

冷数据集中的流媒体存储系统节能方法

柴云鹏 杨楠

(中国人民大学数据工程与知识工程教育部重点实验室 北京 100872)

(中国人民大学信息学院 北京 100872)

摘要 快速增长的数据量使流媒体存储系统在降低能耗方面遇到严峻的挑战,而针对流媒体应用而进行的节能方法的研究还很少。在对流媒体存储系统的节能问题进行详细建模的基础上,定量地分析了数据分布对节能和流媒体服务质量的影响,发现只有冷门数据的分布对节能的贡献最大,而且对服务质量的负面影响最小。提出了一种冷数据集中的节能方法,亦即将最冷的一部分数据集中存储在部分磁盘上来优化节能效果。实验表明,冷数据集中的节能算法在保障服务质量的前提下,使流媒体存储系统的节能效果获得了显著的提升。

关键词 节能,冷数据,流媒体,绿色计算

中图分类号 TP334.5 **文献标识码** A

Cold Data Concentration: Energy Saving Method for Streaming Media Storage Systems

CHAI Yun-peng YANG Nan

(Key Laboratory of Data Engineering and Knowledge Engineering, MOE Renmin University of China, Beijing 100872, China)

(School of Information, Renmin University of China, Beijing 100872, China)

Abstract Streaming media storage systems have encountered serious challenges due to the rapid accumulation of the data volume. However, the researches on saving energy specially for streaming media applications are few. First of all, based on detailedly modeling the energy conservation problem in streaming media storage systems, we analyzed data layout's impacts on energy saving and quality of service (QoS) quantificationally. It was found that cold data has the most contribution to energy saving, and has the least negative impact on QoS. Secondly, a new energy saving algorithm called cold data concentration (CDC) was proposed, which puts the coldest data on part of the disks to optimize the energy saving effects. Finally, simulations validate that CDC saves much more energy than traditional energy saving methods with the guarantee of QoS.

Keywords Energy saving, Cold data, Streaming media, Green computing

1 引言

近年来全世界的数据量呈爆炸性增长。根据 IDC 的估计,数据年增长率接近 60%,即每经过约一年半的时间,全球的数据量就会翻一番。因此,存储系统的能耗越来越高,已经成为备受关注的一个问题。仅 2008 年一年,西欧所有的数据中心的电费开销即高达 49 亿欧元^[1]。而互联网上新兴的流媒体服务增长更为迅速,例如 YouTube 的视频年增长率为 123%,而 Facebook 的视频量每年增长高达 239%^[2],都远远高于 60% 的数据年平均增长率。因此,流媒体存储系统中的节能问题非常重要和紧迫。

虽然针对一般存储系统节能问题的研究已经积累了一些优秀的成果,但是专门针对流媒体应用进行节能研究的还很匮乏。传统节能方法主要是针对访问集中的热门数据部分或者整个数据集进行面向节能的重新分布,但是我们对流媒体存储系统的节能问题建立模型时发现,对流媒体存储系统节

能影响最大的部分其实是访问概率非常低的冷门数据,访问热度较高或者中等的的数据对节能的贡献非常小,而且会带来服务质量显著下降等问题,针对这些数据实施节能的数据分布效果不好,而且会得不偿失。所以在流媒体存储系统中节能的重点应该是优先对冷门数据进行处理。

本文提出了一种冷数据集中(Cold Data Concentration, CDC)的节能方法,它通过对节能贡献最大的冷门数据进行重新分布,一方面可以优化节能效果,另一方面也利于保障流媒体服务的服务质量。实验表明,冷数据集中的节能算法相对于已有的节能方法有了显著的提升,是一种在流媒体应用环境中非常高效的节能方法。

2 流媒体存储系统的节能模型

2.1 磁盘能耗模式

磁盘一般可分为 4 种状态:工作、空闲、待机和关闭。在前两种状态下,磁盘都处于全速旋转中,能耗比较高,而且比

到稿日期:2011-10-14 返修日期:2012-02-17 本文受中国人民大学科学研究基金项目(中央高校基本科研业务费专项资金)资助。

柴云鹏(1983-),男,博士,讲师,主要研究方向为分布式计算、节能存储系统,E-mail:ypchai@ruc.edu.cn;杨楠(1962-),男,博士,副教授,主要研究方向为数据挖掘、flash 存储系统等。

