

一种基于 MBRC 值的关系型数据库负载能耗预测模型

杨德先 孙 华 于 炯 国冰磊
(新疆大学软件学院 乌鲁木齐 830008)

摘 要 数据库负载的能耗解析与建模是构建节能的绿色数据库的基础。针对数据库负载的高能耗问题,将 SQL 语句消耗的系统资源(CPU 和磁盘)映射为时间代价与功率代价,为数据库负载构建能耗预测模型。首先,根据负载的系统资源消耗模式,计算负载的功耗代价;然后,根据负载资源消耗产生的时间代价,为负载构建动态能耗预测模型;最后,利用 MBRC 值的设置对预测模型的准确度进行深入的研究。实验结果表明,所构建的预测模型能够对数据库负载的能量消耗进行较准确的预测,预测模型的准确度研究有助于在不同的系统环境配置下提升动态能耗预测模型的稳定性与精确度。

关键词 绿色计算,关系型数据库,绿色数据库,SQL 能耗建模,MBRC

中图分类号 TP315 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2017.07.029

Relational Database Energy Prediction Model Based on MBRC

YANG De-xian SUN Hua YU Jiong GUO Bing-lei
(School of Software, Xinjiang University, Urumqi 830008, China)

Abstract Power modeling and power profiling of database workload are the foundation of building energy-aware DBMS. To solve the high energy consumption problem of database workload, this paper proposed a method of building energy prediction model for database workload by mapping resource consumption of the main components (CPU and Disk) to their power and time cost. Firstly, power cost of workload is estimated according to their resource consumption pattern. Secondly, energy cost of workload is estimated by combing the power cost and time cost derived from the resource consumption pattern. And finally, a deep research of the model prediction accuracy is conducted by using the MBRC settings. Experimental results show that, the proposed energy prediction modeling method can make accurate estimation of workload energy cost, and the research on accuracy of model prediction can improve the reliability and accuracy of the energy modeling method under different system settings.

Keywords Green computing, Relational database, Green database, SQL energy modeling, MBRC

随着移动互联网、物联网、云计算、基于位置的服务等新技术和应用的快速发展,数据产生的规模和速度呈指数型增长。作为信息的重要载体,数据中心的数量也随之增长,建设趋势也向大型化发展,其能源消耗和运营成本问题日益突出。

根据文献[1],2011年数据中心能源消耗占美国电网总量的2%,并且还将呈现每5年增长一倍的态势。中国的数据中心能耗也高速增长,2011年我国数据中心总耗电量达到700亿千瓦时,数据中心能耗占全国能源耗电总量的5%,是美国的3倍^[2]。未来5年,我国对数据中心流量处理能力的需求将增长7~10倍,机房面积再增加一倍才能满足需求,因此数据中心的能耗问题将更加突显。

另一方面,数据中心能源效率普遍低下,能源浪费巨大。据工信部统计,目前中国的数据中心的平均 PUE 值在 2.2~

3.0 之间,而实际能耗可能远远高于这一数字。对企业而言,数据中心电费已成为很大一笔开支,这将大幅侵蚀企业的经营利润。文献[3]指出服务器工作3年的用电成本将超过服务器最初的采购成本。与此同时,随着近几年国内外的互联网厂商纷纷兴建各自的数据中心以提供各种云存储或计算服务,能源消耗及温室气体排放问题也愈显突出。据统计^[4-5],能耗已成为数据中心运维的最大成本,并且成为了云计算技术普及与发展的最大障碍。面对日渐增长的能源消耗,绿色环保节能的数据中心越来越受到人们的关注。

在传统的数据中心,数据库系统负载占用着大量的服务器资源(包括 CPU、磁盘 I/O 等),成为数据中心能量消耗的主要组成部分^[6]。由于仅靠节能的硬件和操作系统无法从根本上解决这一问题,因此无论是从降低数据中心的运营成本

到稿日期:2016-06-04 返修日期:2016-09-27 本文受新疆维吾尔自治区自然科学基金资助项目(2015211C263)资助。

杨德先(1991—),男,硕士生,CCF 学生会员,主要研究方向为绿色计算、网络安全,E-mail: yangdexian_xj@sina.cn;孙 华(1977—),女,博士,副教授,CCF 会员,主要研究方向为网络安全与云计算;于 炯(1964—),男,博士,教授,博士生导师,主要研究方向为云计算、大数据与网络安全;国冰磊(1991—),女,博士生,主要研究方向为云计算、绿色计算。

