

基于网格服务的半连接查询优化算法研究

温慧明¹, 宫晓辉², 焦 洋¹

(1. 煤炭科学研究总院, 北京 100013;

2. 中国电力科学研究院, 北京 100096)

摘 要:随着网格计算技术的快速发展,其应用领域在不断扩大,然而,跨网络分布式数据的联合查询往往成为性能的瓶颈,因此文中从提高分布式数据的联合查询的效率出发,以网格计算这一新型 web 体系结构为技术平台,研究基于网格服务的查询优化目标、查询优化对象和查询优化策略。文中采用了应用非常广泛的启发式搜索算法来缩小策略空间这一优化策略,并在该策略的基础上提出了一种基于网格服务的 2-way 半连接查询优化算法,最后在网格计算环境下对该算法与一般的全连接算法做了实验对比,证明基于网格服务的 2-way 半连接查询优化算法可大大缩短查询响应时间。

关键词:网格计算;查询优化;2-way 半连接

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2012)09-0123-04

Research of Semi-join Query Optimization Algorithm Based on Grid Service

WEN Hui-ming¹, GONG Xiao-hui², JIAO Yang¹

(1. China Coal Research Institute, Beijing 100013, China;

2. China Electric Power Research Institute, Beijing 100096, China)

Abstract: With the rapid development of grid computing, it has been applied into a larger area, however, distributed data joint inquiries across the network often becomes the bottleneck of performance, therefore, in order to improve efficiency of the joint inquiries of distributed data, based on the new grid computing web system structure, research query optimization goal, query optimization objects and query optimization strategy based on grid services. It adopts the optimization strategy uses the very extensive heuristic algorithm to narrow this optimization strategy space, on the basis of this optimization strategy, present a 2-way semi-join query optimization algorithm based on grid service, finally, in the grid computing environment, make the contrast experiment of the algorithm and the general all connection algorithm, and it proves that 2-way semi-join query optimization algorithm based on grid services can greatly shorten query responding time.

Key words: grid computing; query optimization; 2-way semi-join

0 引言

网格计算是近年来国际上兴起的一种重要信息技术,它将高速互联网、高性能计算机、大型数据库、传感器、远程设备等融为一体,实现了计算资源、存储资源、信息资源、知识资源等的全面共享,消除了信息孤岛和资源孤岛^[1-3]。

网格平台上的信息集成问题一直是网格计算领域一个关键问题,网格信息集成中的分布式查询处理优化受到了特别的关注。已经有许多为网格应用而开发的调度器。例如 AppLeS (Application Level Schedu-

ler)^[4]和 Condor^[5],其中 AppLeS 是一个高性能的调度器,Condor 是一个具有集中调度模块的高吞吐量系统,它通过使用代理去匹配任务,找到指定的资源来执行。但它们都忽略了一些重要的性能因素,例如:数据传输等。所以提高查询执行的性能需进一步研究。

1 基于网格服务的查询优化策略

1.1 查询优化目标

查询优化包括两类:针对查询执行代价的优化和针对查询响应时间的优化。但它们考虑因素的偏重不一样,查询响应时间对一个组织机构而言,往往就代表着执行代价,所以,查询响应时间是最终要关心的核心问题。

1.2 查询优化对象

在基于网格的优化机制中,数据访问操作对应网

收稿日期:2012-01-16;修回日期:2012-04-28

基金项目:国家“863”高技术发展计划项目(2007AA12Z203)

作者简介:温慧明(1983-),男,硕士,研究方向为图形处理、数据库、煤炭信息化。

格中某个数据服务的数据查询。网格中单个数据服务是自治的,它自行处理提交给它的查询请求,所以不是优化的对象。而跨网格节点的操作是最费时间和空间的操作。因此在优化基于网格服务的数据联合查询处理策略时,就以跨网格节点的操作作为查询优化的研究对象。

1.3 查询优化策略

在执行跨网格节点的操作时,由于网络连接执行服务节点的负载和网络传输性能的差异必然导致一个很大的查询策略选择空间。计划产生器会根据相关的数据统计信息、系统运行信息(系统和网络的负载)和通信代价找到一个最优的查询执行策略^[6-9]。

文中采用启发式搜索算法缩小策略空间^[1,6,8],启发原则是连接代价最小的连接操作。为了适应网格环境的动态性,以动态的方式使用启发式的搜索策略^[1,6,10]:首先,计划产生器^[6,11]在查询操作数最底层的连接操作节点中选取连接估算代价最小的一个去执行,该操作执行完后,需要将运算后的数据集所在的网格连接执行服务节点,既当作网格连接执行服务节点也当作网格数据服务节点。这时,计划产生器需要更新查询操作树,重新对底层连接操作进行代价估算,根据连接代价最小的原则决定下一个连接操作。以此类推。

