

支持四层结构的电子商务数据管理模型的研究

王金伦¹ 张玉龙² 顾明¹

(清华大学计算机科学与技术系 北京100084)¹ (浙江大学计算机科学与工程系 杭州310027)²

Research on Four-Tiered Data Management Model in E-Commerce

WANG Jin-Lun¹ ZHANG Yu-Long² GU Ming¹

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084)¹

(Department of Computer Science and Engineer, Zhejiang University, Hangzhou 310027)²

Abstract E-Commerce is one of the applications of Internet that have mostly influenced and changed people's life in the whole world, and it is even more popular today. Lots of new technologies, such as Java Server Pages (JSP), Extensible Markup Language (XML) and Workflow, have been implemented in E-Commerce. With more functions provided, the burden of Web server is overwhelming. A four-tiered client/server architecture proposed in this paper was used to reduce the burden of the Web server by adding an application server between Web server and database server. And in this model, mobile Agents were used to "connect" the application server and database server.

Keywords Extensible Markup Language (XML), Four-tiered architecture, Mobile Agent

1. 引言

今天,因特网已经培植成为一个全球市场,众多企业供应商和消费者通过因特网作为介质交换各种各样的数据。因特网不但成为广告载体,而且成为商品和服务集中的全球市场,从电子书到传统出版物,从旅游计划到金融服务。总体上电子商务可以分成两大类:企业-消费者(B2C)商务和企业-企业(B2B)商务。数据管理^[1]主要提供了数据处理功能,例如数据集成、数据建模、数据结构、数据查询、数据传输、数据存储等,成为电子商务的最重要的组件之一。

电子商务应用程序通常运行在动态的分布式环境中,需要处理大量异构信息源。而且由于数据应用侧重点的差异,大多数在线数据源都以半结构形式或者无结构形式提供信息,例如HTML页面。最近,可扩展标记语言(XML)作为因特网数据表示和交换的新标准已经浮出水面,成为推动Web发展的强有力的新技术。应用于电子商务的XML是一种使能技术,使企业相互操作和理解文档,表格和消息成为可能。XML成为加快网络经济和基于因特网全新商业模式实现的关键因素之一。

企业应用XML技术具有以下显著优点:

(1)XML是创建自描述数据的标记规范。计算机和人类容易理解XML的语法,因此企业可以用这种广为理解的方式描述提供的服务。

(2)XML是一个平台,独立于应用程序的数据格式,这样企业可以通过不同的内部商务系统交换文档,表单和消息。

(3)可以减少重新输入数据时出现的错误,因为网关负责对数据进行转换。

(4)能够处理商务处理过程的频繁变化,而不必花费巨大的工程费用。

(5)在保证原有系统投资的基础上,可以同最新的因特网技术结合。

许多大公司在它们的电子商务策略中提供了数据管理机制。它们多数计划采用XML管理电子商务数据。然而我们发现,现在被许多电子商务站点采用的三层客户/服务器体系不能很好地满足电子商务的需求,因为Web服务器的负载越来

越重。在三层客户/服务器体系中,Web服务器不但必须处理客户端的请求,而且必须处理数据管理事务。这样,随着Web服务器功能的增加,Web服务器的负担会越来越重。

2. 体系模型

图1是本文采用的电子商务体系模型。在该模型中,我们采用了新的四层客户/服务器体系,而不是传统的三层客户/服务器体系。该模型与三层体系的区别在于在传统的数据库服务器和Web服务器之间增加了应用服务器。同时应用服务器与数据库服务器之间的通信通过移动代理来实现。

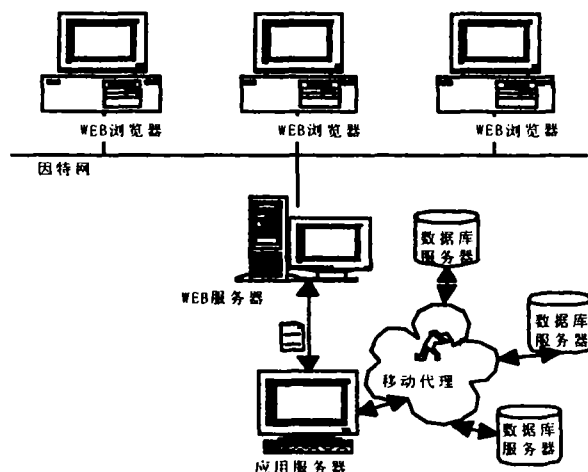


图1 采用的四层体系结构

模型中的Web浏览器是支持可扩展标记语言/可扩展样式语言(XML/XSL)的浏览器。在本层,Web浏览器处理客户的电子商务,是客户的作业平台。浏览器利用电子表格来标准化客户数据的输入。这些输入的数据是原始数据,包括电子支付数据等等。然后浏览器将这些原始数据传递给Web服务器。Web服务器检查这些原始数据,如果发现任何错误,拒绝进行进一步处理。Web服务器根据请求组装这些原始数据,形成XML文档,然后将这些文档传递给应用服务器。应用服务器分解XML文档,调用移动代理将数据存储到数据库。