还是降低能耗进而保护环境的角度出发,关注数据库系统设计过程中的能耗问题以及研究面向可持续发展的低功耗的节能数据库,都具有显著的社会意义与应用前景。

1 相关研究

为解决 IT 系统日益暴露出的高能耗、低效率问题,学术界与工业界分别从硬件和软件两个层次去解决 IT 系统的能耗问题^[7]。在硬件层面上,主要采用动态电压调节技术(Dynamic Voltage Scaling, DVS)^[8]和处理器电压/频率控制机制(Processor Voltage/Frequency Control, PVC)^[9]。DVS 是一种功率管理技术,通过动态调节硬件的电压来达到节能的目的。PVC 机制通过动态调整处理器的电压与频率,使得处理器处于低功率的状态,避免系统长期运行在峰值性能下,从而达到节能的目的。在软件层面上,主要是在操作系统和虚拟机层面上^[10-11]来解决能耗问题。文献[10]综述了在 ECOsystem, nemesis OS, Linux 等系统上采用系统资源管理、任务调度与适应等方法对操作系统进行节能处理的方法;Xen, Vmware, KVM 等虚拟机监视器利用虚拟机的广泛的在线监测、动态合并与迁移、布局优化等技术在物理服务器之间控制系统资源,从而达到节能的目的。

大多数研究工作集中在通过改进硬件及系统层^[12-14]来实现对数据库能耗问题的优化,软件层面的研究却相对较少。目前,软件层面的研究主要考虑两方面:首先通过为数据库负载构建能耗预测模型,对负载能耗进行运行前的预测,达到为数据库估算能耗的目的;进而为数据库负载构建查询评价模型,将其与能耗预测模型嵌入数据库查询优化器中,指导优化器为负载挑选低功耗的执行计划并执行,最终达到构建节能的绿色数据库的目的。

Xu 等^[15-20]致力于构建一种功率监控型的数据库管理系统,通过建立功耗预测模型和成本评估模型来精确估算查询计划的能源成本,构建了一个能耗感知的查询优化框架。通过框架自身的电力成本估算模块和计划评价模型,使数据库系统运行在由数据库管理员指定的(DBA-specified)性能/能源权水平下。框架中的电力成本估值器可以在编译时准确地估计查询计划的电力成本;查询评估引擎可以根据 DBA 所需的权衡来配置系统的关键参数。

Rodriguez 等^[21]提供了一种利用数据库现成的统计数据估算数据库负载的功率和能源成本的方法。其利用选择性因素、元组的大小、数据列和关系基数等统计数据,通过多元线性回归确定重要参数,实现了较准确的成本模型,只需从功率计上获取系统的总功率,而不需要单独测量各个硬件组件。

杨良怀等^[22]基于线性回归方法,利用 CPU 每个 core 的利用率、执行频率以及硬盘利用率来预测整机的功耗。从执行核粒度考查 CPU 执行频率与利用率对系统功耗的综合影响,避免了在多核架构下整体 CPU 的活动信息掩盖单个核使用情况的问题。实验结果表明,系统功率与执行频率成正比相关。

国冰磊等^[23]提出了一种基于结构化查询语言(SQL)资源(CPU 与磁盘)消耗的最小单位的数据库动态功耗模型。

该模型对服务器功耗进行解析,将主要硬件(CPU 与磁盘)的资源消耗映射成功率消耗,采用多元线性回归方法拟合模型的关键参数,实现了较精确的功耗建模。

文献[20,23]与本文的工作最为相关。本文与已有工作的不同之处在于:首先,在构建功耗预测模型的基础上对数据库负载的能量消耗进行建模,完善了能耗预测模型的构建;其次,基于数据库磁盘访问类型(单数据块读操作和多数数据块读操作)的不同,探讨了 MBRC(db_file_multiblock_read_count)值的设置与预测模型稳定性之间的重要关系,为构建节能的数据库奠定了基础。

2 数据库能耗预测模型

2.1 模型的理论基础与设计

在传统关系型数据库中,用户通过 SQL 语句操作数据库,上层应用程序对数据库的操作也最终转化为 SQL 语句对数据库的操作。SQL 语句执行过程中产生的能耗将直接影响数据库的能源消耗。用户在提交一个 SQL 语句后,查询优化器会为该条语句产生一系列的查询计划,并在保障性能的前提下为 SQL 语句构建能耗预测模型,挑选满足用户需求的节能的查询计划,以达到为数据库节能的目的。因此,对 SQL 语句的执行过程进行能耗建模,是构建节能数据库的首要基础。

SQL 语句在执行过程中会消耗系统资源(CPU、磁盘、内存、网络等),这些相关部件产生的能耗总和构成了 SQL 语句的能耗,其中最主要的资源消耗组件是 CPU 和磁盘。通过实验所用数据库服务器进行测试,得到如表 1 所列的数据。

