

基于 Web 标准的信息港研究

尹立杰^{1,2}, 段 岩¹, 李双成¹

(1. 中国地质大学, 北京 100083;

2. 石家庄经济学院, 河北 石家庄 050031)

摘 要:数字城市是以 Internet 为平台的, 实现空间信息互操作为核心的城市信息管理系统, 却受到跨行业应用的基础数据不同组织形式制约。数据的组织与表现松耦合关系是实现空间信息互操作的重要环节, 也是适应语义网络计算标准的要求。Web 标准将页面行为与表现分开, 为信息港实现资源的一体化管理, 满足用户按需的 3A (Anybody, Anytime, Anywhere) 可视化服务, 达到开放、共享根本目标提供了架构和实现的途径, 同时隐含了制约信息港门户性能的关键问题和解决途径。

关键词:数字城市; Web 标准; XML; 信息港

中图分类号: TP393.4

文献标识码: A

文章编号: 1673-629X(2008)05-0098-03

Study of Info Port Based on Web Standards

YIN Li-jie^{1,2}, DUAN Yan¹, LI Shuang-cheng¹

(1. China University of Geosciences, Beijing 100083, China;

2. Shijiazhuang University of Economics, Shijiazhuang 050031, China)

Abstract: Digital city is a city information system with spatial data on Internet has been restricted by the heterogeneous data formats come from different organizations. Semantic Web is going to become next-generation Web demand data are loose coupling between presentation and behavior to make these data more accessible. Web standards achieve to separate data behavior from presentation and afford frames and methods to accomplish the digital city's services anytime and anywhere with opening and sharing. In addition, what can be predicted is the standardization on metadata is the key technique for digital city infrastructure data which has restricted the info port capability.

Key words: digital city; Web standards; XML; info port

0 引言

数字城市(又称为信息港、数码港等)是数字地球的重要组成部分, 是一个基于网络环境的城市信息, 特别是空间信息互操作的服务体系。它的任务是利用现代高科技手段, 充分采集、整合和挖掘城市各种信息资源(特别是空间信息资源), 建立面向政府、企业、社区和公众服务的信息平台、信息应用系统以及政策法规保障体系等。随着空间信息技术的发展, 空间数据归属不同机构, 分布在不同的场所, 彼此孤立的信息孤岛成为客观存在的事实。OGC 组织 (OpenGIS Consortium) 认为空间信息互操作^[1]是要克服异构处理环境和空间信息所引起的空间数据转换障碍以及访问分布资源的障碍, 使得一组软件构件能够协调工作。目前,

网络计算技术的发展为空间数据服务的集成提供架构模型, 基于 Web 服务^[2-6]的体系结构成为首选。因此, 数据的组织与表现的松耦合便成为实现不同格式的空间数据互操作, 实现空间信息服务集成的瓶颈问题。

由于用户终端和服务需求的多样性, 信息港作为针对公众访问的门户窗口, 对用户界面的导航功能、信息智能化、交互性操作、跨平台机制等方面有着特殊的要求。实现服务的机制也随着应用规模的扩大、新技术的出台而变化。传统的门户页面内容评价性能的标准显然是不够的, 而且适合文本信息显示的表格布局将数据的表现和行为绑定更制约了终端应用的灵活性。行业间基础数据的不同组织形式(不仅是文件格式, 还有数据模型的不同)不仅制约了数据间的交互操作, 更制约了用户的便捷访问。对现有网络计算技术进行扩展的下一代语义 Web^[7], 需要丰富的语言和标准去规范 Web 开发和描述 Web 内容, 数据的组织更趋向于资源的本体化。Web 标准将页面的结构、表现和

收稿日期: 2007-08-26

基金项目: 河北省科技计划项目 (052135148)

