

# ASNLIS 系统数据库平台转换

程 军

(中国科学院文献情报中心 北京 100080)

**【摘要】** 主要讲述了中国科学院文献情报中心自主开发的 ASNLIS 图书馆自动化系统后台数据库的转换过程。介绍了操作系统和数据库的选型,以及转换中遇到的问题和解决方法。

**【关键词】** 图书馆自动化系统 Linux PostgreSQL

## The Database Transfer of ASNLIS System

Cheng Jun

(The Documentation and Information Center of CAS, Beijing)

**【Abstract】** ASNLIS is a library automation system that developed by the documentation and information center of CAS. This article uses ASNLIS as an example and introduces the database transfer of it. It includes the choices of operation system, database system, and the solutions of some problems in the procedure.

**【Keywords】** Library automation system Linux PostgreSQL

### 1 问题提出

中国科学院文献情报中心计算机室开发的 ASNLIS 系统(图书馆自动化系统),所使用的后台数据库是 MS SQL Server 6.5。开发过程完毕以后,在推广过程中出现了以下问题:

• MS SQL Server 运行环境要求比较高,必须有 Pentium II 档次的机器,较好的配置,并运行 Windows NT 操作系统。

购买正版 Windows NT 操作系统和 MS SQL Server 需要的较多经费,应用部门较难负担。

为此,我们希望能够找到一种较好的解决方法。

### 2 解决方案

#### 2.1 操作系统的选择

在计算机飞速发展的今天,新技术层出不穷。在长期的技术跟踪中,我们发现 Linux 的出现,对于广大的小型用户来说,无疑是天上掉下的一个“馅饼”,它是一个完全免费的 Unix 操作系统,现在世界各

地有数十万自愿者在为这个充满魅力的操作系统的发展贡献自己的才能。它所具有的优点有:

(1)免费软件。用户可以从 Internet 网上下载,也可以很便宜地购买到 Linux 的磁盘或 CD-ROM。

(2)提供全部源代码。Linux 的最伟大的一个特性就是它的全部源代码都是免费公开的,这包括整个系统核心,所有的驱动程序,开发工具包以及所有的应用程序。这样做的好处是:排错变得非常容易,因此系统的健壮性要强于大多数商业操作系统。

(3)强大的网络功能。Linux 在最新的核心内包含的基本协议有 TCP, IPv4, IPv6, AX. 25, IPX, DDP (Appletalk), NetBEUI, Netrom 等等。可以提供 Netware 的客户机和服务器,也可以和 Windows 系统共享资源。

(4)能够提供众多的网络服务。在 Linux 上,可以设置 WWW 服务器、代理服务器、FTP 服务器、BBS 服务器、新闻服务器。现在已经有许多网站使用 Linux 做各种服务。

(5)丰富的软件。Linux 大概是应用软件最多的 UNIX 操作系统了,每时每刻网络上都有 Linux 的新软件开发出来,可以预见在 Linux 上用户的工作会越来越方便。

(6)其它优点还有:支持的硬件非常多,只要不太冷僻的设备,在 Linux 中都可以找到相应的驱动程序;支持多种文件系统,几乎所有现有的文件系统都可以装载为系统的一个目录;支持多个虚拟控制台,实现了多平台、多用户、多任务等。

#### 2.2 数据库系统的选择

选好操作系统后,就可以选择它所支持的数据

库系统了。Linux 平台上的免费数据库系统主要有：PostgreSQL、MySQL、MiniSQL。经过调查发现，三个系统差别不大，但 PostgreSQL 的优势有：

- 开发时间更长，已有 15 年历史了，在经历了长时间的考验之后，拥有数百万的用户；
- 更完善，支持 ISO SQL, ANSI SQL/98, SQL/92 和 ANSI SQL/89；
- 它同样也是开放源码开发模式，完全免费；
- 可在 UNIX、PC 及 Linux 等多种平台下运行，很多人甚至把他看成是免费的“ORACLE”；
- 支持大对象，并且支持历史记录的保存与检索；

因此，我们感觉如果要真的开发一个大型数据库应用，PostgreSQL 是一个很好的选择。

### 3 实施步骤

选择完操作系统和数据库平台之后，我们就开始着手进行实验。Linux 操作系统也有很多版本，我们测试了 Xteam Linux、红旗 Linux、还有 Red Hat Linux，感觉还是 Red Hat Linux 使用比较方便，附带功能较多。安装过程如下：

#### 3.1 Red Hat Linux 的安装

##### (1) 机器配置

CPU	Cyrix 166
内存	32MB
硬盘	1.2GB
主板芯片组	SIS 5513
显示芯片组	SIS 5596
显示器	14 寸普通彩显
网卡	D-Link 220P