较接近,所以可以认为功率为 P_h ; 待机状态下磁盘不旋转, 功率 P_l 很低; 关闭状态下不耗能。但是磁盘的不旋转和全速旋转两种状态之间的切换需要消耗一定的时间和能量。磁盘节能方法的原理是尽可能在磁盘没有访问的时间段内将其置于低功率的待机状态, 但一般不会在磁盘无访问时立即将其待机, 而是在一个固定的阈值时间段(例如 15s)内没有访问, 才会将其待机, 这样可以避免磁盘频繁地切换。阈值一般设定为磁盘的收支平衡时间(Break-even Time, BT), 它指处于正常工作状态的磁盘在 BT 时间内的能耗应该与磁盘启停一次的能耗相等。

假设磁盘在服务时间 ST 内处于待机状态下的时间为 $t_{standby}$, 全速旋转和待机状态之间的切换次数分别为 N_d 和 N_u , 切换能耗为 E_d 和 E_u , 切换时间为 t_d 和 t_u , 那么磁盘在服务时间 ST 内的节能值为

$$SE_i^{ST} = (P_h - P_l) \cdot t_{standby} - (N_d \cdot (E_d - P_h \cdot t_d) + N_u \cdot (E_u - P_h \cdot t_u)) \quad (1)$$

2.2 流媒体存储系统节能建模

流媒体服务和其他服务的一项主要区别就是用户在线持续的时间比较长。在每个用户的服务期内, 流媒体服务器并不是持续不断地推送数据, 而是每隔一定周期推送一定量的数据, 使用户可以流畅地观看视频; 而且可以为尽可能多的用户同时提供流媒体服务。以流媒体存储系统中某个磁盘 i 为例, 对其节能值进行建模。假设单位时间内到达磁盘 i 的流媒体用户请求数为 λ_i , 则连续两次用户访问的时间间隔 $t_{interval} = 1/\lambda_i$ 。如果流媒体用户的平均会话长度为 L , 根据 $t_{interval}$ 和 λ_i 值所处的不同区间, 将磁盘 i 的节能问题分为几种具体的情况:

(1) 当 $t_{interval} \leq L$, 即 $\lambda_i \in [1/L, +\infty)$ 时, 两个连续的用户访问会重叠或相连, 因此磁盘的实际待机时间为 0, 节能也为 0。

(2) 当 $L < t_{interval} \leq L + BT$, 即 $\lambda_i \in [1/(BT + L), 1/L)$ 时, 如图 1(a) 所示, 在磁盘进行状态切换之前的新的用户请求就会到达, 因此 $t_{standby} = 0$, $SE_i = 0$ 。

(3) 当 $L + BT < t_{interval} \leq L + BT + t_d$, 即 $\lambda_i \in [1/(BT + L + t_d), 1/(BT + L))$ 时, 如图 1(b) 所示, 在磁盘进行状态切换过程中新的用户请求到达, 由于磁盘进行状态切换的平均功率和空转功率近乎相等^[6], 因此可以近似认为 $t_{standby} = 0$, $SE_i = 0$ 。

(4) 当 $t_{interval} > L + BT + t_d$, 即 $\lambda_i \in [0, 1/(BT + L + t_d))$ 时, 如图 1(c) 所示, 这种情况下磁盘有机会节能, $t_{standby}$ 的值为 $t_{interval} - L - BT - t_d$, 而 $SE_i = t_{standby} \cdot (P_h - P_l)$ 。

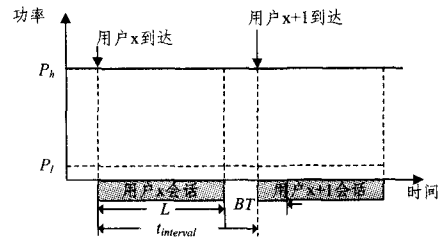
流媒体应用中用户到达的规律一般符合泊松分布, 这就等价于任意两个用户访问之间的时间间隔 $t_{interval}$ 符合指数分布, 其概率密度函数如式(2)所示。

$$f_{t_{interval}}(t) = \begin{cases} \lambda_i e^{-\lambda_i t}, & t \geq 0 \\ 0, & t < 0 \end{cases} \quad (2)$$