表 1 实验服务器各组件的功率消耗

组件	静态功率/W	动态功率/W
CPU 型号:英特尔 Xeon(至强) E3-1225 v3 @ 3.20GHz 四核	25	59
硬盘型号:希捷 ST1000DM003-1CH162 (1 TB / 7200 转/分)	2	2
内存型号:8GB (海力士 DDR3 1600MHz)	4	4
系统整机	35	108

表 1 列出了服务器中各组件的静态功率(无负载运行时的功率)和动态功率(部件的峰值功率减去静态功率)。从表 1 中可知,在对系统各组件进行压力测试时,内存的动态功率非常小,因此本文的能耗预测模型主要考虑 CPU 与磁盘的能源消耗。众所周知,读取磁盘的速度相对内存读取而言要慢得多,在 SQL 执行过程中,若数据缓存在内存,则不会再产生磁盘 I/O,因此实验未显式地把内存部分的功耗加入到模型构造中,而将其作为系统静态功率进行处理^[23]。实验利用多元线性回归模型为负载构建能耗预测模型用于预测负载的能量消耗,并利用最小二乘法求解预测模型中的参数。下面从 CPU 和磁盘两方面来分析能耗。

(1)CPU 能耗。在执行 SQL 语句时会消耗 CPU 资源,数据处理过程中执行的每一步操作都会依据处理量与处理方式的不同请求 CPU 执行一定数量的指令,将 SQL 语句执行过程中消耗的 CPU 指令数设为 CPU 总指令数(Ins)。为方便计算,CPU 总指令数以万为计量单位。针对不同型号的

CPU,其指令功率能力(也称为指令处理能力,即每瓦特可以完成的指令数,*Wins*)存在较大的差异。实验对数据库进行初步的功耗建模,并用一系列训练集训练出 *Ins_watt* 的值。

设每万条 CPU 指令消耗的功率为 *Ins_watt*,则 CPU 动态功耗的计算公式如下:

$$P_{cpu} = Ins * Ins_watt = Ins * (10000/Wins) \quad (1)$$

(2)磁盘能耗。当出现索引范围扫描、索引唯一扫描、索引完全扫描以及有 ROWID 访问表等操作时,会使用单数据块读操作;当出现全表扫描和快速完全索引扫描时,会出现多数据块读操作。在执行 SQL 语句的过程中,设磁盘产生单数据块读操作的总次数为 *Sin_total*,磁盘产生多数据块读操作的总次数为 *Multi_total*。

设一次单数据块读操作消耗的功率为 *Sin_watt*,则磁盘 IO 类型为单块读时,磁盘动态功耗的计算公式如下:

$$P_{disk} = Sin_total * Sin_watt \quad (2)$$

设一次多数据块读操作消耗的功率为 *Multi_watt*,则磁盘 IO 类型为多块读时,磁盘动态功耗的计算公式如下:

$$P_{disk} = Multi_total * Multi_watt \quad (3)$$

由以上可得,当 SQL 语句的 IO 类型为单块读时,执行 SQL 语句产生的系统动态功率为:

$$Active_power = Ins * Ins_watt + Sin_total * Sin_watt \quad (4)$$

当 SQL 语句的 IO 类型为多块读时,执行 SQL 语句产生的系统动态功耗为:

$$Active_power = Ins * Ins_watt + Multi_total * Multi_watt \quad (5)$$

基于以上的动态功耗公式,可得到混合 IO 类型环境下执行 SQL 语句产生的系统动态能耗:

$$Active_energy = Ins(Ins_watt * Ins_time) + Multi_total(Multi_watt * Multi_time) + Sin_total(Sin_watt * Sin_time) \quad (6)$$

其中,*Ins_time* 是 CPU 执行一万条指令消耗的时间代价,*Multi_time* 是磁盘产生一次多数据块读操作消耗的时间代价,*Sin_time* 是磁盘产生一次单数据块读操作消耗的时间代价。

2.2 模型的实现

实验总体环境的描述如表 2 所列。

表 2 实验总体环境

项目	描述
操作系统	Windows Server 2012
数据库管理系统	Oracle 11g
能耗数据测量	HOP1 待机功率测试仪 USB 智能版 (型号:HP-8713/2A 功率误差值±0.01W)
能耗数据采集	用电监测仪数据分析系统 v1.0
能耗相关单位	功率:W,能耗:J,时间:s
数据采样频率	1Hz
实验机 CPU	CPU 型号:英特尔 Xeon(至强) E3-1225 v3 @ 3.20GHz 四核
实验机内存	内存型号:8GB(海力士 DDR3 1600MHz)
实验机硬盘	硬盘型号:希捷 ST1000DM003-1CH162 (1 TB / 7200 转/分)

