般使用 QUORUM 机制,保证 $W+R>N$ 即可。像 Cassandra、Dynamo 等都属于这种设计^[5]。

2.3 数据和集群管理

把集群中数据的元信息以及数据以及集群的管理信息进行分离,数据节点主要根据其元信息管理数据的增删改查,数据节点统一由集群的服务进行管理,集群的管理服务也是分布式,每个管理节点可能是主备机制,也可能是对等机制。

数据节点启动时,向管理节点汇报服务信息,后续每隔一段心跳时间汇报一下数据节点的状态(包括数据状态、服务健康状态等),若某个数据节点挂掉,心跳不能继续进行,故客户端连接时候,抽取心跳时间间隔内的服务器进行访问即可^[6]。

针对心跳时间间隔内的服务器突然不可用,客户端可以采取一定优化措施。有些分布式系统为了元数据的统一管理和方便查询,将元数据信息甚至配置信息统一用新的分布式方案进行部署管理。像 Tair、hyperdex 等属于这种设计方式。

3 一种基于汉明码的容错技术的设计与实现

3.1 基于纠错码的容错技术

纠错码(error correcting code)是在接收端能自动地纠正数据传输中所发生差错的码。纠错码的基本思路是在所有的由发送符号组成的序列中,仅挑出其中一部分作为信息的代表向信道发送,并使得所挑出的这些序列之间有尽可能多的差异。每个被挑出的允许发送的序列被称为一个码字,而码字的总合称为码。在发送端把信息变换成码字的过程称为编码;在接收

端从接收到的信号判定所发码字、从而恢复信息的过程称为解码(或译码)。在解码时,若收到的信号不是码中的一个码字,则可以肯定在传输中出现了差错,从而着手对差错进行纠正。纠错的方法是找到与接收到的信号最接近的码字,并将其判定为发送信号。

3.2 汉明码的原理分析

在信息理论中,汉明距离表示两个等长字符串在对应位置上不同字符的数目,我们以 $d(x, y)$ 表示字符串 x 和 y 之间的汉明距离。从另外一个方面看,汉明距离度量了通过替换字符的方式将字符串 x 变成 y 所需要的最小的替换次数。

对于二进制串 a 和 b 来说,汉明距离等于 $a \text{ XOR } b$ 中 1 的数目,我们又称其为汉明权重,也叫做 population count 或 popcount。长度为 n 的二进制字符串通过汉明距离构成了一个度量空间(metric space),我们称其为汉明立方(Hamming Cube)。采用汉明码纠错,需要按以下步骤来进行:计算校验位数→确定校验码位置→确定校验码→实现校验和纠错。

3.3 基于汉明码的容错技术实现

基于汉明码的纠错技术,很多学者已经开展了较为深入系统的研究。并且为了进一步改善性能,动态汉明码和扩展汉明码纠错技术相继被提了出来。本文使用文献^[7]中的扩展汉明码动态纠错机制。

3.3.1 纠错冗余的确定

使用容错技术进行纠错,本质上就是在发送端对所发送的数据添加一定程度的冗余,冗余量的大小有着明确的计算方法。由于数据在计算机中存储、发送、计算等都是使用二进制的方式,因此使用二进制数据最小单位“位”来表征数据容量和冗余的大小。假设数据总量的位数为 n ,冗余量的位数为 i ,则冗余量能够容错的数据量为 2^i ,则应该满足 $2^i > n+i$ 即可。假设数据总量位数为 8 位,则 i 至少为 4,数据总量位数为 32 位,则 i 至少为 6。由此可见,数据总量位数越大,则冗余位数占数据总量位数的比值越小,即单位冗余能够纠错的数据量越大,或者说纠错效率越高。

3.3.2 总体设计

采取面向对象方法,设计纠错模块、监督模块、发送接收模块、实时显示模块、计算模块共五个模块,每个模块定义一个类。为了充分发挥不同语言的技术优势,本程序使用 C# 与 Matlab 混合语言编程的形式,充分利用 C# 语言在可视化界面设计、网络编程、面向对象、流程化实现等方面的优势,利用 Matlab 语言在大数据计算与存储方面的优势,即内层的计算和数据

