

一种分布式编码算法在云存储中的应用

李琳琳¹, 王庆超¹, 姚 超¹, 田 野²

(1 第二炮兵工程大学 信息工程系, 陕西 西安 710025; 2 中国电子科技集团 第十五研究所, 北京 100083)

摘 要: 在云存储中为保证数据的完整可靠, 必须采用冗余备份的方式存储数据. 为了更好的发挥存储节点的作用, 在分析现有的冗余方式基础上提出了一种分布式编码算法. 在新的冗余方案中编解码操作在各存储节点并行完成, 而不需要将整个数据文件集中进行编解码, 充分发挥了分布式系统优势. 通过实验仿真验证了方法的可行性, 大大提高了数据存储速率, 提高了数据可用性, 改善了系统性能.

关键词: 云存储; 数据冗余; 分布式编码; 纠删码

中图分类号: TP301.6

文献标识码: A

文章编号: 1000-7180(2014)02-0077-03

The Application of a Distributed Coding Arithmetic in Cloud Storage

LI Lin-lin¹, WANG Qing-chao¹, YAO Chao¹, TIAN Ye²

(1 Department of Command Automation, Second Artillery Engineering College, Xi'an 710025, China;

2 The Fifteenth Institute, China Electronics Technology Group Corporation, Beijing 100083, China)

Abstract: The method of redundancy was general be used to ensure the integrity and availability of data in cloud storage. This paper introduced a distributed coding arithmetic after analyzing those method existing. In the new arithmetic, the operation of coding and decoding was doing in the storage nodes, avoiding coding and decoding for the whole data file simultaneous. Through making full advantage of distributed system, the velocity of data storage was heighten consumedly, the availability of data was also improved, the performance of the storage system was enhanced.

Key words: cloud storage; data redundancy; distributed coding; erasure code

1 引言

云存储是云计算的存储部分, 当云计算处理 and 运算的核心是大数据量的存储和管理时, 云计算就必须部署大量的存储设备. 云存储作为云计算的底层支撑, 通过多种技术的融合, 将存储设备虚拟化为弹性、透明、可伸缩、易扩展的存储资源池, 并将存储资源池按需分配给授权用户, 授权用户既可以通过网络对存储资源池进行访问和管理, 并按一定标准收取费用.

为保证云存储的数据可靠, 必须采取一定的数据冗余措施, 以保证在发生数据损坏或丢失时, 及时恢复数据. HDFS^[1] 和 Google FS^[2] 均采取了直接复

制方式, 将文件块直接复制 3 份, 保证了数据可靠, 但对存储空间占用大; 文献[3]中提出将 RS 纠删码应用到云存储中, 提高了数据可用性, 但编码速率较慢; 文献[4]中对 RS 算法进行改进, 使用异或运算进行乘除操作取消了 GF 域, 一定程度上改善了时间性能; 文献[5]中提出使用 Tornado 码算法进行数据编码, 但对编码原理缺乏详细的介绍.

2 分布式编码方式

针对纠删码方式编解码速率慢, 不能满足数据实时存取的缺点, 本文提出了一种基于分布式编码算法的冗余方式. 在新的冗余编码方案中存储节点不仅具有存储数据的功能, 具备了对部分数据的编

收稿日期: 2013-03-23; 修回日期: 2013-05-19

基金项目: 国家“八六三”资助项目(2011701AA0213)

解码处理功能,通过将编解码任务分配到众多存储节点的方式,充分发挥了分布式系统的性能优势.

2.1 算法总体设计

在本文提出的分布式编码方案中,数据存取共分为以下几个步骤:

S1:文件分割:将数据文件 D_0 分割为大小为 b 的 n 块,即 $D_0 = [d_1, d_2, \dots, d_n]$;

S2:生成级联二分图;

S3:编码:按照级联二分图对 n 个数据文件块进行编码,将编码产生的数据块和原数据块分布地存储在各存储节点上;

S4:解码:对于损坏的存储节点或数据块,从相邻节点获取数据并解码,重建数据;

S5:读数据:访问原始数据节点,获取数据.

其中步骤 S1, S2, S4 为数据存取必须经历的阶段,而步骤 S3 仅当发生故障时用来重建数据使用.下面对其核心步骤 S2、S3 以及 S4 进行介绍.

