

李大农 董 慧

数字图书馆软件开发平台的一种理想选择^{*}

摘 要 J2EE 为数字图书馆的软件开发提供了一种理想平台,为广泛开展数字图书馆实验系统的研究、最终汇集为成熟的系统创造了基本条件。它最大的好处是可以充分利用成熟的工业化标准,暂时避开数字图书馆建设中许多复杂的标准问题。参考文献 7。

关键词 数字图书馆 软件平台 J2EE 组件

分类号 G250.76

ABSTRACT In this paper, the authors introduce J2EE, which is an ideal platform for software development of digital libraries and can provide basic environments of the development of pilot and real digital libraries. Among the many advantages, it can fully utilize mature industrial standards and avoid many complicated issues of standards. 7 refs.

KEY WORDS Digital library. Software platform. J2EE. Component.

CLASS NUMBER G250.76

1 数字图书馆软件开发中的平台选择问题

选择软件开发平台要能适应现在的需求,便于将来的扩展。为了实现网络化、社会化、国际化,软件开发平台必须能支持一定的标准。

标准问题涉及许多方面。中国数字图书馆标准化工程提出了包括系统硬件平台、软件平台、元数据、数据质量、资源建设、与图书馆自动化管理系统的兼容性等 12 个方面的内容^[1]。每个方面都有其

复杂性。目前的实际情况与“统一标准”相差甚远。仅就元数据标准而言,已经有很多理论和实践方面的研究,但实践中仍然未能统一。为了真正实现元数据的互操作,在内容结构、语法表示、语义定义、标识格式、转换方式方面还有许多细节有待澄清^[2]。实际工作中标准问题是很难处理的。它与软件的生命力紧密关联,软件开发者不能不慎重对待。标准问题是制约数字图书馆发展的难题之一。

软件平台的选择还应考虑到便于同行之间的分

服务与活动一项在不同的数字图书馆中区别很大,所以不做细致划分,具体的下一级单位的统计内容由相关专家进行设计。

3 结束语

数字图书馆标准体系的制定是一个庞大的工程,涉及到各个方面的协调统一。统计标准只是这个庞大体系的一部分,标准的制定要做到不断完善以及与其他标准的协调还需要做出更大的努力。

我国的数字图书馆建设已经全面展开,但是很多方面的相对落后状况应当引起我们的重视。标准的制定就是很重要的一方面。我们提出这个关于数字图书馆统计标准的初步设想,希望与大家共同探

讨,研究、制定出更加具体、适用的统计标准,为我国数字图书馆建设做出贡献。

参考文献

- 1 ANSI/NISO Z39.7 - 1995 Library Statistics.
- 2 台湾图书馆统计标准. CNS13151
- 3 蔡丽萍. 图书馆统计工作刍议. 河南图书馆学刊, 1997, 17(3)
- 4 李春田. 标准化概论. 北京:中国人民大学出版社, 1982

黄 奇 南京大学信息资源管理南京研究基地主任、副教授,南京大学学科建设办公室副主任。通讯地址:南京大学。邮编 210093。

邵艳丽 李 伟 南京大学信息管理系硕士研究生。通讯地址同上。

(来稿时间:2002-12-10)

* 本文系国家自然科学基金项目“数字图书馆相关关键技术研究”(批准号:00BTQ004)的成果。

工合作与竞争。理想的竞争不是让用户对复杂、完整的系统作整体取舍,而是能互相取长补短,以便最终组合、开发出更好的系统。

2 J2EE 简介

J2EE(Java 2 Platform Enterprise Edition)是开发分布式和以服务器为主的企业软件的一种统一的发展平台。J2EE定义了一套标准来简化N层企业级应用的开发,与传统的互联网应用程序模型相比有着不可比拟的优势^[3~4]。J2EE基于组件技术,以Java2平台的标准版为基础,同时提供了对EJB(Enterprise Java Beans)、JSP(Java Server Pages)、XML等技术的全面支持。J2EE平台由一整套服务、应用程序接口(API)和协议构成。它降低了开发基于Web的多层应用的费用和复杂性,同时提供集成现有应用程序的能力。有很多著名软件厂商支持J2EE规范,并相应地推出支持该规范的应用服务器。可以预计,J2EE将成为商业开发的热点。

