

云计算环境下大型电气数据库存储性能分析与优化

在云计算环境下,电气数据通过云的形式存储于大型电气数据库中,大型电气数据库的存储性能关系电气数据的吞吐能力和数据资源的调度能力,研究电气数据库的云存储优化模型意义重大。传统的云计算环境下的大型电气数据库的云存储构建模型,采用用户行为特征资源负载均衡方法进行云存储系统任务调度,存在电气数据库存储资源浪费,云计算网格空间分配不均等问题。提出一种基于云计算网格空间资源信息流时间尺度分析的大型电气数据库的存储优化模型设计方法,构建云计算环境下大型电气数据库系统模型,设计云格空间电气数据资源调度模型,分析电气数据信息流的时间尺度特性,设计资源尺度伸缩算法实现电气数据库存储调度均衡,提高大型电气数据在数据库的云存储和数据调度性能。仿真实验表明,采用该算法进行大型电气数据库云存储设计,极大提高数据资源利用率,减少数据调度任务执行时间,优化了数据库的云存储性能。

李进生/武汉软件工程职业学院 杨东陵/武汉供电设计有限公司



李进生/副教授

关键词/Keywords

云计算·
大型电气数据库·
云存储·
调度·

云计算是近年来随着计算机技术和网络技术迅速发展起来的一种先进的计算技术,在云计算中,可以基于互联网进行相关服务的增加、使用和交付,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。云计算是通过使计算分布在大量的分布式计算机上,支持用户在任意位置、使用各种终端获取应用服务。在电气数据调度研究领域,电力电气分配和电力调度的智能管理系统通常建立在云计算环境中,电气网络利用各种通信方式进行各种信息传送和处理,产生大量的电气网络数据,电气数据通过云存储的形式存储于大型电气数据库中。大型电气数据库的存储性能

关系电气数据的吞吐能力和数据资源的调度能力,研究电气数据库的云存储优化模型意义重大。然而电气数据以大数据散布的形式分布在云存储系统中,由于云存储节点的不规则性和自组网性,导致存储环境状态多变复杂,存储性能不好,对云计算环境下的大型电气数据的云存储性能分析和研究是当前的一道难题和热点^[1]。

传统方法中,对云计算环境下的大型电气数据库的云存储构建模型,采用用户行为特征资源负载均衡方法进行云存储系统任务调度,存在电气数据库存储资源浪费,云计算网格空间分配不均等问题^[2-3]。对此,相关文献进行了算法改进设计。其中,文献[4]采用谐波振荡算法对大型电气数据的干扰特征进行滤波器设计,无法有效实现对通信信号的线性反馈和谐波抑制,抗干扰性能不好,没能实现有效的存储优化;文献[5]提出一种基于遗传正交加权约束的大数据库集成和抗干扰算法,构建大型电气数据库的特征信息提取模型,并进行云存储调度算法设计,但算法结构层次复杂,具有计算量大的缺点;文献[6]采用基于块匹配的大型电气数据库资源调度方法实现数据访问,在弱信号强干扰环境下干扰抑制能力差,算法稳健性不好。综合传统方法分析可知,传统的云计算环境下的大型电气数据库的云存储构建模型,采用用户行为特征资源负载均衡方法进行云存储系统任务调度,存在电气数据库存储资源浪费,云计算网

湖北省规划办课题(2013B256)。

格空间分配不均等问题^[7]。

针对上述问题, 本文提出一种基于云计算网格空间资源信息流时间尺度分析的大型电气数据库的存储优化模型设计方法。首先在云计算环境下, 对大型电气数据库进行数据库系统模型构建, 设计云格空间电气数据资源调度模型, 提取电气数据信息流的时间尺度特征, 基于资源尺度伸缩算法进行电气数据调度均衡设计, 提高大型电气数据在数据库的云存储和数据调度性能, 最后进行仿真实验进行性能验证, 展示了本文算法在提高电气数据库存储能力和数据资源调度能力方面的优越性能。

