

网格环境下的数据库访问和集成

石 柯

(华中科技大学计算机科学与技术学院, 武汉 430074)

摘 要: 为了集成网格环境中的数据库资源, 促进网格应用支持现有数据库的访问, 提出一种基于服务的数据库访问和集成系统(GridDBAdmin)。GridDBAdmin 为用户提供了虚拟的全局逻辑数据库视图, 支持用户使用现有的 SQL 语言同时访问多个数据库。系统由元数据服务和网格虚拟数据库服务构成。其中元数据服务负责发现含有用户所需数据的数据库, 网格虚拟数据库服务提供全局逻辑视图, 通过分布式查询机制将用户的 SQL 请求分解到具体的数据库中并进行结果合并。对基于 Globus 和 OGSA-DAI 工具包开发的原型系统进行了测试, 得到了较好的结果。

关键词: 数据网格; 数据库访问与集成; 服务; 元数据; 分布式查询

Database Access and Integration on the Grid

SHI Ke

(School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074)

【Abstract】 To integrate existing database resources into the grid and encourage wider use of databases in a grid setting, a service-based database access and integration system, GridDBAdmin, is presented. GridDBAdmin provides a virtual global logical database view for the users and enables grid application to query existing multiple databases using SQL. The system is composed with MetaData Service(MDS) and Grid Virtual Database Service(GVDS). MDS searches the metadata catalog to find the physical databases that contain the required data. GVDS delegates global logical database and converts SQL query into search conditions to the physical databases, which uses a distributed query processing scheme to join multiple physical databases. Performance evaluation based on the response time shows the efficiency of the method. The system is fully implemented on the top of the Globus toolkit and OGSA-DAI.

【Key words】 data grid; database access and integration; service; metadata; distributed query

1 概述

近年来, 随着数据密集型应用的飞速发展, 应用数据呈爆炸式的增长, 现有的数据管理体系结构、方法和技术已经不能满足高性能、大容量分布存储和处理的需求。数据网格(data grid)^[1]就是根据这种需求提出的一种新的数据管理和存储架构, 是网格技术在数据管理方面的应用和实现。它将地理上分布、异构的多种数据资源, 通过高速互联网络连接并集成起来, 屏蔽底层异构的物理资源, 形成单一的逻辑视图, 实现资源共享和协同工作, 为用户提供虚拟的数据访问、数据存储、数据管理和处理环境。但是目前已有的比较成熟的数据网格都侧重于文件级别的数据访问, 如SRB^[2]。

目前大量应用仍然依赖数据库系统实现相应的数据存储、访问和组织, 大量各种各样的数据以不同的格式存储在不同的数据库中。为了促进数据网格的发展和普及, 网格应用必须支持使用数据库技术进行数据的存储、访问、组织等一系列活动, 而且这些活动必须遵循网格的标准, 为用户提供统一、方便、透明的访问接口。

现有的网格环境下支持异构数据库访问的系统如OGSA-DAI^[3]、Spitfire^[4]等, 仅仅提供了符合OGSA规范的基本服务, 不支持复杂的数据库查询命令。商用方面Oracle推出的Oracle 10g和IBM推出的DB2都宣称可以支持网格应用, 但其实都只是提供了一个原有数据库接口的网格包装, 还不能达到支持数据的广域分布和访问。

为了实现数据库系统的网格化, 本文提出了一种基于服务的数据库访问和集成系统(GridDBAdmin)。

2 GridDBAdmin 的系统结构

GridDBAdmin 系统基于 Globus 和 OGSA-DAI 提供的服务接口, 实现了对网格环境下数据库资源的统一管理, 屏蔽了底层物理数据库的分布性和异构性, 为用户提供了一个虚拟的逻辑数据库视图, 实现了对数据库资源的透明访问。OGSA-DAI 提供了网格环境下数据库资源的分布透明性和异构透明性, GridDBAdmin 系统在此基础上, 进一步解决了数据库资源的命名透明性和并行查询透明性等更高层次的要求, 实现了元数据管理、服务发现、数据连接、分布式查询处理等功能。

