

ESB 数据操纵控制器^①

邓庆彬^{1,2}, 廉东本², 李文博², 田月²

¹(中国科学院大学, 北京 100049)

²(中国科学院沈阳计算技术研究所, 沈阳 110171)

摘要: 本文基于企业服务总线, 从多数据库支持、数据操作请求调度、数据操作的安全性、数据库的访问优化四个方面考虑, 对 ESB 系统内部模块对数据库的访问操作进行分析和研究, 提出数据访问控制层, 即 ESB 数据操纵控制器。数据操纵控制器的目的是将 ESB 内的数据操作限制在尽可能小的指令范围内, 这样提高了其数据操作的安全性和高效性。并且数据操纵控制器支持多种数据库的访问和操作, 使得 ESB 系统的产品通用性更强。

关键词: SOA; 企业服务总线; 数据操作; 多数据库支持

Data Manipulation Controller of Enterprise Service Bus

DENG Qing-Bin^{1,2}, LIAN Dong-Ben², LI Wen-Bo², TIAN Yue²

¹(University of Chinese Academy of Sciences, Beijing 100049, China)

²(Shenyang Institute of Computing Technology of Chinese Academy of Sciences, Shenyang 110171, China)

Abstract: This research which is based on the Enterprise service bus considers from multiple databases support, data operation request scheduling, data operation security, database access optimization four sides. On this basis, we analysis and research the mechinsm of communication between the interal modules of the Enterprise service bus and the database, abstract a data access control layer, namely ESB data manipulation controller. Data manipulation controller improves the security of data manipulation by limiting the data manipulation of ESB within a small range of instructions, In addition, data manipulation controller's multi-database support gives product realization of ESB system more flexibility.

Key words: SOA; enterprise service bus; data manipulation; multiple database support

ESB(Enterprise service Bus)即企业服务总线^[1], 为解决企业信息化过程中产生的信息孤岛问题提供了良好的解决方案。企业服务总线是一个轻量级的基础架构, 提供了即插即用的企业功能^[2], 能够对各种技术和应用系统提供支持, 具有很强的灵活性和可扩展性, 是目前企业理想的应用系统集成平台^[3]。企业服务总线提供的功能有: 服务请求者和提供者之间传输协议的转换、服务请求者和提供者之间消息格式的转换、在企业服务总线服务间对消息路由等。

企业服务总线在进行产品化实施时, 可能面临不同的企业和 IT 环境, 需要使用不同的数据库系统, 单一的数据库支持制约了 ESB 的产品化实施的灵活性。

企业服务总线在监听到消息访问时, 将会对消息包进行解析, 这时需要查找数据库中的协议格式, 同时在对消息路由时又要查询数据库中的路由表。由于企业服务总线模块间是多线程并发操作, 大量数据库访问会影响企业服务总线的服务效率和数据库安全, 同时对数据库访问的操作分散, 不利于对数据库的访问操作控制。因此, 将对数据库操作的指令限制在较小的范围内成为提升企业服务总线的数据库访问安全的关键。

针对上面对数据库访问产生的问题, 本文提出企业服务总线数据操纵控制器, 将企业服务总线中数据库访问层对数据库的访问集中到数据操纵控制器中管理, 通过数据操纵控制器的多数据库支持提高企业服务总

① 基金项目: 国家水体污染控制与治理科技重大专项(2012ZX07505)

收稿时间: 2013-08-28; 收到修改稿时间: 2013-09-25

线产品的灵活性,数据操纵控制器的内存数据库访问模式提高企业服务总线的数据库访问速度.

1 数据操纵控制器框架设计

如图 1,数据操纵控制器是一个单独的功能模块,在数据访问模块与数据库之间增加一个缓冲,为企业服务总线系统内其他各功能模块提供数据库的访问和操作,方便对数据库访问的控制,提高系统实现时的灵活性.企业服务总线数据操纵控制器图 2 所示.