大型电气数据库云存储系统模型构建及特征描述

1. 大型电气数据库云存储模型及相关问题描述

大型电气数据库云存储技术将大量的数据分布到多个服务节点进行缓存分析, 对数据库查询指向性信息特征进行挖掘和数据资源调度, 在云计算环境下, 用户节点通过云服务器进行对等的网络连接, 实现云资源的存储、服务和多媒信息调度。该模式大多应用在非网络环境, 一般以日常实体资源模型为主, 基于各大电气系统网络进行自主服务, 不涉及其他平台领域。大型电气数据库云存储三层模型从上到下分别是: I/O, USB和磁盘层。通过Cloud-P2P融合模型, 云服务提供商将多源信息资源放在代理上, 代理服务器则根据用户需求进行服务信息搜索, 用户最终利用代理信息检索结果选择服务商并获取其服务^[8], 大型电气数据库云存储模型总体构建设计和实现过程如图1所示。

图1设计的云计算环境下的大型电气数据库云存储模式采用分布式代码执行方式, 将任务代码分布到多个数据节点, 实现分布式执行。执行完后, 结果统一返回到客户端, 从而有效避免大型电气数据库中的缓存数据的移动与传输。为了实现大型电气数据库的云存储资源调度优化, 根据云存储任务调度系统中每个节点的感知范围和感知任务的不同, 执行信息具有差异^[9-12]。假设 $V=$

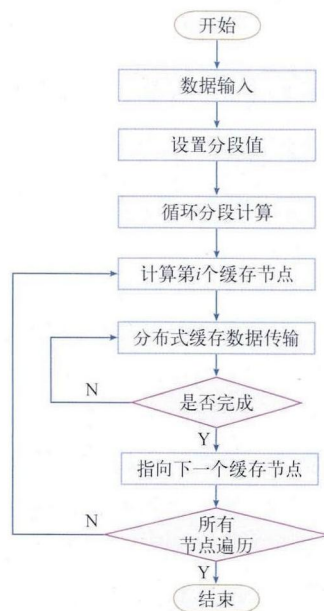


图1 大型电气数据库云存储模型

$[v_1, v_2, \dots, v_n]$ 表示任务调度向量, 对云存储系统来说, v_i 代表第*i*个任务的数据由 v_i 的值代表的资源节点提供。为了保证云存储性能和质量, 云节点长期稳定地在网络上运行, 系统在进行多源资源访问和信息融合中, 由此构建云计算环境下多源信息资源的大型电气数据库云存储模型, 为下一步实现存储优化提供模型基础。

2. 云计算环境下大型电气数据库的数据特征分析

在上述云存储模型构建的基础上, 分析云计算环境下大型电气数据库的数据特征, 设计云格空间电气数据资源调度模型, 进行电气数据信息流的时间尺度特性。在此, 需要给出大型电气数据库的云网格空间分布模型, 示意图如图2所示。

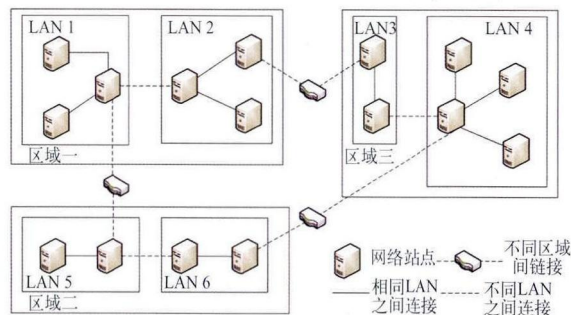


图2 大型电气数据库的云网格空间分布模型

图2对应到三层网络模型中, 将云存储Web

网络环境的网络拓扑结构分为三层：第一层连接互联网的存储数据的存储元件，进行大型电气数据的集成和特征分析；第二层由本地局域网（LAN）组成，在考虑到云计算环境涉及了大量多源异构资源，进行大型电气数据库数据的存储；第三层是会话层，实现对计算资源和数据集的信息交互。根据上述描述，进行电气数据资源特征分析，提取电气数据信息流的时间尺度特征，大型电气数据库的云存储资源信息特征为

$$s_m(t) = \cos\{2\pi f_0[t + \tau_m(\theta)]\} \quad (1)$$

式中， θ 为数据分类中的偏移角度； f 为频率； t 为周期。

云计算网格中节点通过频繁收集新的网络拓扑信息，计算大型电气数据库的云存储资源调度方案并传送给每个节点，并按下式更新抽样概率分布参数

$$\hat{\sigma}_{k+1} = \hat{\beta}_k \hat{\sigma} + (1 - \hat{\beta}_k) \hat{\mu}_k \quad (2)$$