2 基于网格服务的半连接查询优化算法

文中采用 2-way 半连接^[5,6,12]的思想,将不参与连接的或无用的数据记录在传输到网格连接执行服务节点之前就在网格数据服务端过滤掉,尽量在网络中只传输参与连接的数据。

但在实际应用中,由于它们之间的网络性能和系统统计信息不近相同,采用半连接不一定就能降低网络中的传输费用。所以半连接查询策略的选择需要依据查询执行传输代价的估算。本章第一节将会对代价估算模型进行详细介绍。

2.1 代价估算模型

在面向远程通信的网格查询系统中 TC (通信代价)对查询响应时间的影响占有主导地位,所以只考虑通信代价 TC 。一般来说,一次传输的通信代价粗略估算为如下的公式:

$$TC(X) = C_0 + C_1 \times X \quad (1)$$

其中, X 为数据的传输量,通常以 bit 为单位计算; C_0 为两站点间通信初始化一次所用的时间,它由通信系统确定,近似一个常数,以秒为单位; C_1 为传输率(传输速度的倒数),即单位数据传输的时间,以秒/bit 为单位。

用 C_{ij} 表示从站点 I 到站点 J 的传输率, $Size(R)$ 表

示关系 R 一个记录的平均大小, $Card(R)$ 表示关系 R 的元组个数, $Size(B)$ 表示关系 S 的属性 B 的大小, $Val(B[S])$ 表示关系 S 里 B 属性上不同值的个数, $Val(B[S])_m$ 和 $Val(B[S])_{nm}$ 分别表示 S 在属性 B 上的投影和 R 在 A 属性上的投影匹配的个数和不匹配的个数, $Val(dom(A))$ 和 $Val(dom(B))$ 表示属性 A 和 B 所在值域的大小, R' 和 S' 分别表示被半连接缩减之后的数据集。

由上面的通信代价的估算方法得先过滤 R 后过滤 S 方式的 2-way 半连接的通信代价如下:

$$TC_1 = 3C_0 + C_{12} \times Size(B) \times Val(B[S]) + C_{21} \times Size(B) \times \min(Val(B[S])_m, Val(B[S])_{nm}) + \max(C_{13} \times Size(R) \times Card(R'), C_{23} \times Size(S) \times Card(S')) \quad (2)$$

由于半连接的不对称性还存在另一种方式的费用:

$$TC_2 = 3C_0 + C_{21} \times Size(A) \times Val(A[R]) + C_{12} \times Size(A) \times \min(Val(A[R])_m, Val(A[R])_{nm}) + \max(C_{13} \times Size(R) \times Card(R'), C_{23} \times Size(S) \times Card(S')) \quad (3)$$

这两个半连接程序的费用可能有所不同,对它们进行比较就可以选定一个较优的半连接方式来处理 R 和 S 的连接操作。在这里可以假设两个站点之间的网络传输率是相同的,即: $C_{12} = C_{21}$, 经分析 $Size(A) = Size(B)$, 则如果 $Val(A[R]) > Val(B[S])$, 必然有:

$$\min(Val(B[S])_m, Val(B[S])_{nm}) < \min(Val(A[R])_m, Val(A[R])_{nm}) \quad (4)$$

此时 $TC_1 < TC_2$, 即第一种方式的半连接较优, 否则第二种较优。

为了说明是否用 2-way 半连接的算法思想来处理连接操作, 可以将其与全连接操作的费用进行比较。 R 和 S 两关系的全连接操作费用为:

$$TC' = C_0 + \max(C_{13} \times Size(R) \times Card(R), C_{23} \times Size(S) \times Card(S)) \quad (5)$$

若 $TC_1 < TC'$, 即

$$2C_0 + C_{12} \times Size(B) \times Val(B[S]) + C_{21} \times Size(B) \times \min(Val(B[S])_m, Val(B[S])_{nm}) + \max(C_{13} \times Size(R) \times Card(R'), C_{23} \times Size(S) \times Card(S')) < \max(C_{13} \times Size(R) \times Card(R), C_{23} \times Size(S) \times Card(S)) \quad (6)$$

