

大数据视野下云存储数据可靠性技术研究¹

李胜嘉

南京工程学院 211167

摘要: 【目的】探究云存储数据可靠性技术,降低云数据中心数据丢失的风险,节省成本及能源。【方法】尝试通过使用镜像和纠删码技术来提高云计算中大数据应用程序的数据可靠性,并进行全面比较。【结果】云存储在为大数据应用程序提供服务方面发挥了主要作用,因为它提供了具有成本效益的按需服务。【结论】在大数据时代,提高对数据故障的容错能力会带来各种挑战,随着故障已成为云存储系统中的常态,各种容错机制已在云存储系统中采用,以提高数据可靠性。

关键词: 大数据; 云存储; 数据可靠性; 镜像; 纠删码

Research on cloud storage data reliability technology from the perspective of big data

Li Sheng-jia

Nanjing institute of technology 211167

Abstract:

Objective: Explore cloud storage data reliability technology, reduce the risk of data loss in cloud data center, and save cost and energy. **Method:** Try to improve the data reliability of big data applications in cloud computing by using image and erasure code correction technology, and make a comprehensive comparison. **Result:** Cloud storage plays a major role in serving big data applications because it provides cost-effective on-demand services. **Conclusion:** In the era of big data, improving the fault tolerance of data faults will bring various challenges. As faults have become the norm in cloud storage systems, various fault tolerance mechanisms have been adopted in cloud storage systems to improve data reliability.

Key Words: Big data; Cloud storage; Data reliability; Mirror image; Erasure code

1 引言

大数据时代下,数据量的增长远超存储设备扩容升级的速度。每周,新浪微博仅需约 60TB 的空间即可存储用户的文字、照片等信息,而弹幕视频网站 bilibili 每天则需要 1PB 的新存储空间来存放 2 亿用户每分钟上传的累计约 200 个小时的视频^[1]。国际数据公司 (IDC) 的第六次年度研究表示,到 2020 年,世界数据存储量将每两年翻一番,互联网中大约 40% 的数据将由云计算提供存储或处理^[2]。由此可见,云计算提供了一种经济高效的方式来推动互联网服务和数据处理业务的发展,而云存储为存储大数据提供了合理的可扩展性,并且有助于

¹ 本文为南京工程学院挑战杯项目 (编号: TP20190006) 资助研究成果; 国家级大学生创新创业训练计划项目 (编号: 201911276018Z) 资助研究成果。

处理不同种类的数量和速度稳定增长的数据。由于云存储是建立在商用服务器和磁盘驱动器的基础之上，因此，故障的产生不可避免。例如，facebook 在 2009 年受硬盘故障的影响丢失了约 10%的用户数据^[3]。故障是云存储系统中的常态，因此在数据恢复期间提高数据可靠性并保持系统性能是在云计算上部署大数据应用程序时最重要的挑战之一。

云存储中的数据故障通过各种数据冗余技术来处理，较为常见的冗余技术有镜像和纠删码技术。镜像是一种简单的数据冗余机制，该技术会复制相同的数据并将其存储在云存储系统上的多个位置，如果所需数据在一个磁盘中失效时，则会从下一个可用的磁盘找到该数据^[4]。纠删码技术是一种更复杂的数据冗余技术，该技术通过创建奇偶校验数并将其与原始数据一起存储，同镜像技术相似，当请求的数据不可用时，则可以从奇偶校验数据中重建^[5]。存储纠删码与镜像技术相比显著降低了成本，因此采用纠删码技术可以减少数据存储的硬件需求，并为数据中心节省了大量成本和能源。

本文尝试研究通过使用镜像和纠删码技术来提高云计算中大数据应用程序的数据可靠性。由于这两种技术各有利弊，笔者首先讨论了科研工作者对上述技术的研究现状，然后将两种技术进行全面比较并讨论其面临的问题与挑战，最后，对云存储可靠性未来的发展趋势进行展望。

2 相关研究

2.1 云存储系统

云存储系统由网络连接的众多存储设备组成，它分为具有存储虚拟化功能的网络附加存储（NAS）和存储区域网络（SAN）类型的分布式存储。存储虚拟化是一种从应用程序中抽象物理存储并将逻辑存储映射到物理存储的技术，可以将存储设备网络视为单个存储设备，并且用户可以访问信息，而不考虑物理位置和存储模式。