2.2 生成级联二分图

(1) 由系统依据数据重要程度确定冗余度 R (在这里冗余度 R 表示为编码块与原数据块数量的比值);

(2) 解方程组

$$\begin{cases} n\beta^{(k+1)} = n^{1/2} \\ (1 - \beta^{(k+1)}) / (1 - \beta) = R + 1 \end{cases} \quad (1)$$

计算扩展层数 k 和编码比例 β ;

(3) 对第 0 级联层构建边;

① 第 1 层创建 $\beta \times n$ 个节点;

② 对第 1 层每一节点任取 3 条边引至第 0 层的 3 个不同节点;

③ 对度为 1 的 0 层节点增加一条边;

(4) 重复(3),将 $\beta \times n$ 依次用 $\beta^2 \times n, \beta^3 \times n \dots$ 替换,用类似的方法构建其余 k 级二分图.

2.3 编码与解码过程

编码算法使用级联二分图的形式描述各节点之间的关系,如图 1 所示. D_0 为原始数据文件块,按照二分图描述的各节点之间的关系,对 $D_0 = [d_1, d_2, \dots, d_n]$ 进行编码产生 $\beta \times n$ 个校验块 D_1 ,依次类推直至 D_k . 二分图定义了不同层级间数据块的关系,编码时上层节点将数据块推送至与之相关的下层节点,下层节点对收到的数据块进行异或运算,产生本层的校验块,并作为数据块推送至下层节点,直至最后一层.

解码算法与之相反,通常由高层开始,首先由 D_k 和可用的 D_{k-1} 恢复出完整的 D_{k-1} ,然后依次得到

完整的 $D_{k-2}, \dots, D_2, D_1$,最后获得完整的数据文件.若已确定 D_i 完整,则可直接从 D_i 开始解码.若原始数据文件完整,则整个解码过程均可省略,用户直接读取数据文件.当然若系统检测到某个存储节点失效或其数据损坏时,由相邻节点及时对其恢复.

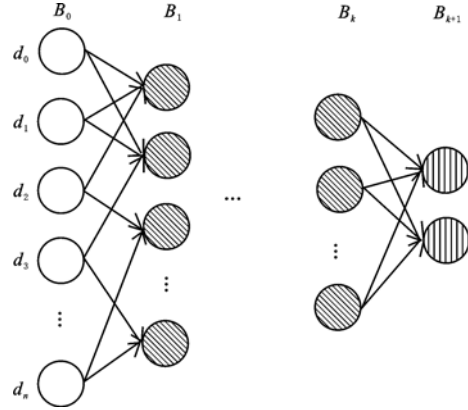


图 1 级联二分图

3 数据可用性评估与分析

由于节点失效属于小概率事件,因此通过实验的方式有很大的偶然性,很难得到准确的数据可用性,因此只能通过数学建模的方法求解.

根据文献[6]中的研究表明在不编码直接复制方式下, α 表示复制份数,则在不编码直接复制方式下的数据可用性 A_1 可表示为

$$A_1 = (1 - (1 - p)^\alpha)^n.$$

在传统 RS 纠删码方式下, m 表示编码冗余块,则 RS 纠删码方式的数据可用性 A_2 为

$$A_2 = \sum_{i=n}^{n+m} C_{n+m}^i p^i (1 - p)^{n+m-i}.$$

因为解码过程是由高层依次向底层进行,因此分析可用性也依此顺序.假设任意时刻存储系统的存储节点的平均可用性为 p .由于高层节点的存在,保证了低层节点数据的可靠,因此不同层级节点所存放的数据可用性并不一致.假设第 i 层节点的数据可用性为 p_i 则

$$p_i \geq (1 - ((1 - p) p_i^2 p_{i+1} + p)) ((1 - p) p_i^2 p_{i+1} + p) + ((1 - p) p_i^2 p_{i+1} + p) \quad (3)$$

$$p_{k+1} = p \quad (4)$$

依次迭代,求取 p_0 ;

所有数据分块均可用才能保证整个数据文件可用,数据可用性用 A_0 表示,因此:

$$A_0 = p_0^n.$$

由式(3)和式(4)可以看出 A_0 的计算结果与 p 和 k 有关,而 k 由 n 和 R 共同决定,因此 A_0 的最终表

达式包含 p, n 和 R . 但由于其计算过程相当繁琐, 本文在此不对其可用性的完整表达式进行手工推导, 而使用计算机迭代计算出具体数值.

表 1 中列出了在不同冗余度不同文件长度的情况下本文算法与 RS 算法的数据可用性对比, 表 2 中则表示本文算法与直接复制方式的对比情况.

