

面向数据网格的动态复制选址策略及其应用

冯永¹, 李志国^{1,2}, 钟将¹, 叶春晓¹, 邓伟¹

(1. 重庆大学计算机学院, 重庆 400030; 2. 上海宝信软件西南研发中心, 重庆 400041)

摘要: 动态复制技术对于提高数据网格的性能非常重要。鉴于目前效果较好的动态复制策略均为单选址算法, 对于延迟较大、分布较广的网格存在很大局限性, 提出3种多选址的动态复制策略, 将其转化为经典的数学问题求解。并给出了多选址动态复制策略在远程教育资源管理中的应用。

关键词: 数据网格; 动态复制; 复制选址; 远程教育

Dynamic Replica Placement Strategies and Applications for Data Grids

FENG Yong¹, LI Zhi-guo^{1,2}, ZHONG Jiang¹, YE Chun-xiao¹, DENG Wei¹

(1. College of Computer Science, Chongqing University, Chongqing 400030;

2. Southwest Research Center, Shanghai Baosight Software Corporation, Chongqing 400041)

【Abstract】 Within data grid environments, dynamic replication is a general mechanism to improve performance and availability for distributed applications. As the experimentally efficient strategies are based on single-location algorithms, they are not suitable for geographically broad grids with long latency. To address such issues, three multiple-location strategies are proposed, which are converted into classical mathematic problems that can be solved by some famous approximation algorithms. The strategies are suitable for resource management in the remote education.

【Key words】 data grid; dynamic replication; replica placement; remote education

数据网格(data grid)的概念来自网格(grid)。它是网格技术在数据管理方面的应用和实现, 是为了在网格环境下, 建立透明访问异构数据资源的新体系结构。在数据密集型的大规模分布式协作应用中, 数据的用户群分布广泛, 数据本身也分布在不同位置上。为了减少访问数据的时间, 可先从别的存储位置复制一部分数据到本地机; 或在多个位置存放某一数据。当某计算节点需要这些数据时, 可以从访问时间最短的存储节点上获取所需数据。即, 复制数据可以降低远程访问该数据的网络延迟及带宽消耗, 还可以增强网络的负载平衡, 提高数据的安全性、可靠性, 因此数据复制在数据网格中十分重要^[1]。

1 动态复制策略

动态数据复制对复制的创建、删除和管理都可动态且自动地进行。网格环境下, 由于数据量、用户量非常大且网格环境的动态性, 只有动态创建副本才能够满足网格环境的要求。在运行时, 复制系统根据网格系统的应用特征以及用户的访问特征等自动将数据项复制到不同节点上, 以实现网格系统的整体性能优化。

动态复制的核心是动态复制策略, 文献[2]综述了6种策略: 无复制策略, 最优客户策略, 逐点复制策略, 平凡缓存策略, 逐点复制结合平凡缓存策略和快速传播策略。

由于以上几种策略均未考虑用户的动态行为, 因此文献[3]提出将用户的请求(经常动态变化)与网络状态结合的多种动态复制策略。其效果最好的一种复制选址算法 *MinimizeExpectedUtil* 如下: 对每个节点 i 计算其使用期望 $EU(i)$, 然后在所有的节点中选择使用期望最大的节点作为复

制对象存放的站点。当节点总数为 n 时, 源节点标号为 0, 节点 i 的使用期望 $EU(i)$ 定义如下:

$$EU(i) = \sum_{j=1}^n \text{MinReplicaDistance}(j) * \text{Request}(j)$$

其中, $\text{MinReplicaDistance}(j)$ 表示节点 j 到节点 i 和 0 两节点距离中的小者, 距离定义为路由表中的时间延迟; $\text{Request}(j)$ 表示节点 j 在一定时间内对文件或者对象的请求次数。

2 多选址策略

算法 *MinimizeExpectedUtil* 是一个单选址算法, 在用户地理分布较广、人数较多、网格延迟较大的场合下, 有必要引入多选址算法, 使得复制站点越多, 用户的访问延迟越小, 全局网络带宽占用也越小。但复制站点越多也会导致复制站点上被占用的磁盘空间越大, 复制管理越复杂。因此, 有必要在两者间寻求一个平衡点。