如果客户想查找数据库中的数据,或者以电子表格的形

式浏览数据,那么将 XML 形式的请求发送给 Web 服务器。Web 服务器将请求发送给应用服务器。应用服务器根据客户请求调用适当的移动代理在数据库中查找所需要的数据。在数据返回以后,应用服务器根据请求重新将数据组装,形成 XML 文档,然后通过 Web 服务器将文档发送给浏览器。这样客户可以通过浏览器看到所需的数据。

现代电子商务的发展已经经历了两个阶段:专用电子商务发展阶段;基于组件的电子商务发展阶段。如今正处于第三个发展阶段。电子商务开发方法的进步有助于提高应用程序开发的生产效率和改进系统的互操作性,扩展性和可靠性。尽管本文采用的四层客户/服务器体系比三层体系更加复杂,但是应用服务器的引入,减轻了 Web 服务器的负载。应用服务器的引入使得系统扩展性更好,可靠性更高,而且增加功能时不会产生性能瓶颈。此外,数据库连接池技术也可以用来减少数据管理的时间。

与传统的客户/服务器分布式系统相比,移动代理支持的分布系统具有7个主要优点:减轻网络负载,克服网络时延,支持协议封装,异步和自治执行,动态适应,自然异构以及健壮性和错误容忍等^[4]。本文的模型中,移动代理为应用服务器提供了透明的运行环境。

3. 技术

3.1 Web 服务器

Web 服务器的主要功能是向 Web 客户提供服务,一方面根据浏览器请求将原始数据组装成为 XML 文档,传递给应用服务器进一步处理,另一方面将应用服务器反馈回的响应 XML 文档传递给浏览器。由原始数据组装成的 XML 文档主要包含两方面的数据:动作和客户输入的原始数据。

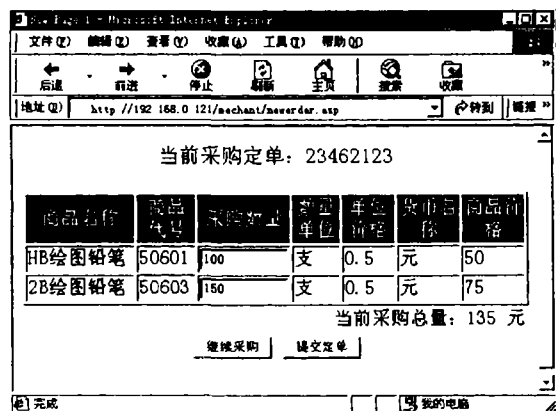


图2 浏览器上订单表格

```
<?xml version="1.0"?>
<action name="neworder" order_id="23462123" op="insert">
<order>
  <customer-id>23</customer-id>
  <row>
    <prod-id>50601</prod-id>
    <prod-num>100</prod-num>
  </row>
  <row>
    <prod-id>50603</prod-id>
    <prod-num>150</prod-num>
  </row>
</order>
</action>
```

图3 Web 服务器产生的 XML 文档

图2是电子商务中典型的关于订单的电子表格。Web 服务器可以利用已经定义的标签用 XML 来标记图2中的数据,

使原始数据结构化^[5]。图3是在订单提交后,Web 服务器产生的 XML 文档。文档中动作的名字是根据客户的动作定义的。应用服务器将根据文档中的动作处理文档中的数据。在应用服务器解析了图3所示的 XML 文档后,产生移动代理,让移动代理将新的订单数据存储到数据库中的订单表。

3.2 应用服务器

在四层电子商务数据管理模型中,应用服务器是核心,起着极其重要的作用。最近中间件技术已经成为 Web 应用的趋势之一。模型中的应用服务器也是一种中间件。

应用服务器主要管理所有的移动代理,同时进行数据事务处理,包括数据集成,数据查询和数据重构。应用服务器成为连接 Web 服务器和数据库服务器的桥梁。应用服务器接收 Web 服务器的 XML 形式的请求,为 Web 服务器提供相应的服务,例如事务处理,安全控制,性能调整等等。应用服务器提供的服务是通过它管理的移动代理实现的。应用服务器只需要解析 Web 服务器的请求,根据应用逻辑调派相应的移动代理,让移动代理处理请求。当移动代理处理完请求后,将处理结果传递给应用服务器。应用服务器将处理结构重新组装,形成 XML 文档传递给 Web 服务器。