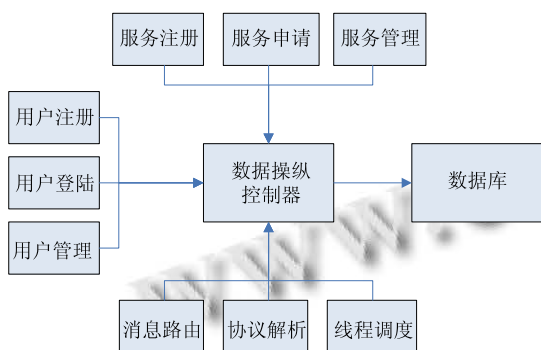


图 1 企业服务总线访问数据库示意图

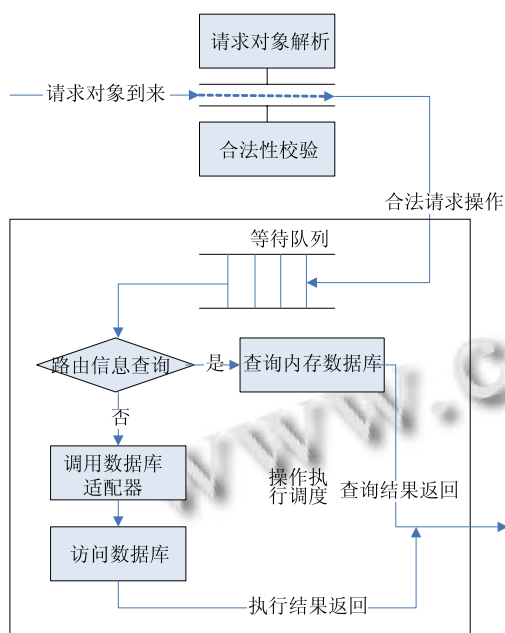


图 2 数据操纵控制器示意图

1.1 数据库支持问题

在进行软件项目的开发设计时,所采用的开发环境和数据库通常都是确定的,并不存在对多种数据库的支持需求.但是,若一个软件以一种产品的形式出

现,需要适应不同的企业和 IT 环境时,多数据库的支持将是非常重要的需求.

数据操纵控制器在企业服务总线的数据库访问层和实际的数据库之间构筑一个缓冲,对数据库访问层提供统一的数据访问操作方式,屏蔽实际的数据库操作接口的差异,对不同的数据库实现不同的适配器.

当企业服务总线的数据库系统出现变化,或者在实现分布式的企业服务总线,在各个企业服务总线分布点采用的数据库系统不相同,不必修改企业服务总线数据库访问层的数据库操作的代码,只需添加对应的数据库的适配器.

1.2 数据操作请求的调度

数据操纵控制器接收到的操作请求都是对数据库的操作,应当使请求的操作尽快执行;当有多个数据库操作请求时,在有限的系统资源条件下,使请求尽快的被响应,需要一个好的调度策略;当两个或多个数据库操作请求的数据操作有冲突时,例如,两个请求对同一个表分别进行读和写操作,数据操纵控制器应该正常运行并保证正确的数据操作结果.

1.3 数据操作的安全性问题

非法的数据操作可能会造成数据的泄露或丢失,甚至对数据库造成损坏,为保证数据操纵控制器可以实现安全的数据操作,对于收到的数据库操作请求,对其进行合法性校验.这可以在操作请求未到达数据库时就先将非法的请求拦截处理掉,减少对数据库的非法访问,提高数据库的安全性.

1.4 数据库的访问优化

企业服务总线的最频繁的操作就是进行消息路由,对收到的请求进行转发或者将收到的响应转发给请求者,消息路由需要查询存储在数据库中的服务注册信息.频繁的消息路由意味着会有对数据库中服务注册信息的大量查询访问,查询数据库需要访问外存,这会增加企业服务总线消息路由时的时间开销,降低企业服务总线对服务请求的响应速度,制约企业服务总线的性能.

内存容量的快速增长对数据库管理系统有着深刻的影响.在某些场合,将整个数据库放进内存是可能的,正常的查询处理可以完全脱离硬盘^[4].现在计算机的内存配置通常在 2-4GB,甚至更大,企业服务总线的路由信息查询是对数据库中服务注册信息的读取,将企业服务总线的服务注册信息表放置在内存中是可行的.

企业服务总线服务注册信息常驻内存, 在进行消息路由时不必再访问硬盘上的数据库信息, 将路由信息获取的时间从对硬盘访问的时间级减小到对内存访问的时间级, 大幅提升企业服务总线对请求的响应速度和处理能力.

2 ESB数据操纵控制器的关键技术

2.1 多数据库的支持

数据操纵控制器通过对实际数据库的操作分为三个阶段来实现多数据库的支持: 构造数据库操作请求对象、对构造的数据库操作请求对象解析、生成对应的实际数据库的执行语句并执行.