GridDBAdmin 采用 Web Service 体系结构, 以服务的方式向 OGSA 中的其他网格服务或上层应用提供统一的服务接口。GdbAdmin 的逻辑结构分为 2 层: 第 1 层是核心层, 基于 OGSA-DAI 提供的 GDSR(Grid Data Service Registry)、GDSF(Grid Data Service Factory)和 GDS(Grid Data Service)服务接口, 提供了 2 种类型的服务——网格虚拟数据库服务和

基金项目: 国家自然科学基金资助项目(60125208)

作者简介: 石 柯(1973 -), 男, 副教授、博士, 主研方向: 并行分布式处理, 普适计算

收稿日期: 2007-04-30 **E-mail:** keshi@mail.hust.edu.cn

元数据服务；第 2 层是面向用户提供的逻辑数据库服务使用界面和接口。其中网格虚拟数据库服务(Grid Virtual Database Service, GVDS)主要提供对数据库资源的访问、优化和调度，集成分布在异构数据库资源上的数据，实现对数据的统一操作；元数据服务(MetaData Service, MDS)为系统提供全局资源的信息服务，提供元数据管理和数据库服务发现等功能，其结构如图 1 所示。

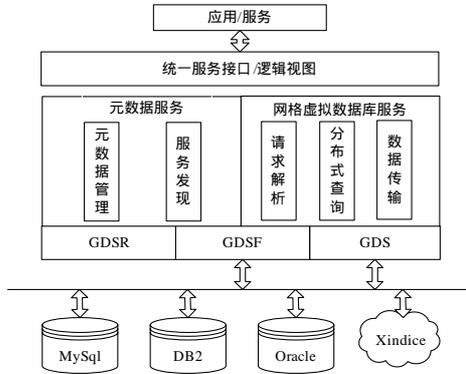


图 1 GridDBAdmin 系统结构

GVDS 为用户提供对数据库资源的访问和管理功能。当用户提交请求时，GVDS 对用户的请求进行解析并转发给 MDS，调度到合适的 GDSF，访问该 GDSF，创建一个能够满足用户请求的 GDS，为用户提供各种数据库操作和管理服务。GVDS 获取数据以后，采用 SOAP 协议直接向客户端发送数据。

目前 GridDBAdmin 系统能够集成网格环境下分布、异构的各种数据库资源，例如 MySQL, DB2, Oracle 等关系数据库系统和 Xindice 等 XML 数据库系统，为用户提供一个统一的数据访问和管理的接口。通过 GdbAdmin 提供的接口，异构的数据库资源可以被看成是一个统一的逻辑数据库视图，用户使用标准的 SQL 语言就可以实行相应的数据查询和存取。

3 系统工作流程

GridDBAdmin 支持用户采用类似于传统数据库的方法对逻辑数据库进行访问，系统自动地将用户的请求映射到对底层数据库资源的操作，其访问过程如图 2 所示。

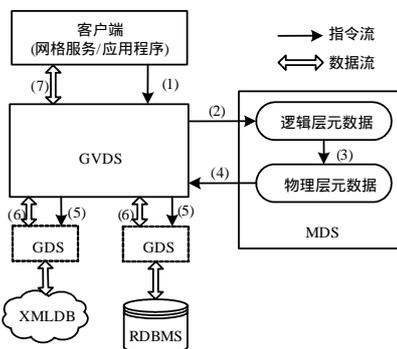


图 2 GdbAdmin 工作流程

在图 2 中，各流程所完成的操作如下所示：

(1)客户端(其他网格服务或上层应用程序)在逻辑数据库视图中选择满足应用需要的逻辑数据库，向 GVDS 提出数据请求。

(2)GVDS 接受用户的请求，并对用户的请求进行解析，解析后发送给 MDS 进行处理。

(3)MDS 根据请求中所引用的逻辑数据库(表)名查询元数据库，获取逻辑数据库元数据值，将其映射到数据库服务元数据。

(4)查找数据库服务元数据，选择能够产生满足用户需求的 GDS 的 GDSF 返回给 GVDS。

(5)GVDS 访问相应的 GDSF，请求产生 GDS，并将用户的请求发送给该 GDS。

(6)GDS 执行用户的请求，将结果返回给 GVDS。

(7)GVDS 收集返回的结果，并对结果进行处理加工后返回给客户端。

4 关键技术

系统的关键技术包括元数据服务和分布式查询，分别解决如何找到用户需要的数据所在的数据库系统和如何将用户的请求转换成具体的数据库查询命令并进行结果合并。

4.1 元数据服务(MDS)