作者简介: 尹立杰 (1970-), 女, 吉林长春人, 博士研究生, 讲师, CCF 会员, 研究方向为系统开发与信息处理、地学信息工程。

行为分离处理,主张基于 Internet 应用的界面输出分层,很好地解决了传统方式的弊端和语义功能,满足开放和共享的 Web 应用的根本要求。

1 Web 标准

Web 标准并不是某一个标准,是由 W3C(The World Wide Web Consortium)起草和发布一系列标准的集合。描述的是以 XML 为核心的数据组织形式,通过页面结构、表现和行为分离处理,实现 Web 应用体系结构中页面规范化开发技术标准。

1.1 W3C 技术架构

W3C 的目标是通过 XML 提供信息交换的手段,实现 Web 的语义化。换言之就是通过 RDF(Resource Description Framework)和 OWL(Ontology Web Language)语言使得 Web 上的信息内容更加容易理解、更便于交换和共享。其技术架构如图 1^[7]所示。

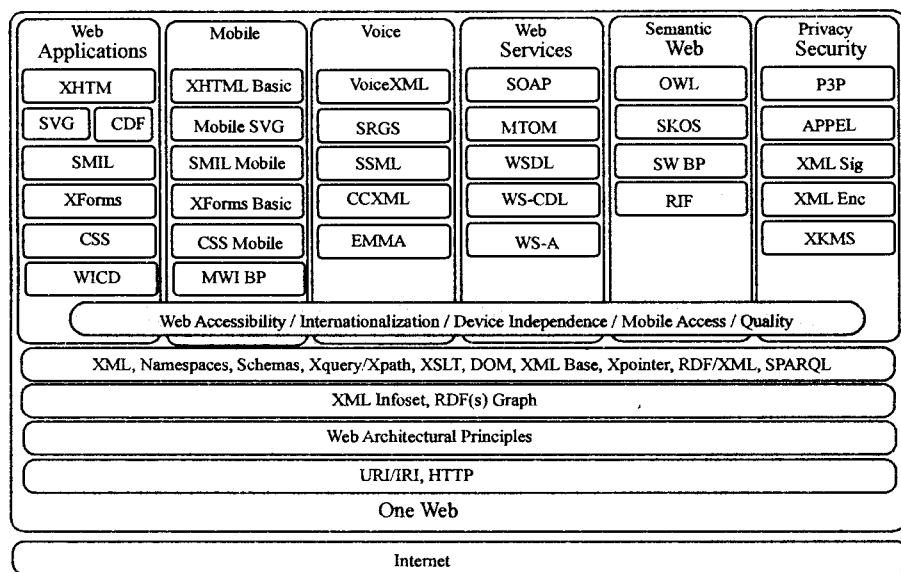


图 1 W3C 技术架构图

在 HTTP 应用层协议上以 XML 数据组织形式为基础,将 Internet 应用层的体系结构扩展成两层 Web 应用模型,通过逐层应用的万维网基础框架结构实现 Web Accessibility(Web 可访问性)、Mobile Access(移动访问)、Internationalization(国际化)、Device Independence(设备独立) and Quality Assurance(质量保证)的性能要求(如图中的阴影区域)。从底至顶依次为:

* URI/IRI, HTTP

* Web Architectural Principles

* XML Infosets; RDF(S) Graphs

* XML, Namespaces, Schemas, XQuery/XPath, XSLT, DOM, XML Base, XPointer, RDF/XML, SPARQL

* Web Applications, Mobile, Voice, Web Services, Semantic Web and Privacy

以此架构实施的 Web 服务能够实现面向商务信息、面向用户的、面向设备的和面向系统的应用需求。不仅囊括了数字城市领域应用的需求,并与未来网格计算模式架构相匹配。遵循 W3C 的技术架构,实现无障碍网页、国际化、设备无关和质量管理等主题应用,将成为信息世界中有高稳定性、可提升和强适应性的基础框架。

1.2 Web 标准的标准语言

网页开发对应的结构、表现和行为标准语言是:

* 描述网络数据结构的标准语言,如 XHTML(The Extensible HyperText Markup Language,可扩展标识语言)和 XML(The Extensible Markup Language,可扩展标识语言)等。在 HTML4.0 的基础上,用 XML 的规则对其进行扩展,得到了 XHTML。建立 XHTML