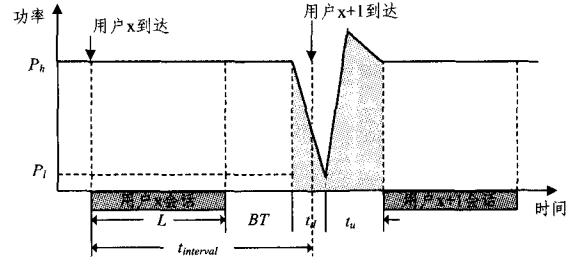
因此, 可以求出 SE_i , 如式(3)所示。

$$\begin{aligned} SE_i^{ST} &= \frac{ST}{t_{interval}} \cdot \int_0^{+\infty} ((P_h - P_l) \cdot f_{t_{interval}}(t) \cdot t_{standby}) dt \\ &= ST \cdot \lambda_i \cdot (P_h - P_l) \cdot \int_{L+BT+t_d}^{+\infty} (\lambda_i e^{-\lambda_i t} \cdot (t - L - \end{aligned}$$

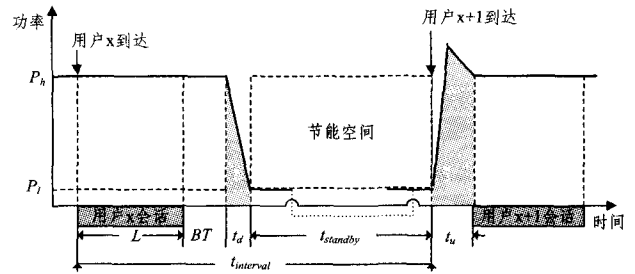
$$\begin{aligned} &BT - t_d) dt \\ &= ST \cdot (P_h - P_l) \cdot e^{-\lambda_i(L+BT+t_d)} \end{aligned} \quad (3)$$



(a) $\lambda_i \in [1/(BT+L), 1/L)$



(b) $\lambda_i \in [1/(BT+L+t_d), 1/(BT+L))$



(c) $\lambda_i \in [0, 1/(BT+L+t_d)$

图 1 流媒体存储系统中单个磁盘节能过程示意图

因此, 设 N 为流媒体存储系统的磁盘数, λ 为整个存储系统的用户到达率, 可以进而求出整个流媒体存储系统在服务时间 ST 内的节能值为

$$SE^{ST} = \sum_{i=0}^{N-1} SE_i^{ST}, \text{ s. t. } \sum_{i=0}^{N-1} \lambda_i = \lambda$$

$$SE_i^{ST} = \begin{cases} 0, & \lambda_i \in [1/(BT+L+t_d), +\infty) \\ ST \cdot (P_h - P_l) \cdot e^{-\lambda_i(L+BT+t_d)}, & \lambda_i \in [0, 1/(BT+L+t_d)) \end{cases} \quad (4)$$

2.3 流媒体存储系统服务质量建模

流媒体服务中服务质量非常重要, 因为服务质量变差会使用户由于忍受不了长时间的等待和频繁的中断而退出, 这就意味着服务的失败。一般流媒体服务的服务质量由启动延迟和播放过程中的抖动这两个指标来衡量, 而二者都是由每一次数据包传输的延迟而产生的, 因此了解延迟产生的原因是对服务质量建模的关键。一般来说, 主要有两种延迟: 一是由于个别磁盘过载而产生的延迟 $D_{overload}$, 二是由于磁盘在工作 and 待机状态间切换造成的延迟 D_{switch} 。

对于过载延迟 $D_{overload}$ 来说, 磁盘 i 上的平均在线用户数为 $\lambda_i \cdot L$ 。假设每个磁盘的服务容量是 Cap_i , 那么磁盘 i 的 $D_{overload}$ 可由式(5)来计算, 其中 $f(x)$ 是一个在 $x \geq 0$ 区间内的单调增函数, 而在 $x < 0$ 时, 值为 0。

$$D_{overload}^{ST} = \sum_{i=0}^{N-1} D_{overload,i}^{ST} = \sum_{i=0}^{N-1} f(\lambda_i \cdot L - Cap_i) \quad (5)$$