MDS 提供基于元数据的服务发现和选择机制，主要包括元数据管理和元数据服务发现 2 大功能模块。元数据管理采用分层的分布式结构，如图 3 所示。网格中的数据库资源被划分为多个逻辑域，每个域都有一个元数据服务器(Domain MetaData Service, DMS)和本域内的 GDSR 相连，存储和管理本域内元数据，提供本域内的数据库系统的操作。整个 GdbAdmin 系统中有一个中央元数据服务器(Global MetaData Service, GMS)，用来建立到各个域的 DMS 的索引。

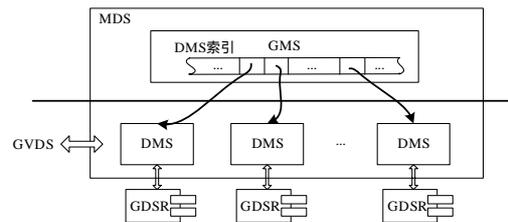


图 3 MDS 结构

中央元数据服务器 GMS 由索引服务模块(Index Service, IS)和元数据库组成，元数据库中存储了索引信息和系统内每个域的元数据服务器的地址，索引信息记录了系统内所有逻辑数据库到域元数据服务器的索引。当某个域的 DMS 不能提供可用的数据服务时，就会将请求转交给 GMS，由 GMS 定位到其他域的 DMS。

每个域的元数据服务器 DMS 由服务发现模块(Service Discovery, SD)、服务日志、域元数据管理模块(Domain MetaData Management, DMM)和元数据库组成。元数据库中存储了域内所有逻辑数据库、数据库资源的元数据。DMM 模块从 GDSR 获得注册到本域内的所有 GDSF 的句柄，访问 GDSF 的服务数据元素，抽取描述数据库资源的信息，存储到本域的元数据库中。服务日志记录了本域内数据库资源提供的数据库服务活动，SD 模块用于从服务日志中发现满足用户需求的服务。

GVDS 可以向系统内的任意一个 DMS 发送服务请求，DMS 在接收到服务请求后，优先搜索服务日志，查找是否有满足用户需求的历史服务记录，如果找到匹配的记录，则直接向 GVDS 返回可用的服务。只有在匹配服务日志失败的情况下，DMM 才会查找元数据库中的元数据。通过搜索元数据库，DMM 可以查到当前域内任意一个或多个可用的数据库服务，并且保证不存在的服务不会被搜索到。如果搜索元数据库仍没有找到可用的服务，DMS 就会将用户的请求转发

给 GVDS, 由 GVDS 选择另一个域的 DMS 为 GVDS 服务。

4.2 分布式查询处理

GVDS 实现网格环境下动态的分布式查询处理过程, 将若干个分布的数据库动态地组合起来完成相应的查询。因此, 查询计划生成程序必须通过元数据服务得到这些数据库的详细信息, 才能生成合适的查询计划。

分布式查询处理模块的结构如图 4 所示。命令分析模块接受用户的 SQL 查询命令, 进行语法检查, 确定用户访问的数据是否可用。该过程需要和 MDS 交互。若用户需要的数据库因自身原因或网络原因变得不可用, 或用户提交的要访问的数据不存在(MDS 没有其元信息)时, 能够及时返回错误信息, 既缩短了相应时间也节约了系统资源。如果该命令执行成功, 会生成一个用户查询命令的内部表达(SQL 语句用一棵查询树表示)给下一个模块。

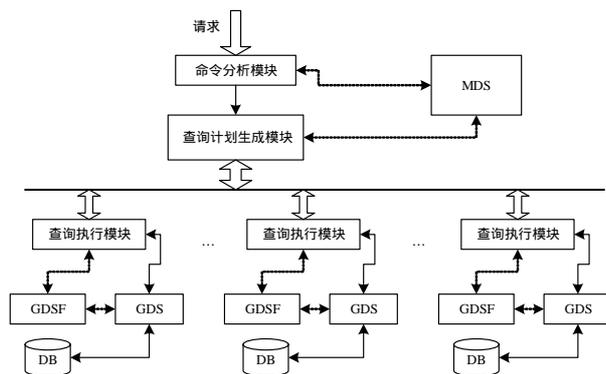


图 4 分布式查询处理模块

查询计划生成模块获取上一步得到的查询命令的内部表达式, 根据一定的优化算法做优化处理, 分解成若干子查询, 然后把得到的子查询发给查询执行模块执行。如果有需要, 负责控制下面的查询执行的同步过程。具体的查询优化算法见文献[5]。

