

基于 web 服务的内容管理系统构建

谢翠萍 赵 云 向 函 (广东医学院 信息工程教研室 广东 东莞 523808)

摘要:内容管理是企业信息化管理的重要环节,也是协同商务的重要组成部分。讨论了当前内容管理系统的现状,针对当前企业内容传递和发布方式落后,费用高且不能及时反馈市场,与其他信息系统结合性差等情况,介绍了一个结合 web 服务的松散耦合的内容管理系统实现框架,该系统系统具有良好的应用集成能力。

关键词:内容管理系统 web 服务 内容 UDDI

1 引言

在当今信息化时代,企业信息内容快速膨胀,管理起来非常困难,很多企业的内容传递和发布方式落后,费用高又不能及时反馈市场,而企业对信息的同步一致性、发布的及时性要求又较高^[1],为改变现状,企业只有建立跨平台、破除孤岛的内容管理系统(Content Management System, CMS),才能真正提升信息应用的效益,内容管理对企业内部而言,有利于信息共享及增值利用;对于外部来讲,它是协同商务的组成部分。

CMS 是企业中所有系统的基础,它所实现的功能模块应能替大多数信息管理类的应用提供完善的基础功能支持,大大减少开发人员的重复工作,也能更好适应需求的变化。松散耦合的系统结构是今后的发展方向,而传统的 CMS,从内容的创建到管理、挖掘、发布是一个紧密耦合的整体,系统各部分很难独立运行,各子系统间相互协调性很差,系统的扩展性也差,且它们一般都作为一个独立的应用系统出现,较少考虑与企业的其它应用系统间的相互整合协作等问题,新型的内容管理技术应从整个企业角度,考虑其在企业、组织中的作用、功能以及与企业其它系统间的相互配合关系,消除应用(如 ERP、SCM、CRM 等)之间妨碍内容应用的界限,从而真正有助于企业整体性能的优化。这些问题可以使用 Web 服务技术来解决,本文将结合 Web 服务技术提高各子系统之间的耦合性。

2 Web 服务技术标准

Web 服务是近几年发展起来的新一代 web 技术,被称为计算机 IT 技术的第三次革命^[2]。它提供了一个企业内部或企业间各独立系统松散耦合的标准接口,

可灵活实现跨厂商、跨平台、跨语言应用。Web 服务技术标准主要包括:

2.1 HTTP

HTTP 是 web 上最主要的协议,Web 整个层次构架的协议/标准都与 HTTP 紧密结合。如,XML 一般通过 HTTP 传输,SOAP 和 HTTP 进行绑定。HTTP 为 Web 服务部件通过 Internet 交互奠定了协议基础,并具有穿透防火墙的良好特性。

2.2 XML

在 Web 服务模型的核心中,XML 在任何形式的通信中都扮演着通用连接格式的重要角色,它还是其他 Web 服务标准的基础。它是一种元语言,除了易于建立和易于分析外,主要优势在于提供开放式的数据交换功能,它使用户可以很简单地在网络上发送消息并保证在格式转换过程中不丢失任何信息,而不需预先知道对方使用何种系统。

2.3 SOAP 技术

SOAP 是为分布式环境中对等地交换结构化和类型化信息的轻量级协议,它的交换格式使用 XML,SOAP 协议本身并不定义任何应用语义,它只是定义了一种简单交换信息的机制。它通过一个模块化的包装模型和对模块中特定格式编码的数据的重编码机制来表示应用语义。

2.4 WSDL 技术

WSDL 标准用于描述网络服务及其访问信息。它定义了一种绑定机制,用于附加协议、数据格式、抽象消息或用于定义服务位置的端点组。它将 Web 服务描述为能够进行消息交换的服务访问点的集合,客户端可以通过这些服务访问点对包含面向过程调用的服

务进行访问,可以看作是 Web 服务与外界沟通的桥梁。

2.5 UDDI 信息模型

它是一套基于 Web 的、分布式的、为 Web 服务提供的信息注册中心的实现标准规范,同时也包含一组使企业能将自身提供的 Web 服务注册以使得别的企业能够发现的访问协议的实现标准。通过 UDDI,人们可以发布和发现有关某个公司及其 Web 服务的信息。UDDI 包含有关公司服务的接口信息,通过一套基于 SOAP 的 XML API 调用,用户可以在设计和运行时与 UDDI 进行交互以发现技术数据,从而调用和使用这些服务。

3 基于 web 服务的 CMS 架构

3.1 CMS 的总体框架模型

本 CMS 是围绕内容流提出的,内容从不同的数据来源接收过来,再进行分类、创建、编辑、审核,存储,传递和内容分析,最后再发布。本文设计的目标是搭建一个 CMS 框架,使开发人员能够在这个平台上轻松实现内容的创建、储存、发布、应用、更新等应用,并能很容易的整合到商务应用中去。

