

科技部科技基础性工作专项资金重大项目 研究成果

项目名称：我国数字图书馆标准规范建设

子项目名称：我国数字图书馆标准规范发展战略与基本框架

项目编号：2002DEA20018

研究成果类型：研究报告

成果名称：数字图书馆支撑技术领域标准规范的现状和发展

成果编号：CDLS-S01-007

成果版本：总项目组推荐稿

成果提交日期：2003年7月

撰写人：牛振东（中国数字图书馆有限责任公司）

师雪霖（北京理工大学）

叶成林（中央党校图书馆）

项目版权声明

本报告研究工作属于科技部科技基础性工作专项资金重大项目《我国数字图书馆标准规范建设》的一部分，得到科技部科技基础性工作专项资金资助，项目编号为 2002DEA20018。按照有关规定，国家和《我国数字图书馆标准规范建设》课题组拥有本报告的版权，依照《中华人民共和国著作权法》享有著作权。

本报告可以复制、转载、或在电子信息系统上做镜像，但在复制、转载或镜像时须注明真实作者和完整出处，并在明显地方标明“科技部科技基础性工作专项资金重大项目《我国数字图书馆标准规范建设》资助”的字样。

报告版权人不承担用户在使用本作品内容时可能造成的任何实际或预计的损失。

作者声明

本报告作者谨保证本作品中出现的文字、图片、声音、剪辑和文后参考文献等内容的真实性和可靠性，愿按照《中华人民共和国著作权法》，承担本作品发布过程中的责任和义务。科技部有关管理机构对于本作品内容所引发的版权、署名权的异议、纠纷不承担任何责任。

《我国数字图书馆标准规范建设》课题组网站 (<http://cdls.nstl.gov.cn>) 作为本报告的第一发表单位，并可向其他媒体推荐此作品。在不发生重复授权的前提下，报告撰写人保留将经过修改的项目成果向正式学术媒体直接投稿的权利。

我国数字图书馆标准规范发展战略与基本框架

目 录

1. 概述	1
2. 计算技术	1
2.1 并行计算.....	1
2.1.1 PVM (Parallel Virtual Machine) 标准	1
2.1.2 MPI标准 (Message Passing Interface) 标准	2
2.2 分布式计算.....	2
2.2.1 公共对象请求代理体系结构.....	3
2.2.2 Microsoft的COM/DCOM组件模型	4
2.2.3 SUN的EJB组件模型	5
2.3 移动计算.....	8
2.4 网格计算.....	8
3. 网络技术	10
3.1 IPv6.....	10
3.1.1 IPv6 (Internet协议第六版) 规范	11
3.1.2 IPv6 与IPv4 的比较	12
3.2 宽带网络.....	12
3.2.1 IP over ATM	13
3.2.2 IP over SONET/SDH.....	13
3.2.3 IP over WDM	15
3.3 无线网络.....	16
3.3.1 IEEE 802.11 标准.....	17
3.3.2 IEEE 802.11a标准.....	17
3.3.3 IEEE 802.11b标准.....	17
3.4 网络存储.....	20
3.4.1 NAS	20
3.4.2 SAN	21
3.5 信息资源检索协议.....	22
3.5.1 ISO 10162 和ISO 10163	23
3.5.2 Z39.50.....	23
3.5.3 ISO 23950	24
4. WEB技术	24
4.1 基本WEB技术.....	24
4.1.1 HTML.....	25
4.1.2 HTTP	25
4.1.3 MIME	25

4.1.4	URL.....	25
4.2	动态WEB技术.....	26
4.2.1	Microsoft的.NET平台.....	26
4.2.2	SUN的J2EE平台.....	27
4.3	主流WEB服务器.....	31
4.3.1	IBM的WebSphere.....	31
4.3.2	Microsoft的IIS.....	31
4.3.3	BEA的Weblogic.....	31
4.3.4	Apache服务器.....	32
4.3.5	Tomcat服务器.....	32
4.3.6	Raisin服务器.....	32
4.4	WEB知识发现.....	32
4.4.1	Web知识发现的任务.....	33
4.4.2	Web知识发现方法.....	33
5.	数字对象和数据库.....	33
5.1	字符编码.....	33
5.1.1	ISO/IEC10646.....	33
5.1.2	GB18030-2000.....	34
5.2	多媒体信息编码.....	35
5.2.1	PEG.....	36
5.2.2	JBIG.....	38
5.2.3	MPEG.....	38
5.2.4	MHEG.....	40
5.3	数据元素.....	40
5.3.1	ISO/IEC 11179 系列标准.....	42
5.3.2	ISO/IEC TR 20943 系列技术报告.....	43
5.3.3	业务对象登记框架.....	43
5.4	标记语言.....	44
5.4.1	SGML.....	44
5.4.2	HTML.....	45
5.4.3	XML.....	45
5.5	元数据.....	46
5.5.1	数据资源描述的层次.....	46
5.5.2	元数据的标准规范.....	46
5.5.3	数据登记系统.....	47
5.5.4	元数据管理机制.....	47
5.6	数据库技术.....	47
5.6.1	分布式数据库.....	47
5.6.2	多媒体数据库.....	49
5.6.3	海量数据库.....	50
5.6.4	数据仓库.....	50
5.7	数据挖掘.....	51
5.7.1	数据挖掘的理论基础.....	51

5.7.2	商用数据挖掘系统.....	52
6.	信息安全	52
6.1	私钥加密.....	53
6.2	公钥加密.....	54
6.3	数字签名.....	55
6.4	认证技术.....	56
6.4.1	Public Key Infrastructure	56
6.4.2	X.509	56
6.4.3	Kerberos	56
6.5	数字版权管理.....	57
6.6	数字水印.....	57
6.7	安全传输标准.....	58
6.7.1	IPSec.....	58
6.7.2	Secure HTTP	58
6.7.3	Secure Sockets Layer	58
6.7.4	Transport Layer Security	59
6.7.5	Wireless Transport Layer Security	59
6.7.6	Secure Electronic Transaction	60
6.7.7	Secure Wide Area Network	60
6.7.8	Secure Shell	60
6.7.9	Secure Mail.....	61
6.8	XML安全机制.....	61
6.8.1	XMLEnc	61
6.8.2	XMLSig.....	61
7.	软件工程	62
7.1	软件开发技术.....	62
7.1.1	面向对象.....	62
7.1.2	面向主体.....	63
7.2	过程管理方法.....	66
7.2.1	ISO/IEC 12207	66
7.2.2	能力成熟度模型.....	67
7.2.3	ISO9000.....	68
8.	结束语.....	69
	参考文献	69
	附录.....	72

1. 概述

数字图书馆可非正式地定义为有组织的信息馆藏以及相关服务，信息以数字化形式保存，并通过网络进行访问。定义的核心在于说明信息是有组织的。数据一旦经过系统化的组织，便成为数字图书馆的馆藏。数字图书馆包含各式各样的可用数据，供不同用户使用，规模可大可小，并可使用各类计算机设备和相关软件。但是所有数字图书馆都有共同的特点：信息在计算机内得以组织并通过网络加以利用，并带有选择信息、组织信息、存储信息和发布信息的程序。在某些方面，数字图书馆可能与传统图书馆非常不同，但是在另一个方面它们又极为相似。

建立数字图书馆的主要原因是人们相信数字图书馆能够比过去的模式更好地传送信息^[7]。传统图书馆是社会的重要组成部分，但还欠完美。当前，计算机和网络已经改变了人类彼此交流的方式。对某些学科，专家学者从连接于网络的个人计算机上得到的服务会优于直接造访图书馆。以往仅供专业人士利用的信息现在可以为大众所利用。从一台PC上，用户可以参考世界上其他计算机中存放的资料。另一方面，印刷文档在人类文明中扮演着信息存储和传递的重要角色，这种地位不会消失，只是在逐渐演变。

建设数字图书馆的最终目的是实现数字资源的共享，这就离不开统一的标准。标准的选择恰当与否直接影响到系统的实现以及将来运行和维护。数字图书馆的发展应该在共同认可的标准规范约束下，保障数字资源的最大可用性、系统之间的或操作性和集成性。

数字图书馆的支撑领域涉及到计算技术、网络技术、Web 技术、数字对象和数据库、信息安全等。本文对上述领域中的主要标准进行了分析比较，同时探讨了各领域标准的发展趋势。

2. 计算技术

2.1 并行计算（Parallel Computing）

随着信息资源的不断膨胀，对计算机的处理能力、速度提出了更多的要求。而目前芯片处理速度的提高远远不能满足这一需求，因此并行计算成为解决这一瓶颈的有效途径。并行计算是实现高性能、高可用计算机系统的主要途径。并行计算技术经过多年的发展，虽然取得了显著的进步^[18]。目前，并程序开发和移植的成本较高国际上并行计算工业标准或事实上的标准主要有PVM、MPI两种。

2.1.1 PVM（Parallel Virtual Machine）标准

PVM（并行虚拟机）是这样一个软件系统，它可以使多个异构计算机进行协作和灵活的并发计算。每个计算机可以是多处理器计算机、向量超级计算机、图

形工作站或者是标量工作站，这些计算机通过诸如以太网、FDDI等各种网络连接起来。PVM支持软件在用户配置池中的任何计算机上执行，并且为并发程序提供统一的、强大的计算环境^[1]。

通过调用 PVM 类库，用户的程序可以访问 PVM。这样用户可以控制程序组件的执行位置。PVM 系统透明地处理消息路由、数据格式转换等在异构网络环境下必须进行的操作。

2.1.2 MPI 标准 (Message Passing Interface) 标准

该标准是由很多并行计算用户和厂商制订的，定义了处理器之间基于消息的通信规范。采用MPI标准可以保证并行程序的可移植性。MPI标准定义了第三方软件开发应该遵循的标准框架^[2]。

和 PVM 相比，MPI 具有如下优势：

- a. MPI 有更多的实现方式；
- b. MPI 定义了第三方的实现规范；
- c. MPI 完全支持异步通讯；
- d. MPI 集群具有更好的稳定性、效率和确定性；
- e. MPI 对消息缓冲池的管理更好；
- f. MPI 同步有效地保护了第三方用户程序；
- g. MPI 更易于移植。

2.2 分布式计算 (Distributed COMputing)

原先的客户/服务器体系结构是为局域/广域、企业级网络而设计的，它能为某些模式相对固定的应用和相对稳定的数据需求提供服务。随着网络技术和面向对象技术的发展，全球的服务器和数千万的用户正在互相连接起来，这就要求当前的计算机体系发生变革以适应这一变化，从而满足一个几乎完全无序的超大规模的网络环境。分布式对象技术就是在这样的环境中产生的。

分布式对象技术的基本概念之一是所谓构件 (COMponent) 概念。构件可以跨越平台、网络、语言、应用程序、工具、硬件而运行。随着Internet的发展，我们可以预见新型的应用体系结构正在产生，其生成、创建、开发、运行和管理的方式及其功能都与局域网的应用程序不同。它将彻底改变生产软件的传统模式，目前普遍认为在Internet网上开发未来的软件，应当以构件为基本单位，软件开发可以边在网上订购所需的构件，一边构造目标软件。而从网上购得的构件，由COM和CORBA这样的标准软总线提供运行所需的各种服务。分布式对象技术追求的目标是软件无缝连接，即插即用。构件不仅可重复使用和提高效率，还可以节省硬件资源，如果采用了构件程序设计思想，可以随时更换有问题的模块，以降低其维护成本。

目前分布式对象技术正处在工业标准形成阶段。它包括有Microsoft公司提出的对象构件模型COM（COMponent Object Model）、OMG（Object Management Group）提出的CORBA（COMmon Object Request Broker Architecture）以及由Sun公司推出的JavaBeans。这三个标准类似于构件运行的软总线。

2.2.1 公共对象请求代理体系结构（CORBA）

公共对象请求代理体系结构（CORBA）是一个语言和平台中立的规范体，它用于建立分布式对象应用程序。CORBA 是通过一组标准描述的，这些标准是由一个称为对象管理组织（OMG）的组织创建的。

CORBA代表着一种标准的模型，利用它可以创建服务能够发布给远程客户的对象和组件。CORBA使用了一种标准的通信模型，利用这种模型以异构的语言实现并且运行在异构的硬件和操作系统平台之上的客户和服务端可以进行交互^[8]。

CORBA 应用程序以这样一种方式建立：应用程序与通信代码的细节大体上是隔离的。事实上，CORBA 应用程序是这样定义的：从 CORBA 客户的角度来讲，CORBA 服务器的分布式本质可以是完全透明的。公共对象服务规范（CORBA services）、公共设施体系结构（CORBA facilities）和 CORBA 业务对象都是建立在 CORBA 之上的标准，它们提供了十分丰富的分布式通信服务和框架。CORBA 接口定义语言（IDL）提供了一个中型的语言机制，它利用已经存在的一些把 IDL 映射为 Java 和把 Java 映射为 IDL 的标准来定义分布式对象接口。CORBA 对象现在有能力通过引用或通过值的方式被传递。

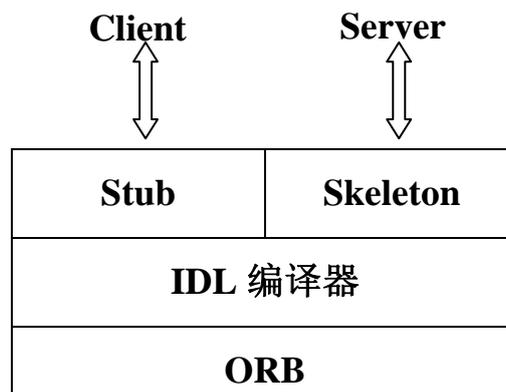


图 1 CORBA 组件模型

CORBA 组件模型的底层结构为 ORB。一个 CORBA 组件采用 IDL 进行描述。CORBA 提供了 IDL 到 C、C++、Java、COBOL 等语言的映射机制--IDL 编译器。IDL 编译器可以生成 Server 方的 Skeleton 和 Client 方的 Stub 代码，通过分别与客户端和服务端程序的联编，即可得到相应的 Server 和 Client 程序。图 1 展示了 CORBA 组件模型。

CORBA 同时提供了一系列的公共服务规范——COSS，其中包括名字服务、

永久对象服务、生命周期服务、事务处理服务、对象事件服务和安全服务等。它们相当于一类用于企业级计算的公共组件。此外，CORBA 还针对电信、石油等典型的应用行业提供了一系列的公共设施。

CORBA 是一种语言中性的软件组件模型，可以跨越不同的网络、不同的机器和不同的操作系统，实现分布对象之间的互操作。

2.2.2 Microsoft 的 COM/DCOM 组件模型

Microsoft的组件对象模型（COMponent Object Model, COM）和分布式组件对象模型（Distributed COMponent Object Model, DCOM）分别提供了建立依赖于Windows平台的非分布式组件和分布式组件的方法^[9]。基于Java的DCOM客户和服务端可以被建立并且位于Microsoft Java虚拟机（MSJVM）内。MSJVM提供了对二进制COM和DCOM库的链接以及与Windows平台进行交互。

Microsoft 的 COM 组件模型一直在很稳定的发展中，舍弃繁杂的 OLE 细节后，COM 才真正奠定了 Windows 组件模型的核心，真正的开始可以提供只做一切逻辑对象。

Microsoft 引入了组件对象模型 COM，形成了 COM 对象之间实现互操作的二进制标准。COM 规定了对象模型和编程要求，使 COM 对象可以与其他对象相互操作。这些对象可以用不同的语言实现，其结构也可以不同。基于 COM，微软进一步将 OLE 技术发展到 OLE2。其中，COM 实现了 OLE 对象之间的底层通信工作，其作用类似于 CORBA/ORB。不过此时的 COM 只能作用在单机 Wintel 平台上。在 OLE2 中，也出现了我们今天熟知的拖-放技术以及 OLE 自动化。

同时，微软在 VB 中引入了可以嵌入任何可视组件的通用模型 VBX。VBX 的主要局限在于它并不是一个开放的结构，也没有为第三方软件开发商提供 VBX 集成的标准。最后，微软将上述思想集中在一起，以 COM 作为组件通信框架。VBX 也发展为 OLE 控件 OCX 的形式。DCOM 是 COM 在分布计算方面的自然延续，它为分布在网络不同节点的两个 COM 组件提供了互操作的基础结构，而所有以 OLE 为标志的技术如今也已挂上了 ActiveX 标志。

DCOM 是 Microsoft 与其他业界厂商合作提出的一种分布组件对象模型（Distributed COMponent Object Model），其发展经历了一个相当曲折的过程。DCOM 起源于动态数据交换（DDE）技术，通过剪切/粘贴（Cut/Paste）实现两个应用程序之间共享数据的动态交换。对象连接与嵌入 OLE 就是从 DDE 引伸而来的。

在 DCOM 的客户端，DCOM 客户与 DCOM 库通信以创建 DCOM 对象接口的一个实例。DCOM 库接着查询 Windows 注册表提取如何创建和寻找这个对象的信息。这一信息可以使 DCOM 库与服务控制管理器（Service Control Manager）进行交互。DCOM 库查询服务控制管理器以激活一个远程 DCOM 对象实例。然后服务

器端的服务控制管理器使用 DCOM 库创建被请求对象的一个实例。

DCOM 开始提供远程访问和分布式计算以及对象回收机制，让 COM 组件模型能够提供企业级计算的能力。不过在 DCOM 的时代，客户端仍然是通过 Proxy/Stub 直接和 COM 对象互动，还未达到像 EJB 组件模型一样由虚拟服务器控管，以提供系统服务等功能。但是 Microsoft 很快地在 MTS1.0 中正式加入了这个功能，至此 COM 组件模型才能够顺利地加入企业核心服务，例如 Object Pooling, Role-Based 安全权限和交易管理等功能。严格地说，在 MTS 出来之后，COM 组件模型才有资格成为关键性系统的核心组件模型，也因为 MTS 才有后来的 Microsoft DNA 架构。在 Windows 2000 中，MTS 正式成熟演进到了 COM+1.0，除了把 MTS 调整得和操作系统更契合外，最重要的进步是把执行效率大幅提升。

DCOM 实际上是一个自底向上增长的组件模型。也就是说，DCOM 是从 COM 发展而来，COM 是从 OLE 发展而来，它们都是从当前的 Windows 平台体系结构发展而来。COM 提供了一个组件模型，DCOM 允许用户建立 Microsoft Windows 平台相关的分布式服务，所以，DCOM 客户和服务器对 Windows 平台就有一个双重的依赖。

从 CORBA 的观点来看，我们可以粗略地说，ActiveX 控件与 DCOM 的关系相当于 CORBA 组件与 ORB 的关系。当然，按照微软一贯的产品开发逻辑，微妙的思想都退到了幕后，而提供给开发者的是一个以 Wizard 方式生成各种应用的可视化开发环境。在公共服务方面，微软提出了自己的事务服务器 MTS (Microsoft Transaction Server) 和消息队列服务器 MSMQ (Microsoft Message Queue Server)。前者与 CORBA 对象事务服务目标类似，后者则是为了保证应用之间进行可靠的消息通讯和管理。此外，微软在网络安全方面也有一整套实用的解决方案。

2.2.3 SUN 的 EJB 组件模型

按照 Sun 和 Javasoft 对 Java 的界定，Java 是一个应用程序开发平台，它按照高性能、可移植、可解释的原则，提供面向对象的编程语言和运行环境。Java 计算的本质就是利用分布在网络中的各类对象共同完成相应的任务^[3]。例如 JavaApplet 可按用户的需求从服务器上动态地下载到客户机的浏览器上，完成 HTML 页面的动态变化。

Java 对于软件组件的观点与 CORBA 中的组件观点存在一定的区别。在 CORBA 中，CORBA/ORB 相当于一根软总线，组件可以即插即用。也就是说，从 CORBA 的观点看来，所有组件的地位相当，完全是一种平行的关系。而在 Java 中，软件组件是能够进行可视化操作的可重用软件，它满足一定的特征要求，并可以根据需要进行定制和组装。

Java 的软件组件称为 JavaBean，或者简称 Bean。按照 Javasoft 给出的定义，Bean 是能够在构造工具中进行可视化操作的可重用软件。JavaBean 的组件模型包

