

异构系统的信息协同机制的研究^①

Information Collaborative Mechanism of Heterogeneous Systems

王翔宇 刘高嵩 龙军 (中南大学 信息科学与工程学院 湖南 长沙 410075)

摘要: 目前异构系统的信息交互主要是通过 web 服务或 CORBA 技术建立点对点的通信来完成, 由于其系统间功能和结构紧密耦合, 缺少灵活性, 因此很难解决多系统的信息异构和系统异构问题。本文提出一种新的机制, 采用 web 服务来完成系统间的相互连接, 通过自定义消息机制完成异构系统间的信息转换, 以解决信息不一致、信息丢失、信息冗余等问题, 并使用订阅/发布的模式动态组织系统间的逻辑结构, 以解决通信紧密耦合的问题。此机制对解决多异构系统间的信息协同, 多系统间的信息一致性更新, 广域网内的分布式通信和多个异构系统集成等问题都有很好的参考价值。

关键词: 信息协同 消息机制 web 服务 异构系统

1 引言

为了充分利用现有资源, 大型系统之间的整合、大量基础信息的共享与一致性维护已经成为发展趋势^[1], 然而其存在系统架构模式不同, 数据库选择各异, 信息异构, 信息不一致等问题, 另外其业务流程、信息结构和信息内容存在着大量的相同或相似之处, 要实现各个异构系统的无缝整合、信息的充分共享和信息的一致性更新, 必须为多个功能相互独立而业务逻辑比较紧密的异构系统提供一个大任务量的信息交互服务机制。以保证在各个系统独立运行的情况下, 各个系统之间能够提供相关的服务, 并能根据业务需要动态的组织各个系统完成此业务, 当某个系统在信息更新时, 其主动地向对更新信息感兴趣的系统发送信息更新请求, 让其它系统更新相应的信息; 或某个系统主动向其它系统发出信息请求, 搜集其它系统的相关信息, 并在本地集成。本文中称此为异构系统的信息协同机制。

2 信息协同机制的体系结构

广域网上异构系统整合后具有协同性(为了解决某个共同感兴趣的任務, 各个系统通过协作的方式, 完成不同的任务或一个大的任务中的不同阶段)^[2,3]、任

务的随机性(在异构系统之间, 有些任务是随机的, 并且任务量的大小可能是不可预知的)、数据的异构性(异构系统之间的数据源使用各自独立的数据格式、元数据)、数据的分布性(异构系统之间的数据源在不同的系统中, 数据具有分布的特性)等特征, 由于发起任务的随机性、不同任务协同性和信息分布性, 则在组织系统时, 要求这个协同系统的系统结构非常灵活^[4], 能够满足根据一个具体的协同任务来组织其结构。本文借鉴网络拓扑模型, 采用以层次结构为基础的、局部可以动态组织(局部允许出现环)的模型作为异构系统信息协同的逻辑上的基础架构模型, 其拓扑模型见图 1。

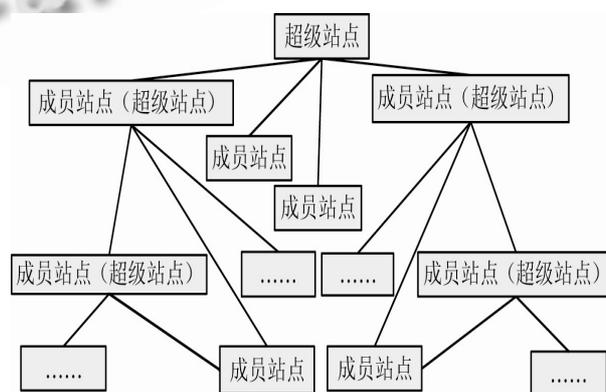


图 1 信息协同逻辑拓扑结构

① 基金项目:国家自然科学基金项目(60873081)