通过估算 $Card(R')$ 和 $Card(S')$ 来判断上面的公式是否成立可以变相地借用 SDD-1 算法^[4]里采用的收益费用分析的思想进行判断。如果收益大于费用则 2-way 半连接程序是最优的, 否则不是。下面结合 2-way 半连接思想介绍半连接选择因子及与收益费用

分析相关的几个概念:费用(Cost)、收益(Benefit)和效率(Efficient)。在这里采用先过滤 R 后过滤 S 的2-way半连接方式。

(a) 半连接选择因子:以 $R \propto S$ 为例,那么选择因子为 S 在属性 B 上的投影后属性集的个数和 S 的元组个数的比值,记:

$$SF_{sj}(R \propto S) = \text{Val}(B[S]) / \text{Val}(\text{dom}(B)) \quad (7)$$

其中 A 和 B 是连接属性,则:

$$\text{Val}(\text{dom}(A)) \equiv \text{Val}(\text{dom}(B)) = \text{Max}(\text{Val}(A[R]), \text{Val}(B[S])) \quad (8)$$

(b) 费用(Cost):半连接比全连接在网络中多传输的数据量,记:

$$\text{Cost} \leq \frac{3}{2} \text{Val}(B[S]) \times \text{Size}(B) \quad (9)$$

(c) 收益(Benefit):就是将关系 R 和 S 变为了 R' 和 S' 后,网络中传输量的缩减数量。由于 R' 和 S' 采用并行传输的方式传到 $node3$ 上去执行连接,所以其总收益为缩减前关系的最大传输量和缩减后关系的最大传输量之差。总的收益记为:

$$\text{Benefit} = \text{Max}(\text{Size}(R) \times \text{Card}(R), \text{Size}(S) \times \text{Card}(S)) - \text{Max}(SF_{sj}(R \propto S) \times \text{Size}(R) \times \text{Card}(R), SF_{sj}(S \propto R) \times \text{Size}(S) \times \text{Card}(S)) \quad (10)$$

(d) 效率(Efficient):

$$\text{Efficient} = \text{Benefit} - \text{Cost} \quad (11)$$

上面所谈到的费用和收益都是假设网络中站点之间的距离和发送/接收性能相当、传送率相同,在网格环境中每两个节点间网络传输率不近相同,而网格节点的动态性导致两个站点之间的网络负载是动态变化的,所以考虑2-way半连接算法的费用和收益时就需要把每两站点之间的传输率和网络负载也考虑其中。用 C_{ij} 表示站点 I 和 J 之间网络固有的传输率,也就是在假定负载为零的情况下网络的传输率,用 NL_{ij} 表示 I 和 J 之间的网络负载因子(例如20%),那么网格环境下 I 和 J 之间的网络传输率为 $C_{ij}/(1 - NL_{ij})$ 。修正的费用和收益公式如下:

$$\text{Cost}' \leq \frac{3}{2} \text{Val}(B[S]) \times \text{Size}(B) \times C_{12} / (1 - NL_{12}) \quad (12)$$

$$\text{Benefit}' = \text{Max}(\text{Size}(R) \times \text{Card}(R) \times C_{13} / (1 - NL_{13}), \text{Size}(S) \times \text{Card}(S) \times C_{23} / (1 - NL_{23})) - \text{Max}(SF_{sj}(R \propto S) \times \text{Size}(R) \times \text{Card}(R) \times C_{13} / (1 - NL_{13}), SF_{sj}(S \propto R) \times \text{Size}(S) \times \text{Card}(S) \times C_{23} / (1 - NL_{23})) \quad (13)$$

如果 $\text{Efficient}' = \text{Benefit}' - \text{Cost}' > 0$,则说明2-way半连接程序存在好的优化效果,可以被用来减少

通信代价,否则将直接进行全连接操作。

2.2 算法描述及流程图

通过对半连接代价估算的分析得适合网格查询系统的2-way半连接算法描述如下:

Step1:

(1) 获取 $\text{Val}(A[R])$ 和 $\text{Val}(B[S])$ 的值,并比较大小,如果 $\text{Val}(B[S])$ 小则采用先过滤 R 后过滤 S 的2-way半连接方式,否则过滤顺序相反;

(2) 计算 $SF_{sj}(R \propto S)$ 、 $SF_{sj}(S \propto R)$ 、 Cost' 和 $\text{Benefit}'$,最后得出效率 $\text{Efficient}'$;

(3) 判断如果第二步得出的 $\text{Efficient}' > 0$,则采用2-way半连接优化算法对该次连接操作进行优化,转Step2。否则不采用2-way半连接算法,转Step3。