这样在模型中,Web 服务器成为同客户交换信息的接口,它只需要负责接受客户请求,组装客户原始数据和请求,形成 XML 文档,或者将包含处理结果的 XML 文档发送给客户,在这个过程进行简单处理。最终的请求处理由应用服务器调派移动代理完成。这样大大地减轻了 Web 服务器的负载,基本上消除了 Web 服务器可能产生的性能瓶颈。

3.3 移动代理

移动代理是一种特殊的代理,一般情况下需要特定的运行上下文运行^[6]。移动代理在运行期间在虚拟机之间迁移而不需要中止程序的执行。该模型中,应用服务器和数据库服务器拥有移动代理运行所需要的运行上下文。在运行上下文中,委托代理可以执行任何合法的动作。

模型中的数据库服务器与三层客户/服务器模型中传统的数据库服务器不同。它主要包含数据库管理系统和移动代理运行需要的上下文。应用服务器调派的移动代理将根据它承担的任务执行所需要的数据库操作。这个方案极大地提高了系统的可扩展性。应用服务器只需要为特定的任务向数据库服务器分派不同的代理。当新类型的任务出现,应用服务器只需要产生新代理,而不需要重新编写数据库服务器代码。每个数据库服务器都有它知道的数据库服务器的列表。当移动代理登陆数据库服务器,在搜索本地数据库的同时,它也读取本地数据库服务器列表,将它与来自应用服务器的列表进行比较。如果代理发现了新数据库服务器地址,它将向列表中添加这个服务器。然而从应用服务器的角度来说,它或许不知道某些数据库服务器,但是它能够访问这些服务器上的数据。

移动代理技术为分布式数据管理提供了更强大和更灵活的解决方案。下面是模型中两种主要代理 FunAgent 和 FunMAgent 的简要介绍。其中 FunMAgent 负责管理 FunAgent。当然,如果需要更多的服务,我们可以很容易地向系统增加新的移动代理,而且只需要修改现存系统的少量代码。

FunAgent 代理主要根据应用服务器的请求完成数据库的数据操作,由 FunMAgent 来调派。当它到达目的数据库服务器,首先搜索本地数据服务器列表。如果发现新的数据库服务器,它将把这个服务器加到搜索路由中供以后访问。当结束了对本地数据库的访问,FunAgent 将查找结果通过消息发

送给调派它的 FunMAgent。如果在它的搜索路由中还有别的数据库服务器,它将自己分派到下一个数据库。当代理搜索完所有路由,它将自己和发现的数据分派给应用服务器,在到达应用服务器后释放自己。在大多数时间里, FunMAgent 驻留在应用服务器上,好像静态代理。它产生并调遣与任务相关的代理(也就是 FunAgents)到目的地,然后通过消息同这些代理进行通信,管理任务的处理过程。它负责判断任务是否结束或者失败。它也通知外部的 FunAgent 继续进行来尝试结束中断的任务。图4是 FunMAgent 和 FunAgent 的关系图。

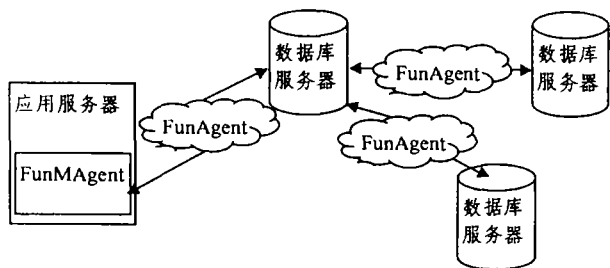


图4 FunMAgent 和 FunAgent 关系图

结论 本文中,我们综述了四层客户/服务器体系结构以及用到的技术。由于这仅仅是个模型,所以我们没有讨论数据管理中的数据安全问题,尽管数据安全在电子商务中是极为重要的。将来电子商务中还会有许多新技术得到实现。Web 应用的趋势之一是使用中间件来提高系统的可扩展性和效率。模型中的应用服务器也是一种中间件。

参考文献

- 1 Lee M ,Rigney S . Data Management & Storage Techonogy . the Survivor's Gide to 2001 ,2000 ,11:33~35
- 2 Widom J . Data Management for XML. IEEE Data Engineering Bulletin , Special Issue on XML ,1999 ,22(3):44~52
- 3 http: // www. soc. staffs. ac. uk/~ cmryl/tefpapers/ecommerce. pdf
- 4 Lange D B,Oshima M . Serven Good Reasons for Mobile Agents. Communications of the ACM,1999 ,42(3):88~89
- 5 Harold E R . XML Bible . IDG Books Worldwide ,Inc . 1999. 60 ~70
- 6 Aridor Y ,Oshima M . Infrastructure for Mobile Agents : Re-quirements and Design . Mobile Agents-Second International workshop,MA98,Lectures Notes in Computer Science ,1998,14 (37):36~49

