

## 基于 XML 的实时海洋数据 传输网络系统

姬鹏<sup>1</sup>, 张承慧<sup>1</sup>, 孙汝鹏<sup>2</sup>

(1. 山东大学 控制科学与工程学院, 山东 济南 250061; 2. 山东大学 计算机科学与技术学院, 山东 济南 250061)

**摘 要:** 近年来, 在海洋灾害预测、预报等海洋信息产品的开发中, 对准确、大量、连续、实时的海洋要素监测数据的依赖性日趋增强。同时, 由于海洋数据的复杂性, 数据来源的多源性, 数据平台异构, 数据格式不规范等原因, 增大了数据集成的难度, 降低了数据使用的效率。文章提出了用 XML 作为传输媒介, 结合 Socket 网络通信技术, 建立实时海洋数据传输网络系统, 实现了海洋数据集成与交换格式的统一, 屏蔽了影响海洋数据集成和使用效率的因素。文章对 XML 在海洋信息产品开发中的应用做出了有益的探索和尝试, 旨在促进 XML 在我国海洋信息领域的应用。

**关键词:** XML; Socket; TCP; 数据传输; 数据交换

**中图分类号:** P731 **文献标识码:** A **文章编号:** 1003-0239(2006)2-0045-07

### 1 引言

近年来, 随着海洋灾害预测、预报技术的快速发展, 以及其它海洋信息产品的大量涌现, 对准确、大量、连续、实时的海洋要素监测数据的需求日趋强烈。在迅速发展起来的可扩展标记语言 XML (Extensible Markup Language) 具有支持多种数据格式的能力, 能够有效地解决异构平台、不同数据格式之间的数据交换问题, 许多行业已将其应用于行业数据交换和数据发布领域。海洋 XML 是一种基于 XML 的封装海洋数据的新兴技术, 它具有提供存储、传输、显示和发布海洋数据的高效手段的潜力。对于现存的海洋数据格式和海洋信息系统以及未来种类繁多的海洋信息产品, 海洋 XML 都将提供良好的支持, 极大地提高海洋信息产品的开发效率, 为各种海洋信息产品提供开放的数据交换接口标准。目前, IOC (Intergovernmental Oceanographic Commission) 组织已在 IODE (International Oceanographic Data and Information Exchange) 中成功地应用 XML 处理数据交换问题。欧美发达国家业已将 XML 在海洋信息领域中的应用成功地拓展到数据交换、数据处理及数据存储等方面。鉴于 XML 在我国海洋信息领域的应用尚处于起步阶段, 本文结合 XML 技术和 Socket 网络编程技术, 提出了一种实时海洋数据传输网络系统的实现架构, 为海洋监测台站与海洋数据中心之间的数据传输和数据交换提供了一种开放、

收稿日期: 2005-11-09

基金资助: 山东省科学技术发展计划重点项目 (2004GG2205108)

作者简介: 姬 鹏 (1975-), 男, 博士研究生, 研究方向为控制理论、海洋环境监测技术。

高效、可靠的解决方案,同时也为海洋 XML 在我国海洋信息领域的应用和发展进行了有益的探索和实践。

## 2 系统描述

基于 XML 的实时海洋数据传输网络系统是一个将海洋监测台站现场连续监测的海洋环境要素、数据按照系统定制的 XML 数据文档规范格式化后,通过 INTERNET 实时地传送至海洋数据中心,再按照系统定制的 XML 数据文档规范进行原始数据恢复,然后使用恢复后的海洋环境要素原始数据对海洋数据中心数据库进行实时更新的海洋信息传输系统。在基于 XML 的实时海洋数据传输网络系统中,要求网络通信连接可靠,传输数据无丢失,实时数据完整。TCP 协议可以充分满足传输系统可靠通信的要求。同时,基于 Socket 的数据通信,可以隔离网络的文件访问功能和一般的 Web 访问功能,有效地提高了传输系统的安全性。因此,本文选择了 TCP 协议以及基于该协议的流式套接口,结合 XML 进行 Client/Server 模式下的基于 XML 的实时海洋数据传输网络系统的设计(系统结构见图 1)。