为了排除功率监控软件对采样数据准确度的干扰,通过

双机通信的方式对能耗数据进行精确的采集。实验测试集采用 TPC-H 标准中的 8 张表(CUSTOMER,LINEITEM, NATION,ORDERS,PART,PARTSUPP,REGION 和 SUPPLIER)和 22 条复杂的 SQL 语句(SQL 实例见图 1)。所使用的训练集为 10GB 大小的数据库,测试集为 1GB 大小的数据库。TPC-H 是由 TPC(Transaction Processing Performance Council)事务处理性能委员会公布的一套针对数据库决策支持能力的测试基准,通过模拟数据库中业务相关的复杂查询和并行的数据修改操作来考查数据库的综合处理能力。TPC-H 中数据库大小的缺省值是 1GB,通过调节比例因子 SF (Scale Factor)可以控制生成的数据库的大小。

```
select
  l_returnflag,l_linestatus,
  sum(l_quantity) as sum_qty,
  sum(l_extendedprice) as sum_base_price,
  sum(l_extendedprice * (1-l_discount)) as sum_disc_price,
  sum(l_extendedprice * (1-l_discount) * (1+l_tax)) as sum_charge,
  avg(l_quantity) as avg_qty,
  avg(l_extendedprice) as avg_price,
  avg(l_discount) as avg_disc,
  count(*) as count_order
from
  lineitem
where
  l_shipdate <= ADD_MONTHS('1998-12-01',3)
group by
  l_returnflag,l_linestatus
order by
  l_returnflag,l_linestatus;
```

图 1 TPC-H select 查询语句示例(Q1)

为了减小误差,采用 PL/SQL+动态 SQL 的方式使 SQL 语句无缝嵌套在 PL/SQL 代码段,每条 SQL 在清空共享池和缓存的情况下连续执行 100 次,记录能耗数据,并获取各项参数的平均值。

实验中,系统需绕过缓存从磁盘直接读入数据块。因为从磁盘读取数据块和从缓存读取数据块是两个不同的操作,这需要定期清除驻留在共享池(shared_pool)及缓存(BUFFER_CACHE)中的数据,所以每提交一次 SQL 语句到客户端(SQL Plus)都需刷新缓存(BUFFER_CACHE)及共享池(shared_pool)。

为验证预测模型的准确性,在 DBMS 独占系统资源下,执行训练集(见图 1),收集相关统计数据,并利用能耗采集软件收集负载的能耗数据。在 MATLAB 中利用多元线性回归模型进行线性拟合,可得到如下公式:

$$Active_energy = (0.191 * Ins + 93.530 * Multi_total + 14.512 * Sin_total) * 10^{-4} \quad (7)$$

式(7)中参数的含义与式(6)相同。

图 2 示出了利用本文构建的能耗预测模型对 22 条 select 语句进行能耗预测的结果。由图 2 可知,能耗预测模型能够

良好地预测负载的能耗,除第 18 条 query 语句的预测误差高于 10%,其他误差都低于 10%;且通过分析残差数据可知,负载能耗预测的平均误差在 6%以内,能够基本接近能耗的真实值,为构建节能的数据库奠定了基础。

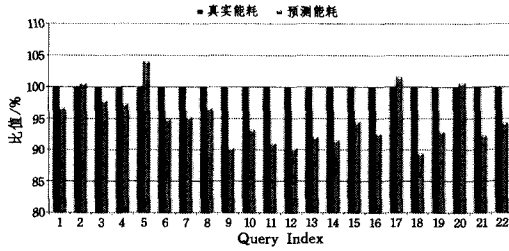


图 2 TPC-H 能耗预测图

3 预测模型的准确度分析

在整理第 2 节中负载的能耗数据时发现,随着多块读操作比例的增加,预测模型的准确度呈不同程度的下滑趋势。从实验的 22 条 SQL 语句中提取 8 条实验结果明显的 SQL 语句,以反映多块读操作数量对预测模型准确度的影响,如表 3 所列。

表 3 TPC-H 单块读-多块读操作比例/%

TPC-H 序号	多块读操作比例	单块读操作比例	模型预测误差率
Q2	0	100	2
Q22	49.50	50.50	2.5
Q19	66.67	33.33	3.1
Q5	80.90	19.10	3
Q20	85.60	14.40	3.3
Q21	88.89	11.11	5.77
Q7	90.90	9.09	6.4
Q8	93.33	6.67	7.2