式中， $\hat{\beta}_k$ 为云存储系统的时间尺度机器学习因子； $\hat{\sigma}$ 为目标的归一化时间向量。

在进行云格大型电气数据库的云存储资源调度中，各空闲时间片长度分别记作 $x_1, x_2, \dots, x_{m+1}$ ，可以得到如下方程

$$x_1 + x_2 + \dots + x_{m+1} = T + t - mt \quad (3)$$

式中， T 为总的计算时间长度； m 为云格空间维数。

计算大型电气数据库的云存储资源的计算节点进行时间尺度均衡，得到云计算网格空间中大型电气数据库的云存储资源尺度伸缩与时间平移示意图如图3所示。

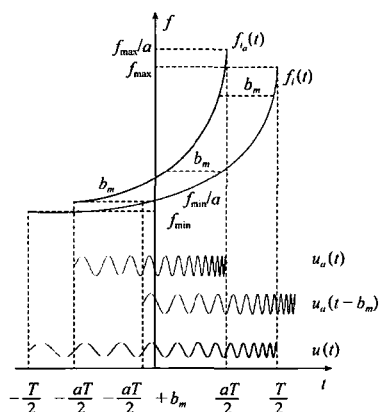


图3 云计算网格空间中大型电气数据库的云存储资源尺度伸缩与时间平移示意图

图3中，每个任务节点 $v \in V$ 都有一个延迟属性 $\text{delay}(v)$ ，对于尺度为 a 的大型电气数据库的云存储资源任务请求信息，以 t/a 代替 t ，得到特征尺度向量空间和空闲时间片的长度。由此实现了云计算环境下大型电气数据库的数据特征分析，下一步将以此为基础设计资源尺度伸缩算法实现电气数据库存储调度均衡。

云资源尺度伸缩和大型电气数据库云存储性能优化

1. 云资源尺度伸缩算法的提出

在上述模型设计和特征分析的基础上，结合传统的云计算环境下的大型电气数据库的云存储构建模型，采用用户行为特征资源负载均衡方法进行云存储系统任务调度，存在电气数据库存储资源浪费，云计算网格空间分配不均等问题。进行算法改进，本文提出一种基于云计算网格空间资源信息流时间尺度分析的大型电气数据库的存储优化模型设计方法。设计关键技术描述如下：在大型电气数据库的云存储设计中，假设 T 为每个大型电气数据库的云存储资源为所有用户任务服务一轮花费的时间，每个时间片都可以被用来处理任意用户请求任务，云存储资源综合负载为

$$R = w_1 C_i + w_2 D_i + w_3 M_i + w_4 N_i \quad (4)$$

在云网格系统中进行电气数据调度运行的一个时间跨度记为

$$h(\tau_i, t) = \sum_{i=1}^{N_m} a_i e^{j\theta_i(t)} \delta[t - \tau_i(t)] \quad (5)$$

式中， $\delta[t - \tau_i(t)]$ 为周期跨度； $e^{j\theta_i(t)}$ 为方向角跨度； a_i 为时间节点数据。

假定网络结构模型为某种几何空间 S ，在计算主机 h_1 和 h_2 的网格间距离，采用尺度伸缩算法，得到资源调度的尺度特征为

$$x(t) = \frac{1}{\sqrt{T}} \text{rect}\left(\frac{t}{T}\right) \exp\{-j[2\pi K \ln(1 - \frac{t}{t_0})]\} \quad (6)$$

式中， $\text{rect}(t) = 1, |t| \leq 1/2$ 。大型电气数据库的云存储资源信息流的调频规律为双曲函数，即

$$f_i(t) = \frac{K}{t_0 - t} \quad |t| \leq \frac{T}{2} \quad (7)$$

式中, $K = T f_{\max} f_{\min} / B$; $t_0 = f_0 T / B$; f_0 为算术中心频率; $f_{\min}$, $f_{\max}$ 分别为最低和最高频率。