假设在服务时间 ST 内一共有 C 个用户, D_{switch} 由磁盘 i 的启动次数 N_u^i 和每次磁盘启动时在线用户数这两个因素决定。每次磁盘 i 启动过程中会有 $\lambda_i \cdot t_u$ 个新到达用户, 而每个用户的平均延迟是 $t_u/2$, 所以每次启动的平均延迟是 $\lambda_i \cdot t_u \cdot t_u/2$ 。因此, $D_{switch, i}^{ST}$ 等于

$$D_{switch, i}^{ST} = \frac{\lambda_i \cdot N_u^i}{2C/t_u^2} \quad (6)$$

而由图 1(b)和(c)可知, 每个 $t_{interval}$ 内磁盘完成一对关闭和开启操作, 所以 N_u^i 的值为 $ST \cdot \lambda_i$ 。这样, 在 ST 内整个流媒体存储系统的延迟为

$$D^{ST} = \sum_{i=0}^{N-1} D_{overload, i}^{ST} + \sum_{i=0}^{N-1} D_{switch, i}^{ST}, \text{ s. t. } \sum_{i=0}^{N-1} \lambda_i = \lambda \quad (7)$$

$$D_{overload, i}^{ST} \propto \begin{cases} (\lambda_i \cdot L - Cap_i), & \lambda_i \cdot L \geq Cap_i \\ 0, & \lambda_i \cdot L < Cap_i \end{cases}$$

$$D_{switch, i}^{ST} = \begin{cases} 0, & \lambda_i \in [\frac{1}{BT+L}, +\infty) \\ \frac{ST \cdot t_u^2 \cdot \lambda_i^2}{2C}, & \lambda_i \in [0, \frac{1}{BT+L}) \end{cases}$$

3 冷数据集中的节能方法

3.1 节能原理

流媒体存储系统节能的目标是尽可能使节能值 SE^{ST} 最大化, 而延迟 D^{ST} 尽可能小, 限制条件是所有 N 个磁盘的用户到达率之和等于一个给定值——系统实际的用户到达率。所以, 节能算法的关键在于通过数据分布使所有磁盘上的用户到达率达到一种合理的分布, 使 SE^{ST} 尽可能大, D^{ST} 尽可能小。

在节能方面, 由式(4)可知, 当 λ_i 大于等于一定阈值 $1/(BT+L)$ 时, 磁盘无法节能; 而当 λ_i 小于 $1/(BT+L)$ 时, SE^{ST} 与之 λ_i 间是指数函数的关系, 随着 λ_i 的增大, SE^{ST} 迅速下降, 只有当 λ_i 的值为 0, 或接近 0 时, 磁盘 i 才能高效地节能。

在延迟方面, 由式(7)可知, 切换延迟和 λ_i 是幂函数的关系, 当 λ_i 增大时, 延迟增加更为迅速。不产生延迟或令延迟很小的方式有两类: 一是令 λ_i 大于 $1/(BT+L)$ 来消除切换延迟, 但同时令其小于 Cap_i/L 来避免产生过载延迟; 另一类是令 λ_i 接近甚至等于 0, 这样既不产生过载延迟, 切换延迟也非常小, 甚至为 0。

综合节能和延迟两个方面, 关键在于使节能目标磁盘的 λ_i 值变得非常小, 甚至等于 0, 这样既能够保持节能的高效率, 同时也不会产生明显的延迟。而随着 λ_i 的增大, 节能值按指数规律迅速下降, 因此中等程度负载的磁盘既不能带来明显的节能, 同时在服务质量方面的副作用又十分明显。也就是说, 节能的关键在于将冷门的, 即访问概率最低的流媒体内容集中起来存放到某些磁盘中, 使这些磁盘的用户到达率达到非常低的程度才能高效地节能, 而不要将冷门内容分散到很多磁盘上, 使这些磁盘都只是维持中等程度的负载。而且, 只围绕冷数据进行数据的重分布, 数据迁移的规模能得到有效控制, 迁移能够迅速完成, 这样可以延长磁盘节能的时间长度。