含组件和容器两个基本要素，这一思想在 ActiveX/DCOM 技术中同样存在。作为一种典型的组件模型，JavaBean 具有属性、方法、事件、自我检查、定制和永久性等 6 个方面的特征。其中前 3 种特征（属性、方法、事件）是面向对象的组件必须满足的基本要求，属性和方法保证 Bean 成为一个对象，而事件可以描述组件之间的相互作用以及组件与容器之间相互感兴趣的事情。通过事件的生成、传播和处理，组件相互之间关联在一起，共同完成复杂的任务。后三种特征（自我检查、定制和永久性）主要侧重于对 JavaBeans 组件性质的刻画。内省用于暴露与发现组件接口。使用内省机制，可以使组件的使用者了解到组件的属性、方法和事件。由于一个组件通常是具有一定性质和行为的对象的抽象，它往往有很大的通用性。为了在一个具体的应用环境中使用组件，必须对组件进行定制。JavaBeans 的定制通常在一个可视化生成工具中进行，通过组件的内省机制，发现组件的属性、方法和事件，然后利用生成工具提供的属性编辑器实现定制。永久性是将组件的状态保存在永久存储器中并能够一致恢复的机制。Java 通过序列化（Serialize）实现定制组件的永久性存储，通过反序列化可以实现组件状态的恢复。

JavaBean 组件的本地活动是在与其容器相同的地址空间内进行的。在网络上，JavaBean 组件可以以三种方式进行活动：

JDBC 使 Bean 组件能够访问 SQL 数据库。Bean 可以实现给定数据库中的表操作，完成相应的业务逻辑；JavaRMI（远程方法调用）使分布在网络不同地址上的两个组件之间实现互操作。组件之间的调用方式采用经典的 Client/Server 计算模型；JavaIDL 是一个 Java 版的 CORBA/ORB。通过 JavaIDL 可以实现一个 JavaBean 和一个 CORBA 服务之间的互操作。基于 JavaIDL 的 Java 组件互操作模型完全等同于 CORBA 的思想，只不过具体的编程语言采用 Java，而 CORBA/ORB 选择了 JavaIDL。

远程方法调用机制 RMI 是构成 Java 分布对象模型的基础结构。RMI 系统包括桩/框架层、远程引用层和传输层。目前，RMI 的传输层是基于 TCP 实现的，将来的 RMI 体系结构建立在 IIOP 协议之上，可以实现 Java 技术与 CORBA 技术的深层融合。应用层建立在 RMI 系统之上。

最近两年，Java 又提出了企业 JavaBean（EJB）的思想，其结构完全采用基于软件组件模型的分布对象计算体系。

企业 JavaBean 各组成部分的含义为：JDBC（Java Database Connectivity）：基于 SQL 标准 Java 数据库连接，其基本功能和设计与 ODBC 相似；JavaRMI（Java Remote Method Invocation）：Java 远程方法调用；JNDI（Java Naming and Directory Interface）：Java 名字与目录服务；JavaIDL：Java 和 COBRA 之间的连接；JTS（Java Transaction Service）：Java 事务管理服务；JMAPI（Java Management API）：Java 网络管理 API；JMS（Java Message Service）：Java 消息传递服务。

用 CORBA 的观点来看，企业 JavaBean 中包括了分布组件的基础结构，也包括了各类公共服务组件。并且由于 Java 与生俱来的跨平台性和语言的一致性，使其成为软件组件模型的一个有力的竞争者。EJB 定义了一种服务器端组件模型，它允许商务对象的开发，并可以从一种品牌的 CTM 转移到另一种品牌的 CTM。组件 (bean) 代表一个简单的编程模型，它允许开发者集中于商务目的。EJB 的服务器 (遵从 EJB 规范的 CTM) 则负责将组件生成分布式对象，管理各种负物，如事务、持久性、并发性和安全性等。

EJB 提供了一种标准分布式组件模型，它可以大大简化开发过程，并允许在一个供应商的 EJB 服务器上开发和配置的 bean 能轻易地被配置到其他供应商的 EJB 服务器上。

SUN 最近推出了新一版的 EJB2.0 功能规格，很快被 BEA 和 Borland 的 BES 实现出来。SUN 在 EJB2.0 种提出了许多先进而复杂的功能，目的是为了大幅度强化 EJB 做为企业核心组件架构的本钱，以便准备和 Microsoft 下一代的 .NET 组件在企业系统中竞争。

从 SUN 定义 EJB 的规范开始就展现了和 COM 不一样的概念，EJB 一开始便非常重视 Design Pattern 和组件种类，例如 Session Bean 和 Entity Bean 各自负责不同的角色。再借助 Java 的 Garbage Collection，因此提供了成为企业信息系统组件必要的基础。但是在 EJB1.x 中所有的 Bean Instance 之间的调用都是使用 Remote Interface 的方式，因此在许多的应用方面付出了较大的成本，导致一些情形中执行效率不佳的状况。因此在 EJB1.x 中发展出了许多的 Design Pattern 来改善这种现象，例如鼓励使用 Coarse-Grained 对象，以减少网络 Round-Trip 的 Bean 之间的调用。

在 EJB2.0 种 SUN 终于改善了这个问题。提出了 Local Interface。所谓 Local Interface 指当 Bean Instance 是在同一个 EJB Container 之中时，EJB Container 可是使用 Local Interface 调用来代替使用 Remote Interface 调用，这样可以增加 3 倍以上的对象激活效率。另外 SUN 加入 Local Interface 功能的重要原因也是为了支持 EJB2.0 种的大幅度强化的 CMP (Container-Managed Persistence) 功能。

典型的 EJB 应用模型如图 2 所示。

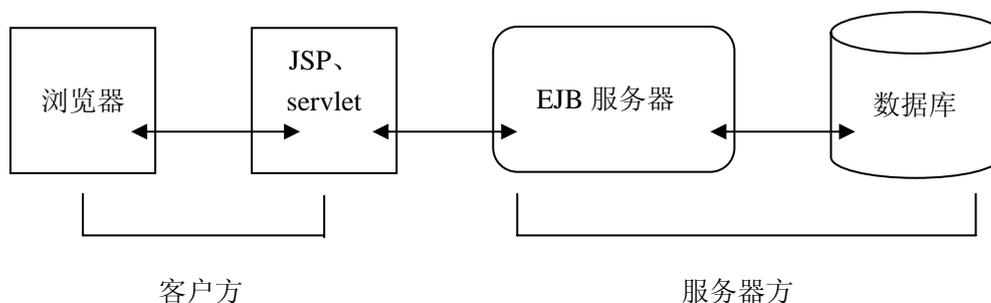


图 2 客户方和服务方

N 层体系结构或三层体系结构是对标准两层客户/服务器模型的扩展，它是通过客户端和数据库之间增加一个多线程应用服务器而实现的。Enterprise JavaBean 就在这个服务器上，而且应用的业务逻辑也在这个服务器上执行。

2.3 移动计算

“移动计算”是现代计算机通信领域中一个引人注目的新课题，它利用计算机技术和电信技术为用户提供移动的计算环境和新的计算模式，其作用在于将有用、准确、及时的信息传递到传统的企业边界之外，使企业保证其数据可提供给在任何时间、任何地点需要它的雇员和用户。随着企业在IT方面竞争的日渐激烈，以及移动计算硬件平台的技术改进和价格的不断下降，例如，顶级的手持PC在美国的市场价格已经降到了500美元左右，企业对移动计算的需求将会稳步增长。

移动计算技术可应用在如下方面：无处不在的工作范围，甚至是山谷、野外；随时随地办公，如汽车办公；产品推销；现场数据采集；股市行情分析；战场医疗；实时指挥控制；即时导航等。今年以来，国际上的一些著名的IT市场研究机构纷纷发布了市场研究报告，“移动计算”被认为是对未来最有影响的四大技术方向之一（其余为：网络基础设施，电子商务和软件重用）。据预测，到2000年，75%的商业用户将要求有一个远程或移动解决方案。

“移动计算网络（或称为游牧计算网络）”指主机或局域网可在网中漫游的计算机网络，它支持包括通话、寻呼信息、高速数据、视频在内的多媒体业务，是实现个人通信的途径之一。基于Internet的移动计算网络也称“移动Internet”，是目前该领域研究的焦点。

移动计算被认为是对未来最有影响的四大技术方向之一（其余为：网络基础设施，电子商务和软件重用）。据预测，到2001年，75%的商业用户将要求有一个远程或移动解决方案^[20]。

2.4 网格计算

网格计算（Grid COMputing）是伴随着互联网技术而迅速发展起来的，专门针对复杂科学计算的新型计算模式。这种计算模式利用互联网把分散在不同地理位置的电脑组织成一个“虚拟的超级计算机”，其中每一台参与计算的计算机就是一个“节点”，而整个计算是由成千上万个“节点”组成的“一张网格”，所以这种计算方式叫网格计算。这样组织起来的“虚拟的超级计算机”有两个优势，一个是数据处理能力超强；另一个是能充分利用网上的闲置处理能力。简单地讲，网格是把整个网络整合成一台巨大的超级计算机，实现计算资源、存储资源、数据资源、信息资源、知识资源、专家资源的全面共享^[66]。

目前，网格计算不仅在学术界、研究领域进行着深入的研究与实验，同时也得

到了来自产业界诸如IBM、HP、Microsoft、NTT、Intel、SGI和Sun等各大公司的巨资支持与商业应用开发。

网格计算系统一般由网格硬件、网格操作系统、网格界面、网格应用4层基本结构构成，其最突出的特点是资源共享、协同工作和开放性标准。也正因为如此，网格计算目前研究发展的主要障碍便是标准协议的建立和体系结构的确定。

就像TCP/IP协议是Internet的核心一样，构建网格计算也需要对标准协议和服务进行定义。目前，包括全球网格论坛（Global Grid Forum）、研究模型驱动体系结构（Model Driven Architecture）的对象管理组织（OMG）、致力于网络服务与语义WWW研究的W3C，以及Globus.org等标准化团体都开始了对网格计算的标准制定工作。

2002年7月，OMG、W3C、GGF等标准化组织与来自学术、商业领域的人士出席了“软件服务网格研讨会”，加快全球大网格标准的制定。接着，另一开放源代码网格标准组织——Globus也集会研究通过广域网联接的高性能计算的基础设施问题。Globus目前正致力于开发标准的网格架构和其他技术。

迄今为止，网格计算还没有正式的标准，但在核心技术上，相关机构与企业已达成一致：由美国Argonne国家实验室与南加州大学信息科学学院（ISI）合作开发的Globus Toolkit已成为网格计算事实上的标准，包括Entropia、IBM、Microsoft、COMpaq、Cray、SGI、Sun、Veridian、Fujitsu、Hitachi、NEC在内的12家计算机和软件厂商已宣布将采用Globus Toolkit。作为一种开放架构和开放标准基础设施，Globus Toolkit提供了构建网格应用所需的很多基本服务，如安全、资源发现、资源管理、数据访问等。目前所有重大的网格项目都是基于Globus Toolkit提供的协议与服务建设的^[68]。

我国的许多行业，如能源、交通、气象、水利、农林、教育、环保等，以及涉及科研、开发、教育的诸多部门、单位和企业，对高性能计算网格及信息网格的需求是非常巨大的。这些需求可简单地归纳如下：

在资源数据方面，国家在过去几十年已经花费了大量财力收集各种地下、地面、大气的资源数据；采用的手段包括地下勘探、地面人工测绘和监测，航空遥感，卫星遥感等。做数据收集的部门包括国土局、测绘局、水利部和防总、林业部、核工业部、地矿部、石油部、科技部（3S）、环保局、建设部、气象局、海洋局、中科院等。这些数以PB计的数据是国家的宝贵财富和资产，但目前由于零碎地分散在很多地方，共享困难，利用率很低。将这些数据组织起来放在网络上供全国使用将有很多好处。

在高性能计算方面，由科技部支持创办的五个国家高性能计算中心已经运行二年。他们一致提出急需升级计算设备。计算网格的建立有利于这些中心共享资源，在诸如基础和应用基础科学研究、汽车制造、大型水电工程、石油勘探、气

象气候、航空、交通、金融、医疗卫生等领域发挥更大的作用。

在生物信息方面，我国的生命科学界对网络有强烈呼声，已经把建立网络（称为生物信息学中心）的配套研究列为“十五计划”自然科学基金委的头等任务。石油行业将对信息技术的需求分为四类：信息应用指的是企业管理、数据库/Web, Intranet等；勘探应用指的是地震资料处理、钻井数据的收集和处理（含测井、录井）；开发应用指的是油床模拟等；设计应用指的是地面建筑、管道等的CAD。在这四类应用中，勘探和开发是计算密集型。

在环境保护方面，已经建立了若干数据收集网和计算机网。现在的问题是从这些途径得到的宝贵的数据没有统一上网共享和管理，使用极不方便，利用率很低。网络可以解决这些问题。在媒体信息方面，我国的大众媒体（报纸、通信社、电视、电台、纪录片、出版社）包含大量的信息，但缺少有效的机制让全国人民快捷、方便廉价地获得自己需要的信息。网络建立后，可以逐步提供缓存服务、目录服务、索引服务、数据挖掘和分类等多种新的应用形式。

所有上述计算网络及应用网络的建立，都需要网络系统软件进行统一的管理。网络系统软件将为广大用户提供方便的使用环境，并为各种应用网络的建立提供基础和保障^[67]。

3. 网络技术

3.1 IPv6

IP 第 6 版（IPv6）是继 IP 第 4 版（IPv4）以后，Internet 协议的一个新版本。IPv6 的出现引起了世界重要研究机构和公司的重视。目前 IETF 正在制定大量的 IPv6 相关标准，包括地址结构、域名解析、安全、自动配置、邻居发现、路由协议等方面，同时为了对 IPv6 协议特性进行研究并积累 IPv6 组网经验，IETF 于 1996 年建立了全球范围的试验床（Testbed），称作 6Bone。6Bone 是一个虚拟的网络，以隧道（tunnel）的方式通过基于 IPv4 的互联网实现互联。1998 年底，面向实用的全球性 IPv6 研究和教育网（6REN）开始启动，建立了物理的以 ATM 为中心的 IPv6 洲际网络。

1998 年 6 月我国国家教育科研网 CERNET 也加入了 6Bone，并于同年 12 月成为其骨干成员。CERNET 建立了 IPv6 试验床并在 IPv6 领域在中国开展了许多开拓性的研究。从 1999 年底，CERNET 与 Nokia 合作，启动了 Internet6 计划，准备首先在中国的若干高校搭建 IPv6 网络，形成一个大规模的 IPv6 研究和试验网络^[19]。

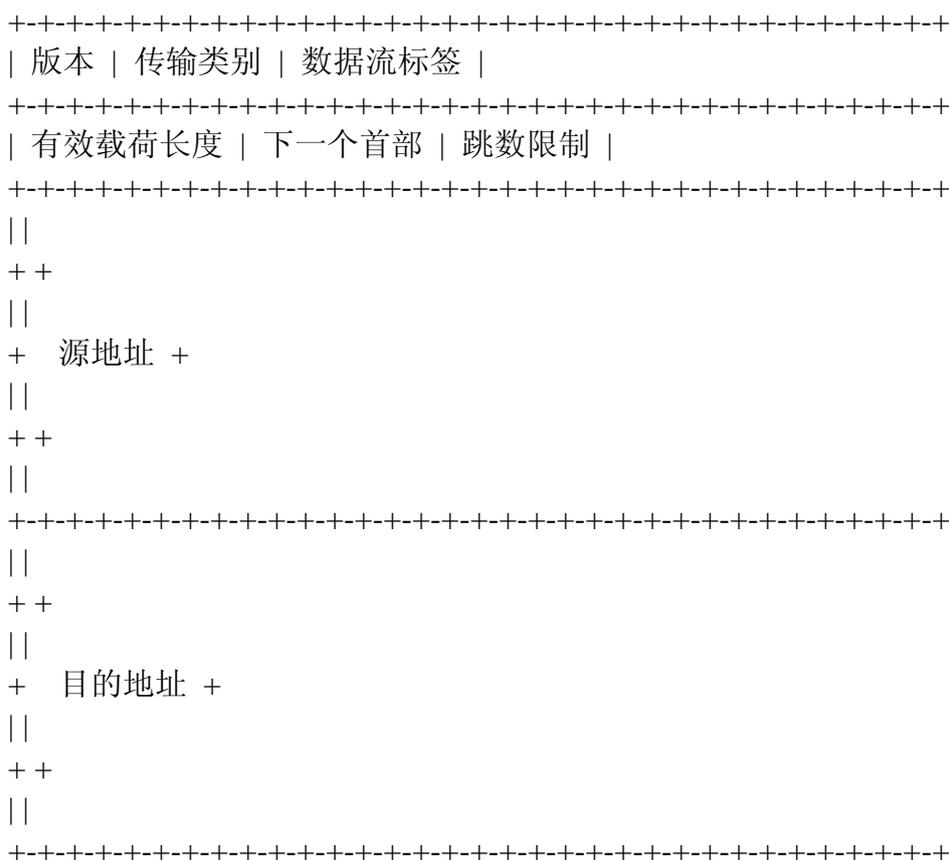
目前 FreeBSD、Solaris、Linux、Unix 上都已经有了 IPv6 协议栈的实现，同时许多大厂商宣称，即将在产品中支持 IPv6，如 Cisco、Nortel Networks、Sun、Microsoft 等。以 Cisco 为例，其路由器操作系统 IOS 从版本 12.1 已开始支持 IPv6，并且 IOS

的后续版本将加以改进，提高 IPv6 的性能，而且 硬件平台也将采用支持该协议的配置。Microsoft 也已经提供了 Windows NT 和 Windows 2000 平台的 IPv6 协议栈，并且即将在其流行的浏览器 Internet Explorer 中加入 IPv6 的支持。

当然，IPv6 的最终实现还需要很长时间。据国际著名 ISP 和权威人士估计，2003 年 IPv6 网络将进入大规模实施阶段，之后 IPv4 和 IPv6 将保持长时间共存，并最终过渡到 IPv6。如何实现 IPv4 向 IPv6 的平稳过渡是 IPv6 首先必须解决的问题，同时也是成功的关键。而且，在技术层面上，IPv6 目前只完成了初步的标准化工作，其中很多新特性在 IETF 仍处于研究、制订标准草案阶段，比如 DNS 发现、流标识的语义、IPv6 在 3GPP 中的应用、API、域间组播路由、以及地址分配等。IPv6 在标准化方面的进展对实现大规模的商业化应用非常重要。

3.1.1 IPv6 (Internet 协议第六版) 规范

1998 年 12 月公布的 RFC2460 (IPv6 规范)^[4]中定义的 IPv6 首部格式如下所示：



各部分长度说明如下：

- (1) 版本：4 比特，Internet 协议版本号 = 6。
- (2) 传输类别：8 比特
- (3) 数据流标签：20 比特
- (4) 有效载荷长度：16 比特，无符号整数。IPv6 有效载荷长度是以八位组为单位，

在这个包中 IPv6 首部后面的其余部分的长度。（注意，扩展首部将被认为是有效载荷的一部分，计算在长度里。）

(5) 下一个首部：8 比特，选择器。标识紧接在 IPv6 首部后面的下一个首部的类型。使用与 IPv4 协议字段相同的数值。

(6) 跳数限制：8 比特，无符号整数。在每个传输此包的节点处递减 1。如果跳数限制减为零，就抛弃此包。

(7) 源地址：128 比特，包的制作者的地址。

(8) 目的地址：128 比特，包的预期接收者的地址（如果存在路由首部的话，可能不是最终的接收者）。

在 IPv6 里，可选的网络层信息在一个独立的首部编码，放在包中 IPv6 首部与上层协议首部之间。有这样几个为数不多的扩展首部，每个首部由不同的“下一个首部”的值来标识。一个 IPv6 首部可以携带零个、一个或者更多的扩展首部，每个扩展首部由前一个首部中的“下一个首部”字段标识。

3.1.2 IPv6 与 IPv4 的比较

由 IPv4 到 IPv6 的改变主要集中在以下几个方面：

(1) 地址容量的扩展

IPv6 把 IP 地址的大小从 32 位增至 128 位，可以支持更多的地址层次，更大数量的节点，以及更简单的地址自动配置。组播路由的可缩放性改进为给组播地址增加一个“范围”字段。又定义了一个叫做“anycast”的新的地址类型，用于把包发送给一组节点中的任意一个。

(2) 首部格式的简化

一些 IPv4 首部字段被删除或者成为可选字段，减少了一般情况下包的处理开销以及 IPv6 首部占用的带宽。

(3) 支持扩展和选项的改进

IP 首部选项编码方式的修改导致更加高效的传输，在选项长度方面更少的限制，以及将来引入新的选项时更强的适应性。

(4) 数据流标签的能力

加入一个新的能力，使得那些发送者要求特殊处理的属于特别的传输“流”的包能够贴上“标签”，比如非缺省质量的服务或者“实时”服务。

(5) 认证和保密的能力

为支持认证，数据完整性以及（可选的）数据保密的扩展都在 IPv6 中说明。IPv6 版的 ICMP 是所有 IPv6 应用都需要包含的。

3.2 宽带网络

随着通信、计算机、图象处理等技术的进步，计算机网络正在向宽带高速的

方向发展，人们对网络的带宽及速率提出了更高的要求，促使网络由低速向高速、由共享到交换、由窄带向宽带方向迅速发展。目前对于主干网来讲，各种宽带组网技术已经或正在迅速发展中日臻成熟和完善。