的目的就是实现 HTML 向 XML 的过渡。

* 描述表现标准语言,如 CSS(Cascading Style Sheets,层叠样式表)、XSL(The eXtensible Style Language)和 XSLT(The eXtensible Stylesheet Language Transformations)等。XML 将数据作为一个信息的基类型来存储,通过表现语言转换成不同的应用格式。使用 XSL,一个样式单文件 XSLT 控制 XML 文本信息选择的数据节点并转换为多种格式(如 HTML, WML, PDF, PostScript 等)的输出。CSS 控制

节点的显示格式。

* 描述行为标准语言,如 DOM(Document Object Model,文档对象模型)、ECMAScript 和 SAX(Simple Application for XML)等。DOM 是一组对象的集合,以分层对象模型访问存储在 XML 文档中的信息,为 Web 开发者提供了访问站点中的数据、脚本和表现层对象的方法,实现对 XML 数据的操纵功能。

由于表现、结构与行为的分离,使得遵循 Web 标准开发的门户网站,代码的简洁,易于系统维护和改版升级。对访问者而言页面输出效果不再受到的硬件设备和软件平台的制约,文件下载与页面显示速度更快,能够通过样式选择定制自己的表现界面等。

2 信息门户平台的需求

数字城市的建立需要城市空间信息基础设施和行业空间数据工程的支持,由于针对应用群的宽泛而使得应用领域也相当广泛。图 2 是以用例模式描述的实现数字城市数字化、网络化、智能化与可视化等应用的服务架构模型,是由用户界面层、业务逻辑层和数据源三层架构组成的 Web 服务系统。

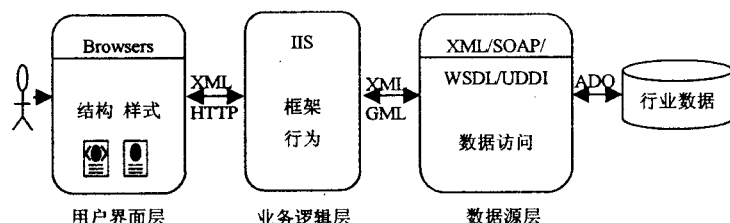


图 2 基于 Web Service 的信息港系统结构

XML 作为一种用来描述数据的标记语言,只定义文档内容而不涉及具体表现。文档中的数据以树型结构统一描述,具有以扩展性、结构化语义以及平台无关性的特点,成为网络数据传输和交换的主要载体。在跨行业领域数据交换上具有优势,即使是不同数据格式和不同数据模型间的数据交互操作。

XML 文档描述资源的组织形式在跨格式(数据存储格式、数据模型)数据的交互访问功能的优势,在 OGC 组织倡导下针对地理空间数据的描述语言规范陆续出台:

* GML (Geography Markup Language): 基于 XML 的空间信息编码标准,针对地理特征(feature)将封装的空间对象的空间数据和图形解释清楚分离。2000 年 5 月,OGC 推出了基于 XML DTD (Document Type Definitions, 文档类型定义) 和 RDF (Resource Description Frameworks, 资源描述框架) 的 GML 1.0 版。2001 年 2 月,OGC 推出完全基于 XML Schema 的 GML 2.0 版。2003 年 2 月,增加了对复杂的几何实体、拓扑、空间参照系统、元数据、时间特征和动态数据等的支持,更适合描述现实世界问题的 GML 3.0 版正式发布。

* XMML (eXploration and Mining Markup Language)^[8]: 是 GML 地学应用,意在实现空间信息的数据挖掘。

* GeoSciML (Geosciences Information Markup Language, <http://www.opengis.net/GeoSciML>) (Marcus, 2005)^[9]: 地球科学信息标记语言更强调地球科学信息的通用性描述,重用了 XMML 和其他相关地球科学软件中的组件,扩展了 GML Schema。