MBRC 是多块读操作一次读取的数据块数量,由用户指定 MBRC 的值。理论上,MBRC 值决定多块读操作的次数(I/O 次数),但是用户指定的理论值与真实情况下的 MBRC 值有不同程度的差异,并且会影响预测模型的精确度,因此当多块读比例增加时预测模型的精确度下降。为进一步验证多块读操作比例与预测模型准确度之间的关系,设计如下测试集(测试数据集是上文中 TPC-H 标准测试中的 8 张表,数据量共 1GB,约 900 万条数据),分别测量只存在单块读操作和只存在多块读操作时预测模型的准确度。实验测试例句如表 4 所列,实验结果如图 3 所示。

表 4 单块读-多块读操作的 SQL 语句示例

序号	数据操作类型	I/O 类型	SQL 例句
1	索引范围扫描	单块读	select /* + INDEX (CUSTOMER SYS_C0011165) */ C_CUSTKEY from SCOTT.CUSTOMER where C_CUSTKEY between 5526 and 1300000;
2	索引完全扫描	单块读	select /* + INDEX (NATION SYS_C0011113) */ N_NATIONKEY from SCOTT.NATION;
3	全表扫描	多块读	select /* + FULL(LINEITEM) */ * from SCOTT.LINEITEM;
4	快速完全索引扫描	多块读	select /* + INDEX_FFS (ORDER SYS_C0011168) */ O_ORDERKEY from SCOTT.ORDER;

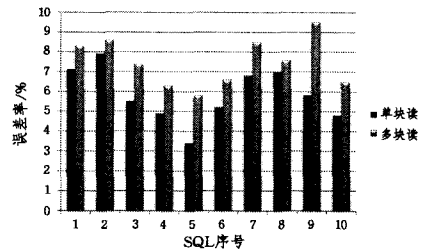


图 3 单块读-多块读预测模型的准确度测试

结合表 3 和图 3 可得,单块读预测模型的预测准确度要明显高于多块读的预测准确度。为进一步分析包含多块读操作的 SQL 能耗比包含单块读操作的 SQL 能耗更难预测的原因,利用 10046 事件对测试所用的 SQL 所有语句进行跟踪,并利用 TKPROF 对 trace 文件进行格式化,以方便分析。Oracle 的 10046 事件可以跟踪应用程序所执行的 SQL 语句,记录一条 SQL 或者 PL/SQL 包的运行情况。通过 10046 事件可将 SQL 执行的整个过程输出到一个 trace 文件中,通过读取 trace 文件来了解在该 SQL 执行过程中 Oracle 执行哪些操作。TKPROF 是 Oracle 自带的一个命令行工具,主要作用是将原始的跟踪文件转换为格式化的文本文件。基于此,实验获取所有 SQL 语句的执行信息,在解读 trace 文件时,发现如图 4 所示的情况。

```
wait #7; nam='db file scattered read' ela= 1491 file#=5 block#=18363
blocks=16 obj#=52884 tim=1337624821997106
wait #7; nam='db file scattered read' ela= 5184 file#=5
block#=18443 blocks=10 obj#=52884 tim=1337624822005413
wait #7; nam='db file scattered read' ela= 32363 file#=5
block#=18571 blocks=16 obj#=52884 tim=133762482 2041788
wait #7; nam='db file scattered read' ela= 1930 file#=5
block#=18587 blocks=16 obj#=52884 tim=1337624822044227
wait #7; nam='db file scattered read' ela= 345 file#=5
block#=18603 blocks=16 obj#=52884 tim=1337624822045165
wait #7; nam='db file scattered read' ela= 3712 file#=5
block#=18619 blocks=16 obj#=52884 tim=1337624822047555
wait #7; nam='db file scattered read' ela= 58 file#=5
block#=18635 blocks=16 obj#=52884 tim=1337624822048219
```

图 4 跟踪文件示例

图 4 示出的是包含多块读操作的 SQL 语句执行时的信息,block 代表读取的数据块号,blocks 代表读取的数据块数量。MBRC(db_file_multiblock_read_count)即执行一次多数据块读操作时从磁盘读取的数据块数量。通常情况下,操作系统的最大 I/O 会影响 Oracle 的 MBRC 设置,当最大 I/O 为 1MB 时,MBRC 的值为 1MB/db_block_size。实验中,数据库服务器设定的 db_block_size(数据块大小,在数据库创建时被指定,数据库一旦创建,该值就不可更改)的值为 8kB,即上述实验中系统允许的 MBRC 最大值为 128,图 4 对应的 MBRC 值为 16(即一次磁盘 I/O 读取 16 个数据块)。由图 4 得知,设置 MBRC 后,系统在执行负载时并不一定能保证其值的有效性,即实际执行过程中并不是每次都读取规定数量的数据