就技术上来言，所谓的宽带（Boardband）是指在同一传输界质上，可以利用不同的频道进行多重的传输，并且速率在原级（1.54Mbps）以上。对应宽带的是基带（Baseband），也就是无法利用不同频道进行多重传输的技术，这与传输界质无关，T3或光纤所能提供的带宽再高，依然是基带而非宽带。

现有的IP网是由一些使用IP协议的设备组成的局域网，在现有的电信网上传输，从而组成的广域网。为了在电信网上传输IP业务，主要采用的技术有：IP over ATM（POA）、IP over SDH/SONET（POS）和IP over WDM等。

3.2.1 IP over ATM（POA）

IP与ATM相结合的技术有重叠和集成两种模型。重叠模型的主要思路是IP在ATM上传送，IP的路由协议由IP路由器来实施，ATM网络需要实施ATM路由协议，在寻址方式上需要进行地址解析，在计费方面相对复杂，但在标准化方面较为完善，对于广播和多播业务效率较低。采用重叠模型的实现技术包括LAN仿真（LANE）、传统的IPOA（RFC1577）、ATM上的多协议规范（MPOA）等^[21]。

集成模型将IP协议与ATM层集成在一起，使得ATM交换机与IP路由器成为一体，不需要进行地址解析，从而简化了ATM选路功能，但在计费方面较为复杂，在标准化方面还不够完善。目前，集成模型的产品仅能支持IP应用，不能够为IP以外业务提供传送能力，从而限制了提供多业务的公用宽带网络对该技术的选择。集成模型的代表产品有Ipsilon公司的IP交换机、Cisco公司的标记交换机（Tag Switch）、Ascend公司的IP导航器（IPNavigator）等。

3.2.2 IP over SONET/SDH（POS）

在同步光纤网络（SONET/SDH）上直接传送IP业务是一种比较理想的选择，特别是在核心骨干网上^[22]。SDH是ITU-T标准，为从155Mb/s（STM-1）到9.9Gb/s（STM-64）的光接口提供了速率、格式和光参数定义。除了提供高容量的链路，SDH传输系统还提供了：

- （1）简单和标准化的传输运行、管理和维护（OAM&P）能力
- （2）标准化的保护倒换结构提供了高生存性/可靠性的网络性能
- （3）成熟的标准使得多厂商产品互联和互操作性能良好

IETF规范的IP over SDH/SONET是将IP数据包通过采用点到点协议（PPP），映射到SDH/SONET帧上，按各次群相应的线速率进行连续传输。IP over SDH/SONET也叫Packet over SDH/SONET或PPP over SDH/SONET，它保留了IP面向非连接的特性。

PPP协议是一个非常简单的OSI第二层网络协议，标头只有两个字节，没有地址信息，只是按点到点顺序，是非面向连接的。PPP协议可将IP数据包切成PPP帧，以满足映射到SDH/SONET帧结构上去的要求。

IP over SDH/SONET技术的实现需要高速路由器和PPP协议，采用的仍然是传统路由器的逐包转发方式。这种方法的基本思路是将路由计算与包的转发分开，采用缓存技术、硬件芯片快速处理技术、以ATM信元交换矩阵作为路由器内部体系构架的交换路由技术，将路由器包的逐包转发速度控制到与第二层交换的速度相当。它无须利用广域网上的ATM交换机来建立虚电路VC。目前不少网络设备公司已推出基于IP over SDH/SONET技术的交换路由器产品，如Cisco千兆位交换路由器GSR 12000、Ascend千兆位路由转发器GRF等。ITU-T建议的IP over SDH是将IP通过LAPS直接映射到SDH低阶或高阶虚容器中，具有简单、可靠、高效的特点，是很有发展前途的一种技术。

SDH上的IP更确切的说应该是在IETF RFC1619中描述的SDH上的IP/PPP/HDLC。IP数据报被封装进点对点协议（PPP）分组中。在IETF RFC 1661中描述的PPP提供了多协议封装、差错控制和链路初始化控制特性。

由PPP封装的IP数据报通过使用RFC 1662中描述的高级数据链路控制(HDLC)成帧，然后字节同步映射入SDH同步净荷包封(SPE)中。HDLC的主要功能是描述或划分通过同步传输链路的PPP封装的IP数据报。划分通过一种叫做字节填充的技术来实现。每个HDLC帧由字节标志0x7e开始和结束。在传输端，监视HDLC帧的标志序列和出口序列。如果在HDLC帧中的信息区域的任何地方出现标志序列，都被改为序列0x7d 0x5e。同样地，出现出口序列0x7d，就被转换为0x7d 0x5d。在传输的接收端，填充模式被移走并用初始的区域来代替。另外，在没有数据报需要传送的空闲期间，把HDLC标志模式作为帧间填充。

对于IP骨干网，目前许多ISP开始转向SDH上直接传送IP的技术。从目前的IP骨干网来看，考虑数据报尺寸的因素，平均的ATM开销约为25%，而同样情况下，SDH上直接传送IP的开销只有约2%。除了开销效率外，SDH上直接传送IP，其业务使用了一对路由器间的所有的SDH链路带宽，即使是业务量只需要一部分，这一点需要权衡利弊。SDH上直接映射IP的另一个原因是其具有的可扩展性。大多数ISP的骨干路由器运行在OC-3（155Mb/s）和OC-12（622Mb/s）链路上。随着IP业务的增长，目前许多ISP已经把他们的骨干路由器链路升级至OC-48（STS-48c）。

ATM的分段和重组（SAR）功能随着接口速率的提高变得十分复杂。目前，达到OC-12速率的接口可以使用ATM SAR芯片，而OC-48C的接口则开始直接使用SDH接口。这样，SDH上直接传送IP将是满足ISP Internet骨干网超过OC-12（622Mb/s）的容量扩展需求市场的首选技术。

3.2.3 IP over WDM

波分复用（WDM）技术是未来光纤通信的一种主要复用方式。光纤通信以其无与伦比的优势已成为现代通信网的主要通信手段。在过去的10年中，光纤通信的容量提高了100倍，但还远没有发挥光纤带宽的巨大潜力。随着人们对通信的需求，特别是IP技术的发展产生的对网络带宽的巨大市场需求，要求有更先进的传输方式。随着通信技术的发展，传统的时分复用方式的传输速率已达10Gb/s，已接近其理论极限，迫切需要有新的复用方式，波分复用能够极大地扩展单根光纤通信的能力，它将是未来光纤通信的一种主要复用方式^[23]。

IP over WDM的基本结构是：光纤直接连光耦合器，耦合器把各波长分开或组合，其输入和输出端都是简单的光纤连接器，它把原波长内的数据送给SONET/SDH设备或高性能路由器。IP over WDM最大的优点之一是可以适应IP数据业务的不对称性，不同波长上数据速率可以不同。另一特点是光纤环的两侧都能使用，使路由可获得全部带宽，在传送大量的突发数据时就不需要缓存，也不会有包丢失，只有当光纤断裂时才会发生包丢失。网内有工作光纤和保护光纤，保护光纤中的空闲频带在业务高峰时也可用来传送数据，不会引入抖动、时延或包丢失。恢复工作在IP层而不在物理层上完成。在光纤断裂时，对抖动和时延敏感的实时业务可给以高的优先级。使用工作光纤和保护光纤的另一优点是可以建立“直通”或“旁路”波长。

在IP over WDM场合，高速路由器通过光ADM或WDM耦合器直接连至WDM光纤，光纤内各波长是链路层互连的。高速路由器取代传统的基于电路交换概念的ATM和SONET/SDH电交换和复用设备，成为关键的统计复用设备，用作主要的交换/选路设备，由它控制波长接入、交换、选路和保护。因此IP over WDM是一个真正的链路层数据网，可以通过指定波长作旁路或直通连接，网络的业务工程可以只在IP层完成。由于使用了指定的波长，在结构上更灵活，并具有向光交换和全光选路结构转移的可能。密集WDM系统能够对电信业务进行光复用，一部分波长被指定用于高宽带IP光网，即IP over WDM；另一部分波长被指定用于ATM光网，IP over ATM；还有一部分波长被指定用于传统的SONET/SDH，即P over SONET/SDH，它可能用来集中和传送传统的IP网业务。

构成IP over WDM的网络的部件包括：激光器、光纤、光放大器、光耦合器、光分插复用器、光交叉连接器和转发器等。非零色散位移光纤（NZDSF）最适合WDM，因其色散的非线性效应小。高性能激光器是WDM系统中最贵的器件。光放大器主要采用EDFA，EDFA多数都是宽带的，能同时放大WDM的所有波长，因对平坦增益的要求很高，在经过6个左右光放大器之后就需要做一次电放大，在OC-48速率，光跨度可达400km，在OC-192可达250km。光耦合器是用来把各波长组合在一起和分解开来的，起到复用和去复用的作用。在长途WDM系统有电再生

中继器，再生分三类：R1，R2和R3。转发器用来变换来自路由器或其它设备的光信号，并产生要插入光耦合器的正确波长光信号。光分插复用器、光交叉连接设备在长途WDM系统中被广泛使用。光交换机可以使ADM和交叉连接设备作动态配置。

目前，国外已有多家营运公司宣布要建设自己的IP over WDM网络。比较典型的有：1998年8月美国Sprint和Trontier要建设的OC-48的“IP over WDM”，以及有120多个成员的加拿大科研、工业、教育促进网于1998年8月15日宣布要与贝尔加拿大组成财团建设的CA*net3。后者计划总投资超过1.2亿美元，参加该项目的有Cisco Systems Canada、JDS Fitel、Newbridge Networks和Nortel等公司。CA*net3采用32波长，最高容量为40Gb/s，超过目前世界上任何因特网的速度，据称下载2个半小时的电影“泰坦尼克号”只要1/15s。

Sprint对IP over WDM进行了实验，实验是在实验室进行的，其目的是为了研究同步、接口以及OAM&P能力。实验设备由IP路由器、OC-48/STS-48c接口板和DWDM系统组成。DWDM系统的每个信道在单模光纤上传输，间距为100km，每隔100km用一个光放大器对信号进行放大，实验采用40个信道，对其进行层叠实验，这样总的线路长度达20,000km。实验结果令人振奋，经过16h的误码率测试，没有记录到一个误码，系统的抖动测量表明在这种超长距离上运行良好。实验结果表明在很长的DWDM间距上可直接在物理层上实现IP传输。

1998年4月，Cisco、Ciena、Lucent、NTT、AT&T、3COM、Bellcore、HP、Qwest、Sprint、WorldCOM等网络通信设备公司和运营公司成立了光互联网论坛OIF（Optical Internetworking Forum），以加速光互联网技术的发展。目前，OIF正在和ITU-T、IETF、ATM论坛等标准化组织合作制订有关光互联网的技术规范，但是它关注的不是数据网络或光网络内部的技术问题，而是数据网络和光网络之间的互操作性问题，如光互联网的光网络物理层传输接口（如比特率、不同数据格式的成帧和同步以及光纤的特性）、光网络与数据网络层之间中间层的适配以及中间链路层的管理（包括故障和性能管理、中间层的配置、中间层的保护/恢复、会话管理、计费和安全）等。目前OIF已确定了用于描述光互联网的多协议参考模型，即光互联网重叠模型。

此外ITU-T也在对IP over WDM进行各种标准化研究。

3.3 无线网络

以往，无线局域网发展缓慢，推广应用困难，主要是由于传输速率低、成本高、产品系列有限，且很多产品不能相互兼容。如以前无线局域网的速率只有1~2Mb/s，而许多应用也是根据10Mb/s以太网速率设计的，限制了无线产品的应用种类。针对现在高速增长的数据业务和多媒体业务，无线局域网取得进展的关键

就在于高速新标准的制定，以及基于该标准的 10Mb/s 甚至更高速率产品的出现。这些新标准推动了无线网络的快速发展。

无线接入技术区别于有线接入的特点之一是标准不统一，不同的标准有不同的应用。正因为此，使得无线接入技术出现了百家争鸣的局面。在众多的无线接入标准中，无线局域网标准更成为人们关注的焦点。下面比较了几种最热门的无线局域网标准^[24]。

3.3.1 IEEE 802.11 标准

802.11 是 IEEE 最初制定的一个无线局域网标准，主要用于解决办公室局域网和校园网中，用户与用户终端的无线接入，业务主要限于数据存取，速率最高只能达到 2Mbps。目前，3COM 等公司都有基于该标准的无线网卡。

由于 802.11 在速率和传输距离上都不能满足人们的需要，因此，IEEE 小组又相继推出了 802.11b 和 802.11a 两个新标准。三者之间技术上的主要差别在于 MAC 子层和物理层。

3.3.2 IEEE 802.11a 标准

802.11a 工作在 5GHz U-NII 频带，物理层速率可达 54 Mb/s，传输层可达 25Mbps。采用正交频分复用（OFDM）的独特扩频技术；可提供 25Mbps 的无线 ATM 接口和 10Mbps 的以太网无线帧结构接口，以及 TDD/TDMA 的空中接口；支持语音、数据、图像业务；一个扇区可接入多个用户，每个用户可带多个用户终端。但是，芯片没有进入市场、设备昂贵、空中接力不好、点对点连接很不经济、不适合小型设备。值得庆幸的是，Radiata 的低成本 COMS 无线引擎芯片装置可支持 802.11a。

无线网络标准机构——无线以太网兼容联盟（WECA）将使用 802.11a 标准的无线网络称作“Wi-Fi5”。名称更换和思科等厂商的新设备使用 5GHz 范围内的频率有关。原来 802.11b 标准的 Wi-Fi 设备使用 2.4GHz 范围内的频率。WECA 的主席丹尼斯表示，由于业界不认同，WECA 暂时取消了在 802.11a 设备加贴“Wi-Fi5 Certified”标志的计划。他认为，直接使用 Wi-Fi5 的名字，而不是 Wi-Fi1 (2,3,4)，可以避免引起别人误解。Synergy 调研集团的无线业务分析家称，命名非常好，因为除非有专门无线网络知识，否则搞不清原来的各种名字之间的差异。这给消费者带来了混乱。目前有很多变种的 802.11 标准，具体差别体现在速度和安全性。现在广泛使用的是 Wi-Fi 网络基于 802.11b 标准，但 802.11a 和 802.11g 具有更高的安全性能和速度。

3.3.3 IEEE 802.11b 标准

802.11b 物理层支持 5.5 Mbps 和 11 Mbps 两个新速率。802.11 标准在扩频时

是一个 11 位调制芯片，而 802.11b 标准采用一种新的调制技术 CCK 完成。802.11b 使用动态速率漂移，可因环境变化，在 11 Mbps、5.5 Mbps、2 Mbps、1 Mbps 之间切换，且在 2 Mbps、1 Mbps 速率时与 802.11 兼容。

IEEE 802.11b 从根本上改变了无线局域网的设计和应用现状，满足了人们在一定区域内实现不间断移动办公的需求，为我们创造了一个自由的空间。

IEEE 802.11b 无线局域网的带宽最高可达 11Mbps，比两年前刚批准的 IEEE 802.11 标准快 5 倍，扩大了无线局域网的应用领域。另外，也可根据实际情况采用 5.5Mbps、2 Mbps 和 1 Mbps 带宽，实际的工作速度在 5Mb/s 左右，与普通的 10Base-T 规格有线局域网几乎是处于同一水平。作为公司内部的设施，可以基本满足使用要求。IEEE 802.11b 使用的是开放的 2.4GB 频段，不需要申请就可使用。既可作为对有线网络的补充，也可独立组网，从而使网络用户摆脱网线的束缚，实现真正意义上的移动应用。

IEEE 802.11b 无线局域网与我们熟悉的 IEEE 802.3 以太网的原理很类似，都是采用载波侦听的方式来控制网络中信息的传送。不同之处是以太网采用的是 CSMA/CD（载波侦听/冲突检测）技术，网络上所有工作站都监听网络中是否有信息发送，当发现网络空闲时即发出自己的信息，如同抢答一样，只能有一台工作站抢到发言权，而其余工作站需要继续等待。如果一旦有两台以上的工作站同时发出信息，则网络中会发生冲突，冲突后这些冲突信息都会丢失，各工作站则将继续抢夺发言权。而 802.11b 无线局域网则引进了冲突避免技术，从而避免了网络中冲突的发生，可以大幅度提高网络效率。

802.11b 运作模式基本分为两种：点对点模式和基本模式。点对点模式是指无线网卡和无线网卡之间的通信方式。只要 PC 插上无线网卡即可与另一具有无线网卡的 PC 连接，对于小型的无线网络来说，是一种方便连接方式，最多可连接 256 台 PC。而基本模式是指无线网络规模扩充或无线和有线网络并存时的通信方式，这是 802.11b 最常用的方式。此时，插上无线网卡的 PC 需要由接入点与另一台 PC 连接。接入点负责频段管理及漫游等指挥工作，一个接入点最多可连接 1024 台 PC（无线网卡）。当无线网络节点扩增时，网络存取速度会随着范围扩大和节点的增加而变慢，此时添加接入点可以有效控制和管理频宽与频段。无线网络需要与有线网络互连，或无线网络节点需要连接和存取有线网的资源和服务器时，接入点可以作为无线网和有线网之间的桥梁。

802.11b 无线局域网由于其便利性和可伸缩性，特别适用于小型办公环境和家庭网络。在室内环境中，针对不同的实际情况可以有不同的典型解决方案。

（1）对等解决方案

对等解决方案是一种最简单的应用方案，只要给每台电脑安装一片无线网卡，即可相互访问。如果需要与有线网络连接，可以为其中一台电脑再安装一片有线网卡，无线网中其余电脑即利用这台电脑作为网关，访问有线网络或共享打印机等设备。

但对等解决方案是一种点对点方案，网络中的电脑只能一对一互相传递信息，而不能同时进行多点访问。如果要实现有线局域网的互通功能，则必须借助接入点。

（2）单接入点解决方案

接入点相当于有线网络中的集线器。无线接入点可以连接周边的无线网络终端，形成星形网络结构，同时通过 10Base-T 端口与有线网络相连，使整个无线网的终端都能访问有线网络的资源，并可通过路由器访问 Internet。

（3）多接入点解决方案

当网络规模较大，超过了单个接入点的覆盖半径时，可以采用多个接入点分别与有线网络相连，从而形成以有线网络为主干的多接入点的无线网络，所有无线终端可以通过就近的接入点接入网络，访问整个网络的资源，从而突破无线网覆盖半径的限制。

（4）无线中继解决方案

无线接入器还有另外一种用途，即充当有线网络的延伸。比如在工厂车间中，车间具有一个网络接口连接有线网，而车间中许多信息点由于距离很远使得网络布线成本很高，还有一些信息点由于周边环境比较恶劣，无法进行布线。由于这些信息点的分布范围超出了单个接入点的覆盖半径，我们可以采用两个接入点实现无线中继，以扩大无线网络的覆盖范围。

（5）无线冗余解决方案

对于网络可靠性要求较高的应用环境，比如金融、证券等，接入点一旦失效，整个无线网络会瘫痪，将带来很大损失。因此，可以将两个接入点放在同一位置，从而实现无线冗余备份的方案。

（6）多蜂窝漫游工作方式

在一个大楼中或者在很大的平面里面部署无线网络时，可以布置多个接入点构成一套微蜂窝系统，这与移动电话的微蜂窝系统十分相似。微蜂窝系统允许一个用户在不同的接入点覆盖区域内任意漫游，随着位置的变换，信号会由一个接入点自动切换到另外一个接入点。整个漫游过程对用户是透明的，虽然提供连接服务的接入点发生了切换，但对用户的服务却不会被中断。

早期的 802.11b 无线局域网技术已经在纵向市场应用方面取得成功，例如生产、存货控制和零售点等方面，1999 年已经取得了 4 亿美元的销售额。随着 802.11b

性能价格比实质性的提高，一个全新的横向市场应用将全面展开。企业将可以应用无线局域网作为他们有限局域网的延伸。这一应用将使他们全方位地与公司应用程序和网络外围设备取得连接，从而大大提高雇员在移动中的工作效率。无线局域网技术将首先应用于企业的会议厅和部门办公室，随着其使用的深入，最终将应用于公司的每一个角落。小企业和家庭用户也将使用无线局域网代替有线网络，从而获得无线局域网提供的在“无线”安装和维护方面带来的节约。

3.4 网络存储

近年来，随着Internet和Intranet技术的广泛应用，交易的信息量以令人难以置信的速度迅速增长。而且，电子商务、电子邮件也对数据在可用性、可靠性、可扩展性等方面提出了更加严格的要求。网络存储技术的出现和发展正好构成了对Internet技术的有效支撑。当然，除了Internet技术以外，在大型数据仓库的应用、大型CAD设计系统、图形处理系统、在线事务处理系统、宽带多媒体系统、电信综合业务系统中，网络存储技术都是人们翘首期盼的突破。因此，研制高性能、可扩展、高可用、可管理的以存储为核心的新型网络服务器系统就成为网络服务器系统发展的主要趋势。