基于这种思路, 本文提出了一种冷数据集中的节能方法,

算法的关键有两个方面: 一是将注意力集中在对节能最有贡献的冷门数据上, 将这些很少甚至没有人访问的数据集中存放; 二是有限的节能目标, 即只有那些存储冷门内容的磁盘总的访问概率低于一定阈值时才成为节能目标磁盘, 其他磁盘不以节能为目标, 不进行能耗状态的变化, 而尽可能减少数据块的移动也避免了中等负载的磁盘较为频繁的启停操作导致服务质量下降。

已有的典型节能算法, 例如 MAID(全称为 Massive Array of Idle Disks)^[3]和 PDC(全称为 Popular Data Concentration)^[4], 并不是以冷数据集中设计节能算法, 因此在流媒体存储系统中的节能效率较低。文献[4]指出, MAID 和 PDC 只有在存储系统负载很低的情况下才能有效节能。具体来说, MAID 的原理是在存储系统中设置一些额外的缓存磁盘, 用来缓存访问概率比较高的内容, 这样可以使原有磁盘的负载下降, 从而延长原有磁盘的空闲时间, 进而提高节能的效果。MAID 实际上是一种以热门数据为中心而设计的节能算法, 缓存盘虽然能够减小原有磁盘组整体的访问数, 但是中等的和冷门的数据依然会访问原有磁盘, 这一点没有得到任何改善。所以, 磁盘的负载虽然从整体上降低了, 但是对节能影响最大的冷门数据的访问模式却没有得到任何改变, 磁盘的负载很难降到足够低、能够高效节能的水平, 因此节能效率不高, 对服务质量的副作用也较大。此外, MAID 中额外的缓存磁盘也会增加系统的能耗。

PDC 算法周期性地将所有数据按照流行度从高到低进行排序, 并以此为基础重新安排存储到所有磁盘上, 这样磁盘的热度会从高到低降序分布, 后面的磁盘会有较大的希望进行节能。PDC 实际上是一种面向全部数据的分布而设计的节能算法, 并没有以冷数据或热数据为重点, 这样注意力过于分散到全部数据上, 在所有磁盘上都进行数据的重分布代价过大, 需要进行的数据迁移量太大, 需要花费相当长的时间才能完成理想的数据分布, 因此在实际操作中很难达到算法设计的理想节能效果。相对而言, 我们提出的冷数据集中的节能算法以对节能贡献最大的冷门数据为中心, 进行数据重分布时代价不大, 能够迅速完成面向节能的数据分布。

3.2 节能算法

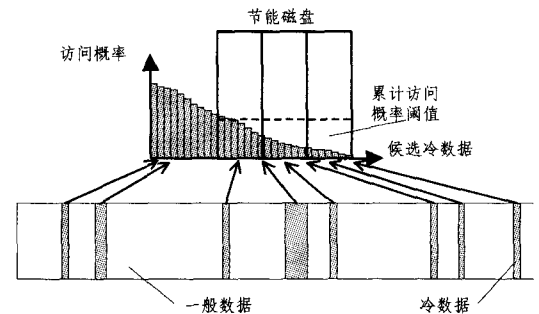


图2 冷数据集中的节能算法原理

冷数据集中的节能算法的原理如图 2 所示, 首先从流媒体数据中按照一定的模式抽取访问概率最低的一部分候选冷数据, 并按照其访问概率进行排序。然后从访问概率最低的候选冷门数据开始填充到磁盘中, 至到某一磁盘上全部数据的累积访问概率大于预设的阈值为止, 之前填充的磁盘

作为节能的目标磁盘。由于这些节能磁盘的总访问概率非常低,因此能够高效地节能,而且对服务质量造成的影响尽可能小。

3.2.1 候选冷数据提取和排序

判断一个流媒体数据块是冷数据有两个条件:一是在过去一个时间窗口内该数据块的访问热度要低于一定阈值;二是该数据块的上一次访问距离现在的时间长度要大于一定阈值。数据块 j 的访问热度计算如式(8)所示,其中 l_{w} 是统计时间窗口的长度。由于很多视频,例如新闻类视频总体的访问次数很高,但是访问基本集中在短时间内,之后就很少人访问了,因此 l_{w} 不能设置得过长; at_j 是数据块 j 在 l_{w} 内被访问的次数; $offset(j)$ 是根据数据块在视频内部的位置而设置的加权值,越靠近视频开头部分权值越大,越靠近视频结尾权值越小,这是因为流媒体用户往往看开头的一部分之后,可能就提前退出而不会持续看完一个视频,尤其是长视频。