查询执行模块把上一步得到的查询执行计划发送到真正的查询执行节点上, 这些查询执行节点通过 GDSF 创建 GDS 访问数据库得到数据, 互相通信协作完成查询, 并把结果返回到用户。GDSF 是用户在本地配置好的访问本地数据库的 Web Services 接口。上述 3 个模块可以分布在不同的节点机上, 通过标准的 Grid Services 接口互相通信、同步。

5 系统部署与测试

目前基于 Globus 4.0 和 OGSA-DAI 已经实现了 GridDBAdmin 的原型系统, 并部署在 3 个数据中心上, 这些数据中心通过校园网相连, 能提供跨多个文献数据库的检索功能。

笔者首先进行了元数据服务的性能测试, 如图 5 所示。在校园网环境中, 查找元数据确定具体数据库的延迟不超过 2 s。在相同的测试环境下, 采用服务日志的平均响应时间明显小于不采用服务日志的平均响应时间, 用户的查询越集中, 这种趋势越明显。不采用服务日志的平均响应时间总在一个固定的范围内波动, 与用户访问查询热区的比例无关; 采用

服务日志时, 用户访问查询热区的比例越大, 服务日志的命中率越高, 服务日志的置换次数越小, 平均响应的时间越短。这说明 MDS 模型中, 采用服务日志的方法是有效的, 并且日志命中率是影响服务发现效率的重要因素。

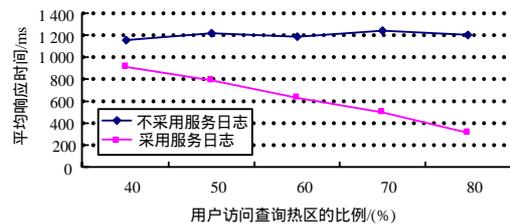


图 5 元数据服务的响应时间

图 6 显示了 GVDS 进行跨多数据库连接查询的处理时间。GVDS 的处理时间超过了 JDBC, 这是因为 GVDS 相应的服务封装带来了较大的系统开销, 但性能仍较好。当条件数量增加时, 处理时间的趋势总体上也在下降, 通过 JDBC 和 GVDS 访问的处理时间的差异也在减小。这是由于具体查询产生的记录数量减少, 从而使数据传输时间减小, 此时的计算时间占处理时间的比例大大增加。

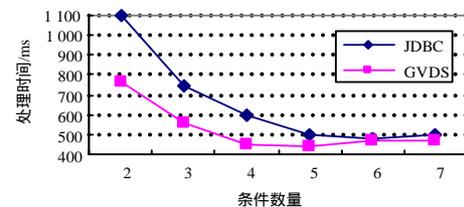


图 6 连接查询处理时间

下一步的研究工作包括 GridDBAdmin 的完善和改进, 如提高连接操作的效率, 在 GVDS 中实现缓存机制以提高用户请求的响应速度, 以及实现一套灵活的数据库访问的授权机制。

参考文献

- [1] Foster I, Vockler J, Wilde M, et al. The Virtual Data Grid: A New Model and Architecture for Data-intensive Collaboration[C]//Proc. of the 15th International Conference on Scientific and Statistical Database Management. Cambridge, USA: [s. n.], 2003.
- [2] Rajasekar A, Wan M, Moore R. MySRB & SRB—Components of a Data Grid[C]//Proc. of International Symposium on High Performance Distributed Computing. San Francisco, USA: [s. n.], 2002: 30.
- [3] Magowan J. A View on Relational Data on the Grid[C]//Proc. of International Parallel and Distributed Processing Symposium. Nice, France: [s. n.], 2003: 90.
- [4] Bell W H, Bosio D. Project Spitfire—Towards Grid Web Service Databases[C]//Proc. of Global Grid Forum 5. Edinburgh, Scotland, UK: [s. n.], 2002: 21.
- [5] 石柯, 林海华. AnyQuery: 网格环境下基于服务的分布式查询处理系统[J]. 小型微型计算机系统, 2006, 27(8): 1503-1505.