以存储为核心的网络服务器系统是一个新型高可用网络服务器系统。系统设计目标是通过采用新型的服务器系统结构以及对系统和数据有效的监控和管理，提供高效、可管理、可扩展、高可用的数据和网络应用服务，包括Web、Email、流媒体等海量信息服务^[16]。

3.4.1 NAS

网络存储服务器（NAS）管理文件的存储和检索更加有效，而且比原来传统的服务器总拥有成本（TCO）更低，它们的操作系统被优化来只做一件事情：非常好地管理和保护文件。根据IDC的资料，NAS是一个网络中心的概念，它取代传统的网络文件服务器，提供一个共享的存储器。对数据访问来说，它是一个瘦服务器，专为文件共享而设计，应用于工作组和小办公室市场。这些产品趋向于只有一个多功能的存储控制器。NAS是一个更加便宜、更加简单，但是容量更大的文件服务器。

NAS独立于PC服务器，扩大了存储的容量。它提供了一个简单的、高性价比的解决方案，网络的存储容量并提供一个高性能，高可用性，高扩展性和低总拥有成本（TCO）。他们的操作系统被优化用来只做一件事情：非常好的保护和管理文件。

NAS服务器在这样的情形下出现：对公司来说，保存的信息是公司的核心资产。例如，在一个直销的网站，E-mail数据库，学校信息系统。通过这种方式，人际的通讯及商业活动均可在网上进行，因此产生大量的数据，需要存储，分析

和保护。

NAS 服务器和传统服务器的显著区别在于：

- (1) 快速联网并投入运行；
- (2) 易于使用；
- (3) 外形紧凑；
- (4) 价格低。

通常是根据它们直接连接到 IP 网络上这一点来定义 NAS 服务器的。实际上，NAS 服务器是作为大企业提出复杂的存储需求的一种彻底、简单的解决方法出现的。今天市场上许多产品基本上都是传统的服务器。它们一般带有一个经过剥离的操作系统，并通常通过硬件 RAID 来管理三个到四个硬盘。这些产品拥有很强的功能和强大的软件支持。尽管它们具有某种程度的可扩充性，但控制部件的初期投资是相当可观的，因此，只有当它的全部功能被用足时，这种系统才“可承受得起”。

NAS 存储装置采用了一种更加明确的方法。NAS 直接连接互联网，并提供了从 150GB 到 800GB 的存储容量。它们将用小的外形尺寸来节省空间，这样就可以以机架的方式安装几十个设备。对于“管理日益增大的存储器”的需求，它们给出了非常直接而又简单的解决方法。除去所有不必要的功能来优化系统。它唯一允许的冗余就是数据的冗余。

到目前为止，NAS 装置因其容量和软件功能的限制，仅应用在工作组和小型企业中。然而，现在借助允许有中央控制点的软件，可管理 NAS 装置的网络，NAS 装置就如同多个数据中心，可以进行分布式的存储。

3.4.2 SAN

SAN（存储区域网）是指独立于服务器网络系统之外几乎拥有无限存储容量的高速存储网络。这种网络采用高速的光纤通道作为传输媒体，以 FC+SCSI 的应用协议作为存储访问协议，将存储子系统网络化，实现真正的高速共享存储。SAN 技术的发展，长期以来基于 100MB/s 的技术（自 1994 年光通道协议制定以来），FC 技术在性能上比起 SCSI、iSCSI、NAS 技术等多种存储技术，具有一定的优势。近一段时间，SCSI 技术也在寻求着自身的发展，在速度方面有了一定的提高。但从总体上讲，SAN 在性能、网络化、容灾等诸多方面仍然具有综合优势。

SAN 的发展经历了三个阶段：

- (1) 多主机的磁盘共享，即存储池共享

这一阶段，多服务器通过 SAN 实现了真正意义上的计算机系统和存储设备之间的相互连接（任何节点到任何节点的连接特性），存储系统与主机系统均形成无限的双向扩展能力。SAN 在 1999 年已基本实现了这一阶段的目标。如 1999 年第一个在国内实现的 SAN 体系—21CN 电子邮箱系统就是一个明显的标志。

(2) 服务器之间的文件系统共享以及存储设备的多服务器平台连接

应用数据的共享归根到底要实现文件系统的并行操作，这实际上是当前大多数操作系统中文件系统的瓶颈。相当多的SAN设备提供商已经成功地实现将存储设备与多种服务器同时跨接，使用户的应用系统变得更加灵活和结构合理。在国内外，越来越多的用户正在享受数据资源真正共享所带来的巨大财富。

(3) 异构(不同平台)服务器之间文件系统共享以及多种存储设备的开放式共存这也是SAN的精髓。实际上，只有真正开放的存储网络才是计算机产业发展的潮流。正像SCSI技术和Ethernet技术的发展所走过的道路，SAN最终将实现不同主机和不同存储设备的真正意义的资源共享。无疑，企业的管理者将是这一目标的最大受益者，企业将充分享受科学技术带来的巨大收益。存储系统将有机地融为一体，存储空间将变得十分广阔，充分体现信息对企业的价值。

虽然SAN战略已经为企业的合理的存储构架提供了可行性解决，并且可以为企业带来可观的经济效益，但仍然有一部分管理者认为：SAN缺乏灵活和开放的互用能力，因此在使用SAN方案上比较犹豫。根据INFORMATION 周刊提供的统计，在调查的600家IT经理中，96%认为他们计划建立存储网络，但当他们需要前进一步时，他们犹豫了，许多人在等待着完整的解决方案，或者选择传统或权宜的方式来解决。

无论如何，对SAN的需求与日俱增。电子商务的巨大需求，数据分析所依托的海量存储，不经过主机的数据直接快速备份，这一切使得SAN的发展不可逆转。人们无需总在等待它的开放标准出台，实际上，当我们在计划海量数据的可管理性和可行性方案时，我们会发现，SAN会比传统的SCSI结构费用低的多。例如在E-mail系统中，如果使用了SAN的设计，人们可以集中和直接管理大量的存储设备，而无需在分布的服务器上建立多个存储设备。同时，人们也可以将多个应用服务器形成集群，提供系统整体的冗余。这一切，几乎无需多少的管理员就可以轻易做到。系统在扩容时，使用SAN将显示巨大优势：系统在线就可以加入新的存储设备，无需停机。

3.5 信息资源检索协议

数字图书馆的最终目的就是给各种用户提供完善、友好的信息检索服务。因此如何在不同机型、不同平台上实现联合服务称为数字图书馆必须考虑的问题。人们要求不仅仅能够查询分散在各地的数据库信息，而且要去不再为不同的登录环境、不同的系统平台所困扰。因此，在互联网的最上层——应用层建立通用的应用层协议将解决这个问题。目前相关的标准有：ISO 10162、ISO 10163 和 Z39.50。其中，ISO 10162、ISO 10163 是国际标准化组织公布的国际标准，建立在 OSI 七层参考模型的基础之上，而 Z39.50 是美国国家标准化组织公布的国家标准，建立

在 TCP/IP 的基础之上。由于 TCP/IP 已经成为了世界通用的网络模型，因此国际标准化组织根据 Z39.50 修订了自己的 ISO 23950 标准，该标准也工作在在 TCP/IP 的基础之上，基本和 Z39.50 一致。

3.5.1 ISO 10162 和 ISO 10163

为了实现开放系统互连之间信息交换的标准化，国际标准化组织一共制订了十类 ISO 标准。其中应用层标准这一大类下又细分了若干小类，其中一小类为“图书馆应用”。ISO10162 和 ISO10163 都是“图书馆应用”这一小类中的标准。在该分类下，除了这来呢标准之外还有 ISO10160 和 ISO10161。而 ISO10162 和 ISO10163 制订了检索协议和检索服务标准。这两个协议于 1988 年 11 月通过了草案，目前已经成为成型标准。

这两个标准为国际信息检索提供了服务定义和协议说明的国际标准。它们的适用范围均局限在系统互连上，主要用于支持信息检索的服务的系统，诸如数字图书馆、公共信息事业、联合目录中心等。它们可以独立使用，也可以联合方式来支持某种需要多种通信服务的图书馆应用^[25]。

3.5.2 Z39.50

ANSI/NISO Z39.50 的全称是 American National Standard Information Retrieval Application Service Definition and Protocol Specification for Open System Interconnection。它于 1988 年推出了第一版，并不断发展完善，成为在全世界范围图书馆自动化领域内应用最为广泛的国际标准 ISO23950。

Z39.50 的第一版是由 NISO 的 D 分会起草的。D 分会的任务是为图书馆的应用编写网络应用层协议，其中也包括信息检索协议。该分会在 Z39.50v.1 通过之后即行解散。1989 年成立了 Z39.50 管理机构，负责标准的修订工作。

1990 年 Z39.50 执行小组 (ZIG) 成立。小组成员有机器生产商、机器销售商、咨询社、信息服务中心和大学。因为他们希望检索各种数据库信息，并为各种检索需求提供服务。数据库的信息类型很多，其中有书信息、全文信息、商业信息、公共服务信息以及新闻等。尽管在开始提出的设计目标只是关于书目信息检索服务，但从 1990 年开始已经将它扩大为信息检索协议。

Z39.50 是一种 Client/Server 体系结构下描述的客户机检索服务器上数据以获得检索结果的数据结构与交互规则的协议，它是网络中的应用层协议，定义了客户机/服务器之间数据交换的标准，Z39.50 的特点有^[26]：

(1) 服务器上的检索结果可以汇集（即可支持多数据库检索），还可以做二次检索以缩小结果集，提高检索精度；

(2) 包含有检索管理功能，例如服务器可以提供正在检索的进度、或者请求客户在检索中提供一些认证信息使一些敏感检索可以进行下去；客户也可以终止

一个正在进行的检索或者重新开始。

(3) 包含有管理检索结果集的措施，如对检索结果进行排序等。

(4) 定义了一种标准的查询语言。

(5) 定义了不同的记录语法，包括书目数据的 MARC 格式的语法、非通用语法 GRS-1 (Generalized Record Syntax One)，它可以用来服务器和客户端之间传送数据。

(6) 定义了一种语言，描述了如何构建将结果集从服务器端传递到客户端的记录。

(7) 定义了一个称为解释 (explain) 的措施，它允许客户端从服务器获得广泛的关于哪个数据库是可访问的、数据库中的哪些点是可访问的等方面的信息。

Z39.50 对信息服务的发展和服务方式产生了重大的影响。在图书馆自动化领域中，各软件生产公司为适应当代社会对开放系统的要求，纷纷在新产品中支持该标准。

3.5.3 ISO 23950

ISO 23950 是 Z39.50 的国际标准版本，其完全和 Z39.50 一致，故不做具体讨论。

4. Web 技术

Web 是计算机历史的巨大成功之一。Web 及其相关技术是数字图书馆快速发展的关键。Web 是分布于 Internet 上各个计算机中信息相互链接而成的集合。这些计算机被称为 Web 服务器。

Web 技术是由位于瑞士的欧洲粒子研究中心 (CERN) 的 Tim Berners-Lee 和他的同事开发的。伊利诺斯大学的 Marc Andreessen 和其他人开发的用户界面 Mosaic 则使它得到普及。Mosaic 在 1993 年发布，随后的几年又有它的许多商业版本，其中使用最广泛的是 Netscape Navigator 和微软的 Internet Explorer。这些用户界面被称为 Web 浏览器或简称为浏览器。

Web 的成功经验可以被简单地总结成：它提供了 Internet 上方便的信息传播方法。不需要帮助或培训，个人就可以自己存取或者发布信息。

4.1 基本 Web 技术

从技术上来看，Web 基于四种简单的技术：超文本标记语言 (HTML)、超文本传输协议 (HTTP)、MIME 数据类型和统一资源定位器 (URL)。每项技术不仅对 Web 而且对数字图书馆中广泛存在的互操作性也都十分重要。

4.1.1 HTML

HTML 是用于描述文档结构和表现的语言。HTML 文件包括需要显示的文本以及用以描述格式和结构的标记。HTML 标记总是被包含在符号“<”和“>”中。大多数的 HTML 标记成对出现。少数标记是自我包容的，不含任何内容。

4.1.2 HTTP

在计算机领域，协议是在计算机之间传送消息的一组规则。一个典型的协议包括所使用的格式描述、各种信息、传送它们的顺序、正确的响应方式、出错条件等等。HTTP 是 Web 浏览器和服务器之间传送消息的协议。

4.1.3 MIME

计算机上的数据文件只是二进制数的集合，必须经过解释才有用。在 Web 和 Internet 上的大量应用中，数据类型用 MIME 模式来描述。MIME 最初是为了描述电子邮件的信息而开发的。它使用通用和专用两部分编码。例如，text/ascii 表示用 ASCII 标记的正文文本，image/jpeg 表示是 jpeg 格式的图像，text/html 表示是 HTML 标记的正文文本。大量的计算机程序使用标准的 MIME 类型集合，其他的数据类型可以使用实验性的标记。

MIME 类型对于 Web 的重要性在于 HTTP 的 get 命令传送的数据都对应一个 MIME 类型。例如文件 index.html 的 MIME 类型是 text/html。当浏览器收到一个这种类型的文件时，它就知道应该把它解释成 HTML 文本并显示出来。

许多计算机系统通过文件名粗略地表示数据类型。MIME 是系统、灵活地记录和传送类型化数据的方法。

4.1.4 URL

URL 提供了简单灵活的寻址机制，使得 Web 能够连接世界各地计算机上的信息。一个简单的 URL 地址：<http://www.nlc.gov.cn/index.html>，这里 http 是协议的名称，www.nlc.gov.cn 是计算机的域名，index.html 是在计算机上的文件名。这个 URL 可以简单地解释为：“使用 http 协议，连接到 Internet 上地址为 www.nlc.gov.cn 的计算机上并取得文件 index.html”。

Internet 中数字对象的标识基准是 URI 命名机制。URI (Uniform Resource Identifier, 统一资源标识符)：作为 Internet 资源的统一的标识符体系，其中“资源”是能够被标识的任何对象。“标识符”是一个对象可以引用的名称，实际定义为符合某种语法规则的字符串。而“统一”则保证了现有不同类型的资源标识符可以采用统一的方式使用，并为新的资源类型提供统一的标识框架。实际上，URI 成为所有标识 Internet 资源的地址和名称的通用集合。

按照早期的定义，URI 由三部分组成：URN (统一资源名：Universal Resource

Name): 作为数字对象的逻辑名称; URC (统一资源属性, Universal Resource Characteristics): 作为描述数字对象的元数据; URL (统一资源定位符: Universal Resource Locator): 作为物理上对数字对象定位和获取的机制。实际应用中URC从来没有得到广泛认同, 取而代之的是各种类型的元数据。鉴于此, URI机制被重新定义为包括URN和URL的集合, 其中URN作为逻辑化的资源名称, 可以是某个机构或联盟对资源的统一命名: 如PURL, 也可以是采用某种命名大纲制订的持久、独立于地址的资源标识符。URL作为物理地址, 代表Internet上某个具体的地址。URN和URL可以相对独立, URI既可以是URL, 也可以是URN^[41]。

4.2 动态 Web 技术

早期的 Web 资源都是基于上述四种基本技术构建的静态站点, 只提供诸如文本、图片、音频、视频等静态信息, 用户无法与服务器进行交互, 因此具有很大的局限性。

随着 Internet 的不断普及, 静态的信息提供方式越来越无法满足用户的需要。人们需要通过 Web 这一廉价、开放的平台让更多的用户可以访问存储在数据库中的动态数据。一旦完成这样的应用, 用户可以更快、更经济地获取自己感兴趣的信息, 服务提供者也可以借此获得更多的用户反馈信息。在这种需求的推动下, 动态 Web 技术就应运而生了。

动态 Web 技术是指利用脚本代码、程序等实现能够与用户交互的动态 HTML 页面, 是对静态 Web 平台的扩展。这种扩展可以分为 Web 服务器端扩展和浏览器端扩展两类。目前, 动态 Web 技术已经成为 Web 发展的主流, 当前常见的商业化动态 Web 构建平台分别是 Microsoft 的 .NET 平台和 SUN 的 J2EE 平台。

4.2.1 Microsoft 的 .NET 平台

Microsoft 的旧的分布式 Internet 体系 (DNA, Distributed Internet Architecture), .NET 平台作为一个全新基础平台的优势, 简单对象访问协议 (SOAP, Simple Object Access Protocol) 和统一描述、发现和集成协议 (UDDI, Universal Description、Discovery and Integration) 之类的下一代技术平台

.NET Enterprise Servers

Microsoft .NET Enterprise Servers (包括 Windows 2000 Server) 为应用程序提供了一整套的服务。这些服务器一同提供了一个开发平台, 可以用这个开发平台来构建可扩展的、可管理的应用程序。例如, Windows 2000 为可扩展的、可靠的 COM+ 应用程序创建了架构; 它也提供了一个架构来处理 n 层的应用程序, 而在其它情况下, 开发人员就不得不自己另外创建这个架构。Windows 2000 与 .NET Enterprise Servers 的集成可以让用户以一种方式来管理它的服务器, 或者把它们看

作是一个服务器组，而不管它们的功能是怎样的。

COM 技术

Microsoft 的 COM 组件模型一直在很稳定的发展中，舍弃繁杂的 OLE 细节后，COM 才真正奠定了 Windows 组件模型的核心，真正的开始可以提供只做一切逻辑对象。

Microsoft 引入了组件对象模型 COM，形成了 COM 对象之间实现互操作的二进制标准。COM 规定了对象模型和编程要求，使 COM 对象可以与其他对象相互操作。这些对象可以用不同的语言实现，其结构也可以不同。COM 技术也是 .NET 框架中不可缺少的组成部分。

开发环境

开发环境升级到 Visual Studio.NET (VS.NET)，它综合了 Microsoft 原有 Visual Studio 6.0 的全部开发集成环境：Visual Basic、Visual C++、Visual J++等，此外还包含了 Microsoft 最新开发的 C#语言。统一的开发环境为应用程序框架的构建提供了便利的工具。

4.2.2 SUN 的 J2EE 平台

J2EE 是一个基于 Java 组件技术的企业应用系统开发规范，它为不同厂商创建平台产品提供了标准，使不同 J2EE 平台产品之间的交互成为可能，J2EE 规范定义了一个基于组件技术的多层企业信息系统的标准平台，旨在简化和规范企业应用系统的开发和部署，J2EE 技术的基础就是核心 Java 平台或 Java 2 平台的标准版，J2EE 不仅巩固了标准版中的许多优点，例如“编写一次，随处运行”的特性、方便存取数据库的 JDBC API、CORBA 技术以及能够在 Internet 应用中保持数据的安全模式等等，同时还提供了对 EJB (Enterprise JavaBeans)、Java Servlets API、JSP (Java Server Pages) 以及 XML 技术的全面支持。其最终目的就是成为一个能够使企业开发者大幅度缩短投放市场时间的体系结构。

J2EE 体系结构提供中间层集成框架用来满足无需太多费用而又需要高可用性、高可靠性以及可扩展性的应用的需求，通过提供统一的开发平台，J2EE 降低了开发多层应用的费用和复杂性，同时对现有应用程序集成提供强有力的支持，完全支持 Enterprise JavaBeans，又良好的想到支持和部署应用，添加目录支持，增强了安全机制，提高了性能。

J2EE 是一个基于组件-容器模型的系统平台，其核心概念是容器。容器是指为特定组件提供服务的一个标准化的运行环境，Java 虚拟机就是一个典型的容器。组件是一个可以部署的程序单元，它以某种方式运行在容器中，容器封装了 J2EE 底层的 API，为组件提供事务处理、数据访问、安全性、持久性等服务。在 J2EE

中组件和组件之间并不直接访问，而是通过容器提供的协议和方法来相互调用。组件和容器间的关系通过“协议”来定义。容器的底层是 J2EE 服务器，它为容器提供 J2EE 中定义的各种服务和 API，容器在 J2EE 服务器之上根据需要提供如目录服务、事务管理、数据访问、消息机制、安全性等扩展服务。

企业开发体系结构

J2EE 采用多层的分布式应用模型，按功能不同划分为不同的逻辑层：客户端机器上的客户层、J2EE 服务器上的 Web 层、J2EE 服务器上的业务层、数据库服务器上的数据层。这种划分也可以看成是前面介绍过的三层结构的扩充。采用多层的方式使得应用具有很强的伸缩性，每一层能够专注于特定的角色和功能。以下是 J2EE WEB 应用中典型的四层结构，如图 3 所示。