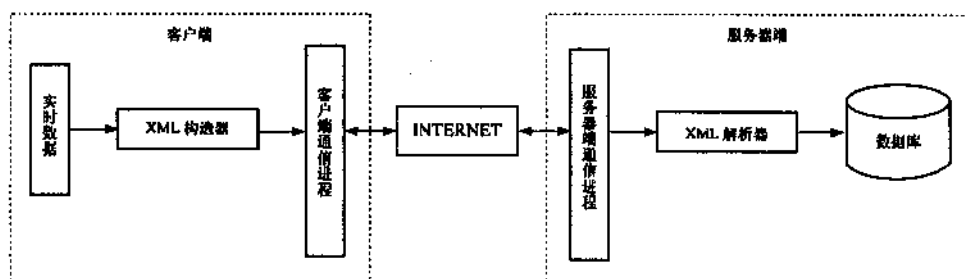


图 1 基于 XML 的实时海洋数据传输网络系统的结构示意图

其中,客户端(海洋监测台站)由实时数据、XML 构造器、客户端通信进程组成:实时数据是海洋监测台站现场连续监测的海洋环境要素数据;XML 构造器将实时数据按照系统定制的 XML 数据文档规范格式化,生成 XML 数据文档;客户端通信进程通过 INTERNET 和海洋数据中心的服务器通信进程建立连接,将 XML 数据文档实时地传送至海洋数据中心。服务器端(海洋数据中心)由服务器通信进程、XML 解析器、数据库组成:服务器通信进程通过 INTERNET 和海洋监测台站的客户端通信进程建立连接,接收海洋监测台站传送的格式化后的 XML 数据文档;XML 解析器将接收到的 XML 数据文档按照系统定制的 XML 数据文档规范恢复原始数据,然后使用恢复后的数据对数据库进行实时更新。

整个系统的核心由一对 XML 构造/解析器和一对 C/S 通信进程组成,下文将会分别详细介绍。

### 3 XML 数据交换机制

XML 发展的首要目标是为处理 Internet 上的电子数据交换提供一种强大的工具,而且也应具备处理大规模数据集的能力。它是一种结构化的描述语言,采用的是树型存储结构,支持深层次的嵌套表达,它的数据集不仅包含数据值,同时也包含数据的含义;XML 是文本格式,而不是二进制格式,易于阅读和编辑修改;XML 相比其它标记语言更简洁,创建和解析 XML 数据文档更容易;XML 是一种开放的标准,具有一套统一的数据格式,使用它意味着在任何平台、任何操作系统、对于任何语言都能很容易实现不同格式数据的相互交换;XML 允许数据提供者描述数据的结构和控制它们的用法。因此,XML 非常适合于作为海洋信息系统的数据交换、处理、存储及发布的标准,表示和封装结构复杂的海洋环境要素数据。下面给出一段描述海洋监测台站属性的 XML 数据文档片段:

```
<STA>
<NAME>LYG</NAME>
<LOCATION>06452</LOCATION>
<LAT>34.78°</LAT>
<LONG>119.43°</LONG>
</STA>
```

要在实时海洋数据传输网络系统中实现 XML 数据文档的交换,系统必须能够创建和解析 XML 数据文档。因此,我们在客户端(海洋监测台站)设计了 XML 构造器,在服务器端(海洋数据中心)设计了 XML 解析器。

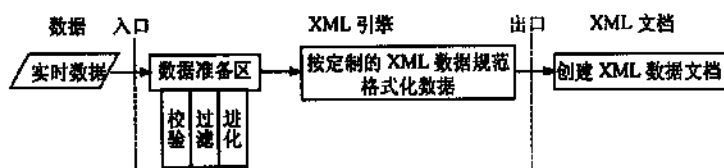


图 2 XML 构造器工作流程

XML 构造器工作在实时海洋数据传输网络系统的客户端,其工作流程如图 2 所示。

首先,海洋台站现场采集到的海洋环境要素实时数据进入数据准备区,数据准备区主要功能包括:数据校验、数据过滤、数据净化。其中,数据校验主要是检查数据的完整性;数据过滤主要是对进入数据准备区的数据按照系统的需求进行筛选,将系统不需要的数据从数据源中剔除,而留下的数据要与系统的需求相匹配;数据的净化处理主要是对进入数据准备区的数据进行正确性判断,将那些数据内容错误、格式错误或类型错误的数据进行修正和净化处理。在数据源端建立数据准备区,进行数据校验、过滤及净化,有助于提高数据质量。

然后,将经过数据准备区处理后的数据按照系统定制的 XML 数据文档规范进行格

式化。这里提到的 XML 数据文档规范实际上是一个 XML 架构。XML 架构是 XML 文档集合的 XML 描述。最好的比喻就是类和对象：XML 架构相当于类，匹配这个架构的任何特定 XML 文档就相当于类的一个实例。要将各种类型和形式的海洋数据封装到 XML 架构中，必须要有一个良好的数据组织结构。这个数据结构应该能够处理多维、多参数数据和时间（连续、离散）、空间（点、线、面）数据；跟踪数据的任何变化；提供数据质量细节、数据管理细节和其他需要描述的细节内容。