2.1 网络负载和用户行为结合的策略

为了在少建复制站点的条件下减少网格中的负载, 引入建立复制站点将产生费用这个约束, 同时, 每个节点根据与复制站点的距离和请求次数也将产生费用约束, 所求的目标

基金项目: 国家“十一五”重大科技攻关基金资助项目(2006BAH02A24-6); 国家社会科学“十一五”规划教育学重点课题基金资助项目(ACA07004-08); 重庆市科委自然科学基金资助项目(2007BB2192); 重庆市高等教育教学改革研究基金资助重大项目(0616001)

作者简介: 冯永(1977-), 男, 博士, 主研方向: 知识管理, 知识发现, 网格计算; 李志国、钟将, 博士; 叶春晓, 副教授; 邓伟, 博士

收稿日期: 2007-07-30 **E-mail:** fengyong@cqu.edu.cn

函数就是选择多个复制站点,使得网络的总费用最少。

这个问题和一个经典的图论选址问题——配送中心选址问题(facility location problem)很类似。配送中心选址问题较为正式的表述应为:在加权无向连通图 $G=(V, E)$ 中,每个节点都要求配送中心配送一定物资,而在节点 i 上建立一个配送中心产生费用 c_i ,节点 j 根据与中心 i 的距离产生费用 c_{ij} ,问如何设立配送中心,使 $\sum y_i c_i + \sum x_{ij} c_{ij}$ 最小。其中, y_i 取0或1,取1表示在节点 i 上建立配送中心; x_{ij} 取0或1,取1表示节点 j 指定到配送中心 i 。在网格中,过去访问频繁的文件或对象往往以后的访问也很频繁。所以可以通过用户在之前一段时间提出的请求次数来代替他以后的请求次数。如果用 d_j 表示站点 j 在一段时间内的访问次数,那么就要求出 $\min(\sum y_i c_i + \sum x_{ij} d_j c_{ij})$ 。

配送中心选址问题和网格复制多选址问题的差别在于:(1)问题中存在至少一个位置固定的源数据站点(复制站点的源节点),而配送中心选址问题中所有的配送中心位置均可变化。(2)在费用函数中加入了访问次数,但是对于文献[4]中的线性规划算法而言,该访问次数为常数,所以,不影响算法的时空复杂性量级。

在网格中,选择复制站点时必须考虑以下因素:网格的拓扑结构和网络状况,复制站点的运行负载及存储终端效率,数据副本尺寸等。为了在尽可能少建复制站点的条件下达到减少和优化全局负载的目标,数据网格中的费用函数可定义如下:站点 i 建立费用 c_i 定义可参见文献[5]的定义或简单定义为从站点 i 到源节点之间的网络延迟;其他站点 j 访问复制站点 i 的费用 $d_j c_{ij}$ 中, d_j 为该站点在一段时间内的请求次数, c_{ij} 为路由表中节点 i 与节点 j 之间的最短路由时间。

复制站点的费用函数不好确定,可以考虑将本问题转化成配送中心选址问题的一个变形问题,则 k -中心选址问题(k -median problem):在配送中心的数目不超过 k 的条件下 $\sum y_i \leq k$,求 $\min(\sum x_{ij} c_{ij})$ 的选址方案。文献[6]研究了在不同的用户需求下,选择多少个复制站点合适这个问题,那么 k 的值可由此确定。目前, k -中心选址问题的最好 ρ -近似算法为4-近似算法^[4],时间复杂度为 $O(n^3)$ 。同理,可以对该算法稍加修改,用于求解增加了用户访问次数和固定源节点的网格复制多选址问题。

2.2 考虑 QoS 的策略

为了保证数据网格中用户的服务质量 QoS,有必要限制用户与服务站点之间的网络延迟。为此,要求在网络的正常情况下,节点 j 与复制中心 i 之间的网络延迟不超过常数 t 。在此条件下,要求算法能够得出最少需建立的复制站点数目及位置。

通过网络的路由表,采用最短路径算法可以得到每个节点 i 在时间 t 内可以到达的节点集合 S_i 。则原问题变为:在 $S_2, S_3, \dots, S_i, \dots, S_n$ 中求最少数目的集合,使它们覆盖除源数据站点(可能不止一个)已覆盖节点集 S_1 外的所有节点。回顾经典的集合覆盖问题(set cover problem): $E=\{1, 2, \dots, n\}$ 为全集,集合 SS 的元素为定义在 E 上的集合,对于每个 $S \in SS$,定义了一个非负的费用函数 $\omega(S)$,目标是求出能够覆盖(cover)全集且具有最少费用的一组集合。