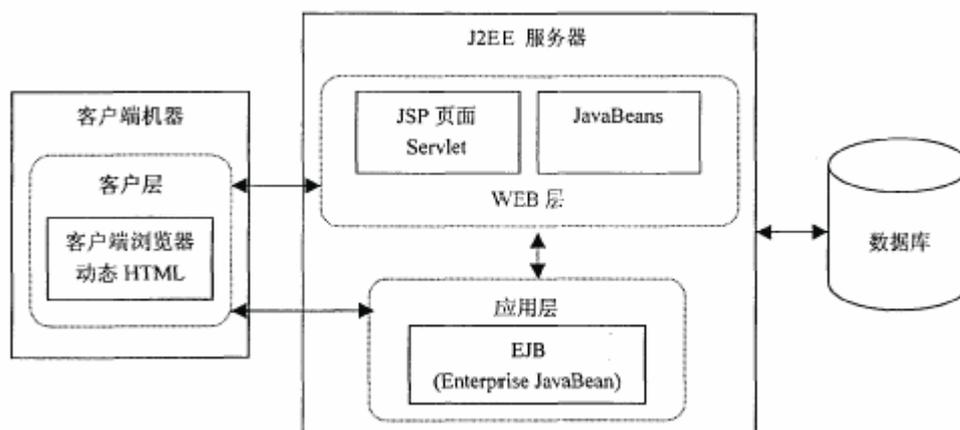


图 3 J2EE WEB 应用的四层结构

J2EE 平台核心技术

J2EE 是由一整套的服务 (Services)、应用程序接口 (API) 和协议 (Protocol) 组成，对开发基于 Web 的多层应用提供了功能支持。一下为 J2EE 的 13 种核心技术：

(1) JDBC (Java Database Connectivity)

JDBC 对于开发者来说是透明的通用接口，为访问不同的数据库提供了一种统一的方法，具有平台无关性。JDBC 定义了四种不同的驱动程序：

①JDBC-ODBC: 通过 JDBC-ODBC 桥, 可以使用 JDBC 来存取 ODBC 数据源。但是需要在客户端安装 ODBC 驱动程序, 也就是说必须安装 Microsoft Windows 的某个版本, 这影响了 JDBC 的平台独立性。

②JDBC-native driver bridge: 提供了一种建立在本地数据库驱动程序的顶层 JDBC 接口, 因此不需要使用 ODBC, 它将数据库的 API 从标准 JDBC 调用转换为

本地调用，影响了 JDBC 的平台独立性。

③JDBC-network bridge: 不再需要客户端数据库驱动程序，它使用网络上的中间服务器来存取数据库。这使得负载均衡、连接缓冲池和数据缓冲等技术的实现有了可能。由于这种方式往往只需要相对更少的下载时间，具有平台独立性，而且不需要在客户端安装并取得控制权，所以很适合于 Internet 上的应用。

④Pure Java driver: 通过使用一个纯 Java 数据库驱动程序来执行数据库的直接访问。

(2) JNDI (Java Name and Directory Interface)

JNDI 为应用提供了一个统一的接口来完成标准的目录操作，如 DNS、LDAP、本地文件系统或应用服务器中对象的访问和操作。

(3) EJB (Enterprise JavaBeans)

EJB 是 J2EE 的核心，它定义了一个可重用的组件框架来实现分布式的、面向对象的商业逻辑，简化了复杂的企业级应用的开发。EJB 的核心思想是将商业逻辑与底层的系统逻辑分开，使开发者只需关心商业逻辑，而由 EJB 容器实现目录服务、事务管理、持久性、安全性、容错性等底层系统逻辑，通过容器提供的协议和方法来相互调用组件。

一个可部署的 EJB 组件包含 3 个部分：

①Remote 接口：定义 EJB 组件中提供的可供用户调用的方法，也就是通常所说的实现商业逻辑的函数或过程，以供远程客户端调用。在 EJB 组件部署到容器的时候，容器会自动生成 Remote 接口相应的实例，即 EJB 对象，它负责代理用户的调用请求。

②Home 接口：定义一组方法来创建的 EJB 对象，查找、定位和清除已有的 EJB 对象。在 EJB 组件部署时容器也会自动生成相应的 Home 对象，该对象负责查找和创建 EJB 对象，返回 EJB 对象的引用给客户。用户利用该引用调用 EJB 组件的方法，得到结果。最后 Home 对象清除 EJB 对象，可以形象地称 Home 接口为 EJB 对象的工厂。

③Enterprise Beans 类是商业逻辑的具体实现类，其可供用户调用的方法在 Remote 接口中定义。根据功能不同，EJB 2.0 规范中定义了三种 Enterprise Beans：会话 Beans (Session Beans)、实体 Beans (Entity Beans) 和消息驱动 Beans (Message-driven Beans)。会话 Beans 分无状态和有状态两种：一般无状态的会话 Beans 模拟商业逻辑；有状态的会话 Beans 通常模拟一个客户会话，它会临时保存客户信息，根据客户要求调用其它 Beans 来存取数据。两种会话 Beans 都不保存状态信息或数据，当客户断开连接或者服务器关闭时，会话 Beans 也随之消失。实体 Beans 模拟商业逻辑，它表示一个数据存储，可以是状态信息或数据库中的一条记录。实体 Beans 在客户断开连接或者服务器关闭后，仍有服务保证其数据得

以保存。消息驱动 Beans 在行为上很像会话 Beans，不同的是仅在需要向这些 Beans 发送消息时才调用消息驱动 Beans。

(4) RMI (Remote Method Invoke)

RMI 协议调用远程对象上方法。使用序列化方式在客户端和服务端传递数据。RMI 是一种被 EJB 使用的更底层的协议。

(5) Java IDL/CORBA

通过 Java IDL 的支持，可发者可以将 Java 和 CORBA 集成在一起，可以创建 Java 对象并使之在 CORBA ORB 中展开，还可以创建 Java 类并作为和其它 ORB 一起展开的 CORBA 对象的 Agent，实现新、旧系统的集成。

(6) JSP (Java Server Pages)

JSP 页面由 HTML 代码和嵌入其中的 Java 代码组成。服务器在页面被客户端请求后执行这些 Java 代码，然后将生成的 HTML 页面返回给客户端的浏览器。

(7) Java Servlet

Servlet 是一种小型 Java 程序，它扩展了 Web 服务器的功能，是一种服务器端的应用，在服务器端驻留并响应客户端请求。Servlet 提供的功能大多与 JSP 类似，不过实现的方式不同，servlet 全部由 Java 编写并且生成 HTML。

(8) XML (Extensible Markup Language)

XML 是一种可以用来定义其它标记语言的语言，被用来在不同的商务过程中共享数据。XML 的发展和 Java 是相互独立的，但是 XML 和 Java 的共同目标都是平台独立性。通过 Java 和 XML 的结合，可以得到一个完美的具有平台独立性的 Web 应用解决方案。

(9) JMS (Java Message Service)

JMS 是一组用户和面向消息的中间件相互通信的 API，它既支持点对点的消息通信，也支持发布/订阅式的消息通信。

(10) JTA (Java Transaction Architecture)

JTA 定义了一种标准的 API，为应用系统提供可靠的事务处理支持。

(11) JTS (Java Transaction Service)

JTS 是 CORBA OTS 事务监控的基本的实现，规定了事务管理器的实现方式。JTS 事务管理器为应用服务器、资源管理器、独立的应用以及通信资源管理器提供了事务服务。

(12) JavaMail

JavaMail 是用于存取邮件服务器的 API，支持 SMTP 和 IMAP，允许在应用程序中以独立平台、独立于协议的方式收发电子邮件。

(13) JAF (JavaBeans Activation Framework)

JavaMail 利用 JAF 来处理 MIME 编码的邮件附件，实现 MIME 的字节流和 Java

对象互相转换。

4.3 主流 Web 服务器

4.3.1 IBM 的 WebSphere

WebSphere是IBM公司推出的一套系列产品，其最终目的是帮助用户建立自己的电子商务平台^[34]。目前WebSphere的最新核心产品是WebSphere Application Server 4.0，最新的配套开发工具为WebSphere Studio 5.0。

WebSphere 提供了开放的 Web 服务功能。现有的贯穿整个 IBM WebSphere 产品的预置功能可以开发和部署开放的 Web 服务。WebSphere 全面地简化了原有资产扩展到开放的 Web 服务的过程，能够最快地部署开放的 Web 服务，并使得企业成本降低，在市场中取得竞争优势。基于开放标准，IBM 已经建立了业内最全面的开放基础架构集合。所有 IBM 的中间件（middleware）都全面支持互联网开放标准，以实现 Web 服务应用的发展。

此外，WebSphere服务器是一个完全基于Java的Web服务器，目前支持Servlet、JSP和EJB等Java技术^[54]。

4.3.2 Microsoft 的 IIS

Microsoft的Internet Information Server作为传统的Web服务器，至今仍然占有很大的市场份额。它主要支持Microsoft提出的ASP技术，目前又推出了最新的开发框架Microsoft .NET系列^[55]。

4.3.3 BEA 的 Weblogic

BEA Weblogic Enterprise Platform是Bea公司出品的应用基础结构平台，该平台以支持EJB的BEA Weblogic Server 7.0 为核心，还包括了BEA Weblogic Portal、BEA Weblogic Integration和BEA Weblogic Workshop，这些技术结合起来，为应用开发、部署和管理提供了单一的、统一的、易用的基础结构平台。BEA的平台提供的应用结构基础简化了信息流、降低了应用管理成本，并使得企业更加灵活，更具效率，实现全面连接^[35]。

BEA WebLogic Enterprise Platform 所提供的应用基础结构的特征：

（1）统一性：BEA WebLogic Enterprise Platform 是唯一融 6 种应用基础结构功能于一体的应用基础结构软件。一套统一的应用基础结构能节省时间和资金，并确保企业软件更和谐地协同工作。

（2）简便性：BEA WebLogic Enterprise Platform 消除了应用基础结构的复杂性。简便的应用基础结构，可让企业软件开发人员和 IT 工作人员更富效率，更少失误。

(3) 可扩展性: BEA WebLogic Enterprise Platform 既能够与已有的遗留系统协作, 也能与未来的新系统协作。可扩展的应用基础结构为企业带来必要的敏捷性, 让企业能够快速地适应新的软件需求, 且避免以往那些导致大量软件项目失败的复杂性。

4.3.4 Apache 服务器

Apache系列服务器是Apache Software Foundation提供的开发源代码程序。Apache Http Server Project提供了可以运行在多种操作系统的版本, 可以在<http://httpd.apache.org/>上免费下载^[52]。

4.3.5 Tomcat 服务器

Tomcat服务器也是Apache Software Foundation提供的开发源代码程序。由于Apache Http Server不支持JSP、Servlet技术, 因此往往需要同时与Tomcat服务器配合使用。Tomcat服务器实质上为一个Servlet容器, 支持Servlet和JSP技术。目前公布的最新版本为Tomcat 5.0, 它实现了Servlet 2.4 规范和Java Server Pages 2.0 规范。与Apache Http Server一样, Tomcat也提供了可以运行在多种操作系统的版本, 可以在<http://httpd.apache.org/>上免费下载^[53]。

4.3.6 Raisin 服务器

Raisin服务器是Caucho公司开发的一种支持JSP、Servlet技术的Web服务器。目前最新的Web服务器为Raisin Server 2.1, 它是一个完全实现HTTP/1.1 的Web服务器, 同时完全支持Sun公司的Servlet2.3 规范及JSP 1.2 规范, 此外还增加了对XML的全面支持。同时, Caucho还推出了Raisin EE 2.1 服务器, 它在Raisin Server 2.1 的基础上增加了对EJB (Enterprise JavaBeans) 组件的支持。目前这两种产品Caucho都推出了分别运行与Windows和Unix操作系统的不同版本^[56]。

4.4 Web 知识发现

目前, Web 已经发展成为拥有 3 亿页面的分布式信息空间, 而且这个数字仍以每 4 至 6 个月翻一番的速度增加。这些海量的、异构的 Web 信息资源中, 蕴含着具有巨大潜在价值的知识。人们迫切需要能够从 Web 上快速、有效地发现资源和知识的工具, 提高在 Web 上检索信息、利用信息的效率。

Web知识发现是一项综合技术, 涉及Web、数据挖掘、计算语言学、信息学等多个领域。从更一般的角度, 可以对Web挖掘做如下定义: Web挖掘是指从大量Web文档的集合C中发现隐含模式p。如果将C看作输入, 将p看作输出, 那么Web挖掘的过程就是从输入到输出的一个映射 $\zeta : C \rightarrow p$ ^[36]。

4.4.1 Web 知识发现的任务

Web 上信息的多样性决定了 Web 知识发现的多样性。按照处理对象的不同，我们将 Web 知识发现分为两大类：内容发现和结构发现。前者指的是从 Web 文档的内容信息中抽取知识，而后者指的是从 Web 文档的结构信息中推导知识。

Web 内容（文本）发现可以对 Web 上大量文档集合的内容进行摘要、分类、聚类、关联分析，以及利用 Web 文档进行趋势预测等。由于 Web 中包含的结构信息处理起来比较困难，因此通常的 Web 搜索引擎等工具仅将 Web 看作是一个平面文档的集合，而忽略了其中的结构信息。Web 结构挖掘的目的在于揭示蕴含在这些文档结构信息中的有用模式。

4.4.2 Web 知识发现方法

在 Web 信息检索的质量方面，目前的搜索引擎很难取得好的效果，所以人们采用不同的方法进行研究以改善当前的状况。

(1) 个人信息过滤

Profile 文件为用户的描述文件，用它来表达用户的个性化要求，从动态资源流中将不必要的信息过滤掉，从而达到筛选资源的目的。该文件的获取主要分为两种，一种为用户主动填写，即用户通过关键词主动设定信息需求，或设定主题，或根据信息结构建立规则集。另一种是监视用户的信息搜索与浏览过程，采用智能主体（Agent）分析用户的偏好，然后根据用户的信息偏好进行信息推荐。

(2) 分布系统研究

一种为分布式处理的搜索引擎，它启动多种集中式搜索引擎查询后，将任务进行分配，对所有返回结果进行核查、整理、综合，而后返回给用户。另一种为 Agent 的协作。任务层 Agent 可以从人那里学习并知道如何选择信息，而信息层的 Agent 有信息结构的知识，层内的主体也可以协作。

5. 数字对象和数据库

5.1 字符编码

字符编码是信息技术应用中最重要技术基础。1980 年我国颁布了第一个汉字编码标准，即 GB2312-80《信息交换用汉字编码字符集 基本集》，奠定了中文信息处理的基础。随着国际间信息交换的扩大，ISO/IEC 制定了多八位编码标准，较好地解决了数字图书馆多文种和大字量的要求^[10]。

5.1.1 ISO/IEC10646 (GB13000)

ISO/IEC10646 建立了一个全新的编码体系，它采用 4 个“八位”（即 4 个字节）编码方式，将世界上各主要文字统一编码。这四个字节分别表示组、平面、行和

字位，每个平面含 65536 个码位空间，其中 00 组 00 平面称为基本多文种平面 (BMP)，编码空间总共 2,147,483,648 个码位(128 组×256 平面×256 行×256 字位)，编码结构见图 4。在基本多文种平面中包括 27,484 个汉字和我国少数民族文字(藏文、蒙文、彝文)等，其中藏文基本集 194 个字符，蒙文基本集 155 个字符，彝文 1215 个字符。

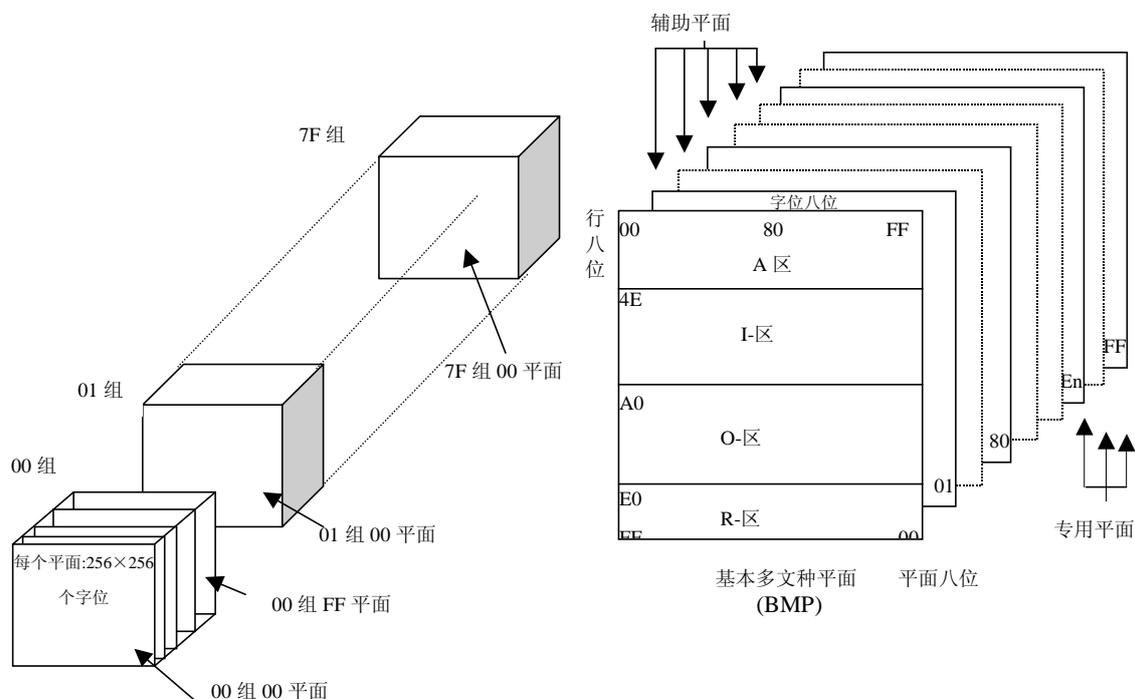


图 4 通用多八位字符集编码空间

已实现的 BMP 系统，其码位空间不足以容纳全世界的文字，为此，ISO/IEC10646 开放了 00 组的 16 个辅助平面，其中第二辅助平面用于汉字字符编码。这样，总的汉字编码空间超过 9 万，基本可以满足汉字的需求。目前，第二辅助平面有 42711 个汉字、329 个汉字部首、1 个汉字变体指示符、12 个汉字描述符，ISO/IEC10646 的汉字总数超过 70,000，覆盖了《康熙字典》、《汉语大字典》等。

5.1.2 GB18030-2000

GB2312 只收录了 6763 个汉字，远不能满足实际需要。为此，考虑到 GB2312 编码体系的延续性和现有资源和系统的有效利用，GB18030 选择了对汉字编码基本集的扩充方案，并且在字汇上与 ISO/IEC10646.1 (GB13000.1) 兼容。GB18030 采用单字节、双字节和四字节三种方式对字符编码，收录了 27,484 个汉字，编码总体结构见图 5。

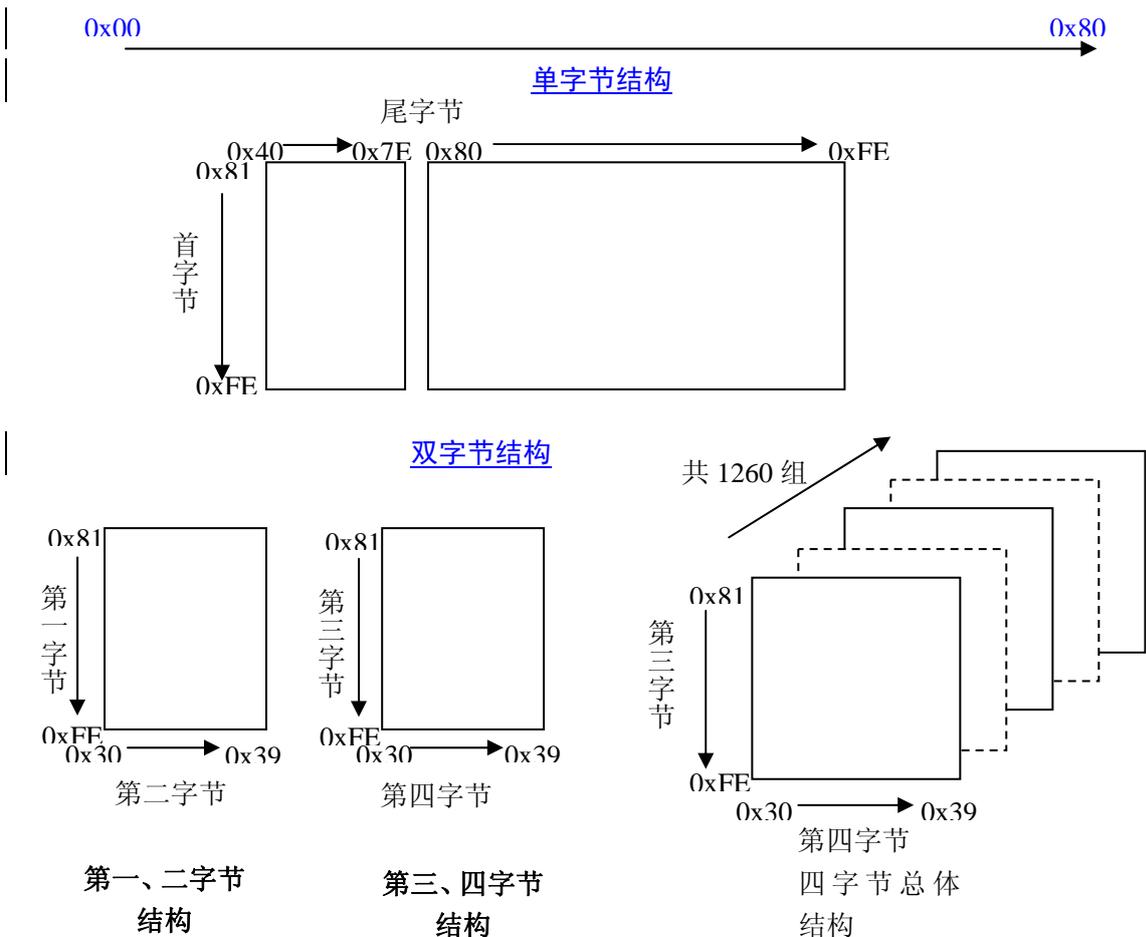


图 5 编码总体结构

鉴于中国数字图书馆的服务对象和广泛的应用领域，建议采用 ISO/IEC10646 (GB13000)，当与符合 GB18030 标准的系统互联时用映射表方式解决转换问题。同时，根据需要考虑采用必要的字型标准，以保证其正确显示和输出。