表 1 本文算法与 RS 算法可用性对比

冗余度 算法	0.4	0.8	1	1.2	
RS 算法	0.974 309	0.999 109	0.999 853	0.999 977	
本文算法	$n=4$	0.948 617	0.998 218 1	0.999 706 1	0.999 954
	$n=12$	0.958 894	0.998 574 5	0.999 764 9	0.999 963
	$n=20$	0.971 739	0.999 019 9	0.999 838 4	0.999 978

表 2 本文算法与直接复制方法可用性对比

算法	冗余度为 1	冗余度为 2
本文算法($n=4$)	0.999 706 195	0.999 999 86
直接复制方式	0.99	0.999

由表 1 和表 2 可以看出, 本文提出的分布式编码算法, 在数据可用性方面优于直接复制方式, 略逊于传统的 RS 纠删码方式, 但在冗余度大于 1 的情况下, 两种方式的可用性均达到 99.99% 以上, 基本无差别.

4 时间性能分析

本文提出的编码算法, 编解码为数据块间的异或运算, 编解码操作由存储节点完成, 不需要将所有数据文件集中进行编解码操作, 因此大大节省了处理时间. 增强了处理的并行性, 算法复杂度与二分图的总边数无关, 仅与二分图的级联深度和节点的最大度数有关, 其算法复杂度为 $O(kb)$, 其中 k 为二分图深度, b 为分块长度. 由式(1)求得 $k = -\frac{1}{2} \log_2 n - 1$, b 为固定分块长度, 因此时间复杂度为 $O(\log n)$.

为了保证数据传输的可靠性, 使用 TCP/IP 协议构建局域网进行数据传输. 存储节点为 15 台配置为 i5 四核处理器, 4 GB 内存, 磁盘转速为 7 000 r/min 的台式机. 图 2 和图 3 表示对于相同数据源分块大小同为 10 kB 的实验结果.

由图 2 和 3 可以看到随数据文件的增大, RS 编码的速率迅速下降, 本文提出的算法的编码时间随数据大小线性增长, 对处理大数据文件有很大优势.

文献[7]中指出 RS 纠删码的时间复杂度为 O

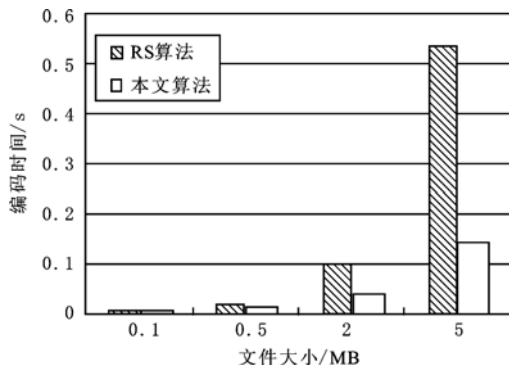


图 2 编码时间对比图

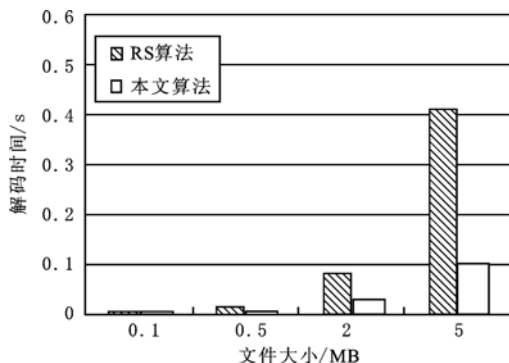


图 3 解码时间对比图

($n \log^2 n \log \log n$), 其编码时间和解码时间均远高于本文提出的算法, 并且当文件较大时, 其速度下降也非常明显, 这也与实验结果相符. 在更大规模的分布式系统中本文提出的冗余方式由于其分布编码的特点将会发挥更大的优势.

5 结束语

为了保证云存储系统数据的安全, 必须采取一系列的冗余措施, 避免数据出现不可恢复的损坏. 本文提出了一种分布式编码算法, 降低了编码算法的时间复杂度, 提高了数据可用性, 改善了系统数据读取性能, 具有算法简单、容易实现的优点, 对于大规模分布式实时存储系统具有极好的应用前景.

参考文献:

- [1] Ghemawat S, Gobioff H. The google file system[C] // Proceedings of the 19th ACM Symposium on Operating Systems Principles. New York: ACM Press, 2003: 19-22.
- [2] Hvachko K, Kuang H, Radia S. The hadoop distributed filesystem[C] // IEEE 26th Symposium on Storage Systems and Technology. Piscataway, NJ: IEEE Press, 2010: 1-10.