块,每次读取的数据块数量(blocks)和 MBRC 的值有不同程度的差异。

在执行多块读操作时,由多块读操作产生的能耗与一次多块读的时间(Multi_time)、一次多块读的功耗(Multi_watt)和进行多块读操作(Multi_total)的次数有关。不同的 MBRC 值设置可能会导致多块读操作有不同的 Multi_time 和 Multi_watt 值。为进一步探究多块读操作的能耗误差根源,通过改变 MBRC 的值来对相同工作负载进行能耗预测,得到图 5 和图 6 所示结果。

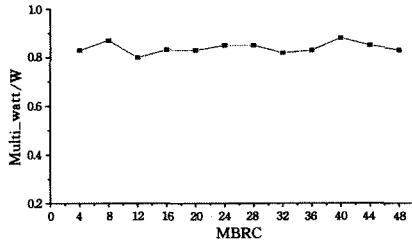


图 5 MBRC 设置对 Multi_watt 的影响

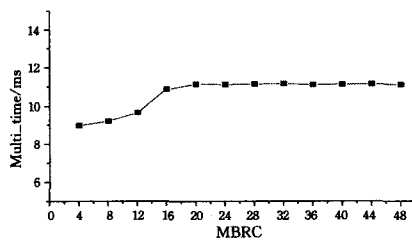


图 6 MBRC 设置对 Multi_time 的影响

由图 5 和图 6 可知,对于相同的负载,改变其 MBRC 值后会得到不同的 Multi_watt 和 Multi_time 值,这解释了第 2 节中多块读操作的能耗误差比单块读操作大的原因。图 5 中,随着 MBRC 值的增大,Multi_watt 的值几乎没有变化,即一次多块读的功耗与一次读取多少个数据块没有直接联系。图 6 中,随着 MBRC 值的增大,Multi_time 有逐渐增大的趋势,但是在 MBRC 取值 20 后,Multi_time 的值趋于平稳,不同的 MBRC 值下 Multi_time 的值没有较大的差异。综上,MBRC 值的设置会影响 Multi_time 的值,与 Multi_watt 之间没有明显的关系。但是对于不同的数据库系统配置和不同的 MBRC 值,需要对数据库系统进行训练,得到满足当前系统设置的稳定 Multi_watt 和 Multi_time 值,才能提高能耗预测模型的准确度,实现精确的能耗预测。实验还对一张数据量为 1GB 的表进行全表扫描(全表扫描 I/O 操作类型为多数据块读操作),并记录在不同的 MBRC 值下的磁盘 I/O 次数与全表扫描所用的时间,如图 7 所示。

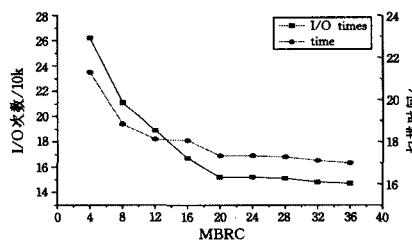


图 7 不同 MBRC 值下的磁盘 I/O 次数与全表扫描时间

由图 7 可知,随着 MBRC 值的增加,磁盘 I/O 次数逐渐减少,MBRC 值和磁盘的 I/O 次数为反比关系,二者的乘积

为测试所用表的大小(1GB)。由图 6 可知,随着 MBRC 的变化,一次多块读操作耗费的时间并没有很大的差异。由图 7 可知,查询所用时间随着 MBRC 值的增加而减少,这是因为随着 MBRC 值的增加,I/O 次数逐渐减少,最终导致查询时间减少。实验结果表明,提高 MBRC 的值可以减少 I/O 请求数,提高数据库的查询性能,使表扫描得更快。

为对不同 MBRC 值下能耗预测的误差做进一步的对比,在 MBRC 分别为 4,6,10,20,30 的情况下对 5 张不同的表进行全表扫描,得到表 5 和图 8 的结果。同时,在 MBRC 值为 4 和 30 时,分别再次对 TPC-H 中 22 条查询语句进行能耗预测,得到图 9 所示结果。设实验中设置的 MBRC 值与实际一次读取数据块总量之间的差值为 MBRC_D_Value([0,1MB/db_block_size))。

表 5 MBRC 的设置与 MBRC_D_Value 值

MBRC	一次 I/O 实际读取数据块数量 (平均值)	MBRC_D_Value (平均值)
4	3.5	0.5
6	5.2	0.8
10	8.5	1.5
20	18	2
30	26.3	3.7