2.1 J2EE平台中的主要技术组件

组件是通过相关的类和文件组装成的具有独立功能的软件单元。J2EE平台包含了大量的组件技术,主要有4种。

(1) Applet是包含在网页里的Java应用程序,通常放在Web服务器中。Applet在客户端浏览器中的Java虚拟机(JVM)上运行。

(2) Servlet是一个Java类,在服务器端运行。它扩展了Web服务器的功能,可以动态地生成Web页面。与CGI Perl脚本类似,作为一种服务器端的应用,它可以被来自Web客户机的请求激活,然后执行。它比传统的CGI效率更高,功能更强大,可移植性更好。Java Servlet对每个请求用一个轻量级的Java线程处理,减少了启动进程所需要的时间:如果有N个并发的对同一程序的请求,处理请求的是N个线程,只需要1份Servlet类代码,而不是像CGI程序的代码那样在内存中重复装载N次。它能够直接和Web服务器交互,并支持能缓冲计算结果,在各个程序之间共享数据、保持数据库连接等功能。

(3) JSP(JavaServer Pages)是一种实现普通静态HTML和动态HTML混合编码的技术。Web页面往往大部分是静态HTML,动态内容只在页面中有限的几个部分出现。但是Servlet等技术总是通过程

序生成整个页面,对静态HTML也要一行一行地写输出代码。JSP可以分别用HTML代码和嵌入其中的Java代码创建这两个部分。当JSP页面被客户端请求访问时,服务器对Java代码进行处理,然后将生成的HTML页面返回给客户端的浏览器。显然,使用JSP中编写静态HTML代码比使用Servlet等技术更方便,而且便于分工完成页面制作中不同性质的任务,例如静态页面设计和用Servlet程序完成的动态内容。JSP技术与微软的Active Server Pages类似,但JSP的动态部分用Java编写,功能更强大,更易于使用,而且便于移植到其他操作系统和非微软的Web服务器上。

(4) EJB(Enterprise JavaBeans)组件是基于标准分布式对象技术、CORBA和RMI的服务器端Java组件。EJB技术把Java组件的概念从客户机域扩展到了服务器域。EJB组件总是分布式的,它依赖于一些基础的Java服务,但使用EJB组件并不需要掌握相关技术的知识。EJB技术的解决方案是将商务逻辑集成到一套EJB组件中。EJB组件可以供不同类型的客户端访问,也可以访问多个异构数据源。EJB组件控制和管理对后台数据库的访问。EJB技术能自动处理复杂的共享数据同步问题,保证数据的一致性与正确性,减少编程工作量。EJB技术还可以提供很好的安全性能,例如可以把用户的ID安全地从Web资源传递到EJB组件,可以授予或禁止特定用户对任何EJB组件或方法的操作权。总之,EJB技术简化了中间件组件的开发,为构建强健的、具有高伸缩性、高利用率的系统提供了骨架。

还有其他一些用Java语言编写的应用程序。

2.2 J2EE平台中的组件容器

J2EE服务器以容器的形式为所有类型的组件提供公用的后台服务,解决安全性、资源共享、持续运行、并行处理、事务管理等复杂问题。有以下7种容器。

(1) Applet容器运行在客户端机器上,支持J2EE应用程序中的Applet组件在Web浏览器中运行。

(2) 应用程序客户机容器运行在J2EE服务器上,管理应用程序客户机组件的运行。

(3) Web容器运行在J2EE服务器上,管理J2EE应用程序中JSP组件和Servlet组件的运行。

(4) EJB 容器运行在 J2EE 服务器上,管理 J2EE 应用程序中 EJB 组件运行。

2.3 J2EE 的核心 API

以下是 J2EE 中几种主要的 API。

(1) HTTP 是一个属于应用层的面向对象的协议,适用于分布式超媒体信息系统。支持客户服务器模式,具有简单、灵活、通信速度快的特点。

(2) JDBC 对开发者屏蔽了一些细节问题,为访问不同的数据库提供统一途径,使开发者可以用标准的结构化查询语言(SQL)实现对数据库的访问。JDBC 也具有平台无关性,即低层数据库改用另一种产品时,JDBC 应用代码不需作任何改动。