根据客户端访问和接口数据方式不同，云存储系统可以分为文件存储，块存储和对象存储。

2.1.1 文件存储

在文件存储中，文件是按层次结构组织排列的。文件的信息作为元数据存储存储在存储系统中，可以通过指定单个文件的路径来访问文件。文件存储为应用程序提供了更高级别的存储抽象，并且可以在不同平台之间安全地传输数据。

2.1.2 块存储

在块存储中，文件分为几个块，并为每个块分配一个地址。应用程序可以访问这些块并将其与块地址合并。存储应用程序保留元数据并使用其来共享数据。

块存储没有任何文件服务器可以授权 I / O，且客户端可以使用元数据直接访问存储设备。块存储拥有较高的性能，但数据传输的安全性得不到保障。

2.1.3 对象存储

在对象存储中，文件和元数据被封装为一个对象，并为该对象分配了一个对象 ID。该对象可以是任何类型，且可以为每个对象分配唯一的元数据，如关联的应用程序对象的类型，保护级别，复制数量和地理位置等。对象存储将存储管理从应用程序转移到存储设备，故可以使用元数据对客户端进行安全的直接数据访问。对象存储拥有较为出色的可伸缩性以支持大数据应用程序，且除镜像技术外，对象存储还支持有效的纠删码技术。

2.2 数据故障

在云存储系统中，许多因素都可能直接或间接导致数据故障，进而导致云服务故障。Sharma 等对云服务故障原因进行了详细调查^[6]，数据故障的主要原因是硬件，软件，网络和电源故障。如图 1 所示，所有服务器故障中的 78%是由于硬盘引起的，其中 5%是由磁盘阵列（RAID）控制器引起的，3%是由于内存引起的，其余 14%是由于其他因素引起的。硬盘是最常更换的组件，它们是服务器故障的最常见原因。

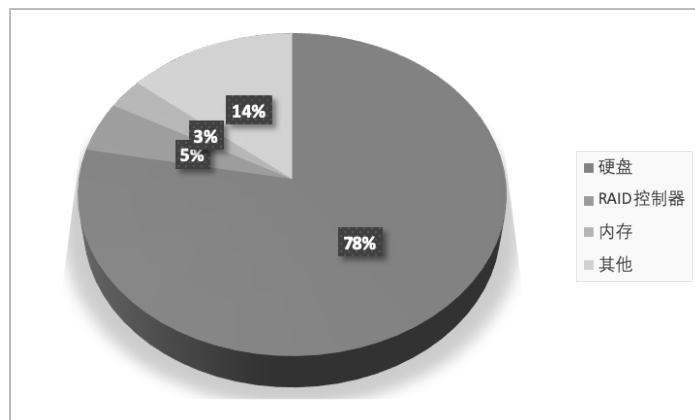


图 1 数据故障原因

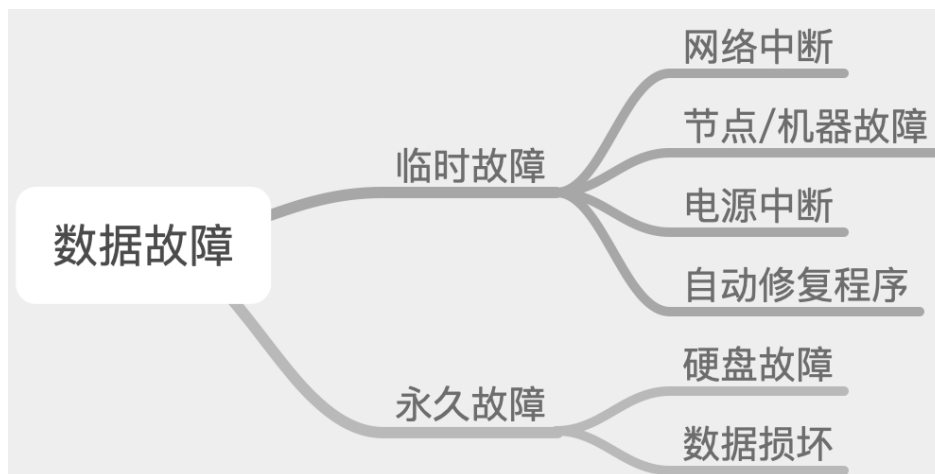


图 2 云存储中的数据故障

如图 2 所示，数据故障可能是临时或永久的。由于网络中断，节点/机器故障，电源中断和自动修复过程而导致的数据不可用是临时性故障，不会导致永久性数据丢失。云存储系统中的各种组件故障会导致永久性和临时性数据故障，磁盘是云存储系统中最重要组件，如果处理不当，磁盘故障会导致永久性数据丢失。云存储系统中的其他组件故障往往仅导致临时中断，一些中断可能会持续数小时，造成巨大的经济损失。以上讨论为提高耐用性和可用性提供了一些启示。以下将讨论云存储系统中采用的各种数据可靠性机制。