为了实现负载最小, 达到大型电气数据库的云存储资源均衡, 基于资源尺度伸缩算法实现电气数据库存储调度均衡。

2. 云存储性能优化实现

基于上述设计资源尺度伸缩算法实现电气数据库存储调度均衡, 提高大型电气数据在数据库的云存储和数据调度性能。假设大型电气数据库的云存储资源与服务调配匹配的任务有效率函数 $E(i, j)$ 为

$$E(i, j) = \begin{cases} \frac{e_{ij} - e(i, j)}{e_{\max} - e(i, j)} & e(i, j) < e_{ij} \\ \frac{e_{ij} - e(i, j)}{e(i, j) - e_{\min}} & e(i, j) \geq e_{ij} \end{cases} \quad (8)$$

式中, e_{ij} 为给定的阈值; $e_{\max}$ 为最大效率值; $e_{\min}$ 为最小效率值。

给定大型电气数据库的云存储资源大数据信息集合中的两条数据轨迹 A, B。对于轨迹 A 上的任意一点 a_i , 计算数据库分区访问的输出概率中第 m 个高斯分量的均值矢量

$$\phi(i) = \arg \min_{j \in B} \|x_i^a - x_j^b, y_i^a - y_j^b\| \quad (9)$$

式中, 云存储数据资源混合数 M 的大小决定了输出概率的准确度, 假设大型电气数据库的云存储资源大数据定位信息在干扰和噪声作用下, 轨迹 A, B 之间的空间有向距离为

$$h(A, B) = \frac{1}{N_A} \sum_{i \in A} \|x_i^a - x_{\phi(i)}^b, y_i^a - y_{\phi(i)}^b\| \quad (10)$$

式中, N_A 为 A 的数据信息流特征分割权向量。A, B 之间的分区方位的协方差矩阵, 计算正交加权约束均衡比, 得到

$$H(A, B) = \begin{cases} h(A, B), & \text{if } h(A, B) \leq h(B, A) \\ h(B, A), & \text{if } h(A, B) \geq h(B, A) \end{cases} \quad (11)$$

找到上式的最小值, 将对应节点的 H_k 作为输出结果, 实现安全性控制下的大型电气数据库的云存储资源均衡, 提高了云格空间大型电气数据库的云存储资源的调度性能, 实现云存储性能优化。

仿真实验与结果分析

为了验证本文算法在实现云计算环境下对大型电气数据库的云存储性能优化调度的性能, 进行仿真实验。采用 10 个计算节点组建云格环境, 对 GridSim 仿真平台中的网格大型电气数据库的云存储资源调度器进行 EFRB 扩充, 构建 100 个任务组成大型电气数据库的云存储资源调度集。每个大型电气数据库的云存储资源调度任务的执行时间为 120 s, 总任务分为 10 个队列。首先进行云存储任务调度系统网络模型设计, 设节点感知半径 $r = 20$ m, 感知偏向角 $\alpha = \pi/4$, 整个云存储任务调度系统执行信息预见范围划分为 16 个虚拟约简检测单元格, 大型电气数据库网络云网格节点辐射分布模型如图 4 所示。

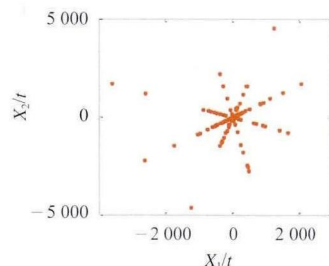


图4 大型电气数据库网络云网格节点辐射分布模型

根据上述仿真环境设计, 构建大型电气数据库云存储系统模型, 并进行数据资源的存储和调度仿真, 首先提取电气数据信息流的时间尺度特征, 提取结果如图 5 所示。由图 5 可见, 采用本文算法进行电气数据时间尺度特征提取, 频间谱峰值明显, 时间-尺度建模分析结果准确, 有效满足了云计算网格空间大型电气数据库的云存储资源分配的均衡设计。