(3) Java Naming and Directory Interface (JNDI) 为访问不同类型的命名和目录服务提供了一致的模型,用于注册、查找业务组件和本地文件系统或应用服务器中的对象。

(4) IIOP 是调用远程对象的协议。它使用序列化方式传递数据,可以同用任何一种与 CORBA 兼容的编程语言开发的客户端进行通信。

(5) Java 接口定义语言(Java IDL)是将 CORBA 和 Java 集成在一起的服务,使用标准的 IDL 实现。该项服务提供了将新的应用和旧的系统集成在一起的一种途径。

(6) JMS 是用于和面向消息的中间件(MOM)相互通信的应用程序接口。它支持点对点和发布、订阅机制。通过在 Java 中使用 MOM,JMS 也做到了独立于供应商。JMS 提供了将新的应用与旧的系统集成成的另一种途径。

(7) Java 事务管理(JTA)用于指定组成一个事务中所有方法间的关系。它可以使容器与事务管理器交互,当客户端激活一个应用程序中的方法时,容器介入并管理事务。JTA 也可以使事务管理器与资源管理器交互。

2.4 J2EE 的优势

(1) 保留现存的 IT 资产。一个以渐进的(而不是激进的,全盘否定的)方式建立在已有系统之上的服务器端平台机制是公司所需求。这种机制便于利用已有的企业信息系统方面的投资,而不是重新制定全盘方案。由于基于 J2EE 平台的产品几乎能够在任何操作系统和硬件配置上运行,现有的操作系

统和硬件也能被保留使用。

(2) 高效的开发。J2EE 允许公司把一些通用的、很烦琐的服务端任务交给中间件供应商去完成。这样开发人员可以在如何创建商业逻辑上集中精力,相应地缩短了开发时间。高级中间件供应商提供状态管理、数据库访问、分布式共享数据对象 CACHE 等复杂的中间件服务。

(3) 支持异构环境。J2EE 能够开发部署在异构环境中的可移植程序。基于 J2EE 的应用程序不依赖任何特定操作系统、中间件、硬件。J2EE 标准也允许客户订购与 J2EE 兼容的第三方的现成的组件,把它们部署到异构环境中,节省了由自己制订整个方案所需的费用。

(4) 可伸缩性。基于 J2EE 平台的应用程序可被部署到各种操作系统上。J2EE 领域的供应商提供了广泛的负载平衡策略。能消除系统中的瓶颈,允许多台服务器集成部署,实现高度可伸缩的系统,以满足未来应用的需要。

(5) 稳定的可用性。J2EE 部署到可靠的操作环境中,支持长期的可用性^[5]。

3 J2EE 是数字图书馆软件开发的一种理想平台

以 J2EE 作为数字图书馆软件开发平台,最大的好处是可以充分利用成熟的工业化标准,暂时避开数字图书馆建设中许多复杂的标准问题。不必在众多的数字图书馆范式中作非此即彼的挑选,可以针对我国的具体情况广泛实践。只要所开发的组件能有效提供必需的服务,将来一定可以纳入新的数字图书馆系统中。即使将来有完全统一的标准,原系统仍可保留成熟部分,而通过重新配置、增删组件甚至仅修改部分组件的接口等方法适应新标准。

除此之外,采用 J2EE 平台还有其他一些技术方面的优势。

(1) 适应数字图书馆的发展趋势。因为 J2EE 本身就是针对分布式应用而提出的,其层次化结构和纳入异构系统的能力正是分布式数字图书馆系统所需要的,其用户界面可利用现有的各种 Web 技术,也必然能适应将来发展起来的新技术。

(2) 数据库的选择具有很大的灵活性。数据库

是数字图书馆资源最基本的保存形式。目前的数字图书馆已经有多种类型的数据库。随着技术的进步,还会出现新形式的数据库。J2EE系统对于底层数据库的选择具有很大的灵活性,便于利用已有的各种资源。如果将来需要扩充、更换数据库,也很容易实现。