2.3 数据可靠性

数据可靠性包括最大限度地提高数据的持久性和可用性。持久性减轻了永久性的故障，可用性减少了临时性的故障。



图 3 云数据中心的故障处理

如图 3 所示，云数据中心使用了各种机制来提高存储系统的容错能力。RAID 阵列，可交换驱动程序和错误校正代码内存（ECC RAM）减轻了硬件故障的影响。RAID 阵列是由几个磁盘组成的逻辑单元，这些磁盘以条带化，镜像和奇偶校验存储数据。可交换驱动程序允许管理员交换在系统保持操作模式时发生故障或预计发生故障的驱动器。ECC RAM 用于通过将奇偶校验位与每个二进制代码相关联来检测和纠正单个位错误。网络故障和断电分别通过网络冗余和双电源处理。

2.4 纠删码技术

在大型存储系统中，纠删码在保护数据免受故障中起着主要作用。在云计算出现之前，纠删码技术用于检测和纠正存储和通信系统中的错误。例如，图 4 表示可以容忍任何两个故障的（4，2）擦除码。用于计算奇偶校验数据的算法可以是标准算法或 Galois Field 算法。在标准算术中，加法作为二进制 XOR 进行，乘法作为二进制 AND。如果单词中的位数为 1，则执行标准算术。当单词中的位数增加时，将使用 Galois 字段算术来计算奇偶校验。在 Galois 字段 $GF(2^n)$ 中，

算术运算被绑定在从 0 到 2^n-1 的有限数字集中;加法是按位 XOR 运算, 乘法则更复杂, 这取决于硬件, 内存和一个字的位数。

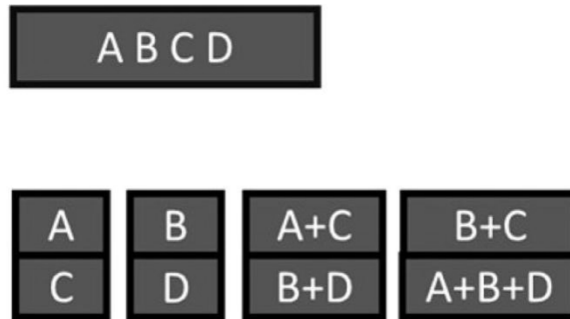


图 4 纠删码技术

2.5 镜像

镜像是在云数据中心使用的最常见的可靠性机制, 它以低延迟和最小带宽消耗来提高可用性和持久性^[7]。发生故障时, 为了保持持久性, 需要在活动磁盘中还原出故障的副本, 该恢复可以主动地进行。镜像中使用的常见方法是静态镜像和动态镜像。

3 镜像和纠删码比较

镜像和纠删码技术是云数据中心使用的重要可靠性机制, 可防止数据出现故障。重要的是要了解这些技术的利弊, 以便在云存储系统中实施最佳技术以提高可靠性并节省资金。关于各种参数的技术分析笔者将在下面详细描述。