##### (2) 准备硬盘

用 PQMAGIC 给一个 1.2G 的硬盘进行了分区。

分区	种类	容量
DOS	主分区	30MB
Linux	主分区	15MB
Linux Swap	扩展分区	64MB
Linux	扩展分区	1150MB

##### (3) 安装系统及应用软件

有了上面的准备以后，安装过程比较简单，只需根据提示进行，其中要注意的事项有：

- 选择自定义安装，在详细设置中注意选中 PostgreSQL 下的所有选项。

- 选中 WWW 服务和 FTP 服务，在管理中比较有用。

##### (4) 后期系统参数配置

##### ① 网卡配置

由于 D-LINK 网卡无法直接检测到，所以只能进行手工使用 Netconfig 进行配置。

##### ② 显卡配置

由于显示内存是共用的系统内存，Linux 检测不到，所以无法配置，因此 Xwindow 也就无法使用，但这并不影响作为服务器进行应用。

#### 3.2 PostgreSQL 的安装

##### (1) 服务器端的安装

服务器端数据库的安装比较简单，因为在 Red Hat Linux 中已经集成了 PostgreSQL。当安装 Linux 时，只需选中此软件包，就会在安装操作系统时，也安装好数据库。

安装完成以后，需要进行以下操作：

##### ① 将数据库服务器加入启动文件，需要运行：

```
chkconfig --add postgresql
```

##### ② 初始化数据库：

```
/etc/rc.d/init.d/postgresql start
```

此操作将会

- 建立数据库用户 postgresql；
- 建立数据库系统目录/var/lib/pgsql/basc，用户的所有数据库将会放在这里；

- 建立模板数据库/var/lib/pgsql/basc/template1；
- 在/var/lib/pgsql/basc 中建立全局类(global classes)；
- 其它一些初始化工作。

##### (2) 客户端的安装

客户端需要安装 PostgreSQL 的 ODBC 驱动和 PowerBuilder 开发环境。

##### ① PostODBC 需要从网上下载，可以查看以下几个站点

- <http://www.postgresql.org>
- <http://www.openlinksw.com>
- <http://www.insightdist.com/psqlodbc>

下载后在 Windows 9X 中安装。

##### ② 安装 PowerBuilder 进行数据库测试，在实际应用中无需安装。

#### 3.3 客户端与数据库的连接

##### (1) 数据库端设置

数据库服务器安装完毕以后，还要修改其中的 pg\_hba.conf 文件，此文件用于限制网络用户对数据库的访问。在文件中加入下面一行参数：

```
host all 159.226.100.0 255.255.255.192 trust
```

这样所有相同网段的用户都可以访问数据库了。

##### (2) 客户端设置

在 PostgreSQL 的 ODBC 中，需要将 Driver 和 Data-Source 中的 ReadOnly 选项取消掉，这样才能对数据库进行写操作。

#### 3.4 数据库从 NT SQL Server 到 PostgreSQL 的转移

做好以上准备工作以后，（下转第 57 页）

# 标记语言及HTML和XML的比较分析

王伟

(东南大学情报研究所 南京 210096)

**【摘要】** 从介绍标记语言的发展历史入手,对HTML和XML分别进行简述和比较分析。

**【关键词】** 标记语言 HTML XML

## The Expression of Markup Language of HTML and XML

Wang Wei

(Institute of Information, South-East University, Nanjing)

**【Abstract】** This paper introduces the history of markup language, and expresses the concept, characters, and structure of HTML and XML.

**【Keywords】** Markup language HTML XML

当今世界是信息管理与共享起着主要作用的网络化信息虚拟世界,其中WWW是帮助人们实现这一虚拟社会的最有力的工具。WWW不仅给计算机领域带来历史性变革,而且为人类开辟了最新获取发布信息的便捷途径。在WWW中,Web发布是其中核心组成部分,而Web发布的实现技术是标记语言。本文从介绍标记语言的发展历史入手,对HTML和XML分别进行简述和比较分析。

### 1 标记语言的发展历史

早在60年代,出现了使用描述性标签的描述编码来描述电子文档格式,其中最具代表性的是IBM公司发明的GML(Generalized Markup Language,通用标记语言)。1980年,ANSI的CLPT委员会和GCA和Genlode委员会在GML的基础上,制定了SGML的草案,1986年,SGML(Standard Generalized Markup Language,标准通用标记语言)成为信息处理、文本和办公系统的国际标准ISO8879。

SGML是一种元语言(Matelanguage),即SGML是描述标记语言的语言。SGML的实质是强调描述性标记,引入文档类型定义(DTD)概念,具有平台独立性,所以,SGML具有结构化、确认性和可扩展性三个特点。但是,庞大、复杂且严格定义的