(3) 具有很好的可扩展性。当馆藏规模、用户规模扩大时,J2EE系统的可伸缩性和可靠性保证了良好的扩展性能。

(4) 有成功的模式可供参考。模式来源于实践,代表着对特定环境中重现问题的专业化解决方案。使用模式可以充分利用已有的经验,避免对同类问题的重新设计,从而更快更好地设计自己的系统。在J2EE实际应用中已总结出很多成功的模式,可供开发者参考^[6]。例如网络的安全性、用户权限管理等问题是所有的数字图书馆实用系统都必须处理的,但这也是其他J2EE应用程序开发时反复研究过、实践过的问题,所以容易找到可供参考的模式,将来甚至可以通过购买现成的组件来解决。

(5) 便于实现软件复用。图书馆应用软件有其自身的特点。实践表明,很少有一种软件安装后不用产销单位过问就能一直顺利运转。软件本身的缺陷常常需要在实际运行甚至长期运行中才能发现^[7],在软件装上运行之后的相当长时间内,还有很多问题需要开发或销售单位解决。存在上述问题的原因,也在于应用软件涉及面宽,未经充分调试。如果以J2EE组件形式开发软件,则与其他应用程序有共性的部分可选用成熟的技术方案或组件,开发者可以专注于图书馆特有的业务逻辑部分,减轻开发负担,提高可靠性。而且对这部分组件可以在调试和应用过程中逐步积累经验,不断改进,使之趋于成熟。

(6) 功能易扩展。未来的数字图书馆将拥有更多形式的信息资源,提供更多更复杂的服务。例如个性化、智能化的信息服务,尽管目前处于起步阶段,将来必定随着相关技术的进步而逐渐发展。建立在J2EE平台上的应用系统可以通过添加组件等方式将新功能纳入其中,而不必对原系统作大的改动。

(7) 便于与其他Web资源的整合。由传统的图书馆资源转化而来的信息资源只是未来数字图书馆信息资源的一个组成部分,来自企业界的信息将是另一重要组成部分。这些资源必然有很大一部分建立在J2EE平台上。采用J2EE平台的系统自然可以与之无缝链接。例如以客户端的形式查询、获取企业系统的有关信息,与其他信息一起作综合处理后,再用适当的形式提交给用户。从软件的角度看,这种过程与本系统内的查询过程无本质差别。

(8) 便于软件开发中的分工、合作与竞争。J2EE平台上的角色分离和组件复用为软件开发的分工、合作与竞争提供了良好环境。现有软件系统各自采用一整套专门技术,互不兼容。尽管这些系统各有特色,却很难把各自的长处结合起来,形成更好的系统。结果导致面面俱到,重复开发,而对于难度较大的技术问题又都难下决心解决。例如对访问权限仅通过有限的IP地址控制,导致用户只能使用指定的计算机组查询。如果开发出专用组件来解决,则所有建立在J2EE平台上的系统都可以采用。因为有人专注于此组件,既可避免重复开发,又便于不断改进提高。

J2EE为数字图书馆的软件开发提供了一种理想的平台,为广泛开展数字图书馆实验系统的研究,以便分工合作,积累各方面的成果,最终汇集为成熟的系统创造了基本条件。

参考文献

- 1 刘锦山. 中国数字图书馆标准化工程建设探析. 现代图书情报技术, 2001(6)
 - 2 张晓林. 分布式数字图书馆机制. 情报学报, 2002(1)
 - 3 曹鸣鹏, 赵伟, 许林英. J2EE技术及其实现. 计算机应用, 2001(10)
 - 4 段靖荒, 林子禹, 万丰. J2EE企业解决方案的平台. 计算机应用, 2001(8)
 - 5 刘湛. J2EE全面简介. <http://adionline.myrice.com/J2ee-all-into.htm>
 - 6 [美]Deepak Alur, John Crupi, Dan Malks 著;牛志奇等译. J2EE核心模式. 北京:机械工业出版社, 2002
 - 7 袁名敦. 关于新一代图书馆应用软件的思考. 现代图书情报技术, 2001(2)
- 李大农 湖北黄冈师范学院物理系教授. 通讯地址: 湖北黄冈. 邮编 438000.
董慧 武汉大学信息管理学院教授, 博士生导师. 通讯地址: 武汉. 邮编 430072.

(来稿时间: 2002-10-10)