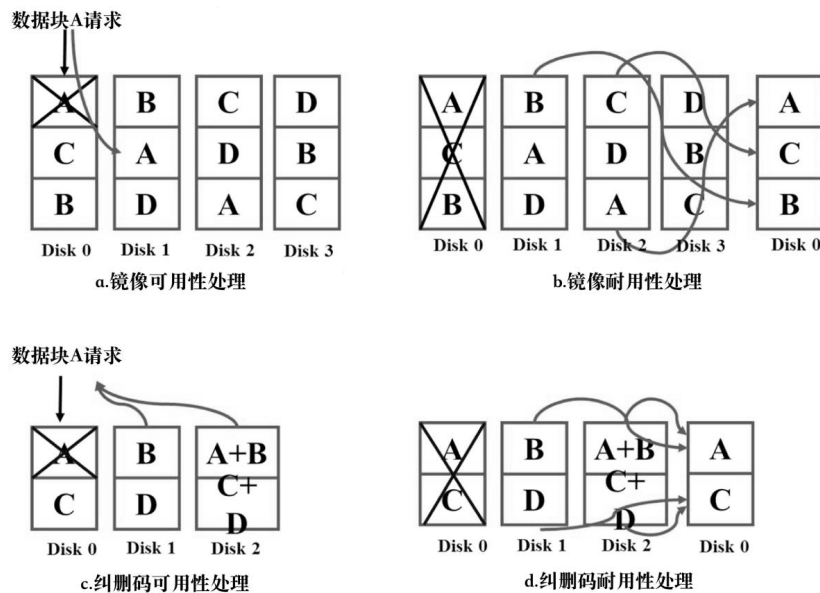


图 5 纠删码技术

图 5 展示了在镜像和纠删码技术中如何处理对故障数据的读取请求, 它还显示了在出现暂时性和永久性数据故障的情况下如何恢复数据。通过将请求重定向

到下一个可用副本即可满足对复制存储系统中故障数据的请求。另一方面，在纠删码存储系统中，暂时不可用的数据是通过动态地从下一个 k 个可用磁盘恢复数据来提供。与镜像存储相比，纠删码存储中的重建涉及更多的磁盘 I/O。例如，在图 5 中块 A 的重建涉及从两个不同的磁盘读取的两个数据块。与镜像相比，这增加了纠删码存储系统中读取请求的等待时间。

表 1 镜像与纠删码的比较

参量	镜像	纠删码
储存费用	高	低
可用性	低	高
耐用性	低	高
故障延迟	低	高
恢复费用	低	高
编解码复杂度	低	高

纠删码系统中永久性故障时的磁盘重建涉及的磁盘 I/O 数量要多于镜像技术。例如，在图 5-b 中，故障磁盘的重建仅涉及三个磁盘访问以重建三个数据片段。然而，在图 5-d 中，故障磁盘的重建涉及四个磁盘访问以恢复两个片段。这增加了纠删码系统中的恢复成本。

纠删码与镜像相比能够更高效地存储，但是在性能上却需要权衡^[8]。在纠删码存储系统中编码数据非常耗时，而对失败的请求可以无延迟地重定向到镜像系统中的下一个可用副本。在纠删码系统中，从下一个可用对象中重建失败的对象会增加读取请求的等待时间。此外，就带宽和磁盘 I/O 而言，重建对象的成本会很高。

表 1 详细比较了镜像和擦除编码，关键字高和低用于表示一种技术相对于另一种技术的优劣。例如，与镜像相比，纠删码有较高的存储效率，增加了持久性和可用性。

4 问题与挑战

4.1 存储效率

采用镜像技术的数据可靠性与存储开销成正比，因此，在不牺牲可靠性的前提下降低存储开销是镜像技术目前面临的巨大挑战。虽然纠删码技术以较高的可靠性降低了存储成本，但是在出现故障的情况下，它仍会增加网络流量和应用程序的延迟。研究表明通过动态镜像策略来节省存储空间的同时仍具有可靠性，而针对纠删码的各种研究都集中于减少网络流量和延迟。由此可见，实现动态冗余

的自动化（例如更改纠删码数据中的副本数量以满足故障和使用高峰期的需要）可以进一步提高存储效率。

4.2 带宽效率

网络带宽始终是分布式存储中的稀缺资源，带宽使用量与分布式存储中传输的数据量成正比。在镜像和纠删码存储系统中，修复故障数据都会消耗大量网络带宽，传统的纠删码比镜像涉及更多的带宽消耗。节点故障会触发云数据中心中的数据恢复并增加纠删码存储系统中网络流量，它们并不能像镜像存储系统那样减少网络流量。即使近些年纠删码技术进行了一些改进，但恢复数据的过程中大量消耗网络流量仍然成为纠删码技术目前面临的最大挑战。由此可见，故障预测技术应得到重视，以此来发现云存储系统中即将发生的数据故障。故障预测技术可以显著减少纠删码存储中的重构次数，从而减少网络流量，提高带宽效率。