如图 3, 与通常的数据库访问方式不同, 这里对数据库的访问并不是直接执行 SQL 语句, 而是由企业服务总线的数据库访问层向数据操纵控制器提交一个数据操作请求对象, 然后数据操纵控制器调用实际使用数据库的适配器等对数据操作请求对象进行解析, 生成实际使用数据库的执行语句并调度执行.

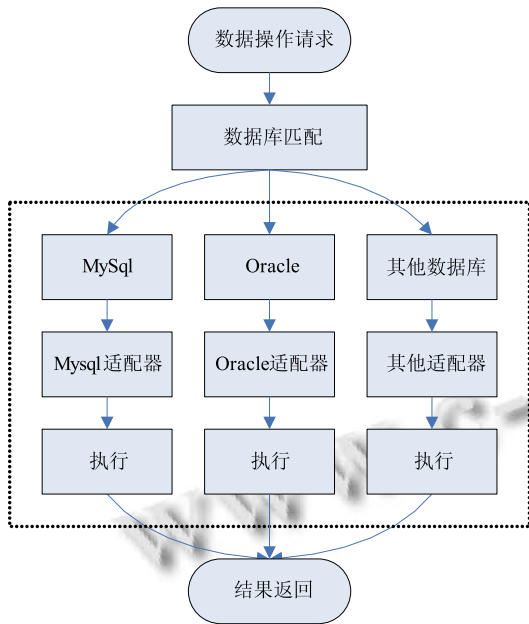


图 3 多数据库支持示意图

对每种数据库实现一个单独的适配器, 适配器提供统一的接口, 根据数据传入的数据操作请求对象, 解析生成符合对应数据库的 SQL 语句. 数据操纵控制器在运行时调用适配器执行数据操作请求, 在增添对新种类数据库的支持时, 增添新数据库的适配器即可, 而不必修改原来的代码.

2.2 数据操纵控制器的调度控制

为了保证多数据操作的完成效率, 数据操纵控制器采用线程池执行数据操作任务. 对于接收到的请求, 设置一个先进先出(FIFO)的等待队列, 新到的请求全部放入等待队列中.

如图 4 所示, 数据操纵控制器的调度执行如下操作: (1)判断等待队列是否为空, 若为空, 则当前没有数据操作请求, 数据操纵控制器没有操作任务可执行, 继续等待操作请求的到来; 若操作队列不空, 取出队首的等待操作, 转入(2)中执行. (2)判断是否有空闲线程, 若有空闲线程, 则转入(3)处理, 否则将其放回等待队列队首等待调度. (3)判断取出的队首等待操作是否和正在执行中的数据操作任务有冲突, 即一个操作的要读(写)的数据是另一个操作正在写(读)的, 也即是