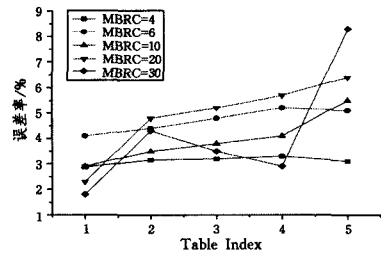


图 8 MBRC 的设置与能耗误差

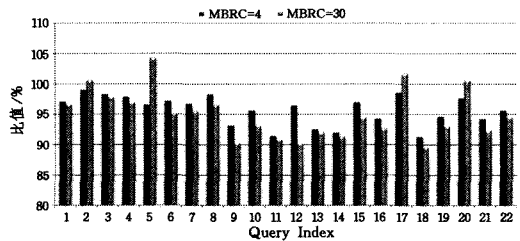


图 9 MBRC 值为 4 和 30 时的 TPC-H 能耗预测图

由表 5 可知,随着 MBRC 值的增大,理论上与实际上一次 I/O 读取的数据块数之间的差值变大(即 MBRC_D_Value 变大);同时仔细观察图 8 可知,在 MBRC 为 4 时,预测模型的误差范围为 2.88%~3.3%,平均误差为 3.126%,相对于其他 MBRC 值,当 MBRC 为 4 时误差比较稳定。此时,MBRC_D_Value 值最小,MBRC 值最接近一次 I/O 操作读取数据块数量的真实值,Multi_total(多块读操作总数)值也最稳定。由图 9 可知,当 MBRC 为 4 时预测模型的预测性能优于 MBRC 为 30 时的预测性能。由此可得,针对不同的数据库配置环境,有一个对应的 MBRC 值符合能耗预测模型稳定性的要求,但这时的 Multi_total 值需要通过大量的测试来获取。综上,多块读能耗预测模型的误差来源于 MBRC 值的设置,同时与系统环境的设置有紧密的关系,在进行数据库能耗建模时,要充分考虑不同数据库环境下能耗预测模型参数之

间的差异,通过大量的实验训练出较稳定的参数值,以提升预测模型的稳定性及准确度,提高模型对数据库负载能耗的预测精度。

结束语 本文以 SQL 语句消耗的 CPU 和磁盘资源产生的时间代价与功率代价为数据库负载构建能耗预测模型,并结合理论与实验论证了能耗预测模型的有效性。在预测模型稳定性的研究中,根据数据库磁盘访问类型(单数据块读操作和多数数据块读操作)的不同,分别探讨了拥有不同磁盘访问类型的数据库负载能耗预测模型的稳定性,深入研究了 MBRC 设置与预测模型稳定性的关系,确定了 MBRC 设置对能耗预测模型准确度的重要影响。在进行关系型数据库能耗建模时,要充分考虑到 MBRC 设置对预测模型准确度的影响,通过大量的实验训练出稳定的预测模型参数,以提升预测模型的稳定性,实现对数据库负载能耗的精确预测。下一步的主要工作包括:1)以能耗预测模型为基础,进一步研究如何利用能耗预测模型为数据库负载挑选节能的查询计划,进而构建绿色数据库;2)研究将本文提出的能耗预测模型嵌入到开源关系型数据库中的实践方法,并对数据库中其他类型的负载(Insert, Update 语句等)进行能耗建模,进一步完善关系型数据库能耗预测模型。