4.3 能源效率

存储系统是云计算中耗能最大的组件之一，数据中心通过提高能源效率可以节省运营成本并有助于保护环境。存储系统的能源效率高度依赖于读/写延迟，Pinheiro 等引入了一种称为转移访问技术的的技术，该技术将存储系统中不同磁盘上的原始数据和冗余数据进行分离^[9]。该技术可以使包含冗余数据的磁盘保持空闲状态，直到出现磁盘故障为止。事实证明，该技术可以节省 20-61% 的磁盘能耗。就纠删码技术而言，存储方面的节能措施专用于数据恢复，反之亦然。因此，在不牺牲存储效率的情况下改善数据重构可以显著提高能源效率。与此同时，减少纠删码中的重构次数可以减少网络流量，并且可以通过季节性备份所选数据块来减少重构次数。

4.4 数据访问延迟

在存在故障的云存储系统中，必须执行解码以重建原始数据。因此，数据访问延迟是云存储系统中面临的最大问题。在镜像的存储系统中，可以通过选择最佳的镜像位置来显著减少访问延迟。大数据应用程序的性能可以通过减少访问延迟来提高，通过访问历史和纠删码中的故障日志激活主动的动态镜像，可以帮助降低访问延迟并减少存储开销^[10]。

5 未来研究趋势

在大数据时代，提高云存储系统对数据故障的容错能力至关重要，以下列出了未来的研究方向。

降低存储开销：开发出主动故障处理和动态镜像技术，以减少镜像的存储开销，而又兼顾了云存储系统的可靠性。

纠删码优化：纠删码在保证可靠性的前提下有助于减少存储开销，但是，这却是以增加网络流量，高磁盘 I / O 和较大延迟为代价。因此，需要一种新型的纠删码技术来减少网络流量，降低磁盘 I / O 和延迟，减少云存储系统的存储开销。

混合技术：由于镜像和纠删码技术各有优劣，因此利用混合技术提高可靠性是一个很有意义的研究方向。

故障预测技术：尽管纠删码减少了云存储系统的能耗，但发生故障时重建数据所需的额外能耗仍然非常巨大。因此，如果算法和技术可以预测数据故障，则可以主动进行数据镜像，避免系统数据的重构。

6 结语

云计算在为大数据应用程序提供服务方面发挥了重要的作用，它为不同的使用需求提供了较多的选择，并且使云存储和计算资源可以根据需求而快速变化。随着数据故障已成为云存储系统中的常态，为了提高云存储系统的可靠性，各种容错机制也已在逐渐采用。由于镜像技术存在巨大的存储开销，纠删码技术因而在大数据应用程序的云存储系统中倍受青睐。但是，如果发生数据故障，重建丢失的数据仍会涉及大量的资源消耗，从而影响大数据应用程序的性能。在本文中，笔者研究了这两种技术的利弊及发展趋势，并提出了提高能源效率降低带宽使用的一些设想，希望上述内容能为科研工作者在提高云存储系统的可靠性上提供一些有价值的参考和启发。

参考文献

- [1] 199IT. 数据永不眠 2.0 看看每一分钟究竟有多少数据产生? [EB/OL]. <http://www.199it.com/archives/213409.html>, 2014 - 04 - 24.
- [2] J. Gantz, D. Reinsel. The digital universe in 2020: big data, bigger digital shadows, and biggest growth in the far east[R]. :IDC iView, 2007(2012).
- [3] 安真. Facebook 数据中心建设模式分析[J]. 中国信息化, 2016(04) :81-87.
- [4] 徐聪. 大数据应用在云计算平台的优化部署与调度策略研究[D]. 清华大学, 2015.
- [5] 王意洁, 许方亮, 裴晓强. 分布式存储中的纠删码容错技术研究[J]. 计算机学报, 2017, 40(01) :23\6-255.
- [6] Y. Sharma, B. Javadi, W. Si, etc. Reliability and energy efficiency in cloud computing systems: survey and taxonomy[J]. Netw. Comput. Appl, 2016, 1(74) :66-8
- [7] Bonvin, N., Papaioannou, etc. Dynamic cost-efficient replication in data clouds. [C]. Proceedings of the 1st Workshop on Automated Control for Datacenters and Clouds. . PP:ACM, 2009. 49-56.
- [8] 李志鹏. 基于纠删码的容错存储系统中数据布局优化[D]. 中国科学技术大学, 2019.
- [9] Pinheiro, E., Bianchini, etc. Exploiting redundancy to conserve energy in storage systems. [J]. ACM SIGMETRICS Performance Evaluation Review., 2006, 34(12) :15-26.
- [10] 唐英杰, 王芳, 谢燕文. 纠删码存储系统中基于网络计算的高效故障重建方法[J]. 计算机研究与发展, 2019, 56(04) :767-778.