5.2 多媒体信息编码

多媒体信息包括文本、图形、动画、图像、音频和视频等。经过数字化处理后其数据量非常大，如果不进行数据压缩处理，计算机系统就无法对其进行存储、传输和交换。因此采取有效的数据压缩（信息编码）是系统处理多媒体信息的前提，而多媒体信息编码标准是核心。目前，比较有影响的标准是由 ISO/IEC JTC1 SC29 标准化小组负责制定的。ISO 是我们所熟知的国际标准化组织（International Standard Organization）。而 IEC 是国际电工协会（International Electrotechnical COMmittee），它是一个负责电工技术标准化的委员会，包括器材、元件及测量方法的标准化。它的工作涉及远程通信技术在有线、电缆、波导及共用天线电视系统中的应用。实际上，有许多工程师既是 ISO 的成员，又是 IEC 的成员。因此，它们的很多标准是一样的。而 JTC1 就是它们的合作领导小组。其中的 SC29 负责

JPEG 2000 系列标准的制定。

SC29 分 WG1、WG11、WG12 三个小组。WG1 是负责 JBIG 和 JPEG 标准的制定。WG11 负责 MPEG 标准的制定。WG12 负责 MHEG 的制定。MHEG 是 Multimedia Hypertext Expert Group 的缩写，即多媒体超文本专家组。

5.2.1 PEG

JPEG (Joint Photographic Experts Group) 是 ISO 和 CCITT 成立的联合专家工作组，致力于研究适用于彩色和单色灰度级的连续色调静止图像的压缩标准，于 1991 年发布了“高质量静止图像压缩编码”，简称 JPEG 标准。该标准采用混合编码方法，定义了两种基本压缩算法：一种是基于空间线性预测技术的无失真压缩算法；另一种是基于离散余弦变换并应用行程编码和熵编码的有失真压缩算法。其有损压缩率为 10:1~100:1，无损压缩率大约为 4:1。

当前最新的标准为 JPEG 2000 标准，这是一个基于小波技术的压缩方法。JPEG 2000 由 WG1 负责。顾名思义，JPEG 2000 是为 21 世纪准备的压缩标准，于 2000 年完成^[12]。

JPEG 2000 以往的 JPEG 标准有一个很大的飞跃。之所以这样说，是因为它有许多原先的标准所不可比拟的优点。

JPEG 2000 能实现 lossless COMpression (无损压缩)。在实际应用中，有一些重要的图像，如卫星遥感图像、医学图像、文物照片等，通常需要进行无损压缩。对图像进行无损编码的经典方法—预测法已经发展成熟，并且作为一个标准写入了 JPEG 2000 中。

JPEG 2000 还有一个很好的优点就是 Robustness to Bit Error (误差稳定性好)。因此，使用 JPEG 2000 的系统稳定性好，运行平稳，抗干扰性好，易于操作。

JPEG 2000 能实现 Progressive Transmission (渐进传输)，这是 JPEG 2000 的一个极其重要的特征。它先传输图像的轮廓，然后逐步传输数据，不断提高图像质量，以满足用户的需要。这在网络传输中有着重大的意义。近年来，世界上的网络用户飞速增长，中国的网络用户也快速增加，上网逐渐成为普通百姓日常生活中的一部分。但是，狭窄的带宽成为网络进一步发展的瓶颈。上网传输的速度很慢，下载一个文件要等待几分钟，而下载一个图片更不能让人忍受，往往半天才只传输了一小片图像。在大多数情况下，用户需要在出现了整个图像之后，才能决定是否需要这个图像。如果用户等待了半天之后，发觉所传输的图像不是想要的，那么，可以想象用户的愤恨了。虽然能通过改进硬件和网络来提高带宽，但是这种改进毕竟有一个极限，而且需要一个很长的过程。有了 JPEG 2000 后，这个问题就可以得到解决了。当下载一个图片时，马上可以看到这个图片的轮廓或这个图片的缩像，然后就可决定是否下载它了。而且，在决定下载的情况下，也可以根据需求和带宽，决定下载的图像质量的好坏，从而控制数据量的大小。

JPEG 2000 另一个极其重要的优点就是 ROI (Region of Interest, 即感兴趣区域)。用户可以指定感兴趣区域, 在这些区域, 可以在压缩时指定特定的压缩质量, 或在恢复时指定特定的解压缩要求。这给我们带来了极大的方便。在有些情况下, 图像中只有一小块区域对用户是有用的, 对这些区域, 采用低压缩比, 而感兴趣区域之外采用高压缩比, 在保证不丢失重要信息的同时, 又能有效地压缩数据量, 这就是基于感兴趣区域的编码方案所采取的压缩策略。基于感兴趣区域的压缩方法的优点, 在于它结合了接收方对压缩的主观需求, 实现了交互式压缩。而接收方随着观察, 常常会有新的要求, 可能对新的区域感兴趣, 也可能希望某一区域更清晰些。

结合 JPEG 2000 渐进传输和感兴趣区域的两大优点, 我们可以解决许多原先不能解决的问题。下面举一个例子。数字化媒体技术的发展, 产生了新的医疗技术, 比如远程诊断和图像建档及通信系统 (PACS, Picture archiving and COMmunication Systems)。在这些应用中, 数字医学图像的传输和存储, 给图像压缩技术提出了新的要求。一方面, 医学图像的数据量往往大过普通图像的好多倍, 在有效地利用通信带宽或节约存储空间的意义, 有损压缩是更期望被采用的; 然而, 另一方面, 对医学图像的任何处理, 都潜在地会给诊断结果带来影响, 如何决定哪些信息可以, 而哪些又不可以被压缩掉? 基于感兴趣区域的图像压缩, 能够在编码过程中结合观察者的主观判断, 是走出这一困境的有希望的途径。有了 JPEG 2000, 我们就可提出一种对医学图像的感兴趣区域进行交互式传输编码的方案。首先, 对图像进行最基本的压缩。然后, 根据图像接收者要求的感兴趣区域, 以及这些区域解码恢复后要达到的视觉质量, 对感兴趣区域进行特定压缩率的编码。由于 JPEG 2000 渐进式压缩特点, 远端不必等待所有的数据都传到, 即可从传输的码流中解压出逐步清晰的图像; 用户在观察中可以多次指定新的感兴趣区域, 编码过程在已经发送的数据基础上继续编码, 而不需要重新开始。

JPEG 2000 一个最大改进是它采用小波变换代替了余弦变换。余弦变换是经典谱分析的工具, 它考察的是整个时域过程的频域特征, 或者整个频域过程的时域特征, 因此对于平稳过程, 它有很好的效果, 但是对于非平稳过程, 它的不足是显而易见的。小波变换是现代谱分析工具, 它既能考察局部时域过程的频域特征, 又能考察局部频域过程的时域特征, 因此即使对于非平稳过程, 它也是强有力的工具。

近年来离散小波变换在包括压缩在内的图像处理与图像分析的各个领域得到了广泛应用。这主要因为小波的时频局域化使它在信号分析中有着优良的性质, 而且由于它对高频成分采用由粗到细渐进的时空域上的取样间隔, 从而能像物理上自动调焦看清远近不同景物一样放大任意细节。因此, 小波分析被誉为数学上的显微镜, 是构造图像多分辨率表示的有力工具。它的快速算法又使它如虎添翼。

它的多分辨率分析提供了我们进行渐进式压缩的基础。

小波在空间和频率域上的局域性，是统计意义上的局域性。这里说的局域性，指的是一个变换系数实际牵涉到的图像空间范围是局部的。因而，要完全恢复图像中的某个局部，并不需要所有的编码都被精确保留，只需要对应于它的一部分编码没有误差，就可以了。所以，我们能实现无损压缩和感兴趣区域压缩。

JPEG 2000 还考虑了人的视觉特性，增加了视觉权重和掩膜。这样在不损害视觉效果的情况下，大大提高了效率。

5.2.2 JBIG

JBIG (Joint Bi-level Image Experts Group) 是由美国国家标准局和许多大公司联合成立的专家组。它最初发布的版本为JBIG1，当前最新的标准为JBIG2，该小组已经制定并发布了IS 11544 (ITU-T T.82) 标准，它是一个黑白二值文件的无损压缩标准^[13]。JBIG2 具有如下特点：

(1) 大大提高了压缩性能，同样图像压缩后，比 Group 4/MMR 小 3 到 4 倍，比 JBIG1 小 2 到 4 倍。

(2) 使用了针对文本、半调色和二进制图像的专用压缩算法。

(3) 提供了有损压缩和无损压缩两种方式。

(4) 可以实现多张文档并发压缩。

(5) 格式灵活。

(6) 解压缩性能好。使用了特定的编码模式，可以使图像在软件中以 25 千万像素/秒的速度解压缩。

5.2.3 MPEG

MPEG (Moving Picture Experts Group) 是ISO/IEC JTC1 SC29 的工作组，致力于研究活动视频的压缩标准，简称MPEG标准。MPEG压缩算法除了对单幅图像进行了与JPEG算法类似的编码外(帧内编码)，还利用图像序列之间的相关特性除去帧间图像冗余，保证了在较高的图像视觉效果的前提下有较大的压缩比，一般可达 30~100 倍。MPEG具有随机访问、音像同步、可编辑性及灵活的视频窗口格式等特点，压缩后的图像可以在ISDN、LAN等网络上传输^[11]。

MPEG-1 标准

MPEG-1 标准是运动图像专家组 1993 年发布的数字存储运动图像及伴音编码标准，主要用于在 CD-ROM 上存储同步和彩色运动视频信号，1998 年的新版标准包括系统、视频、音频、一致性测试和软件仿真五部分。它是一个通用标准，既考虑了应用要求，又独立于具体应用之上。视频部分为 1.5Mb/s 活动图像压缩编码算法，对于带宽为 1.5Mb/s 的位流，能够获得可接受的图像质量。该算法帧内编码采用二维余弦变换、自适应量化、行程编码、变字长编码和 DPCM 技术，帧间编

码采用运动补偿预测和运动补偿内插技术。MPEG-1 对于较低的传输速率、窄带宽的应用（如单速 CD-ROM）是相当完善的，并通过插值可处理大于 352×240 的画面。MPEG 应用的数字存储媒体包括光盘（CD-ROM）、数字录音带（DAT）、磁盘和可写光盘，通信网络包括 ISDN 和 LAN 等。MPEG-1 标准是 VCD 标准的核心，利用 MPEG-1 音频第三层的 MP3 音乐格式已在互联网中流行。

MPEG 在压缩比为 6:1 时具有较高质量，但要达到更高的压缩比会产生较大的失真，且分辨率较低。由于其算法复杂，使得压缩和解压缩费时。

MPEG-2 标准

MPEG-2 于 1995 年成为国际标准。1999 年的新版标准包括了系统、视频、音频、一致性测试、软件仿真、DSM-CC 的扩展、先进音频编码（AAC）、系统解码器实时界面扩展等部分。它是针对 HDTV 和 DVD 等制定的运动图像及其伴音编码标准，是 MPEG-1 的扩充和完善。MPEG-2 的主要特点是：MPEG2 解码器兼容 MPEG-1 和 MPEG-2 标准；其视频数据速率为 4~15Mb/s，基本分辨率为 720×480 ；可以 30:1 或更低的压缩比提供具有广播级质量的视频图像；允许在画面质量、存储容量和带宽之间作出选择，在一定范围内改变压缩比（可高达 200:1）。MPEG-2 适用于包括大屏幕和 HDTV 在内的高质量电视和广播。原计划 MPEG-3 是支持 HDTV 的，但由于技术问题取消了 MPEG3，而代之 MPEG-2 标准的子集—MPEG-2 High-1440Level。

MPEG-4 标准

MPEG-4 “音频、视频对象的编码”完成于 1999 年，包括了系统、视频、音频、一致性测试、参考软件等。MPEG-4 与 JPEG、MPEG-1 和 MPEG-2 有很大的不同，它为多媒体数据压缩编码提供的是一种格式、一种框架，而不是具体算法，以建立一种更自由的通信与开发环境。MPEG-4 的目标是：支持多种多媒体的应用，特别是多媒体信息基于内容的检索和访问，可以根据不同的应用需求现场配置解码器。编码系统也是开放的，可以随时加入新的有效的算法模块。

为了支持对视频内容的访问，MPEG-4 采用了基于对象的压缩编码方法，它把图像和视频分割成不同的对象分别处理。基于对象的压缩编码，除了能提高数据压缩比，还能实现许多基于内容的交互功能。它可以广泛的应用于基于对象的多媒体存取、网上购物和电子商店、远程监控、医疗和教学等。

MPEG-7 标准

MPEG-7 “多媒体内容描述界面”出台于 2001 年，包括系统、描述定义语言、视频、音频、多媒体描述框架、参考软件、一致性测试七个部分，其目的是提供一套标准的描述工具，以实现多媒体信息的创建、交换、检索和重用。传统的关系数据库对结构化数据能够很好地实现存储、管理和检索等功能，对于非结构化

信息尤其是对多媒体信息，其功能却显得极为不足。因此，需要定义一种标准的描述规则和描述语言，对这些信息进行封装，并规定描述界面与检索机制，实现数据资源的全球交互和数据管理的灵活性。

MPEG-21 标准

各种不同的多媒体信息分布式地存在于全球不同的系统上，要想通过异构网络有效地传输这些多媒体信息，必然需要综合地利用不同层次的多媒体技术标准，因此，需要一个综合性的标准来加以协调，即 MPEG-21 “多媒体框架”。MPEG-21 的主要研究目标是：讨论是否需要和如何将协议、标准、技术等不同的组件有机地结合起来，讨论是否需要新的规范。MPEG-21 的基本框架要素包括内容创建、内容识别与描述、内容产品、内容发布、内容管理与使用、知识产权管理与保护、财政管理、用户的隐私权、终端和网络资源抽取以及事件报告等。

此标准的制定工作刚刚起步，首先开展了景像、技术与策略、数字项声明、数字项标识与描述的研究，逐步为多媒体信息的用户提供透明有效的电子内容传输、电子交易和使用环境。可以预见，MPEG-21 将在未来的电子商务应用中发挥重要作用。

总之，这些标准可为数字图书馆多媒体数据压缩和基于内容检索的数据应用提供更为通用的平台。我们在转换这些国际标准时对中文处理和具有中国文化特征的内容标识给予了重视。

5.2.4 MHEG

MHEG是Multimedia and Hypermedia information encoding Experts Group的缩写，即多媒体与超媒体信息编码专家小组。它是ISO/IEC/JTC1/SC29/WG12 和 CCITTSG8/Q11 组成的联合工作组的研究成果。该标准是用于实时交换的多媒体和超媒体信息对象的表示法^[14]。

MHEG 首先将声、文、图、形等基本信息单元抽象为“多媒体与超媒体信息对象”（简称MH对象），然后提出了“信息技术——多媒体与超媒体信息对象的编码表示”，并于1993年作为ISO CD13522标准草案，简称MHEG标准。该标准的目标是为各个领域开发出来的多媒体应用确立一个通用基础，因而它特别注重交互性和多媒体同步、实时表现、实时交换、最终形式表现等几方面。MHEG 在标准的设计中采用了面向对象的方法，但它对于标准的实施却并不是必须的。

5.3 数据元素

“数字化”运作的对象是“数据”。在与计算机交往中，人们大量涉及到的数据是以字符形式表现的数据。由于人文、地域、经济、政治等等复杂因素，即使使用同种文字的人群中也普遍存在着用不同字符组合（“词”）表达相同概念的现

象。然而，在人与计算机和计算机与计算机之间是以计算机资源浪费为代价解决这个差异问题。为了尽量避免这种浪费，数据元素标准在世界上应运而生。

在二十世纪八十年代以前，数据元素标准的制定主要是遵循一般标准的制定方法和程序——提出标准课题项目 征求意见 立项 产生各级草案 征求意见 审查批准 发布等等，时间跨度比较大。数据元素标准的“年产量”不到两位数。

由于信息技术应用的迅速扩展，特别是互联网的迅速扩展，对数据标准化的需求也随着“数字化”建设的推进而大大增强，相应的标准数据元素数量激增。常规标准化程序已经无法满足这种急速增长的需求。面对着数量巨大的有待标准化的数据元素，在各行各业纷纷制定自己急需的数据元素标准的同时，国际上提出了支持大规模快速制定数据元素标准的方法。这些方法受到广泛重视，并且被一些国家和国际纳入标准化活动范畴，制定并发布了相应的标准化文件。其中的突出代表是美国国防部的 8320. 1-M-1《数据标准化规程》和国际标准化组织的 ISO/IEC 11179《数据元素规范和标准化》系列标准[修订版可能的名称为《数据管理和交换 – 元数据注册系统 (MdR)》]。这些标准化文件的主要目的在于支持在短时间内完成大量数据标准的制定，换言之，迅速生成大量“元数据”，以满足数据高效率交换和使用的要求。这些标准化文件所规定的其实就是随时成批制定数据元素标准的方法。

为了描述这些方法，这些文件提出和引用了一系列概念，例如，数据元素 (data element，也翻译为“数据元”)、标准数据元素 (standardized data element)、元数据 (metadata)、属性 (attribute)、实体 (entity) 等等。在国际标准 ISO 11179 中，这些术语的定义如下。

- ◆ 数据元素：数据单位；通过一组属性来规定这些单位的定义、标识、表示形式和可能的值。
- ◆ 标准数据元素：在数据元素注册系统内鉴定认可供优先使用的数据元素。
- ◆ 元数据：定义和描述其他数据的数据。
- ◆ 属性：对象或实体的特性。
- ◆ 对象：有形或无形世界的任何部分。
- ◆ 实体：人们感兴趣的任何有形的或无形的事物以及事物之间的联系。

在 ISO/IEC 11179 中，把数据元素看成是由三个部分构成的：对象类别 (object class)、特性 (property) 和表示形式 (representation)。

在信息系统环境中，数据元素是实体 (或者对象) 及其属性的表示形式。数据元素名一般由实体名和属性名组成。例如，某个单位的实体“员工”，“员工”这个实体的属性有 (例如)“姓名”、“识别号码”、“住址”。如前所述，元数据是定义和描述其他数据的数据。在这个例子中，“员工”是实体元数据，“姓名”、“识别号码”、“住址”等是属性元数据。按定义，数据元素是数据单位，通过一组属

性来规定这些单位的定义、标识、表示和可能的值。在这个例子中，表示“员工”这个实体及其属性的数据元素就有“员工姓名”、“员工识别号”、“员工住址”。描述具体员工的数据由若干相关的数据元素组成，例如“员工姓名 + 员工识别号 + 员工住址”就形成关于员工的数据，三个数据元素构成的数据。

元数据也是数据，可以采用类似于在数据库中存储和检索一般数据的方法在总库中存储和检索。把元数据当作检索数据的关键词，显然将非常便于使用数据。同样，通过在线数据传播可以随同数据一起提供元数据。数据用户可以通过首先查看元数据来确定选取所需要的数据。