收稿时间:2008-09-28

在图 1 所示的信息协同逻辑拓扑结构中，每一个站点既可作为成员加入到高层的站点中，也可作为超级站点管理下层的成员，从而构成一个多级层次结构；因此，在每个节点保存了自己关注的站点的信息和和其它站点的注册信息。每个超级站点与其直接管理的站点构成一个自治管理域，自治域之间也采用同样的方式相互加入对方，可以使整体体系结构具有较好可扩展性(其具体过程在 3, 4 节中将讲述)。在自治域内采用主动服务的形式，当某个节点信息更新或有数据请求时，主动地向自治域的其他成员发送相关信息，实现广域网信息协同。

3 信息协同机制中异构系统之间结构形成与运作过程

多个异构系统之间以注册的方式，部署成一个以树型结构为基础的可动态调节的拓扑结构，此拓扑结构形成的过程是通过订阅/发布机制完成的，图 1 中的成员节点(即订阅节点)向其它任何一个节点(发布服务器)发送请求，被请求节点记录成员节点的地址和验证信息，并生成信息转换表，指定相应的消息处理代理，并授权允许成员节点访问，此节点就成为了超级节点，成员节点接到注册响应消息,完成注册。

为组织多个异构系统间协同性业务逻辑，采用注册/发布机制来组织其逻辑结构，就会形成一个以发起对象为根节点的树状层次拓扑图。

多个系统间在自治域内进行的通信和信息协同过程是：信息在“发布服务器”上被处理完成后，该服务器生成“提交发布”任务，用主动协同的方式，将信息以消息的形式送至“订阅服务器”；订阅服务器接收到消息后，送入消息队列，根据加权优先队列算法取出消息，取出消息后调用消息处理器处理相应的消息，同时检查本机上是否有订阅该消息的其他订阅者，生成一个“转发消息”的任务，使更新消息发布到次一级的“订阅服务器”，通过这种模式，从而实现异构系统之间的信息协同、信息共享。其实现过程如图 2 所示。

4 基于异步消息和web服务的点对点信息协同机制

根据信息协同机制特征和异构系统信息协同的基础逻辑架构模型，提出任意两站点间的层次构架模型

如图 3 所示。

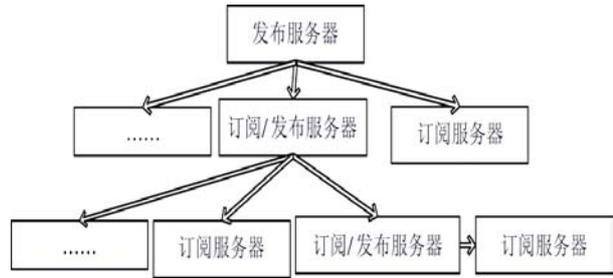


图 2 主动信息协同示意图

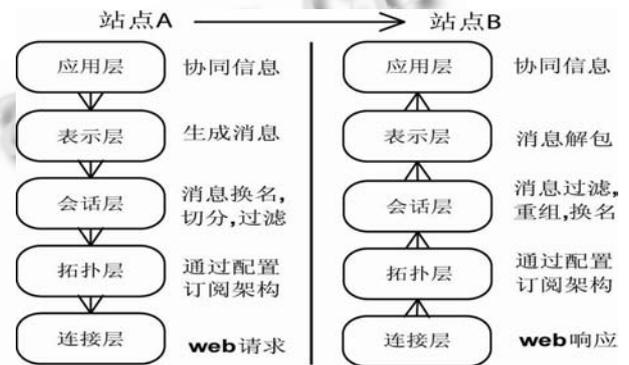


图 3 A 点对 B 点信息协同模型

图 3 所示点对点层次构架模型，借鉴了 Java 中异步消息机制优点，结合协同平台系统的系统异构、信息异构和跨互联网通信的特征，对消息模式进行改造而建立了基于互联网的异构系统之间的点对点层次模型。改进的异步消息机制，在各个站点上发送和接收异步事件消息，并且在消息模型中，各个站点上创建一个消息的队列，用于实现事件消息的存储和转发，从而可以实现事件消息的判断、接收和处理^[9]。Web 服务是服务交互的基本架构，是用来屏蔽各个异构系统的系统异构性。