最后，生成 XML 数据文档，交付客户端通信进程，等待传输。

XML 解析器工作在实时海洋数据传输网络系统的服务器端，其工作流程如图 3 所示。

首先，将服务器通信进程接收到的 XML 数据文档进行格式校验，确认 XML 数据文档的合法性。然后按照系统定制的 XML 数据文档规范从合法的数据文档中提取数据，并将提取的原始数据送入数据恢复区。

数据恢复区的主要功能包括：数据标准化、数据质量确认、数据的过滤与匹配。其

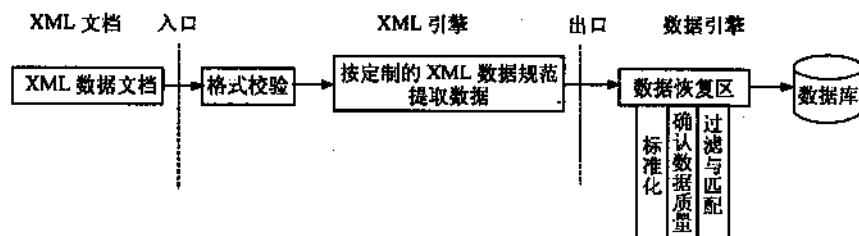


图 3 XML 解析器工作流程

中，数据标准化主要是将数据名称、数据结构有差异的数据进行统一处理，保证数据进入数据库后表征意义的一致性；数据质量的确认主要对数据来源、数据精度等进行确认，以保证基于数据库的应用的可靠性、稳定性、准确性；数据的过滤与匹配是对数据进一步按需划分，满足不同应用的需求。

最后，用经过数据恢复区处理的数据对海洋数据中心数据库进行实时更新，以满足其它海洋信息产品的使用需求。

## 4 Client/Server 模式通信进程实现

客户端与服务器之间的数据通信是由底层的计算机网络通信协议 TCP/IP 来提供的，TCP/IP 协议族为应用层服务提供了两种传输层服务：面向连接的 TCP 协议和无连接的 UDP 协议。UDP 是一种非常简单的传输层协议，它不能保证数据最终到达目的地；与此相比，TCP 提供可靠的传输服务，确认、超时和重传机制确保了应用层数据会被可靠地传送到目的地，同时 TCP 还提供流量控制。

套接口（Socket）最初是由加利福尼亚大学 Berkeley 分校为 Unix 操作系统开发的网络通信接口，后来它又被移植到 DOS 与 Windows 操作系统。Socket 是一个编号接口和通信端点，可以用于连接其它的计算机，并向它们发送和接收数据。Socket 接口本质上

是 TCP/IP 网络的 API 接口函数,Socket 数据传输其实就是一种特殊的 I/O。常见的 Socket 对象类型有两种:流式 Socket 和数据报式 Socket。流式 Socket 是一种面向连接的 Socket,针对面向连接的 TCP 服务应用程序;数据报式 Socket 是一种无连接的 Socket,对应于无连接的 UDP 服务应用程序。

综上所述, TCP 协议和流式 Socket 更适合实时海洋数据传输网络系统的通信要求。实时海洋数据传输网络系统的通信进程包括:客户端通信进程和服务端通信进程,如图 4 所示。