尽管各个组织产生不同类型的数据，对元数据有不同的要求，但是元数据在各个组织之间有许多共同之处，因此可以形成各种普遍性分类方案。例如按系统、应用、管理分类：

- ◆ 系统类元数据：反映程序设计和数据库管理中所需要的物理和逻辑特性，包括诸如文件定位、存储媒体、记录格式、数据库模式、数据词典等等之类的信息；
- ◆ 应用类元数据：为理解或使用数据所需要的信息（含描述性信息），诸如术语定义、收集规程和手段、数据收集后处理等的信息；
- ◆ 管理类元数据：费用、日程安排、预算以及与数据收集程序相联系的和供数据分析员管理用的有关信息。

此外，还可以按照更普遍的归类方式分类，例如，实体元数据、属性元数据、定量元数据、定性元数据、类属元素元数据、类属元素定量元数据等等。具体分类方案可以根据使用目的和范围确定。

自 1991 年美国国防部发布 DoD 8320.1- M-1 以来，围绕这类支持大规模快速制定数据元素标准的活动在国际上迅速展开。ISO/IEC JTC1 SC32 专门成立 WG2（元数据工作组）负责推进这方面的标准化活动，并且于 1994 年发布了 ISO/IEC 11179 系列标准。经过短期试用，WG2 又开始了对 11179 的全面修订。11179 实质上是一种管理工具或方法，用于管理可以共享的数据，主要是元数据。按照 WG2 专家们的设想，11179 如果得到世界各国广泛支持，按照它的方法对数据元素实施标准化并且按照它规定的程序注册元数据（即数据元素标准化的产物），将在世界上建立起一个权威的综合性的关于数据的参考信息的来源。目前，有关的国际标准情况如下：

5.3.1 ISO/IEC 11179 系列标准

- ◆ 修订 ISO/IEC11179-1（数据元素标准化和规范的框架）。该项目将对 ISO/IEC 11179-1 的 1999 版标准进行修订，利用数据建模的思想、概念数据模型、元模型及 UML 技术等来重新确定设定数据元素规范及相关的标准化规程。
- ◆ 修订 ISO/IEC11179-2（数据元素分类法）。该项目将对 ISO/IEC11179-2 的 1999 版标准进行修订，将根据 ISO/IEC11179-3 中的新设想重新研究、确定数据元素的

分类法。

- ◆ 修订 ISO/IEC11179-3 (注册系统元模型)。2001 年 10 月产生了 11179-3 的 FCD 文件, 2002 年初提交 DIS 草案。这个标准是新一代元数据管理标准的核心标准, 它规定了元数据注册系统结构, 描述了元数据的基本属性, 以及注册元数据时的符合性要求。该标准适用于以下活动: 定义和规范元数据注册系统以及元数据注册系统编目; 设计和规范面向应用的数据模型、数据库和数据交换报文类型; 在通信和信息处理系统中实际使用数据; 以及在各种元数据汇集点之间进行交换或引用。
- ◆ 修订 ISO/IEC11179-4 (数据定义格式的规则和指南)。该项目将为统一数据定义的格式提供所需遵循的基本原则, 并为数据元素管理提供相关的指导, 是新一代元数据管理标准的核心标准。它规定了元数据注册的结构, 并提出了元数据注册系统管理所需的规则和规程。
- ◆ 修订 ISO/IEC11179-5 (数据元素命名与标识原则)。这个标准描述了命名和标识结构, 规定了数据标识的注册规则并且就结构化的命名约定给出了指导原则。
- ◆ 修订 ISO/IEC11179-6 (数据元素的注册)。这个标准对数据元素注册管理做了规定。

5.3.2 ISO/IEC TR 20943 系列技术报告

ISO/IEC TR 20943(实现元数据注册系统内容一致性的程序)是系列技术报告。主要是为了支持实施 11179 规定的注册。目前已经产生了 20943-1《第一部分: 数据元素》, 是 JTC1 一级的最终草案。它描述了用于支持在注册系统中实现数据元素及其各项元数据注册的一致性的程序。在对数据元素及其关联的元数据加以抽象(使其概念化)时, 它可以作为用户指南。

5.3.3 业务对象登记框架

“业务对象登记框架”是由日本联合中国、韩国共同提出的新工作项目。该项目将通过软构件登记库的思想来提高软构件(如 Java、CORBA 等中间件)的复用能力。通过该项目研究将建立一种业务对象登记管理机制, 为未来的电子商务、电子政务等应用提供更高效的开发环境和方式。

目前, 国际上数据图书馆使用的许多元数据已经比较成熟, 得到比较广泛的使用, 例如: Dublin Core、CDWA、GILS、EAD 和 TEL。

Dublin Core 元素描述的对象是网络资源。1995 年由美国 OCLC 公司发起, 国际性合作项目 Dublin Core Metadata Initiative 设计, 由参与合作项目的机构共同维护修改。1999 年 7 月发布了 1.1 版。DC 由 15 个元素组成, 每个元素都根据 ISO/IEC11179 定义了 10 个属性。DC 依据其所描述内容的类别和范围可分为三组: 1) 对资源内容的描述; 2) 对知识产权的描述; 3) 对外部属性的描述。

CDWA (Categories for the Description of Works of Art) 是针对描述艺术的需求而设计的, 主要用于博物馆界。CDWA 格式的研究机构是 The Art Information Task Force (AITF)。CDWA 的维护机构是 The Getty Research Institute。CDWA 元数据包括 27 个一级元素, 可描述艺术品的物理形态和数字化图像, 以及艺术品的特征、艺术品的保存与管理方面的内容。

GILS (Government Information Locator Service) 项目是为公众提供可以方便地检索、定位、获取公共信息资源的服务。GILS 体系是一个分布式信息资源利用体系。各政府机构利用 GILS 规定描述自己的信息资源, 建立相应的资源目录和检索系统; GILS 体系的基本构建要素是对具体资源进行描述的元数据, 即 GILS 定位记录。目前 GILS 定位记录的核心元素共有 28 个。用于描述主要来自政府的公用信息资源, 其中包括内容、位置、服务方式、存取方法等。

EAD (Encoded Archival Description) 由美国国会图书馆网络开发&MARC 标准办公室维护。EAD 格式主要用于描述档案和手稿资源, 包括文本文档、电子文档、可视资料和声音记录, 适用于档案资料的管理与服务。

TEI (The Text Encoding Initiative) 是一套用于电子形式交换的文本编码标准。它规定了对电子文本的描述方法、标记定义、记录结构和文本编码方式。TEI 可对元数据 (书目信息) 和内容数据进行描述。

5.4 标记语言

在计算机诞生的早期, 标记语言就被用来描述文本结构和文本的显示格式。今天, 常用的标记语言都属于 SGML 体系。

5.4.1 SGML

1986 年, 国际标准化组织发布了 ISO8879-1986“标准通用标记语言 (SGML)”的正式文本 (我国的相应标准是 GB/T14814-1993), 使 SGML 成为通用的描述各种电子文件的结构及内容的国际标准, 为创建结构化、可交换的电子文件提供了依据。利用 SGML 可以将来源不同的信息, 如 SGML 片断、字处理文件、数据库查询结果、图形文件、视频文件等各方面的信息内容组装在同一个文件中。

SGML 是建立和处理电子化文件的基础性标准。由于它具有内容丰富、可扩充性强等特点, 得到了一些国际组织和机构的认可。一些大型的公司和企业已利用 SGML 定义了公司内部专用的文件格式, 实现组织间对结构化信息的交换和自动化处理, 使得 SGML 在大型的、技术实力强的组织和系统中得到广泛应用。

SGML 不仅是一个标记语言, 它是一个用来定义标记规范的系统。SGML 框架中有一个规范叫做文档类型定义 (DTD)。DTD 是从实体和元素的一般概念上建立起来的。它定义了一类特定文档中哪些实体和元素是合法的, 它声明了文档的基本字符集编码。

SGML 的设计目标是建立一种记录和存储高质量文本的灵活方法。它的灵活性体现在允许文本资料的创建者为他们特殊需要定制 DTD。

SGML 灵活性带来的缺点是需要复杂的软件去处理它。写一个简单的 DTD 的分析程序和表现程序并不难，真正可怕的工作是做一个通用的、可解析任何 DTD、并能结合各种样式表信息在显示器上输出的软件包。因此，完整的 SGML 不适用于有系统互操作性要求的数字图书馆。

5.4.2 HTML

为满足普遍用户群体的需要，尤其是适应 90 年代互联网的推广应用，一种易学、易用的文本描述语言应运而生。它就是推动互联网飞速发展的 HTML，即超文本标记语言。HTML 是 SGML 的子集。HTML 语言简单易用，按照统一规则建立的文本可以让不同的用户利用浏览器进行阅读。HTML 是用于描述文档结构和表现的语言。HTML 文件包括需要显示的文本以及用以描述格式和结构的标记。HTML 标记总是被包含在符号“<”和“>”中。大多数的 HTML 标记成对出现。少数标记是自我包容的，不含任何内容。

但是，HTML 只是网页格式的描述语言，它无法对结构化的数据进行描述和处理，没有提供规范文件以支持应用软件对 HTML 文件进行结构校验。这些都限制了 HTML 在特定领域的发展。

5.4.3 XML

随着互联网内信息的日益膨胀，特别是基于互联网商业信息交换的日益增多，HTML 已不堪重负。虽然可利用 SGML 描述网页，但 SGML 过于庞大、复杂，并且对于网络应用某些选项有些多余，这样就促成了一种兼顾 SGML 的关键功能和 HTML 简洁、易用优点的新的标记语言——XML (eXtensible Markup language, 可扩展标记语言) 的诞生。XML 是 SGML 的另一个子集，它免除了 SGML 的繁琐，但继承了 SGML 具有的可扩展性、结构化及可校验性，从而使 SGML 的优秀品质能方便而直接地被用在 Web 开发上。同时，通过 XML 能够很容易地按照统一格式来组织、传送和处理数据，从而解决目前 HTML 在处理大型的、非常复杂的文件类型方面和有效性验证方面的不足。

中国在制定 XML 国家标准时考虑了中文处理。由于许多国家和地区都存在各自的代码集，非 UTF-8 和 UTF-16 编码的 XML 文本的开始应该有一个带有代码声明的文本声明，如：

```
<?xml version=" 1.0" encoding=" 字符编码名称"?>
```

在文件实体中，编码声明是 XML 声明的一部分。用“UTF-8”、“UTF-16”、“ISO-10646-UCS-2”和“ISO-10646-UCS-4”来表示 ISO/IEC10646 的编码和各种变体；用“ISO-8859-1”、“ISO-8859-2”和“ISO-8859-9”来表示 ISO8859 的编

码等。由于 ISO/IEC646 (ASCII) 是 UTF-8 的子集, 一般的 ISO/IEC646 实体不需要严格的编码声明。但是, 我国目前强制实施的中文编码字符集标准分别是 GB2312 (6763 个汉字)、GB13000.1 (等同于 ISO/IEC10646.1) 和 GB18030 (包括 27484 个汉字), 因此, 在 XML 文件中应加入字符编码声明来指明该文件所采用的中文编码字符集, 如:

```
<? Xml version="1.0" encoding="gb2312"?>
```

除使用以上形式声明编码之外, 还可以使用 XML 元素的 lang 属性来说明使用的文种。如: xml:lang="zh" 表示中文资料。

文种说明很重要。例如, 一个 XML 搜索引擎, 从中可以知道某个 XML 文件应该如何分类, 如何创建索引, 如何使用关键字以及如何排序; 一个联机自动翻译软件, 可以籍此知道启动那种翻译系统; 一份 XML 文件, 可以通过以上说明导出适合各个地区代码集的文件, 等等。例如:

```
<P xml:lang="zh-GB"> “万事如意” 的汉语拼音是: </P>
```

```
<P xml:lang="zh-PY">wan shi ru yi</p>
```

通过以上的标记, 可以清楚地标识 “wan shi ru yi” 不是西文而是中文的拼音。显然, 文种说明的另一个用途是允许在一个 XML 文件中使用不同的文种。

5.5 元数据

元数据 (Metadata) 是关于数据的数据^[42]。元数据不仅是描述基础信息内容对象的工具 (狭义元数据), 而且是一种基本信息组织方法, 为信息系统的各层次内容提供规范的定义、描述、交换和解析机制, 为分布的由多种和多层内容构成的信息系统提供互操作和整合的纽带, 为计算机智能地识别、处理、集成各种信息内容、信息过程和信息系统提供工具, 称为广义元数据。

5.5.1 数据资源描述的层次

数字图书馆中的数字资源对象应该可以灵活地进行聚类、分层或集成, 元数据不仅仅符属于它所相关的数字资源对象, 还可以与任意类目、层次、集合关联, 数字图书馆构成的结构映射为元数据的层次结构, 不同层次的属性具有继承关系, 同时它本身也可是一类数字对象, 可以有它自己的元数据, 这个元数据体系可以通过一定的注册机制提供广域网环境下的资源发现和服务接口, 并有可能进行动态更新^[46]。

5.5.2 元数据的标准规范

元数据作为描述数字对象的数据, 是所有数字信息资源建设项目的重要基础, 因此也是描述体系的重点^[45]。许多描述体系或系统推荐使用一种元数据格式作为核心格式, 允许在核心格式基础上按规范方式进行扩展。例如 CCOP 规定所有项目

或者直接使用DC格式、或者提供所使用元数据格式与DC之间的规范转换。另外在政府信息领域：英国e-GMF（E-Government Metadata Framework）、加拿大TBITS 39.1（Government Online Metadata Standard）、澳大利亚AGIL（Australia Government Information Locator）、欧盟MIReG（Managing Information Resources for e-Government）等都规定在DC基础上建构政府信息元数据格式。

5.5.3 数据登记系统

元数据登记系统（Metadata Registry, MR）是对元数据的定义信息及其编码、转换、应用等规范进行发布、登记、管理和检索的系统^[43]，支持开放环境中元数据规范的发现、识别、调用以及在此基础上的元数据转换、挖掘和复用。

MR的首要作用是建立一个权威、可靠和可持续的元数据规范及应用信息的登记管理机制，从而支持对元数据规范的应用。

MR可分为：单一命名域MR、跨命名域MR和分布式MR。

5.5.4 元数据管理机制

所谓管理机制，作为广义概念，是根据特定的本地化（Localized）和个性化（Personalized）的应用场景（Context）队裸对象（Raw objects）进行选择、组织、解释和表现以形成特定服务的系统方法与过程集合^[44]。

目前，各类机构已经定义了多种管理元数据描述语言及相应机制，其中有些相对比较成熟，有些尚待完善，但已经反映出系统地建设管理元数据的努力。

5.6 数据库技术

数据处理的中心问题是数据管理。数据管理指的是对数据的分类、组织、编码、存储、检索和维护。数据管理随着计算机和软件的发展而不断发展。多年来经历了如下三个阶段：人工管理阶段、文件系统阶段和数据库系统阶段。

数据库技术的萌芽从60年代中期产生，依次经历了层次数据库、网状数据库和关系数据库^[37]。目前关系数据库成为最普及的数据库模式，但是随着计算机应用领域的不断扩大，新的数据库技术不断涌现，在数字图书馆领域中这些新技术也占有重要的地位。

5.6.1 分布式数据库

分布式数据库系统（Distributed Database System, 简称DDBS）的研究始于20世纪70年代中期。由于数据库应用需求的拓展和计算机硬件环境，特别是网络技术的发展，使分布式数据库系统应运而生，并成为计算机技术最活跃的研究领域之一。国际上每年都召开专门会议，研究探讨分布式数据库系统的各类问题及其解决方案。英国国家计算中心专门对分布式数据库做了分析和预测，断言：“分布

式系统，特别是以分布式数据库作为该系统的核心，将称为今后计算机科学发展的主要方向之一。”^[15]

由于分布式数据系统符合当今信息系统应用的需求，符合当今信息组织的管理思想和管理方式。尤其在图书馆领域，不同版本数据库同时并存是普遍的事实。而各图书馆使用不同数据库系统也是屡见不鲜的。因此，分散数据库的协同管理是建立统一资源共享的数字图书馆时必须面对的问题，而且这种情况还将继续下去。对地理上分布的多个数据库，在尽可能少地影响其本地自治性的基础上，构造具有用户所需要的透明性的全局数据库，以支持对各数据库的全局应用和各数据库之间灵活的信息交换和共享，这就迫切需要将分散数据库通过网络连接起来，组成一个分布式数据库。

分布式数据库系统是数据库系统和计算机网络相结合的产物。世界上第一个分布式数据库系统 SDD-1 (System of Distributed Database) 是由美国计算机公司 (CCA) 于 1976 年至 1978 年设计，并于 1979 年在 DEC-10 和 DEC-20 计算机上实现。

20 世纪 80 年代，分布式数据库系统进入成长阶段。一方面因为计算机功能增强而成本下降，使得各行各业都购置了计算机，从而有利于数据的分散处理；另一方面也因为计算机网络技术的发展，降低了数据传输的费用。特别是微型计算机和超级微型计算机的出现和计算机局域网的广泛应用，则为分布式数据库技术研究和实现提供了必要的条件。因此，各国在 DDBS 上都投入了大量的人力、财力和物力，美国，西欧、日本等相继提出规模宏大的 DDBS 研制计划，例如：

- ◆ 德国的斯图加特大学研制的 POREL 系统，历时 11 年，投资 450 万马克。
- ◆ 美国 IBM 公司的 San Jose 研究室研制的 R* 和 System R。
- ◆ 美国加州大学 Berkely 分校研制的 INGRES 和荷兰阿妹斯特丹大学研制的扩展 INGRES，在 Unix/PDP 机上实现。
- ◆ 法国 INRIA 研制的 SITIUS-DELTA 系统和 IMAG 研究中心研制的 MICROBE 系统。

1987 年，关系数据库的最早设计者之一 C.J.Date 提出了完全的分布式数据库管理系统应遵循的 12 条规则，这 12 条规则现已被广泛接受，并作为分布式数据库系统的标准定义。它们是：

- (1) 场地自治性 (Local Autonomy)；
- (2) 非集中式管理 (No Reliance on Central Site)；
- (3) 高可用性 (Continuous Operation)；
- (4) 位置独立性 (Location Transparency and Location Independence)；
- (5) 数据分割独立性 (Fragmentation Independence)；
- (6) 数据复制独立性 (Replication Independence)；

- (7) 分布式查询 (Distributed Query Processing);
- (8) 分布式事务管理 (Distributed Transaction Management);
- (9) 硬件独立性 (Hardware Independence);
- (10) 操作系统独立性 (Operating System Independence);
- (11) 网络独立性 (Network Independence);
- (12) 数据库管理系统独立性 (DBMS Independence)。

分布式数据库系统已有 20 多年的发展历史, 经历了一个从产生到发展的过程, 取得了长足的进步, 许多技术问题被提出并得到了解决。20 世纪 90 年代起, DDBS 已进入商品化应用阶段, 当前, 分布式数据库技术已经成熟并得到广泛应用。一些数据库厂商也在不断推出和改进自己的分布式数据库产品。但是, 实现和建立分布式数据库系统绝对不是数据库技术与网络技术的简单结合, 而是在这两种技术相互渗透和有机融合后的技术升华, 它又产生了很多新的技术。而且, 分布式数据库系统虽然基于集中式数据库系统, 但分布式数据库系统却有自己的特色和理论基础。由于数据的分布环境形成了很大的固有的技术难度, 使得分布式数据库系统的应用别推迟。至今完全遵循分布式数据库系统 12 条规则, 特别是实现完全分布透明性的商用系统还很难见到。

我国对分布式数据库系统的研究约在 20 世纪 80 年代初期开始, 一些科研单位和高等院校先后建立和实现了几个各具特色的分布式数据库系统。如中国科学院数学研究所设计, 由该所与上海科学技术大学、华东师大合作实现的 C-POREL, 武汉大学研制的 WDDBS 和 WOODOBS, 东北大学研制的 DMU/FO 系统等等。它们的工作对我国分布式数据库技术的理论研究和应用开发起到了积极的推动作用。

随着计算机网络技术的飞速发展和广泛应用, 特别是 Internet 的普及, 使得分布式数据库系统的研究和开发变得更加活跃, 分布式数据库系统具有潜在的市场。尤其在数字图书馆领域, 分布式数据库也将是不可或缺的支撑技术之一。

5.6.2 多媒体数据库

传统的数据库管理系统在处理结构化的数据: 如文字、数值等信息方面取得了很大的成功。然而在很多应用领域中包含了多媒体数据和非结构化数据, 传统的数据库管理系统就显得有些力不从心了^[38]。