$$T_j = offset(j) \cdot at_j / l_{w} \quad (8)$$

算法会按照式(8)和上述两个条件对所有视频分块进行筛选,满足条件的冷数据再按照式(8)计算出的热度从大到小进行排序。

3.2.2 数据重分布

在得到排好序的冷数据队列后,算法会从热度最低的数据块开始尝试逐个填满磁盘,如果一个磁盘填满之后,它上面所存储的所有数据块的访问热度之和小于一定阈值时,这个磁盘被选中作为节能磁盘。当队列中没有剩余冷数据,或者填满的磁盘中的累积访问热度大于阈值,那么这个磁盘不能作为节能磁盘,而且这一阶段的节能磁盘选择结束。然后算法会按照当前数据块分布状态和新设定的节能数据分布之间的差异,确定数据块迁移的方案,达成新的数据分布。在冷数据集中的节能算法中,每一固定周期就会实施一次候选冷数据提取、排序和数据重分布。

4 算法评价

4.1 实验平台

为了验证本文所提出的冷数据集中的节能算法的效果,在国际上广泛采用的磁盘存储系统的仿真工具 DiskSim^[5] 上增加了对流媒体服务模式的支持和节能模块,以此作为节能实验的平台。这套系统可以仿真由 10 个 SCSI 磁盘组成的存储系统,所用的磁盘为 IBM Ultrastar 36Z15^[6],主要参数如表 1 所列。

表 1 IBM Ultrastar 36Z15 磁盘参数

磁盘转速	15000 RPM
工作状态功率	13.5 W
待机状态功率	2.5 W
启动时间	10.9 sec.
启动能耗	135 J
停止时间	1.5 sec.
停止能耗	13 J

4.2 实验结果

实验中将本文提出的冷数据集中的节能算法和目前最有代表性的节能算法 PDC 进行比较(文献[4]的实验结论是 PDC 算法的节能效果优于 MAID),并且选择了一组通过算

法生成的流媒体工作负载(视频平均长度为 300 秒,平均用户到达率为 18 用户/分钟),以及一组来自 CCTV VOD 系统的实际流媒体环境下的用户访问记录(视频平均长度为 188.65 秒,平均用户到达率为 10.8 用户/分钟)进行实验,以全面地测试节能算法在流媒体环境下的表现。

图 3 给出了最基本的磁盘能耗管理算法 DPM、PDC 算法和本文提出的冷数据集中的节能算法 CDC 三者分别在两组工作负载下的节能表现。从图中可知冷数据集中的节能方法非常有效,相对于已有算法,在节能方面有了很大的提升。

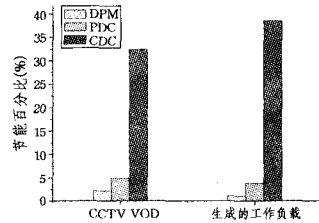


图 3 节能百分比比较

图 4 和图 5 给出了两种节能算法在两组工作负载下的服务质量。其中图 4 展示了启动延迟,图 5 展示了播放过程中的抖动,这两项指标是流媒体应用中最关键的服务质量指标,前者可以衡量流媒体服务的响应速度,后者可以指示整个播放过程的质量和用户感受。由图 4 可知启动延迟与节能百分比是紧密相关的,因为节能更多意味着有更多的机会访问到待机中的磁盘,等待磁盘启动需要花费一些时间。但是相对于图 3 所示的冷数据集中的节能算法在节能方面的提升幅度,其平均启动延迟增加的幅度并不大,整体还在可以接受的范围内。

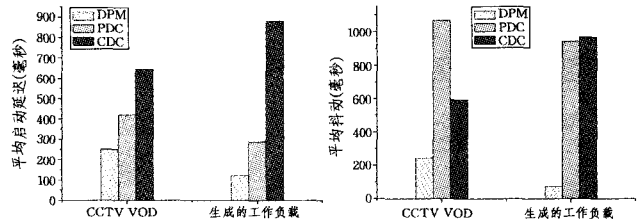


图 4 平均启动延迟比较

图 5 平均抖动比较

播放抖动在用户感受方面的权重要高于启动延迟,因为播放过程中出现断断续续的情况更让用户难以忍受。由图 5 可知,冷数据集中的节能算法在播放抖动方面的表现与 PDC 接近,甚至优于 PDC。这是因为当节能算法集中于冷数据的分布时,需要迁移的数据量小,迁移速度快,对正常用户访问磁盘影响很小;而 PDC 算法要对全部数据进行严格的重分布,迁移数据量大,会拖延正常的用户访问,因此在用户播放过程中更容易出现抖动现象。