(上接第56页)

范,并且受到 IBM、Oracle、BEA 等一大批公司和组织的支持。遵循 EJB 规范使得应用程序开发人员不必了解低层次的事务和状态管理的细节、多线程、资源共享和其他复杂的低级 API,从而专注于业务逻辑的开发,轻松构建分布式的应用程序。基于 J2EE/EJB 的模型由于采用了纯 Java 技术,因此它比 COM+模型更有吸引力的一点是它的跨平台性,即“一次编写,随处运行”。但是也应该看到,由于 java 采用虚拟机机制,实现同样 GIS 功能时,javaBean 的效率比不上 COM+对象。基于 J2EE/EJB 的 WebGIS 的实现框架与基于 COM+的实现框架类似,如图4所示。

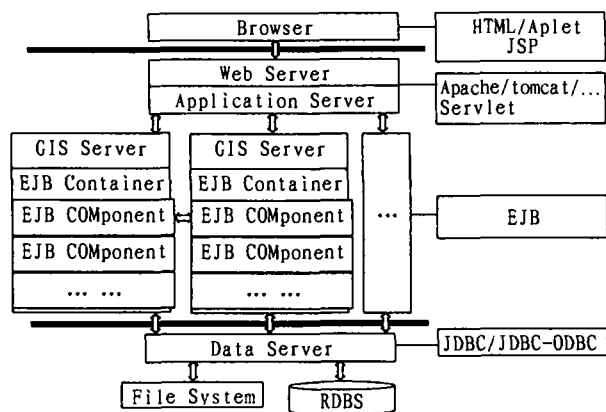


图4 基于 J2EE/EJB 的 WebGIS 实现框架

3. 应用实例

基于 MS/COM+ 的 WebGIS 实现框架,实现了某网站的地图发布、信息查询、最短路径分析、缓冲区分析、专题图制作等功能,图5展示了其逻辑框图。

对网站进行了功能和性能测试,特别是并发用户访问测试,是企业级 WebGIS 系统发布前必不可少的环节。使用商用软件 WebPerformance2.0对网站进行了性能测试,结果显示,由于采用了对象池、即时激活、数据缓存等技术,系统支持400个用户并发访问以及4000用户以每1秒一次的速率持续10分

钟点击放大操作。

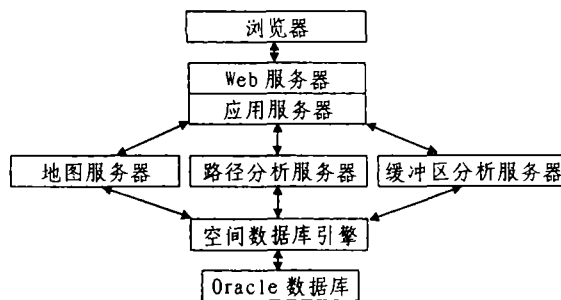


图5 应用实例的逻辑框图

展望 西方经济学界有一个著名的“龙卷风”理论,如果企业把核心技术和业务置于业界发展的中心,就很容易随着龙卷风的抬升而升值;反之,付出再大,也难以打开局面。WebGIS 的发展与计算机和互联网技术的发展密切相关,遵循业已成为国际标准的规范,或者向潜在的规范靠拢,就好比置身龙卷风的中心,会随着这些相关技术的发展自然而然地获得发展甚至突破。可以预计,OpenGIS 规范、XML 规范、各种分布式对象规范以及 Windows.NET 体系结构,将有力推动 WebGIS 发展。

参考文献

- 1 Schell D. Demonstrating the Potential of Open Technology. Open GIS Consortium, Inc.
- 2 The OpenGIS™ Abstract Specification Topic 12: OpenGIS (tm) Service Architecture Version4. 1. Open GIS Consortium, Inc.
- 3 OGC Vision, Mission, Goals, Values. Open GIS Consortium, Inc.
- 4 OpenGIS Web Map Server Interface Implementation Specification Revision 1. 0. Open GIS Consortium, Inc.
- 5 深入理解 COM+ David S. Platt. 清华大学出版社,2000
- 6 边学工,胡瑞敏,喻占武,李德仁. 分布式 GIS 分层体系结构模型的研究. 武汉测绘科技大学学报,2000,25(5):443~447
- 7 袁相儒,龚健雅,林琿,陈莉丽. 异构地理信息处理环境互操作的 Internet GIS 方法. 武汉测绘科技大学学报,2000,24(3):194~198
- 8 韩海洋,黄建明,龚健雅,袁相儒. Internet 下多数据源、超媒体空间信息的分布式调度与管理. 武汉测绘科技大学学报,1999,24 (3):204~208