规范说明令人对SGML望而却步,这使得SGML的使用和推广受到很大的限制。

1990年,Tim Berners Lee采用一组SGML的DTD标记标签作为最初的HTML,并用在Web浏览器的排版中。1992年,HTML1.0的发布表明HTML正式作为SGML的一个小标签集应用于Web中。随着Web应用的不断增长,HTML得到了很大的发展,1997年,HTML4.0提供了更多便于网上用户访问网络的功能。然而,即使是HTML4.0也仅提供的是有限量的标签集,这种标签集的有限性无法满足Web上日益增长的各种多元化信息的描述需求,结果,XML开始登上Web舞台。

XML是SGML的一个严格但简化的子集,它是专门的Web应用标记语言。1996年,W3C编制并推出XML草案,1998年,XML1.0正式发布。作为与SGML同样的国际标准,XML受到越来越多的广泛关注。

### 2 HTML

HTML(Hyper Text Markup,超文本标记语言)是用于创建Web页和Web信息发布的第一个通用语言,它提供跨平台的文档共享。严格讲,HTML是描述型语言,而不是编程语言,它是SGML的一个应用。HTML文档以纯ASCII的形式(即纯文本形式)存储,以标签(Tag)来定义文档

的组织。在HTML文档中,可以嵌入其它对象,如image、audio、video、javascript等,通过URL还能实现Web节点间的超链接。

## 2.1 HTML的功能

到目前为止,HTML已发展到4.0版本。每种新版本都在前版本的基础上增加新功能。HTML0中是最初、且最基本的标签集合,如html、head、title、body、br、hr、p等标签。HTML1.0中增加了页面内部显示图像和加亮文本的标签,如。HTML2.0中增加了表格(Form)标签,实现了页面的动态交互功能,是HTML由被动访问转向主动访问。HTML3.0中又增加了许多新标签和新特性,如显示表(Table)、java applet 环绕图像标签、数学公式、标题显示区、字符位置控制以及制作动态菜单和工具栏等功能。HTML4.0中增加了更多的支持多媒体选项、样式表(css)、脚本语言等。

总之,HTML所提供的功能能够满足许多Web信息发布的需要,如发布在线文档,文档中包含图像、列表、选择框等内容,通过URL实现远程Web节点在线链接,提供交互查询提问单,甚至还嵌入电子表格、视频、音频以及各种应用程序等内容。

## 2.2 HTML的特性

1)通用性:HTML作为WWW中共同的信息描述方式,可以实现不同平台的文档共享。

2)可扩展性:HTML的标签集合可以根据新需求而不断修正或加入有限的新标签符,为实现有限的新功能的扩展提供保障。

3)创建的灵活性:HTML文档是纯文本文件,它可以由UNIX的vi、DOS的edit、WPS、WORD以及专门的HTML编辑器等各种各样的编辑工具进行创建,并在WWW浏览器上都可以运行。

## 2.3 HTML文档结构

HTML的文档标签主要分为八类:格式控制标签、字体控制标签、加载图片标签、定义超文本链接标签、特殊字符标签、颜色背景控制标签、表格标签、交互表格标签。

HTML文档格式:

```
<HTML>
<HEAD>
  <TITLE> </TITLE>; <STYLE> </STYLE>; <
LINK>; <BASE>;
  <META>; <SPOT>; <INSINDEX>; <RANGE
>; <NEXTID>
</HEAD>
```

```
<BODY Bgground="XXXXXXXX" Color="#####"
#">
  <Hx> </Hx>; <P> </P>; <OL> </OL>; <UL
> </UL>; <PER> </PER>;
  <DL> </DL>; <DD> </DD>; <BR>; <HR>; <
LI>; <MENU>
  <B> </B>; <I> </I>; <U> </U>; <SUB> </
SUB>; <SUP> </SUP>;
  <TT> </TT>; <BIG> </BIG>; <SMALL> </
SMALL>
  <IMG SCR="">
  <TAB> </TAB>; <TH> </TH>; <TD> </TD>
  <FORM> </FORM>; <FIG>; <INPUT>; <A> </
A>
</BODY>
</HTML>
```

## 3 XML

XML是eXtensible Markup Language(扩展标记语言)的简写形式,它是Web中的新一代标记语言。XML是SGML的一个简化但严格的子集,它与SGML同属于元标记语言。同HTML相比,XML不是拥有太多固定的标记,而是采用样式表描述规则的方式,允许用户根据需要自我创建自定义标记,创建的标记只需要在样式表中利用规则说明其执行动作就可以了,因此,这样做将能够更大范围地满足Web上的日益增长的对多元化信息描述的需求。