5 相关工作

已有的节能研究针对流媒体应用的非常少,主要是针对通用存储系统的。除了前文介绍的 MAID、PDC 算法外,Carra^[7]、Gurumurthi^[8] 和 Zhu 等人^[9] 基于多速磁盘提出了一系列节能方法。多速磁盘可以工作在不同的转速下,所以当系统负载较低时,多速磁盘可以工作在低速模式下来节约能量。但多速磁盘由于工艺复杂,成本很高,因此到目前为止还没有商品化的产品可用。T. Kgil^[10] 等提出了基于 Flash 缓存层的

(下转第 169 页)

法。该方法基于收集查询历史推测用户的偏好,并根据小枝查询分解得到的查询片段数目的不同,使用不同的方法按照查询片段重要程度的高低对原始查询进行松弛。通过实验验证了该方法可以有效地提高小枝查询的查全率和准确率。在本文近似查询方法中,由查询历史推测的用户偏好是否准确直接影响近似查询结果的正确性,如果用户偏好的推测不准确,则不能保证得到的近似查询结果是用户所需的,因此,用户偏好的推测将仍然是今后工作的研究重点;同时,对查询松弛作限定条件的合并的最小片段数 m 值以及扩展度 e 和松弛度 r 的取值也是将来研究的重点。

参 考 文 献

- [1] 孟祥福,严丽,张文博,等. 基于文档属性单元松弛的 XML 近似查询方法[J]. 计算机研究与发展,2010,47(11):1936-1946
- [2] Polyzotis N, Garofalakis M, Ioannidis Y. Approximate XML query answers [C]//Proceedings of the 23rd ACM SIGMOD International Conference on Management of Data, 2004:263-274
- [3] 金鑫,金远平. 一种改进的基于约束关系的 XML 查询重写算法研究[J]. 计算机研究与发展,2007,44(5):845-852
- [4] Balmin A, Hristidis V, Koudas N, et al. A System for Keyword Proximity Search on XML Database [C]// Proceedings of the 29th Vldb Conference, Germany, 2003
- [5] Spiegel J, Pontikakis E D, Budalakoti S, et al. AQAX: a system for approximate XML query answers [C]// Proceedings of the 32nd International Conference on Very Large Data Bases, 2006:

(上接第 151 页)

节能方法,即利用低能耗的 Flash 存储设备减少高能耗的内存的使用,减少对磁盘的访问,从而降低整体的功耗。但 Flash 存储设备造价较高,而且作为缓存层频繁写入会严重影响设备的寿命,因此目前这类方法也没有广泛采用。另外, Zhu 等人[11]研究了通过调整内存缓存的方法来进一步延迟磁盘空闲时间,从而提高节能效果。EERAID^[12]、eRAID^[13]和 PARAID^[14]等方法研究了如何在磁盘的冗余存储上进一步节能。

结束语 本文在对流媒体存储系统的节能和服务质量建模的基础上,得出磁盘节能随负载降低以指数关系增加的结论,并以此为基础设计并实现了冷数据集中的节能方法。通过仿真实验表明,在保障良好服务质量的前提下,该方法节能效果显著提升,是一种流媒体环境中的高效节能方法。

参 考 文 献

- [1] Nebuloni G. Energy Footprint of the European Server Infrastructure, 2008, and 2009-2013 Forecast [EB/OL]. <http://www.idc.com/getdoc.jsp?containerId=GE11R9>, 2009-10
- [2] Yarow J. Videos on Youtube grew 123% year over year, while Facebook grew 239% [EB/OL]. <http://www.strangelove.com/blog/2010/06/videos-on-youtube-grew-123-year-over-year-while-facebook-grew-239>, 2010-06
- [3] Colarelli D, Grunwald D. Massive Arrays of Idle Disks for Storage Archives [C]// Proc. 2002 ACM/IEEE Conf. Supercomputing. Los Alamitos, CA, USA: IEEE, 2002: 1-11
- [4] Pinheiro E, Bianchini R. Energy Conservation Techniques for Disk Array-Based Servers [C]// Proc. 18th Int'l Conf. Supercomputing. New York, USA: ACM, 2004: 68-78