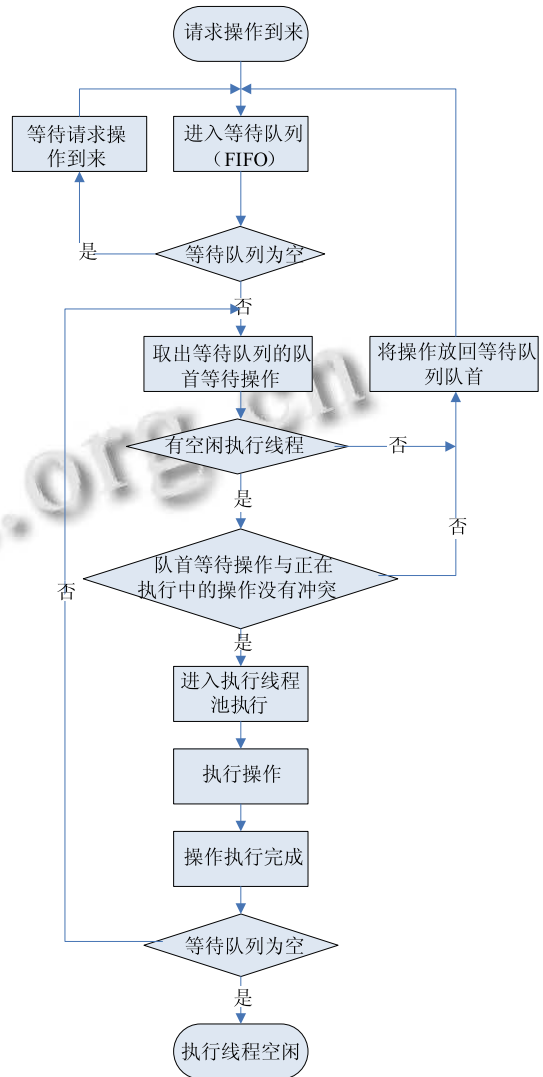


图 4 数据操纵控制器调度流程图

否有数据相关的情况。若没有冲突,则为其分配执行线程执行,否则将其放回等待队列队首等候调度。当有数据操作离开线程池(完成)时,从等待队列中选取队首的操作进行调度。

实行这样的调度策略的原因:(a)数据操纵控制器处理的是对数据库的请求操作,要保持数据访问层原有的对数据库操作的顺序,等待队列为 FIFO,保证先到的数据操作任务先处理。(2)没有数据操作冲突时,处理是多线程并发的,确保请求操作的快速响应和完成。(3)若在没有冲突时将新到的请求放入执行池中,可能会产生等待队列中的任务永远不被调度的情况(线程池一直在满负荷运行),等待队列非空时新请求全部入队,避免产生饿死现象(等待队列中的任务永远不会被调度)。

2.3 数据操作的合法性校验

校验包括两部分:语义校验和权限校验。

语义的校验分析操作请求是否是语义正确的,这部分检查在对数据操作请求对象进行命令解析时进行,若可以正确解析,则认为是语义合法的。

权限校验:根据 ESB 系统各个模块的操作特点的不同,对 ESB 的各个模块赋予角色(ID),针对各个角色提供不同的操作权限(如登陆模块仅具有查询的权限),对于提交的数据操作请求,当其没有语法错误时,进行权限的校验,对越权的操作给予拒绝。

2.4 两种数据库访问模式

数据操纵控制器将企业服务总线数据访问层对数据库的访问分为两种模式:普通模式和内存数据库^[5,6]模式。

2.4.1 普通模式

企业服务总线的用户注册、服务注册、服务申请等功能模块对数据库的访问相对较少,并且这些模块的操作时间相对消息的路由转发要长很多,时间数量级单位为秒,从硬盘访问的时间数量级(毫秒)减小到内存访问的时间数量级(微妙或更小)并不能有效提升用户的使用体验,所以这部分的数据信息仍然通过硬盘数据库访问。

2.4.2 内存数据库模式

将需要频繁访问的企业服务总线的服务注册信息读取到内存并在内存中持久存在,当企业服务总线进行消息路由时,不必再访问硬盘数据库上的服务信息,而是直接查询内存中的服务注册信息,避免了由于访问数据库而引发的硬盘 I/O,将限制企业服务总线消息路由性能的制约条件消除。

数据操纵控制器的两种数据库访问模式的优点:(1)使企业服务总线在消息路由时获得了接近内存数据库的性能:访问频繁的服务信息常驻内存。(2)不会增添太多的实施复杂性:仅服务注册信息存在于内存,访问较少的其他信息仍通过数据库访问,只需要维护内存中服务注册信息和数据库中数据的一致性。

3 结语

本文提出基于 ESB 系统的数据操纵控制器,在数据访问层和数据库之间构建起一个缓冲,将数据操作指令限制在数据操纵控制器的范围内,简化了数据库的访问控制,提高了数据库的安全性;数据操纵控制器对多数据库系统的支持赋予 ESB 系统实现和部署时更高的灵活性,对服务信息访问的内存数据库模式使企业服务总线的消息路由更加快速。

参考文献

- 1 Atkins ME. Choosing the ESB into SOA. 2005.
- 2 罗志伟,汪厚祥,井小沛.基于 ESB 的安全管理模型研究.计算机与数字工程,2011,(6):90-92,97.
- 3 符宁,周兴社,张海辉.基于 JMS 的分布式 ESB 的设计与实现.计算机科学,2007,34(12):118-121.
- 4 王洪海,潘朝华.内存数据库的数据结构分析.现代电子技术,2004,3:96-98.
- 5 李国徽,杨进才.内存数据库查询优化.华中科技大学学报(自然科学版),2003,31(4):21-23.
- 6 谷国栋.内存数据库相关技术的研究与分析.电脑知识与技术,2007,(19):5-6.