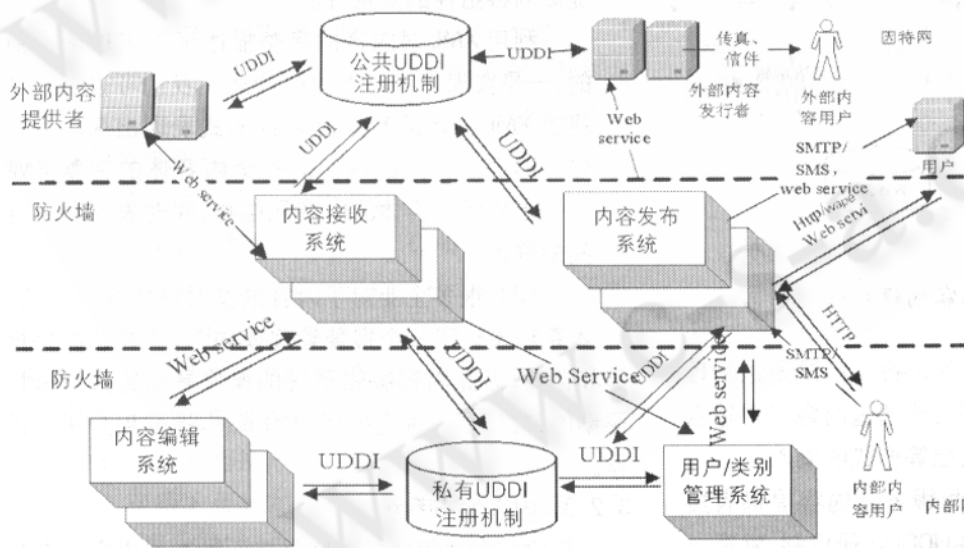


图 1 CMS 的总体框架图

图 1 给出了本文设计的 CMS 总体框架,它是在有多种数据源的异构环境下建立的,在功能上覆盖了内容的“接收—>创建、编辑、产生—>发布”整个生命周期的管理^[3],内容由外部提供者传递给内容接收系统,在内容接收系统对内容进行预处理,然后将处理后的内容传递给内容编辑系统,内容编辑系统再根据

接收的内容创建新的内容,最后通过内容发布系统将内容发布出去,并将内容存储在内容仓库中。

本架构最大特点是结合了 Web 服务技术,Web 服务使系统具有松耦合性、易集成性、服务发布式共享等优势,这使得该 CMS 在架构上更具灵活性,Web 服务架构的客户端程序更容易被开发和维护。

本 CMS 引入 Web 服务有如下优势:

(1) 传统的 B2B 应用使用无标准的架构与商业伙伴连接,缺乏交互性,而 Web 服务提供了一个有利于异构系统交互的开放平台,允许电子商务在 XML 基础上开展,外部内容提供者、内容发行者及客户等都能通过 Web 服务与 CMS 进行交互。

(2) 在企业内部,Web 服务能够在子系统间提供松散耦合的连接,它支持同步和异步消息,异步消息主要用于减少子系统间的依赖性,提高不同实体间通信的可靠性。这些消息提供了支持事件的机制,当相关事件发生时(比如有新的内容到达),这些事件能触发及时的处理,从而提高了系统的及时性。

(3) 对于跨地域的分布式企业,不同部门的系统

提供的各种服务之间的接口很难进行交互,私有 UDDI 注册机制能解决这个问题,我们在企业内部采用私有 UDDI 注册机制,UDDI 技术能管理和描述服务,使用户在一个单一的、开放的和安全的环境中有序地处理工作。

(4) UDDI 注册支持 Internet 上的服务定位。内容接收系统在公共的 UDDI 注册机制注册他们提供的服务,允许外部内容提供者查询这些服务。同样,内容发布系统也在 UDDI