Step2:

(1) 对关系 S 和 R 在其属性 B 和 A 上进行投影得 $\pi_B(S)$ 和 $\pi_A(R)$,将它们暂存本地;

(2) 关系 R 所在的站点获取 $\pi_B(S)$,在该站点上执行 $\pi_B(S) \propto_{B=A} R$,将结果暂存本地,并同时利用 $\pi_A(R)$ 将 $\pi_B(S)$ 分割为匹配的 $\pi_B(S)_m$ 和不匹配 $\pi_A(R)$ 的 $\pi_B(S)_{nm}$;

(3) 关系 S 所在的站点获取 $\text{Min}(\pi_B(S)_n, \pi_B(S)_{nm})$,然后和 S 执行“等于”连接操作或“不等于”连接操作,将连接结果暂存本地;

(4) node3 站点并行获取 R 所在的站点 node1 和 S 所在的站点 node2 上的结果数据集,并在他们的连接属性上执行连接操作,结果返回客户或暂存本地。

Step3:

node3 站点分别从 node1 和 node2 站点上并行获取关系 R 和 S ,然后在他们的连接属性上执行连接操作,并把结果返回给客户或暂存本地。

2-way半连接算法的流程图如图1所示:

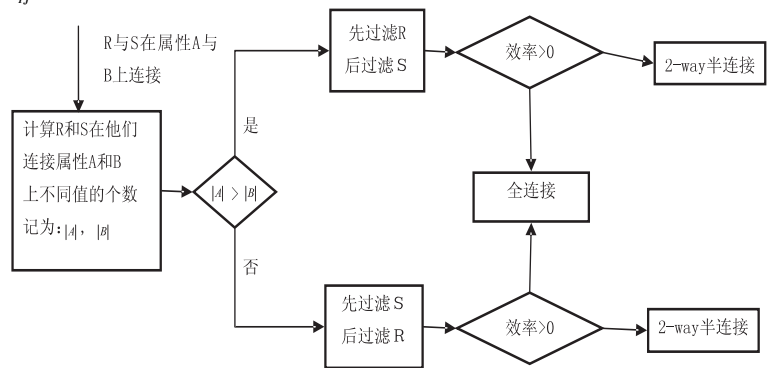


图1 算法流程图

3 实验结果比较分析

用两个分别装有 Sqlserver2005 和 oracle 10g 的数据库服务主机,一个主控主机和一个执行连接操作的

计算主机建立一个小型的网格查询系统测试环境。其中两个数据库主机记作: node1 和 node2。执行连接处理的主机(记作: node3)对外提供连接处理的访问接口,分别在 node1 和 node2 里创建两个关系 R 和 S 。 R 和 S 分别包含两个字段 X 和 $Other$, X 是连接属性,10 个字节; $Other$ 是 50 个字节的数据。node1, node2 和 node3 之间的网络带宽都是 100M/s。对上文中基于网格服务的半连接算法和全连接的算法分别进行下面两种情况的实验测试来分析比较它们的查询响应时间:

(1) 保持 node1 与 node2 之间和 node2 与 node3 之间的网络负载为 50% 不变,使 node2 与 node3 之间的网络负载 NL_{23} 从 10% ~ 90% 之间变化。结果如图 2 所示。

(2) 保持 node1 与 node2 之间和 node2 与 node3 之间的网络负载为 50% 不变,使 node1 与 node3 之间的网络负载 NL_{13} 从 10% ~ 90% 之间变化。结果如图 2 所示:

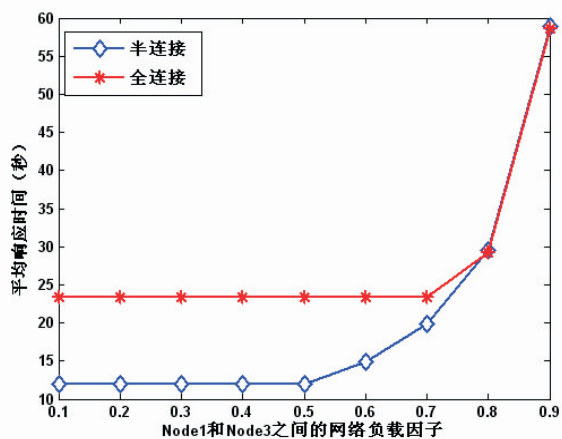
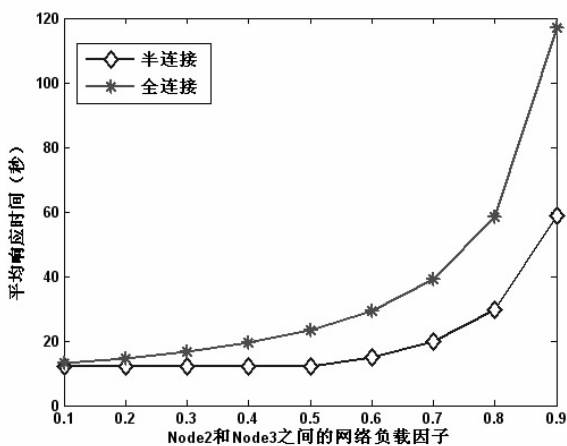