图 3 中，是一个点对点信息协同过程，由站点 A 完成一个信息更新操作后，检测到 B 有订阅信息，然后向站点 B 发送信息处理消息。

应用层包含的是协同信息的生成机制和内部协同信息处理机制，协同信息的生成机制为：

表示层是将协同信息和协同命令与全局消息的转换，此处是协同信息的打包或解包，是解决异构信息的关键部分。打包时主要是消息转换，就是将本地信息结构转换成一个通用信息结构(各个站点保存一个全局信息格式与本地消息格式的对应表，包含字段与字段类型)，通过映射表来进行把本地消息格式转换

成全局消息格式；解包就是打包的逆过程，通过映射表来进行把全局消息格式转换成本地消息格式。

会话层主要是全局消息转换处理机制，还包含消息的切分和过滤机制，消息的切分是解决全局消息信息量过大的问题，把一个较大的消息转换成一定大小的消息包，并按顺序压入消息队列。消息过滤是对重复的消息进行过滤，防止协同消息的重复处理(在一对多信息协同中用到，将在第 5 节进一步论述)。

拓扑层是一个信息协同机制的逻辑拓扑结构组织层。它是通过多点的注册机制，建立信息协同的自治域，多个自治域之间通过注册连成一个协同体系，在运行期间，可以动态注册，在整个协同平台内就会形成一个可以动态架构的层次平台。

连接层是 web 服务建立两个异构系统连接，屏蔽掉异构系统之间的差别，实现跨平台的透明传输，在连接时需要身份验证，通过对消息标识和认证信息加密、解密来识别协同平台内合法消息，阻止非法消息进入，同时通过对消息摘要签名保证平级自治管理单元之间不能伪造消息，此外为协同机制中指定一个统一的消息的生存周期时间，此外还提供一些其它的服务(提供消息注册接口，对拓扑层构建一个可缩回性的系统平台提供接口支持)。

为了避免建立两个节点连接后，只能等待到消息处理完以后才能建立其它的连接的缺陷，在各个系统的站点上建立自己的消息队列，消息队列用来缓冲内、外部消息，然后根据消息的类别进行相应的处理。消息处理原则是将消息放入一个优先级队列中，给不同类型的消息预先定义一个初始的优先级，为了保证消息不会死等，根据其等待时间提高消息的优先级别。在系统中对消息进行分类，为每一种消息都定义了一个消息处理代理，实现协同消息的处理。因此在每一个消息节点中加入一个消息处理代理。消息的结构图如图 4：

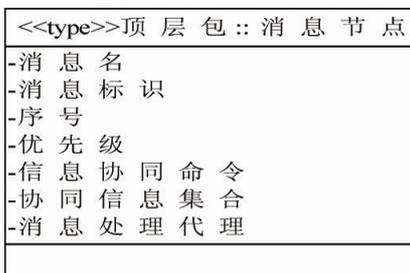


图 4 异步消息结构图

在图 4 中的信息名、消息标识是识别消息时使用，序号是为了消息重组时使用，优先级是消息队列调动时使用，信息协同命令是相应的处理命令(各个异构系统的服务接口)，协同信息是要处理的信息，消息处理代理是用来处理特定消息的处理类。

5 协同机制的消息过滤与恢复策略

为解决在通信中产生信息冗余问题，特别是当信息协同体系结构复杂时，将会出现很多重复消息，会消耗系统资源，如出现循环的转发消息的话，就会造成系统的瘫痪，所有必须要一种较好的策略来解决，下面提供两种策略：

5.1 循环分发检测

在一个有多异构系统组成的协同平台中，如果服务器之间存在循环发布/订阅，那更新消息可能会形成一个回环，极大的消耗系统资源，甚至会造成网络瘫痪。

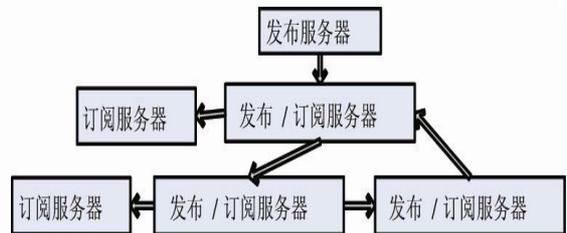


图 5 环形分发路径示意图

在图 5 中发现多台服务之间存在循环订阅/发送关系，通过不断转发到其他订阅服务器，有可能导致信息更新消息在此环上不断被循环转发，为检测此类分发异常，协同机制为自治域中的每一个系统定义一个系统标识，信息经过每一个服务器，均将自身的唯一标识记录在消息中，作为消息的一部分发送给其他订阅服务器；订阅服务器接收到消息时，检测其中是否包含自己的唯一标识，如果存在则表示该信息是通过环形转发路径重复发送来的，因此抛弃该消息。