(下转第 84 页)

oSNR 是评价语音增强算法的最简单的客观指标,代表语音增强算法处理后语音与原始语音之间距离的一个全局指标,可在时域或频域中计算。oSNR 一般分为整体 SNR 和分段信噪比,相比分段信噪比,整体 SNR 更能反应全局噪声衰减、说明整体语音质量情况。本实验中使用整体 SNR 来衡量语音质量。oSNR 定义如下:

$$\text{oSNR} = 10 \times \lg \frac{\sum_{k=1}^K x^2(k)}{\sum_{k=1}^K [\hat{x}(k) - x(k)]^2} \quad (22)$$

式中, k 为采样点数, $x(k)$ 是纯净语音信号, $\hat{x}(k)$ 是增强算法处理后的信号, K 是信号的总长度。oSNR 的值越大,表明语音质量越高。仿真结果表明,本文方法的增强效果优于基本谱减法、几何谱减法,噪声被更好地消除,信噪比得到了进一步提高,尤其是对于输入为低信噪比语音时,本文方法对信噪比的改善效果更加明显。

最后进行主观试听测试,经过实验室六位同学的随机试听,均表示本文方法增强后的语音几乎不存在残留“音乐噪声”和背景噪声,且相比其他两种算法语音的畸变程度更小,话语更清晰,这与 PESQ 的结果相一致。

6 结束语

本文推导了一种利用短时相位关系的谱减法用于对语音增强的公式,并提出了一种结合听觉掩蔽效应的改进算法,客观评价和主观听音测试都表明:在低信噪比环境下,本文提出的方法能更有效地抑制背景噪声及减小语音畸变。

参考文献:

- [1] Loizou P C, Ma J. Extending the articulation index to account for non-linear distortions introduced by noise-suppression algorithms[J]. The Journal of the Acous-

tical Society of America, 2011, 130(2): 986-995.

- [2] Upadhyay N, Karmakar A. A perceptually motivated multi-band spectral subtraction algorithm for enhancement of degraded speech[C]//Computer and Communication Technology (ICCCT), 2012 Third International Conference on. India, IEEE, 2012: 340-345.
- [3] 彭军, 王忠, 刘兴涛, 等. 基于多窗谱相关加权语音增强[J]. 计算机仿真, 2011, 28(3): 142-145.
- [4] Ping W. An improved spectral subtraction algorithm based on auditory masking in voice human-computer interaction[C]//Mechatronics and Automation (ICMA), 2010 International Conference on. China: Xi'an, IEEE, 2010: 1938-1941.
- [5] Cao L, Zhang T, Gao H, et al. Multi-band spectral subtraction method combined with auditory masking properties for speech enhancement[C]//Image and Signal Processing (CISP), 2012 5th International Congress on. India: Guwahati, IEEE, 2012: 72-76.
- [6] Lu Yang, Loizou P C. A geometric approach to spectral subtraction[J]. Speech communication, 2008, 50(6): 453-466.
- [7] Martin R. Noise power spectral density estimation based on optimal smoothing and minimum statistics[J]. Speech and Audio Processing, IEEE Transactions on, 2001, 9(5): 504-512.

作者简介:

万义龙 男, (1989—), 硕士研究生. 研究方向为语音信号处理、语音识别。

张天骐 男, (1971—), 教授. 研究方向为语音信号处理、通信信号的调制解调、盲处理、神经网络实现以及 FPGA、VLSI 实现等。

王志朝 男, (1987—), 硕士研究生. 研究方向为宽带信号的波达方向估计、阵列信号处理。

金静 女, (1988—), 硕士研究生. 研究方向为变速率语音编码、低速率语音编码。

(上接第 79 页)

- [3] 余林琛, 章巍, 林强, 等. RS 纠错码在云存储中的应用[J]. 微电子学与计算机. 2011, 28(8): 234-236.
- [4] Plank J S, Xu L. Optimizing cauchy reed-solomon codes for fault-tolerant network storage applications[C]//Proc. of the 5th IEEE International Symposium on Network Computing and Applications, Cambridge, MA, 2006: 38-46.
- [5] 王意洁, 卢锡城. 基于 Tornado 码的复制算法[J]. 国防科技大学学报, 2004, 26(3): 39-42.
- [6] 周游, 刘鹏, 杨盛祥, 等. 基于私有云的数据冗余技术研究

究[J]. 电脑知识与技术, 2011(1): 16-19.

- [7] 慕建君, 路成业, 王新梅. 关于纠错码的研究与进展[J]. 电子与信息学报, 2002(9): 1276-1281.

作者简介:

李琳琳 女, (1974—), 博士, 副教授. 研究方向为指挥自动化系统工程与计算机应用。

王庆超 男, (1988—), 硕士研究生. 研究方向为云存储与计算机应用。