注册机制注册他们的服务以此发布企业的服务,从而可以获取新的潜在客户并提高现有客户的访问率。

3.2 系统详细设计框架

3.2.1 内容接收系统

图 2 是内容接收系统的框架图,它主要负责内容的接收和预处理。从多种外部数据源自动接收内容到

内容仓库,替代繁复的手工采编工作,并与企业内部的内容源进行内容集成。

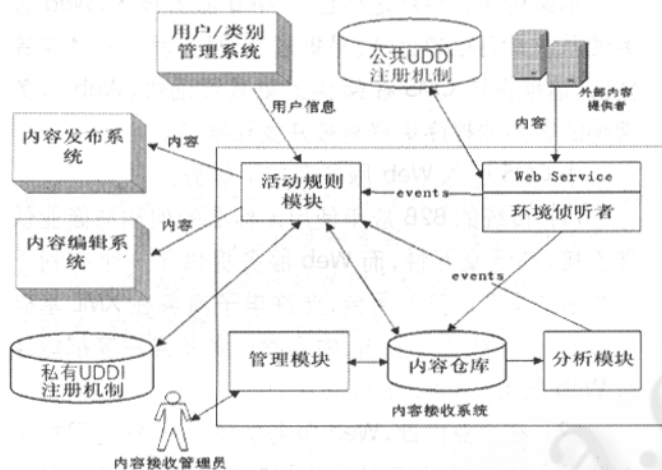


图2 内容接收系统

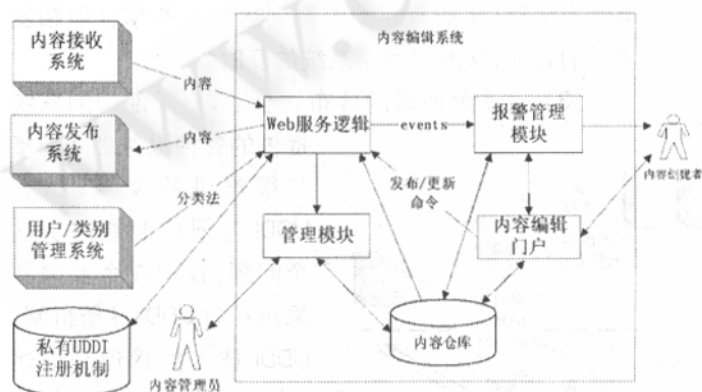


图3 内容编辑系统

(1) 外部内容提供者提供的内容可以通过物理接口输入的图像、多媒体等文件,或通过内容门户输入的内容,还包括 Email、Office 文档等形式的输入。

(2) 环境侦听器通过 Web 服务从内容提供者接收输入的内容。也可通过公共 UDDI 注册机制,发现一些新的内容提供者提供的服务。它使用一个发布-订阅机制,内容提供者通过公共的 UDDI 注册机制发布想要订阅的内容的摘要,然后他们就可调用相关 Web 服务来订阅他们感兴趣的内容,环境侦听器通过调用相应的 Web 服务来接收这些被订阅的内容,收到的内容被存储到内容仓库中,同时发送相应的处理到活动规则模块。

(3) 为了能有效接收内容并产生信息,在设计过

程中以活动规则模块为中心,使用事件驱动方式。它定义了一些规则,详细说明了内容的操作行为限制,这些规则可以被及时执行,相对传统的表格驱动方式,它提供了一个更为灵活的过滤和映射能力。

3.2.2 内容编辑系统

图3是内容编辑系统的框架图,它主要用于创建新内容,并处理接收到的内容。内容创作工具应该很方便用户进行内容创作,并提供良好的封装性,使得内容创建者可以很方便地进行内容创作。我们可使用常用的其它公司提供的编辑工具,如 MicroMedia 开发的一系列的产品, Microsoft Office 等工具来直接编辑内容,也可以通过内容编辑门户采用统一的形式来对内容进行编辑。

(1) 系统对内容的管理是组件级的,内容提交后,系统根据模板生成相应的 XML 文档,并将内容分割为组件存储在数据库中,每一种组件都单独存放并索引,可以单独查询和编辑。比如说可以对图片实现单独索引和查找,新闻中的图片既可在浏览新闻时看到,又可单独进行查询、浏览,这样的管理方式可以提高内容组件的复用性。

利用 XML 对内容的完整描述可以实现两个目的:一是实现内容与表现的分离。我们可很容易地将以 XML 描述的数据以多种形式表现出来。二是建立内容组件间的联系。一个内容体的完整 XML 文档是描述内容体的内容和结构,同时表明各组件间的联系。

(2) 很多企业对于内容的及时性要求非常高,本系统设计了一个报警管理模块^[4],主要功能是报警和监控,因此内容编辑系统的操作是由警报(alert)驱动的,警报是一种提醒用户的信息,提醒用户执行工作。

3.2.3 内容发布系统

接收和创建编辑的内容最终要发布。内容发布就是把服务性的个性化的内容发送到用户的浏览器上。内容发布需要考虑个性化问题,既要考虑用户媒体的不同,还要考虑用户的兴趣,用户还可在一定权限范围内进行模板设计,为用户提供个性化界面。

图4是内容发布系统的框架图,主要功能是发送新的内容给订阅用户,并将内容发布到网上,这是 CMS 与用户接触最多的模块。发布内容有两种情况,一种是 CMS 主动向用户发布内容,让用户查看内容;另一

种是用户向 CMS 发送查询请求, CMS 处理查询请求并将查询结果反馈给用户。

(1) 在接收新内容基础上, 内容传递模块查询用户的注册订阅情况, 再将相关的内容传递给已经订阅且有足够高的权限的用户。内容通过以下方式传递给用户:

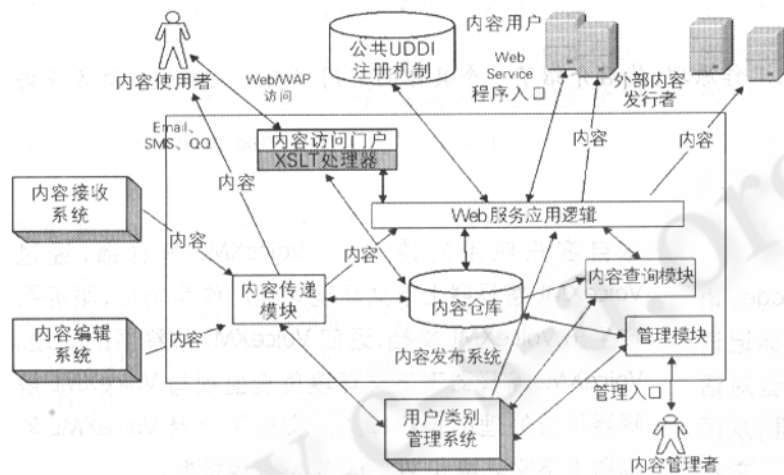


图 4 内容发布系统

- 经由 email、SMS 或 QQ 等交互软件传递。
- 经由 Web 服务传递。有了 Web 服务, 内容发行者可以在无人干涉的情况下通过公共 UDDI 注册机制查寻并自动与系统交互。

(2) 该系统的另一个主要功能是检索和浏览内容, 这也是 CMS 的主要功能之一。CMS 管理的大部分都是非结构化数据, 信息查询是以文档为中心, 而不是常规数据库基于结构化数据为中心的查询。本系统提供的搜索方式是基于 Lucene 的全文搜索。Lucene 是 Apache 组织的 Jakarta 项目之一^[5], 它不是一个完整的全文搜索应用, 而是一个用 Java 编写的全文索引框架工具包, 可很方便地扩充到各应用程序中。本文对其进行了扩展, 使其适合本系统的应用。

(3) 在内容发布系统中, 需要进行模版管理, 根据模版生成发布页面。模版管理提供模板创建, 存储, 修改, 删除等功能。本系统的模板管理使用 XSLT 技术实现^[6], XSLT 模板管理实际是对 XSL 的文件操作。XSLT 文件通过 Xpath 对模版上显示的内容定位, 使得生成的 XML 页面与模版可以以多种形式发布。本系统的内容发布系统实现上是基于 B/S 架构的, 主要利用了 XML 和 XSL 的强大显示功能和灵活的特性, 设计功能

上比可视化工具逊色, 但它使得内容发布系统具有很强的灵活性, 能很好的支持个性化服务。

3.2.4 用户和内容类别管理

该模块主要管理用户的注册信息、订阅的内容和内容的分类模型。

系统的内容是按一定的内容分类模型组织的, 对于连锁企业, 各分公司都采用同样的分类法, 能够更方便对内容进行管理。如果为了内容更好的发布, 还要能够动态的对发布内容进行分类, 并以内容栏目的方式展现在用户面前, 同时还要对每一类内容的发布流程、发布格式等进行相应的定制。

4 小结

目前内容管理还是一个正在发展的领域, 它与企业其它信息系统的集成、共享一直是阻碍企业发展的一个难题, web 服务可灵活实现跨厂商、跨平台、跨语言应用。本文提出了一个基于 Web 服务的 CMS 设计方案, 开发人员在这个平台上能够轻松实现内容的创建、储存、发布、应用、更新等与内容管理相关的应用, 能够很好得融合到商务应用中。

参考文献

- 1 杜义华、张亚, 网站信息管理发布系统设计与应用, 计算机系统应用, 2005.1:9-11.
- 2 应宏、鄢沛, 基于 Web Service 的跨企业应用模型, 计算机工程, 2005, 1.
- 3 H. E. McNay. Enterprise Content Management: an Overview. In Proc. IEEE Intl. Professional Communication Conf., (IPCC 2002), 2002, pp. 396-402.
- 4 D. K. W. Chiu, B. Kwok, R. Wong, E. Kafeza, and S. C. Cheung. "Alert Driven E - Services Management," In Proc. HICSS37, Jan 2004, CDROM, 10pages, and Computing Systems, Minneapolis, Minnesota, 2002, pp.1-15.
- 5 Erik hatcher, Otis gospodnetic, Lucene in action, 2005.
- 6 [美] Kurl Cagle 著, 周生病、肖伟译, XML 高级开发指南, 北京: 电子工业出版社, 2001.