如果 SS 中元素为原问题中的 $S_2, S_3, \dots, S_i, \dots, S_n$,所有费用函数 $\omega(S_i)$ 定义为1,全集定义为除源数据站点(可能不止一个)已覆盖节点 S_1 外的所有节点,由于一般的网格通信是对称的

(即往返时间一致),因此原问题就可转化为集合覆盖问题。

众所周知,一般意义下的集合覆盖问题也是一个经典的 NP 难问题。文献[7]证明了在集合覆盖问题的 ρ -近似算法中 $\rho > (1+O(1))\ln(n)$,其中, n 为节点总数。至今,经典的解集合覆盖问题的 ρ -近似算法为提出的 $O(\ln(n))$ -近似算法。至此,可以采用该近似算法来解决这个考虑 QoS 的多选址问题。

2.3 综合 QoS、网络负载和用户行为的多选址策略

如果在考虑到用户的行为模式的情况下,能尽量减少网络负载并满足用户 QoS 将是一个更为合理的多选址策略。考虑到适应用户的行为模式,通过文献[6]的方法来选择复制站点数目 k 。用 D 表示节点 j 与供应中心 i 的距离 c_{ij} 的上界,则问题可描述为:在 $\sum y_i \leq k$ 且 $\forall x_{ij} x_{ij} c_{ij} \leq D$ 的条件下,求 $\min(\sum x_{ij} d_j c_{ij})$ 的选址方案。用线性规划表示该问题如下:

目标: $\min(\sum x_{ij} d_j c_{ij})$

约束: $\sum_i y_i \leq k; \forall x_{ij} x_{ij} c_{ij} \leq D; \forall i y_i + \sum_j x_{ij} = 1; x_{ij} = 0 \text{ or } 1; y_i = 0 \text{ or } 1$

这个 0-1 整数线性规划问题等价于增加一个常数约束条件的 k -中心选址问题,可以采用前文提到的 4-近似算法^[4],也可用如 CPLEX^[8]等数学包进行求解。

3 复制策略在远程教育资源管理中的应用

在国家发改委的下一代互联网示范工程(CNGI)中,建设开放式教育网络是构建终身学习体系的重要组成部分,而配套的远程教育资源建设是其主要内容。远程教育资源往往是海量的,且分布在不同地域的不同站点上。为了向用户透明地提供一套高效的资源共享、存储管理机制,数据网格成为一种最理想的选择^[9]。为了让用户能够方便和快捷地存取这些资源,对这些资源进行动态复制并发送到便于用户访问的地方就变得相当重要。由于用户分布的地域很广,网络延迟也不相同,因此应该根据网络状态和用户的动态需求来采取不同的复制策略。

对于不同节点数目和延迟的网络,网格的复制管理节点可以根据实际的访问数据确定是否需要提供复制站点,提供单个复制站点或者多个复制站点,具体的流程如图 1。复制选址数的确定可以参见文献[6]提出的方法。

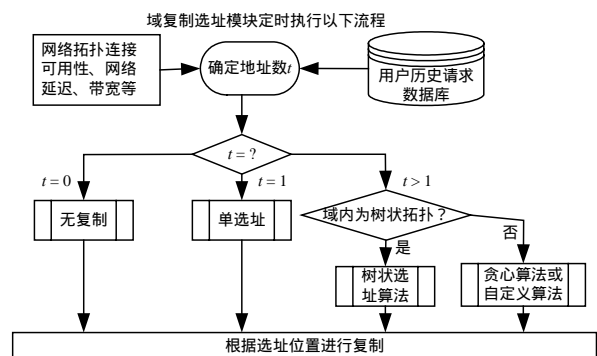


图 1 域复制管理节点的选址流程

文献[6]提出适应用户访问模式的单选址算法,并通过在欧洲数据网格试验床 1 仿真证明了其相较于已有单选址算法在减少网络负载和网络延迟上的高效性。由于在远程教育中用户的访问模式变化较大,因此在单选址时可以采用该算法。

对于多选址算法而言,不同的系统优化指标(如网络负载、网络延迟、存储开销、QoS 要求等)有不同的复制选址策

(下转第 91 页)