图 2 平均响应时间对比图

由上面的两个实验结果可以得出,一般情况下,在动态变化的网格查询系统中采用半连接算法可以大大减小不同网格节点之间的通信代价,缩短查询响应时间,从而提高查询效率。

4 结束语

文中提出了一种新的适应网格环境的 2-way 半连接查询优化算法。实验结果表明了该算法的有效性和高效性。但是有许多问题仍然需要进一步研究。比如说:查询优化器可以通过在进行查询优化所花费的时间和执行所花费的时间之间保持一个平衡来改善查询性能。

参考文献:

- [1] Zheng Xiaoqing, Chen Huajun, Wu Zhaohui, et al. Query Optimization in Database Grid [C]//LNCS 3795; Proc of the 4th Int Conf on Grid and Cooperative Computing. Berlin Heidelberg: Springer-Verlag, 2005: 485-497.
- [2] Luo Yonghong, Chen Tefang, Zhang Yousheng. Distributed query optimization model based on data grid [J]. Journal of Computer Applications, 2008, 28(10): 2553-2557.
- [3] 刘川, 张海齐, 邓鹏飞. 一种基于分布式、异构数据库统一视图的查询优化方法 [P]. 中国专利: CN102163195A, 2011-08-24.
- [4] 邵佩英. 分布式数据库系统及应用 [M]. 第 2 版. 北京: 科学出版社, 2005: 69-92.
- [5] Mokadem R, Hameurlain A, Morvan F. Performance improving of semi-join based join operation through algebraic signatures [C]//Proc of the 2008 IEEE Int Symp on Parallel and Distributed Processing with Applications. NJ: IEEE, 2008: 431-438.
- [6] 于秀霞. 分布式数据库半连接查询优化算法的研究 [C]//International Conference on Broadcast Technology and Multimedia Communication (BTMC 2010). 中国重庆: 出版者不详, 2010: 110-115.
- [7] 高波, 郭常杰, 蒋忠波, 等. 数据库系统和跨数据库查询优化方法 [P]. 中国专利: CN102053975A, 2011-05-11.
- [8] Hu Naijing, Luo Yuan, Wang Yingying. Adaptive evolution of query plan based on low cost in dynamic grid database [C]//Proc of the 9th ACIS Int Conf on Software Engineering. NJ: IEEE, 2008: 411-415.
- [9] Sriprayoonsakul S, Uthayopas P, Jysoo L, et al. Interfacing SC-MSWeb with Condor-G - a joint PRAGMA-Condor effort [C]//Proc of the 4th IEEE Int Conf on eScience. NJ: IEEE, 2008: 570-575.
- [10] De O, Celso H P, Amaral M A. Optimizing database resources on the grid [C]//Proc of the 2008 Int Conf on Grid Computing and Applications. USA: CSREA, 2008: 208-212.
- [11] 张朝阳, 宁洪, 王挺, 等. 多数据库系统查询计划优化研究 [C]//2009 年研究生学术交流会通信与信息技术论文集. 北京: 国防工业出版社, 2009: 70-75.
- [12] 万旺根, 周演飞, 余小清. 一种基于改进型 SDD-1 算法的数据库多连接查询优化方法 [P]. 中国专利: CN102110158A, 2011-06-29.

基于网格服务的半连接查询优化算法研究

作者：[温慧明](#)，[宫晓辉](#)，[焦洋](#)
作者单位：[温慧明, 焦洋 \(煤炭科学研究总院, 北京 100013\)](#)，[宫晓辉 \(中国电力科学研究院, 北京 100096\)](#)
刊名：[计算机技术与发展](#)
英文刊名：[Computer Technology and Development](#)
年，卷(期)：2012 (9)

本文链接：http://d.g.wanfangdata.com.cn/Periodical_wjtz201209033.aspx