### 3.1 XML的特性

XML继承了SGML的三个特性,同时也具有HTML的灵活性和通用性:

1)确认性:应用程序不用外部定义概念就可以检查XML文档结构的正确性和有效性。

2)结构化:XML文档将内容与格式分开描述,并利用样式表中的规则集对所描述的内容文档的格式进行严格的说明,这样,XML的描述就象数据库一样具有了结构性。

3)可扩展性:XML允许用户根据需要自我创建自定义标记,创建的标记只需要在样式表中进行符合格式规则集的说明即可,很显然,这样的扩展性不会象HTML那样受到扩展量的局限。

4)通用性:XML是专门为Web设计的通用标记语言,它比HTML更具有通用性。

5)灵活性:XML的灵活性表现在两个方面,一是XML文档也是纯文本文件,同HTML一样,各种编辑工具创建的XML文档都能被WWW浏览器所显示;二是XML允许自定义标签,这种优势使得HTML应用所无法与XML的应用相比。

### 3.2 XML 的四个主要技术

1)XSL 是 eXtensible Style Language(可扩展样式语言)的简写,它是专门为 XML 文档所设计的高级样式表语言,XSL 包含一个小而严格的规则集,用来说明 XML 文档中标签的行为格式,Web 浏览器只有通过 XSL 的格式翻译,才能阅读 XML 文档的内容。

2)DTD 是 Document Type Definition(文档类型定义)。在 DTD 中,提供包含文档的元素、标记、属性和实体的清单以及其相互关系,这样,用户可以十分方便地为某些特殊应用(例如化学公式、数学符号公式等)专门的标记语言。DTD 分为内部 DTD 和外部 DTD。

3)XLL 是 eXtensible Linkage Language(可扩展链接语言),它提供 XML 文档之间的链接。XLL 由 Xlinks 和 Xpointers 组成,其中 Xlinks 定义如何进行文档间的链接,且链接可以是双向的,也可以是多向的,当然,Xlinks 也可以使用 URL 来实现链接。Xpointers 是一种链接手段,它定义了如何寻址文档,即链接到特定位置的特定文档,甚至是特定文档中的特定部分。

4)XML 应用程序是指“XML 允许创建由标准工具阅读和理解的很多规则形式的语言”,并专门为某特殊领域所用。例如,用户利用化学标记语言(CML),可以将复杂的分子结构、光谱化学分析等数据发布到 Web 上,提供给网上用户阅读。

### 3.3 XML 文档结构

XML 标签有语义和样式两种含义,语义含义便于用户创建自定义标记标签;样式含义是为自定义标记标签提供规则以便使浏览器能够识别并显示自定义标签及其表述的内容。

```
XML 内容文档:xxxx.xml
<? xml version ="1.0" standalone="yes [no]">
<! DOCTYPE document_name [
  <! ELEMENT document_name ANY>
  <! ELEMENT tag_name 1(#PCDATA)>
  .....
  <! ELEMENT tag_name n(#PCDATA)>
]>
<document_name>
<head>
```

```
自定义标签 tag_name(1-m)及文本内容;
</head>
<body>
  自定义标签 tag_name(m-n)及文本内容;
</body>
</document_name>
XML 样式文档:xxxx.xsl
<xsl>
  <rule>
  <root/>
  根元素规则定义说明;
</rule>

  <rule>
  <target-element type="tag_name 1"/>
  自定义标签元素 1 规则定义说明;
</rule>
  .....
  <rule>
  <target-element type="tag_name n"/>
  自定义标签元素 n 规则定义说明;
</rule>
</xsl>
```

## 4 结束语

XML 是新一代 Web 语言,是勾通 Web 与现实世界的里程碑。但是,在未来的几年里,HTML 将仍然存在,而且 HTML 和 XML 之间的相互转换问题将成为 Web 标记语言研究中的一个主题。HTML 和 XML 是相互补充的,HTML 是关于用户显示界面和交互界面的,而 XML 是关于数据和信息的内容描述的。因而,XML 可以向 HTML 文档中增加信息,HTML 则显示以 XSL 格式所表达与描述的数据和信息,结果,HTML 和 XML 将共同为 Web 提供十分丰富的数据和信息的显示机制。

#### 参考文献

1 Elliottc Rusty Harold 著. XML 实用教程. 机械工业出版社

### 《现代图书情报技术》2000年(增刊)征订通知

《现代图书情报技术》2000年(增刊),30万字,已经出版,每册定价:20元(含邮费),如欲订购,请直接将款汇至北京中关村科学院南路8号 邮编100080 本刊编辑部 收,款到后即将发票随刊一并寄出。印数有限,欲购从速!

(编辑部)