5.2 重复分发检测

如果一个订阅服务器从多个分发服务器订阅信息，它从两个发布/订阅服务器订阅消息，在此情况下，发布服务器上信息发生更新时，这些更新将沿两条不同路径到达订阅服务器，访问路径无法检测出此类分发的重复消息，如图 6。

为了避免多路径重复分发进行重复处理的情况，为每一个消息都定义消息标识 GUID，根据协同机制

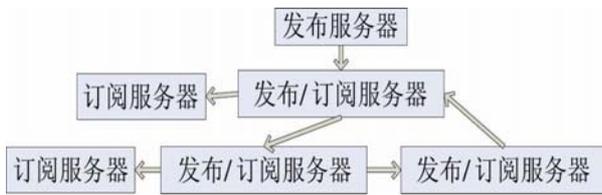


图 6 重复分发路径示意

所组织系统大小，建立一个可以动态配置的缓冲区，当系统启动时维护各自的缓冲区。当生成消息时，本地服务器生成以一个消息标识 GUID 号，而服务器接收到消息时，首先检查其消息标识 GUID 是否存在于缓冲区中，如果存在则表示这是一个重复订阅的信息，并抛弃此消息。通过此种方法，将重复发送的冗余消息降到一个可以接受的水平。

5.3 消息恢复机制

消息恢复机制是为了防止消息在传送时丢失或处理时失败的一种恢复机制，此机制在各个站点建立一个消息日志系统，分别记录本地未完成的协同消息(或命令)和发布时未收到成员站点的回应的消息，记录内容包含消息来源、失败的原因和整个消息。恢复机制还包含一个定期检测线程，定期查看消息日志系统中是否失效消息记录，如果存在记录就恢复该消息，并将其压入消息队列。

6 异构系统的信息协同机制的应用

本文的异构系统信息协同机制，可以应用到广域网信息同步，数据集成，分布式数据搜索。数据集成和分布式数据搜索实现过程是：单结点系统向自治域或相邻自治域发送数据请求命令，通过消息的形式返回请求数据，然后通过消息机制转换成本地的信息格

式，实现数据集成和分布式数据搜索；此机制已用到教育部科技发展中心和部分省份的专家库信息系统中，实现了系统间专家信息一致更新。

7 结束语

本文提出一种异构系统信息协同机制，该机制可以根据自己的业务需要动态调节，非常灵活，并且屏蔽系统之间的异构性，实现异构系统之间跨平台的异步透明通信，解决了信息异构的问题，提供消息检测机制，保证协同平台整体性能。在此基础上可进行复杂业务逻辑架构，实现广域网业务信息同步更新、业务数据集成和协同平台内部的业务数据的分布式搜索，因此具有很好的应用前景。

参考文献

- 1 解吉波,吴华意,龚健雅.基于 XML 的多级异构空间数据库的同步架构.武汉大学学报, 2006,31(5):415-418.
- 2 辛明军,李伟华.基于智能 Agent 的产品协同设计开放式集成开发平台研究.计算机工程, 2002,28(2):59-61.
- 3 史美林,杨光信.一个协同应用开发平台的设计—DICSEX.通信学报, 1999,20(9):63-69.
- 4 Convertino G, Moran T P, Smith B A. Studying activity patterns in CSCW. Conference on Human Factors in Computing Systems, 2006,(5):2339-2344.
- 5 纪波林,王志坚.基于 JMS 体系结构的消息服务技术的应用研究.计算机应用研究, 2003,(11):48-51.