1159-1162

- [6] Cho S, Balke W-T. Building an Efficient Preference XML Query Processor [C]// SAC' 09. USA, 2009: 1585-1586
- [7] Yu Zhi-yong, Yu Zhi-wen, Zhou Xing-she, et al. Handling Conditional Preferences in Recommender Systems [C]// IUI'09. USA, 2009: 407-411
- [8] Mohamed F, Justin J. Toward Context and Preference-Aware Location-based Services [C]// MobiDE'09. USA, 2009: 25-32
- [9] 衡星辰,覃征,邵利平,等. 基于两阶段查询重写的 XML 近似查询算法[J]. 电子学报, 2007, 35(7): 1271-1278
- [10] 徐超,张东. sTwig—一种基于流的 XML 小枝匹配算法 [J]. 计算机研究与发展, 2010, 47: 86-92
- [11] Amer-Yahia S, Koudas N, Marian A. Structure and content scoring for XML [C]// Proc of International Conference on Very Large Data Bases, New York: ACM, 2005: 361-372
- [12] Bruno N, Srivastava D, Koudas N. Holistic Twig joins: Optimal XML pattern matching [C]// Proceedings of the 21st ACM SIGMOD International Conference on Management of Data, 2002: 310-321
- [13] IBM Corporation XML data generator [EB/OL]. <http://www.alphaworks.ibm.com/tech/xmlgenerator>, 2010-08-11
- [14] <http://autos.yahoo.com/>, 2009-12-23
- [15] <http://www.amazon.cn/>, 2010-06-19
- [16] Su W, Wang J, Huang Q, et al. Query result ranking over e-commerce web databases [C]// Proceedings of the 15th ACM Conference on Information and Knowledge Management, 2006: 575-584

- [5] Ganger G, Ganger G, Worthington B, et al. The DiskSim simulation environment (v4.0) [EB/OL]. <http://www.pdl.cmu.edu/DiskSim>, 2009-9
- [6] Hitachi. Ultrastar 36Z15 Datasheet [EB/OL]. <http://www.hitachigst.com/hdd/ultra/ul36z15.htm>, 2003-01
- [7] Carrera E V, Pinheiro E, Bianchini R. Conserving Disk Energy in Network Servers [C]// Proc. 17th Int'l Conf. Supercomputing. New York, USA: ACM, 2003: 86-97
- [8] Gurusurthi S, Sivasubramanian A, Kandemir M, et al. DRPM: Dynamic speed control for power management in server class disks [C]// Proc. 30th Annual Int'l Symp. Computer Architecture. New York, USA: ACM, 2003: 169-179
- [9] Zhu Q, Chen Z, Tan L, et al. Hibernator: Helping Disk Arrays Sleep through the Winter [C]// ACM Symp. Operating Systems Principles. New York, USA: ACM, 2005: 177-190
- [10] Kgil T, Roberts D, Mudge T. Improving NAND Flash Based Disk Caches [C]// Proceedings of the 35th International Symposium on Computer Architecture (ISCA'08), June 2008: 327-338
- [11] Zhu Q, Zhou Y. Power-Aware Storage Cache Management [J]. IEEE Transactions on Computers, 2005, 54(5): 587-602
- [12] Li D, Gu P, Cai H, et al. EERAID: Energy Efficient Redundant and Inexpensive Disk Array [C]// Proc. of the 11th workshop on ACM SIGOPS European workshop. New York, USA: ACM, 2004: 1-14
- [13] Wang J, Zhu H, Li D. eRAID: Conserving Energy in Conventional Disk-Based RAID System [J]. IEEE Transactions on Computers, 2008, 57(3): 359-374
- [14] Weddle C, Oldham M, Qian J, et al. PARAID: A gear-shifting power-aware RAID [J]. ACM Transactions on Storage, 2007, 3(3): 245-260