后两种描述语言由 BGS (British Geological Survey) 提出,遵循 CGI (The Commission for the Management and Application of Geoscience Information) 数据模型和

编码标准。空间信息描述语言功能各有专长,都是 XML 的结构子集,在数据模型层使用 DTD 或 XML Schemas 描述元数据,分别拥有独立标识作用的命名空间。通过元数据与海量数据的关联,保证了数字城市资源的融合处理,跨格式交互访问。

元数据是描述数据的数据,是信息系统建设中对信息资源、控制流程和信息流通等相关信息的规范化的描述信息,是资源的各项特征构成的信息集合。元数据通过描述某种类型资源(或对象, object) 的属性,以及该资源进行定位信息管理,有助于实现数据检索。

XML Schema 文档为元数据的管理和建设提供了标准化的方法和手段,可以避免部门间数据的重复建设,支持部门间的协同工作和管理,从而使得有效的进行信息和服务共享和互操作成为可能。这些对于数字城市的信息资源管理均具有重要的意义。

3 基于 XML 的元数据服务

数字城市是以空间信息为核心的城市信息系统。作为空间资源的有效共享和充分利用一直是空间信息科学界的难题,如何将各种空间信息资源进行汇集和共享,实现一体化的组织和管理,以实现按需服务能力的空间数据管理和信息处理能力的空间信息基础设施,是空间信息网格^[10] (SIG, Spatial Information Grid) 技术(如图 3 所示)的主题。

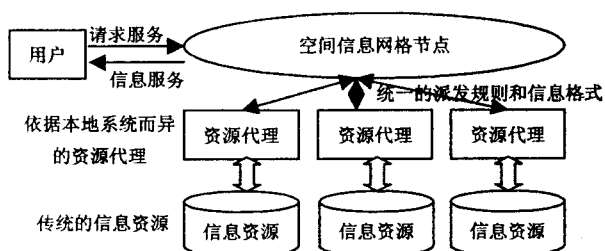


图 3 空间信息网格提供空间信息服务一体化模型

网格就是“动态多机构虚拟组织中的资源共享和协作问题解决”(Ian Foster, 2002), 实现的是临时性的 Web 服务。目前 SIG 服务的理论研究、构建技术、应用方法等诸方面都存在亟待深入的前沿领域,李德仁^[10]、王家耀^[11]等分别从空间数据模型、建模方式和系统结构基础框架上提出了新的技术思路。XML Schema 解决地理数据组织和存储方式的不统一的劣势,重新构建基础数据的不现实性,基于 XML 实现 Web 服务架构的特点,成为运用元数据实现语义 Web 服务的代名词。元数据标准化也成为制约信息港平台

(下转第 104 页)

SNMP 无响应的计算机,若使用同步方式,很多时间浪费在等待超时报文上。所以这里使用异步方式实现。

为了实现这一过程,首先建立一个队列存储所有交换域中的设备 IP,然后对队列中 IP 异步发送 SNMP 请求报文。对于每个应答报文,查看返回的值,若为 2,则把该 IP 放入交换机队列(在这里还有 sysServices 的值大于 2 的应答报文,这是路由器发出的应答报文,通过查找网络层拓扑发现的结构替换序列对队列,找到该 IP 地址对应的 MaxIP,从而找到了相应的路由器标识^[6])。收到第一个超时报文后,等待 1s 并继续接收看是否还有应答报文,1s 后退出等待。

2.6 拓扑更新

上述算法是基于生成树协议来确定交换域内拓扑连接关系的,所以当生成树发生变化时,比如根交换机发生改变或是其它交换机发生故障,拓扑发现都要运行一次,这样才能反映出拓扑的实时性。