表1 模拟实验结果对比表

考察指标	纠错数据量	正确率	速度	占用功耗	报警次数
	1 万字节	95.6%	644 字节 / 秒	250uW	0

转换用 Matlab 语言完成,通过编制 .m 格式函数的形式,生成对应的 .dll 文件,添加到 C# 语言编程系统的命名空间中进行调用。因此另外定义一个语言转换类,用于外层 C# 语言与内层 Matlab 语言的转换。

为了实现对于纠错正确率,纠错效率的实时检测,本文定义 5 个数据库,用 Sql server2015 来实现。分别存储接收数据、便准对照数据、实时检测数据、待显示数据、最终计算数据。为了保证数据库的安全,使用 MD5 算法进行加密。

软件设计整体上要界面清晰、简洁,人机交互性强,算法时空复杂度低,具有较好的健壮性,自动检测异常输入等情况。

3.3.3 纠错模块设计

在纠错模块中,通过系统接收的数据,甄别出待测数据,使用文献^[8]中扩展汉明码进行动态纠错。将处理之前的和纠错之后的数据分别进行保存,并通过关键参数在多窗体之间进行传递,用于实时传输到显示模块中。

3.3.4 监视模块设计

在监视模块中,主要监视的指标为纠错速度、当前纠错量、纠错功耗等,并对其他几个模块的运行正常与否进行密切跟踪,特别是实时显示模块和发送接收模块,跟踪能否正常运转。监视模块设立了报警功能,在出现有明显的错误可能性时,会发出语音报警,提示用户中断当前的操作,进行检查与调试。

3.3.5 计算模块设计

在监视模块中,主要通过对数据量的实时获取,计算当前纠错的总量、正确率、纠错后总数据容量、发送接收效率等。该模块能够实时与相应数据库进行互动操作,实现数据库的增删改查操作,并能够实时将数据库中的变化体现到自身的运行数据中。

3.3.6 发送接收模块设计

在发送接收模块中,主要是进行数据的实时发送与接收,将不同计算机连接成真正的网络,实现“云计算”功能。该模块发送接收功能的实现,通过以 TCP 连接为主,UDP 连接为辅的形式进行。在过多强调数据安全性与传输可靠性的情况下,使用前者进行连接;在过多强调数据传输速度和实现群发功能的情况下,使用后者进行连接。

3.3.7 实时显示模块设计

在实时显示模块中,主要是通过良好的 GUI 界面设计,将实时监测及其他运行数据显示在窗体中,并能够与用户进行简单的互动。该模块设置了异常提示功能,在 3.3.4 所述语言报警的同时,能够使用文字和弹出框的形式,构筑起提示用户异常的双保险。

3.4 实验仿真

实验硬件平台为 Intel Core i5-5200U (2.20G Hz),4G 内存,软件平台为 windows10 (64 位)操作系统, Matlab2016Rb 软件, VS2010 软件, C#3.0 版本。

通过对数据进行模拟操作,得到的最终模拟实验效果如表 1 所示。由此可见,该方法纠错正确率较高,但速度上有待于进一步提升。整个实验过程报警次数为 0,说明系统能够正常运转。

4 总结

基于云计算分布式存储系统架构与容错技术,并没有一成不变的范例,而是根据使用需求的不同可以进行灵活选择。在实践中需要针对不同使用场景和使用目的,按照其自身属性的不同综合选用合适的系统架构与容错技术,不断降低数据安全问题发生的频率,提升大数据存储、挖掘与处理的实际应用效果。

参考文献:

- [1] 王珍.基于云计算分布式存储架构与容错技术探讨[J].网络安全技术与应用,2022,264(12):56-57.
- [2] 李振.基于云计算分布式存储架构与容错技术的研究[J].网络安全技术与应用,2022,253(01):64-65.
- [3] 谭鹏许.云计算环境下安全分布式存储架构与容错技术研究[D].郑州:解放军信息工程大学,2013.
- [4] 徐亮亮.纠删码容错存储系统的编解码流程优化研究[D].合肥:中国科学技术大学,2022.
- [5] 董乾.采用哈夫曼编码技术提高硬件无损压缩效率的算法研究[D].南京:东南大学,2018.
- [6] 蔡媛媛,赵致琢.分布式容错计算描述语言及其应用研究[J].计算机科学,2016,43(05):146-149,168.
- [7] 陈天培,吴校生,强小燕.一种基于扩展汉明码的动态纠错机制[J].电子与封装,2022,22(05):67-71.
- [8] 同[7].