参考文献

- [1] LV T W. A deep analysis of data center energy efficiency in 2013[J]. The world of Power Supply, 2013(6): 7-8. (in Chinese)
吕天文. 2013 年数据中心能效现状深度分析[J]. 电源世界, 2013(6): 7-8.
- [2] 中国 IDC 圈. 未来五年国内数据中心能耗将翻一番[EB/OL]. (2012-03-29) [2015-01-31]. <http://tech.idcquan.com/pro/34910.shtml>.
- [3] BRILL K G. Data Center Energy Efficiency and Productivity [J]. The Uptime Institute-White Paper, 2007(5): 176-184.
- [4] YUN D, LEE J. Research in green network for future Internet [J]. Journal of KIISE, 2010, 28(1): 41-51
- [5] ANDREW L L H, LIN M, WIERMAN A. Optimality, fairness, and robustness in speed scaling designs [J]. ACM Sigmetrics Performance Evaluation Review, 2010, 38(1): 37-48.
- [6] POESS M, NAMBIAR R O. Energy cost, the key challenge of today's data centers; a power consumption analysis of TPC-C results [J]. Proceedings of the VLDB Endowment, 2008, 1(2): 1229-1240.
- [7] MARDAMUTU K, JOON R A L A, JEGATHEESAN J, et al. A Critical Analysis of Green Database System for Energy Efficiency and Green Computing [J]. ACM SIGCOMM Computer Communication Review, 2008, 39(1): 68.
- [8] CHEN J J, KUO C F. Energy-efficient scheduling for real-time systems on dynamic voltage scaling(DVS) platforms [C]//IEEE International Conference on Embedded and Real-time Computing Systems and Applications. IEEE, 2007: 28-38.
- [9] LANG W, PATEL J. Towards eco-friendly database management systems [C]//CIDR. 2009.
- [10] YU J, LIAO B, ZHANG T, et al. A Survey on Energy-Efficient Cloud Storage System [J]. Journal of Frontiers of Computer Science and Technology, 2014, 8(9): 1025-1040. (in Chinese)
于炯, 廖彬, 张陶, 等. 云存储系统节能研究综述 [J]. 计算机科学与探索, 2014, 8(9): 1025-1040.
- [11] LIU L, WANG H, LIU X, et al. GreenCloud: a new architecture for green data center [C]//Proceedings of the 6th International Conference Industry Session on Autonomic Computing and Communications Industry Session. ACM, 2009: 29-38.
- [12] YE R, LI C, LIANG X Y. Energy Storage System in Green Data Centers: Architecture and Management [J]. Journal of Computer Research and Development, 2016, 53(2): 326-340. (in Chinese)
叶冉, 李超, 梁晓晓. 面向绿色数据中心的储能系统: 体系结构和管理方法 [J]. 计算机研究与发展, 2016, 53(2): 326-340.
- [13] WANG J T, LAI W Y, MENG X F. Flash-based Database: Studies, Techniques and Forecasts [J]. Chinese Journal of Computers, 2013, 36(8): 1549-1567. (in Chinese)
王江涛, 赖文豫, 孟小峰. 闪存数据库: 现状, 技术与展望 [J]. 计算机学报, 2013, 36(8): 1549-1567.
- [14] JIN P Q, XING B P, JIN Y, et al. A Survey on energy-aware green databases [J]. Journal of Computer Application, 2014, 34(1): 46-53. (in Chinese)
金培权, 邢宝平, 金勇, 等. 能耗感知的绿色数据库研究综述 [J]. 计算机应用, 2014, 34(1): 46-53.
- [15] XU Z. Building a power-aware database management system [C]//Proceedings of the Fourth SIGMOD PhD Workshop on Innovative Database Research. ACM, 2010: 1-6.
- [16] XU Z, TU Y C, WANG X. Exploring power-performance tradeoffs in database systems [C]//2010 IEEE 26th International Conference on Data Engineering (ICDE 2010). IEEE, 2010: 485-496.
- [17] XU Z, TU Y C, WANG X. PET: reducing database energy cost via query optimization [J]. Proceedings of the VLDB Endowment, 2012, 5(12): 1954-1957.
- [18] XU Z, WANG X, TU Y C. Power-Aware Throughput Control for Database Management Systems [C]//Proceedings of 10th international conference on Autonomic Computing (ICAC). NJ: IEEE, 2013: 315-324.
- [19] XU Z, TU Y C, WANG X. Dynamic energy estimation of query plans in database systems [C]//Proceedings of the 2013 IEEE 33rd International Conference on Distributed Computing Systems (ICDCS). NJ: IEEE, 2013: 83-92.
- [20] XU Z, TU Y C, WANG X. Online Energy Estimation of Relational Operations in Database Systems [J]. IEEE Transactions on Computers, 2015, 64(11): 3223-3236.
- [21] RODRIGUEZ-MARTINEZ M, VALDIVIA H, SEGUEL J, et al. Estimating power/energy consumption in database servers [J]. Procedia Computer Science, 2011, 6: 112-117.
- [22] YANG L H, ZHU H Y. Whole system realtime power profiling and modeling [J]. Computer Science, 2014, 41(9): 32-37. (in Chinese)
杨良怀, 朱红燕. 整机系统实时功率剖析与建模 [J]. 计算机科学, 2014, 41(9): 32-37.
- [23] GUO B L, YU J, LIAO B, et al. Research on SQL Energy Consumption Modeling and Optimization [J]. Journal of Computer Science, 2015, 42(10): 202-207. (in Chinese)
国冰磊, 于炯, 廖彬, 等. SQL 能耗建模及优化研究 [J]. 计算机科学, 2015, 42(10): 202-207.