3 结束语

实验结果表明,以上的拓扑发现算法在准确度和效率上均有很好的效果。

目前,很多网络中通过三层协议交换机连接不同的 VLAN(虚拟局域网),每个 VLAN 都需要运行独立

的生成树协议,但是 802.1D 中的生成树协议并不支持 VLAN。很多交换机厂商开发自己的私有 MIB 变量来支持 VLAN,这就需要对各厂商的私有 VLAN MIB 变量有全面了解才能发现一个 VLAN 的链路层拓扑。目前的工作中,没有对 VLAN 进行研究。这将是下一步需要研究的内容。

参考文献:

- [1] McCloghrie K. Definitions of Managed Objects for Bridges [S]. Internet RFC1286. 1991.
- [2] Breitbart Y, Garofalakis M, Martin C. Topology Discovery in Heterogeneous IP Networks[J]. IEEE INFOCOM, 2000, 26 (1): 265 - 274.
- [3] Lowekamp B, O'Hallaron D R, Gross T R. Topology Discovery for Large Ethernet Networks[J]. ACM SIGCOMM, 2001 (8): 237 - 248.
- [4] Rose M, McCloghrie K. Structure and Identification of Management Information for TCP/IP - based Internets [S]. RFC1155. 1990.
- [5] 康 华,周明全,房鼎益.基于 SNMP 协议的网络拓扑发现和显示[J].微机发展,2003,13(7):60 - 65.
- [6] 熊 英,楚惟善,王春枝.基于 TCP/IP 的网络拓扑发现方法的研究[J].湖北工业大学学报,2005,20(4):53 - 55.

(上接第 100 页)

建设、成为实现以 XML/SOAP/WSDL/UDDI^[3-6]为主干的信息服务的键问题。

4 结束语

数字城市从概念的提出发展到现在仍然处于探索和研究阶段。虽然实现一体化的按需服务的架构的关键技术和标准均以提出和实现,但是在走向工程化的道路上尚有差距。基于 XML 的元数据描述空间和非空间数据成为整合信息港服务实现语义应用的必然,基于 Web 标准的服务体系成为实现 Internet 开放、共享的保障。

针对我国数字城市建设的状况,构建符合 Web 标准的应用平台、元数据标准化、服务标准化建设是加快实施的关键。

参考文献:

- [1] Klopfer M. Interoperability & Open Architectures: An Analysis of Existing Standardisation Processes & Procedures[EB/OL]. 2005 - 05. <http://www.opengeospatial.org/press-room/papers/>.
- [2] 王 浒,李 琦,承继成.数字城市元数据服务体系的研究和实践[J].北京大学学报:自然科学版,2004,40(1):107

- 115.

- [3] Web Service 体系[EB/OL]. 2002. <http://www.techng.com/content.asp?titleid=3797>.
- [4] UDDI 组织. UDDI Technical White Paper[EB/OL]. 2004 - 10. <http://uddi.org/pubs/uddi-tech-wp.pdf>.
- [5] Boubez T, Hondo M, Kurt C, et al. UDDI 程序员 API 规范 V1.0[EB/OL]. 2001 - 03. <http://www.uddi-china.org/pubs/ProgrammersAPI-V1.01-Open-20010327-CN.pdf>.
- [6] W3C. SOAP Version 1.2, W3C Working Draft[EB/OL]. 2001 - 07 - 09. <http://www.w3c.org/tr/2001/WD-soap12-20010709/>.
- [7] W3C Web Home[EB/OL]. 2001. <http://www.w3.org/Consortium/technology>.
- [8] XML project Web Home[EB/OL]. 2004. <https://www.seegrid.csiro.au/twiki/bin/view/Xnm/WebHome>.
- [9] Sen M, Duffy T. GGeoSciML: Development of a generic Geo-Science Markup Language[J]. Computers & Geoscience, 2005 (31): 1095 - 1103.
- [10] 李德仁,朱欣焰,龚健雅.从数字地图到空间信息网格——空间信息多级网格理论思考[J].武汉大学学报:信息科学版,2003,28(6):642 - 650.
- [11] 王家耀,祝玉华,吴明光.论网格与网格地理信息系统[J].测绘科学技术学报,2006(2):1 - 7.