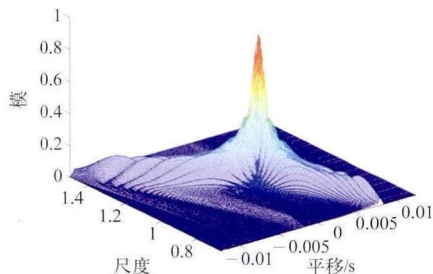


图5 电气数据信息流的时间尺度特征

根据上述电气数据信息流的时间尺度特征提取结果,进行电气数据库存储调度均衡,优化云存储能力。为了对比算法性能,采用本文算法和传统方法,以数据库的电气数据云存储分层融合度为测试指标,得到性能对比结果如图6所示。分析图6结果表明,采用本文算法,进行大型电气数据库云存储设计,提高云存储资源的数据分层融合度,由此可以极大提高数据资源利用率,减少数据调度任务执行时间,优化了数据库的云存储性能。

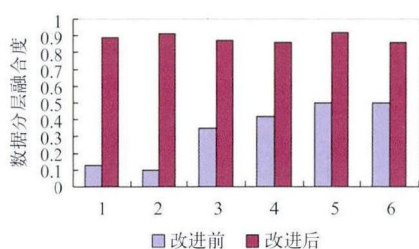


图6 电气数据云存储分层融合度

结束语

电气网络利用各种通信方式进行信息传送和处理,产生大量的电气数据,大型电气数据库的存储性能关系电气数据的吞吐能力和数据资源的调度能力,研究电气数据库的云存储优化模型意义重大。由于云存储节点的不规则性和自组网性,导致对存储环境状态多变复杂,存储性能不好,对云计算环境下的大型电气数据的云存储性能分析和研究是当前一个难题和热点。本文提出一种基于云计算网络空间资源信息流时间尺度分析的大型电气数据库的存储优化模型设计方法,研究结果表明,采用本文方法进行大型电气数据库云存储设计,提高云存储资源的数据分层融合度,由此可以极大提高数据资源利用率,减少数据调度任务执行时间,优化了数据库的云存储性能,在电气电网管理云存储设计中展示了较好的应用价值。

参考文献

[1] Bimal K M, Gholam M A. Differential epidemic model of virus and worms in computer network [J].

International Journal of Network Security, 2012, 14 (3): 149-155.

- [2] Zhu Q Y, Yang X F, Yang L X, et al. Optimal control of computer virus under a delayed model [J]. Applied Mathematics and Computation, 2012, 218 (23): 11613-11619.
- [3] 周涛. 基于改进神经网络的电力系统中长期负荷预测研究 [J]. 电气应用, 2013, 32 (4): 26-29.
- [4] 耿忠, 刘三阳, 齐小刚. 基于非合作博弈的无线传感器网络功率控制研究 [J]. 控制与决策, 2011, 26 (7): 1014-1018.
- [5] 张磊, 王鹏, 黄焱, 等. 基于相空间的云计算仿真系统研究与设计 [J]. 计算机科学, 2013, 40 (2): 84-86.
- [6] 匡桂娟, 曾国荪. 一种基于时分复用的云大型电气数据库的云存储资源管理方法 [J]. 同济大学学报: 自然科学版, 2014, 42 (5): 782-789.
- [7] 曾志, 王晋, 杜震洪, 等. 一种云格环境下可计算大型电气数据库的云存储资源与服务高效调配机制 [J]. 浙江大学学报: 理学版, 2014, 41 (3): 353-357.
- [8] Balasangameshwara J, Raju N. A hybrid policy mr fault tolerant load balancing in Grid computing environments [J]. Journal of Network and computer Applications, 2012, 35 (1): 412-22.
- [9] 杨照峰, 王启明, 吕海莲. 基于任务延迟的云计算大型电气数据库的云存储资源调度算法研究 [J]. 计算机测量与控制, 2014, 22 (2): 499-502.
- [10] Sonmea O, Mohamed H, Epema D. On the benefit of processor coallocation in multicluster Grid systems [J]. IEEE Transactions on Parallel and Distributed Systems, 2010, 21 (6): 778-89.
- [11] 张骏, 田泽, 梅魁志, 等. 基于节点预测的直接Cache一致性协议 [J]. 计算机学报, 2014, 37 (3): 700-720.
- [12] 肖景, 郑秋华. 一种基Petri网的模型调度方法 [J]. 科技通报, 2011, 27 (2): 186-190.

(收稿日期: 2014-11-12) **EA**