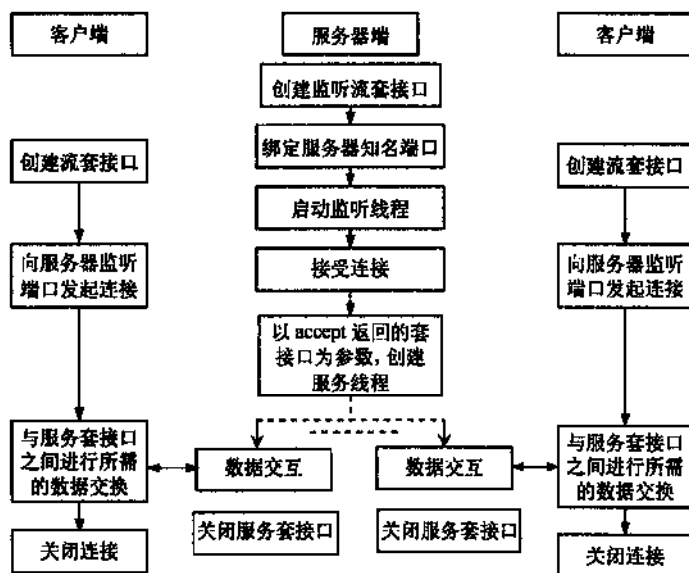


图 4 海洋实时数据网络传输系统的通信进程

其中,客户端通信进程比较简单,大致可分为三步:与服务器端通信进程建立连接;数据传输;与服务器端通信进程断开连接。

而服务器端通信进程由于要同时响应多个客户端通信进程的连接请求,因此它的设计相对比较复杂。根据实际情况,我们选择了相对简单的每客户单线程并发服务器架构进行服务器端通信进程的设计。

基于这种架构的服务器程序流程如下:

- ①创建监听流套接口;
- ②绑定本地的服务器知名端口;
- ③启动监听主线程,使套接口处于监听状态;
- ④接受客户端通信进程连接请求,accept 函数返回服务套接口;
- ⑤以服务套接口为参数,创建服务线程,主线程返回步骤④;同时服务线程进入下一步;
- ⑥服务线程完成客户端通信进程的服务请求,关闭服务套接口并结束服务线程。

TCP 协议和套接口 (Socket) 还可以很好的满足跨系统网络通信的需要。在实际应用中, 实时海洋数据传输网络系统的客户端通信进程可能工作在不同操作系统下, 而客户端通信进程和服务端通信进程也可能工作在不同操作系统下。因此, 本文分别在 Linux 操作系统和 Windows 操作系统下实现了客户端通信进程, 在 windows 操作系统下实现了服务器端通信进程。

## 5 结语

本文将具有简单、开放、可扩充、灵活、自描述等特性的 XML 和基于 TCP 和流式 Socket 的可靠的 Client/Server 通信模式相结合, 提出了一种基于 XML 的实时海洋数据传输网络系统架构, 对 XML 在海洋信息系统中的应用做出了有益的探索和尝试。基于 XML 的海洋实时数据网络传输系统有效地屏蔽了数据来源、平台特性、数据格式等导致数据集成难度大, 数据利用率低的根源, 实现了开放、可靠、高效的数据传输, 有极高的应用价值。XML 在海洋信息领域中的应用前景非常广阔, 但是, 只有制定统一的海洋 XML 标准, 才能使美好的前景变为现实。否则, XML 在海洋信息领域中将会沦为一种普通的数据格式, 无法发挥它的优点。因此, 今后将深入细致地研究海洋数据的 XML 封装技术及基于海洋 XML 的数据展示技术。

### 参考文献:

- [1] 罗军舟, 黎波涛, 杨 明, 等. TCP/IP 协议及网络编程技术 [M]. 北京: 清华大学出版社, 2004, 10.
- [2] Keith Ballinger. Net Web Services 架构与实现 [M]. 北京: 中国电力出版社, 2004, 12.
- [3] Pillich B, Schack C. Next generation ECDIS for commercial and military uses Oceans '02 MTS/IEEE Volume 2, 29~31 Oct. 2002, 2 (s): 1025~1032.
- [4] GB 12460-90 海洋数据应用记录格式.
- [5] 陈继香. XML 在海洋数据服务领域的应用研究 [J]. 海洋通报, 2004, 23 (2):

## THE SYSTEM OF THE REAL TIME MARINE DATA TRANSMISSION NETWORK BASED ON XML

Ji Pen<sup>1</sup>, ZHANG Cheng-hui<sup>1</sup>, SUN Ru-peng<sup>2</sup>

(1. School of Control Science and Engineering Shandong University, Jinan 250061 China; 2. School of Computer Science and Technology Shandong University, Jinan 25006 China)

**Abstracts:** The necessity to access the ever greater quantities of marine data to support a broad range of marine information products, including ocean disaster prediction and forecast, has been increasing rapidly over the recent years. At the same time, the capability of data integration and the data use utilization in marine information fields have not achieved for such needs. This was mainly caused by the complexity of marine data, the multiformity of data source, the difference of the information platforms and the nonstandard format of data. To overcome these problems, authors proposed a structure of the system of the real time marine data transmission network based on XML combined with Socket interface. The system unified the format of data integration and data exchange, shielded the effect factors of data integration and the data use utilization. Authors commit themselves to apply the XML to the development of marine information products, aiming to boost the application of XML in national marine information fields.

**Key Words:** XML; Socket; TCP; Data transmission; Data exchange