多媒体数据库是指存储和管理大量多媒体对象的数据库, 如音频数据、图像数据、视频数据、序列数据以及超文本数据 (包含文本、文本标记和链接)。由于音频视频设备、CD-ROM 和 Internet 的流行和普及, 多媒体数据库系统变得日益常见。典型的多媒体数据库系统包括 NASA 的 EOS (地球观测系统), 各种图像和音频视频数据库, 人类基因数据库和 Internet 数据库。

目前多媒体数据库的研究主要有以下三种途径:

(1) 在现有商用数据库管理系统的基础上增加接口，以满足多媒体应用的需要。

(2) 建立基于一种或几种应用的专用多媒体信息管理系统。

(3) 从数据模型入手，研究全新的通用多媒体数据库管理系统。

第一种途径实用，但是效率很低；第二种途径易于实现，但缺乏通用性，而且可扩展性差；第三种途径是研究和发展的主流，但是具有相当难度。

多媒体数据库要能够对多媒体数据进行内容语义分析，以达到更深的检索层次，即所谓的基于内容的检索（Content-Based Retrieval），具有如下特点：

(1) 从媒体内容中提取信息线索。基于内容的检索突破了传统的基于关键词检索的局限，直接对图像、视频、音频进行分析，抽取特征，使得检索更加接近媒体对象。

(2) 提取特征的方法多种多样。以图像的特征提取为例，可以提取形状特征、颜色特征、纹理特征、轮廓特征等。

(3) 人机交互进行。一般来讲，人对特征比较敏感，因此，使用基于内容检索的系统时，人与计算机相互分工配合进行检索。

(4) 基于内容的检索是一种近似匹配。这一点与传统数据库检索的精确匹配算法有明显不同。

无论从计算机进步的角度还是从信息领域的需求来看，利用多媒体是数字图书馆发展的必然趋势，多媒体数据库则是支持对各种媒体信息管理检索的关键所在。

5.6.3 海量数据库

在信息领域中，随着信息系统功能的增强和用户数的不断扩充，数据量高达几百GB甚至于TB级的超大型数据库应用已经出现。全球各大主流数据库厂商在新产品中无一例外地宣称提供了对VLDB（Very Large Database）——海量数据库的支持，众多数据库支持的数据量已达到TB级^[57]。

5.6.4 数据仓库

数据仓库是从多个数据源收集的信息存储，存放在一个一致的模式下，并通常驻留在单个站点。数据仓库通过数据清理、数据变换、数据集成、数据装入和定期数据刷新来构造。按照W.H.Inmon这位数据仓库系统构造方面的专家的说法，“数据仓库是一个面向主题的、集成的、时变的、非易失的数据集合，支持管理部门的决策过程”^[39]。

数据仓库的优势在于：首先，拥有数据仓库可以提供竞争优势。通过提供相关信息，据此测量性能并做出重要调整，以创造最大效益。其次，数据仓库可以加强生产能力，因为它能够快速有效地收集信息。在此，数据仓库促进了与用户

的联系，因为它跨越所有部门，提供了一致视图。最后，通过一致、可靠的方式长期跟踪趋势，数据仓库可以降低成本。

通常，数据仓库采用三层结构：

(1) 底层是仓库数据库服务器，他几乎总是一个关系数据库系统。使用称作网间连接程序的应用程序，由操作数据库和外部数据源提取数据。

(2) 中间层是 OLAP 服务器，其典型的实现或者是关系 OLAP (ROLAP) 模型，即扩充的关系 DBMS，或者是多维 OLAP (MOLAP) 模型，即一种特殊的服务器，它直接实现多维数据和操作。

(3) 顶层是客户，它包括查询和报告工具、分析工具和数据挖掘工具。

从结构的角度看，有三种数据仓库模型：企业仓库、数据集市和虚拟仓库。

企业仓库 (Enterprise Warehouse)：企业仓库搜集了关于主题的所有信息，跨越整个组织。它提供企业范围内的数据集成，它包含了详细数据和汇总数据。通常，企业数据仓库需要广泛的商业建模。

数据集市 (Data Mart)：数据集市包含企业范围数据的一个子集，对于特定的用户是有用的。其范围限于选定的主题。通常，数据集市可以在低端的部门服务器上实现。

虚拟仓库 (Virtual Warehouse)：虚拟仓库是操作数据库上视图的集合。为了有效地处理查询，只有一些可能的汇总视图被物化。虚拟仓库易于建立。

数据仓库已经在广泛的应用领域使用。几乎每个行业的商务管理人员都使用搜集、集成、预处理和存储在数据仓库中的数据，进行数据分析和决策。

5.7 数据挖掘

近年来，数据挖掘 (Data Mining) 引起了信息产业界的极大关注，其主要原因是存在大量数据，可以广泛使用，并且迫切需要将这此数据转换成有用的信息和知识。简单地说，数据挖掘是从大量数据中提取或“挖掘”知识^[40]。还有一些术语，具有和数据挖掘类似但稍有不同的含义，如数据库知识发现、知识提取、数据/模式分析、数据考古和数据捕捞。

5.7.1 数据挖掘的理论基础

有关数据挖掘的理论基础还没有成熟。坚实的和系统的理论基础对于数据挖掘非常重要，因为它给数据挖掘技术的开发、评价和实践提供一个一致的框架。数据挖掘的理论基础又很多：

(1) **数据规约 (data reduction)：**按照这一理论，数据挖掘的基础是减少数据的描述。

(2) **数据压缩 (data COMpression)：**根据这一理论，数据挖掘的基础是对给定的数据进行压缩，它一般通过按位、关联规则、簇等进行编码实现的。

(3) 模式发现 (pattern discovery): 在这个理论中, 数据挖掘的基础是在数据库中发现模式。

(4) 概率理论 (probability theory): 它基于统计理论。根据这一理论, 数据挖掘的基础是发现随机变量的联合分布概率。

(5) 微观经济观点 (microeconomic view): 它把数据挖掘看作发现模式的任务, 通过数据挖掘来发现对企业决策过程有用的模式。

(6) 归纳数据库 (inductive database): 在这个模式中, 数据库模式看作是由存储在数据库中的模式和数据组成的。

上述理论不是彼此排斥的。建立一个能满足这些要求的定义良好的数据挖掘框架是当前研究的目标。

5.7.2 商用数据挖掘系统

数据挖掘市场还处于起步阶段, 下面对典型的数据挖掘产品做一个简单的介绍。许多数据挖掘系统只提供某一特殊的数据挖掘功能, 下面介绍的系统提供了多种数据挖掘功能和利用了多种知识发现技术。

(1) Intelligent Miner: 这是 IBM 公司的数据挖掘产品, 其特色有两点: 一是它的数据挖掘算法的可伸缩性; 二是它与 IBM DB/2 关系数据库紧密地结合到一起。

(2) Enterprise Miner: 是 SAS 公司开发的产品, 提供多种数据挖掘算法, 它的特色是具有多种统计分析工具。

(3) MineSet: 是由 SGI 公司开发的, 特色是具有强大的图形工具, 用于实现数据和数据挖掘结果的可视化。

(4) Clementine: 是由 ISL 公司开发, 它为终端用户和开发者提供了一个集成的数据挖掘开发环境。

此外, 还有很多商用数据挖掘产品系统和研究模型, 目前这一领域发展很快, 而且处于不断成熟完善中。

6. 信息安全

信息作为一种重要的资源, 在社会生产、生活中的作用日益显示。电脑网络的建立和延伸, 打破了传统的行业、地域和发展空间的概念, 把整个地球罩在一张密密麻麻的信息大网中。由于信息网络国际化、社会化、开放化、个人化的特点, 使它在提供人们“技术共享”、“信息共享”的同时, 也带来了不安全的阴影。信息社会并不安宁, 网上信息的被泄露、篡改和假冒, 黑客入侵, 计算机犯罪, 计算机病毒传播等, 对网络信息形成重大威胁。

由于存在高度分布式的信息供应和用户, 数字图书馆对安全性提出了特殊的应用需求。Internet 并不安全, 人们可以进入网络内部, 观察网络中传递的信息包。

一般说来，安全问题的关键是怎样在不安全的网络上建立安全的应用。

因此，只讲信息应用是不行的，必须同时考虑信息安全问题。在现代条件下，网络信息安全是整个国家安全的重要组成部分，它已成为影响国家全局和长远利益的重大关键问题。

信息安全的概念经历了漫长的历史阶段，90年代以来得到了深化。它包括^[27]：

(1) 信息的保密性：保证信息不会泄漏给未经授权的人。

(2) 信息的完整性：防止信息被未经授权的篡改。

(3) 信息的可用性：保证信息和信息系统确实为授权者所用，防止由于计算机病毒或其它人为因素造成系统的拒绝服务，或者为非法者所用。

(4) 信息的可控性：对信息和信息系统实施安全监控管理，防止非法利用信息和信息系统。

(5) 信息的不可否认性：保证信息行为人不能过后否认自己的行动。

从上述安全性角度出发，通常采用加密系统来保证信息的安全。一个加密系统可以描述为： $S=\{P,C,K,E,D\}$ ；P表示明文空间，C为密文空间，K为密钥空间，E为加密算法，D为解密算法。

6.1 私钥加密

私钥加密是一组加密和解密的密钥相同，而且密钥需要保密的加密方法。私钥加密也称为单钥加密或者密钥加密。常见的私钥加密标准为DES。

数据加密标准DES（Data Encryption Standard）最初是IBM开发的一种私钥加密方法，1977年成为美国国家标准。DES的计算如果仅用软件的实现速度很慢，但对于许多应用来说，还是足够的。目前的个人计算机如果加密1M Byte大约需要一秒钟时间。

DES使用56位的密钥。它将一组数据分割成64位的块，分别单独加密。从56位密钥产生16个更小的密钥。算法的核心是对64位块进行16组连续的变换。解密使用同样的16个小密钥从相反的顺序完成逆向变换。看上去这个算法很简单，但其中却蕴含着很多精妙之处。需要注意的是，算法生成的位模式看上去完全是随机的，不包含任何原来数据和密钥的痕迹^[28]。其基本加密步骤如下所示：

(1) 对64bit的明文单元组进行初始置换IP，如将第1位换到第50位，第50位换到第28位，第28位换到第3位……，并建立置换记录表S，然后将置换结果分为32bit的左右两部分R和L；

(2) 密钥置换，将56bit的密钥分为28bit的左右两部分，每部分分别循环左移1或2位，然后从移动后产生的56位密钥中选出48位作为压缩密钥K缩；

(3) 扩展置换，将单元组右半部分R切分为每4位一小组输入，经过E-盒运算，得到6位的输出结果，即将32bit的R扩展为48位的R扩；

(4) S-盒代替, 将 R 扩与压缩密钥 K 缩作异或运算得到 48bit 的输入数据 R_i , 将 R_s 输入指定的 8 个 S-盒, 经过替换得到一个 32 位 (8 个 4 位分组) 的输出 R_o ;

(5) P-盒置换, 对 R_o 进行动态的 P-盒置换, 并将置换结果 (32bit 数据) 与单元组左半部分 L 异或得到新的右半部分 R_+ , 然后将左、右半部分 L 与 R_+ 交换;

(6) 重复步骤 2~5, 使该组合操作共循环 16 轮;

(7) 将单元组的左、右半部分交换后合并, 并进行与初始置换 IP 相逆的末置换 IP^{-1} , 即得到该单元组的最终加密结果。

DES 的解密算法与加密算法基本一致, 差别在于压缩密钥 K 的产生不同, 二者顺序正好相反, 设加密过程中 16 轮加密的压缩密钥依次为 $K_1, K_2, K_3, \dots, K_{16}$, 则解密过程中的压缩密钥便为 $K_{16}, \dots, K_3, K_2, K_1$ 。DES 的算法复杂难懂, 但却很容易通过计算机实现。

私钥加密的安全性实际上就是密钥的安全性。如果一台计算机希望传递加密的信息到远端的另一台计算机, 它必须找到一种完全安全的方法将密钥传递到远端的计算机。因此, 私钥加密通常用于可信服务间的信息交换。

除了 DES 外, 目前较为流行的对称密钥算法还有 IDEA、FEAL、LOKI、Lucifer、RC2、RC4、RC5、Blowfish、GOST、CAST、SAFER、SEAL 等。

6.2 公钥加密

私钥加密用于网络传输时, 发包和接收包的计算机必须知道密钥。这就提出了一个问题, 就是第一次一个计算机如何将密钥秘密地传到另一方。公钥加密允许所有信息在网上传输, 因此, 它又称为双钥加密或不对称加密。即使每个信息都被截获的情况下, 仍能保证加密信息的安全。

RSA 方法是常见的公钥加密算法。RSA 是 1978 年由 Rivest, Shamir 和 Adleman 共同推出的, 被认为是目前最优秀的非对称密码体制, 它是一种分组密码算法, 其安全性是建立在大整数的素数难以分解上。该算法需要一对密钥, 第一个对外公开, 第二个秘密保存。一个人 A 想将加密的信息传递给另一个人 B, 就使用 B 的公钥加密数据。B 收到数据后, 只有使用自己知道的私钥才能解密[29]。其基本算法如下:

(1) 选择两个大素数 p 和 q (通常为 100 到 200 位的十进制数), 计算: $n = p * q$;

(2) 随机选择加密密钥 e , 要求 e 和 $(p - 1) * (q - 1)$ 互质;

(3) 利用 Euclid 算法计算解密密钥 d , 满足 $e * d = 1 \pmod{(p - 1) * (q - 1)}$, 其中 n 和 d 也要互质;

(4) 数 e 和 n 是公钥, d 是私钥。两个素数 p 和 q 不再需要, 应该丢弃, 不能泄漏;

(5) 加密信息 m (二进制表示) 时, 首先把 m 分成等长数据块 m_1, m_2, \dots, m_i (采用二进制数), 块长 s , 其中 $2^s \leq n$, s 尽可能的大 (s 为小于 n 的 2 的最大次幂);

(6) 加密公式: $c_i = m_i^e \pmod n$;

(7) 解密公式: $m_i = c_i^d \pmod n$ (其中 c_i 为密文分组, m_i 为明文分组)。

RSA 公开密钥算法的主要缺点在于:

(1) 加密/解密速度太慢, 一般而言, 与 DES 相比较, RSA 的软件实现要慢一百倍, 硬件实现要慢一千倍;

(2) 处理数据量小, 最多只能有其密钥的模数大小, 如一个 1024 的 RSA 公共密钥最多只能加密 1013 位数据 (多出 11 位长度要用于编码)。

双钥系统有很多优点, 也有一个重大的缺点。缺点是需要确认使用的公钥九十某人的公钥。通常所又密钥都由一个可信的权威认证机构来生成和确认。权威认证机构生成认证签名信息, 指定公钥给其持有者。只要保证权威认证不被破坏, 整个机制就能有效工作。

6.3 数字签名

所谓数字签名, 是为了接收方能够向第三方证明接收到的消息的真实性和发送源的真实性而采取的一种安全措施, 它的使用可以解决由于发送发不诚实而产生的纠纷, 它可以保证发送方不能否认和伪造信息。数字签名过程包括发送方签名发送和接收方验证签名。因为密钥和数据参与运算的结果, 所以它的特点是动态变化 (根据密钥和数据的不同), 而且签名和数据是不可分割的。数据签名基于散列函数 (一种根据一个计算机文件的内容生成一段固定长度数据的数学函数) 的概念^[30]。

一个常用的散列函数称为 MD5, 它以文件的数位数据为输入, 输出是看上去随机的 128 位的位串。MD5 可以应用于任何长度的文件。哪怕两个文件只有一个位不同, MD5 散列结果也会完全不同。反过来, 如果连个文件又相同的散列结果, 它们不相同的可能性也相当小。因此, 检查文件是否被改变过的一种简单方法就是计算 MD5 散列结果。当文件创建时计算一次, 需要检查的时候再计算一次并和创建时的结果进行比较。如果二者相同, 那么文件肯定相同。

MD5 函数有很多优点, 例如计算长文件石速度很快, 然而就像任何安全设备一样, 始终存在通过逆向工程找到拥有特定散列值文件的可能性。

一个散列值无法给出其计算者的信息。数字签名再保证数字对象的真实性方

面前进了一步：计算一个对象的散列值后，散列值用对象所有者的密钥加密。结合公钥和权威认证，形成一个数字签名。检查散列值之前，用户使用公钥解密数字签名。如果散列值结果匹配成功，就说明对象没有改变，同时也说明数字签名是使用相应的密钥生成的。

不过数字签名也有一个缺点。虽然数字图书馆的用户希望确认对象没有改变，但是他们通常不关心具体的位串，他们所关心的是对象的内容。使用数字签名，即使是一个数位的差别，就会得出内容已经被改变的结论。

6.4 认证技术

上面两节介绍了加密和数字签名。在很多情况下，往往需要一个权威机构发布、管理密钥，这就需要认证技术的支持。而且，随着 Web 技术的不断发展和普及，传统的用户名/密码的认证方式已经不能满足这些新兴应用的安全需求，这就需要认证中心（CA, Certificate Authority）进行密钥分配和安全监测管理，这也是认证技术的一个方面。本节将介绍目前通用的认证（Authentication）标准。

6.4.1 Public Key Infrastructure（PKI）

公钥系统（PKI）是一个安全认证管理体系框架的描述标准^[62]。一个认证管理系统必须具有如下功能：用户注册、分配公钥、给用户如何保护密钥的培训以及公平管理或仲裁。遵照该体系结构，可以建立认证中心（CA），分配公钥，进行安全管理等。

6.4.2 X.509

X.509 是一个认证证书的管理标准，由ITU（国际电信同盟，International Telecommunications Union）的ITU-T Study Group 17 小组负责制定和维护，当前最新的版本是该小组于 2000 年 4 月发布的X.509（2000）。X.509 标准规定了如何将数据进行加密并加上数字签名，它主要目的是促进电子商务的安全性，ITU指出：“我们愿意不断修改X.509，使得它对公钥、认证系统支持地更好，以保证无线通信和基于Web的电子商务更为安全可靠”^[64]。

目前，已经出现了很多支持X.509 的共享软件或商业产品，例如Linux下的OpenSSL服务器^[60]，Bea公司的Weblogic产品等。

6.4.3 Kerberos

Kerberos也是一个认证管理标准，但是在它的发展过程中出现了很多不统一的商业标准，因此目前IETF的Kerberos WG（krb-wg）小组正在比较各种版本，以期制定出一个统一的标准，Kerberos的正式草案已形成RFC 1510 文档^[65]。RFC 1510 为 1993 年 9 月提交，其中给出了Kerberos网络认证系统的协议规范（第五版）。

Kerberos 提供了一种在开放网络中进行身份验证的方法，这种验证独立于主机

的操作系统，不采用主机地址认证，也不需要网络上的主机进行物理上的安全控制。Kerberos 采用第三方认证方式，认证过程如下：客户向认证服务器（AS, Authentication Server）提出请求，要求某个目的服务器的信任状（credential）。AS 将信任状用该客户的公钥加密后传递给客户端。信任状包括两个内容：（1）目的服务器的访问许可；（2）临时密钥（也称为会话密钥）。客户端将访问许可发送到目的服务器，这个访问许可的内容包括客户端身份、会话密钥的副本，这些数据采用目的服务器的公钥加密。此时，会话密钥为客户和目的服务器共享，该密钥用来验证客户，也可以用来验证目的服务器。这样，客户端和目的服务器双方开始通信，通信内容采用会话密钥加密。

6.5 数字版权管理（DRM）

数据版权管理（Data Right Management, DRM）技术，是指数字化内容在生产、传播、销售、使用过程中知识产权保护与管理的技术^[47]。DRM的目标是运用技术手段遏制盗版，保护数字化内容的知识产权。第一代DRM技术主要致力于对数字化内容的安全性与加密技术的开发；第二代DRM技术则扩展到对数字版权的描述、认证、交易、保护、监控、跟踪以及版权持有者相互关系的管理上。

近年来，美国将DRM技术较多地应用在数字化音像产品的网络传播与销售中。由多家著名电子出版商、网络书店及美国图书馆协会等组成的“电子书交换工作组”（EBX Working Group），为保护电子书版权，制定了《EBX系统标准》，该标准详细说明了电子书从出版、发行、销售到实用的全过程中版权保护的技术方法，规定了有关各方的职能和权益^[48]。目前在此基础上成立了OeBF Group（Open eBook Forum Group），成为一个国际化的商业与标准化组织。

6.6 数字水印

数字水印（Digital Watermarking）技术是指用信号处理的方法在数字化的多媒体数据中嵌入隐含的标记^[50]。数字水印具有如下特点：

（1）永久性：一旦嵌入水印，数字水印将随多媒体数据存在，不受 JPEG 压缩和一般的图像处理过程的影响。

（2）高保真：嵌入数字水印的图像、视频、声音等信息保持了极高的品质，水印检测不关系到原图像检测水印过程，也不影响多媒体数据的品质。

（3）多样性：数字水印可以有多种形式如图像、文字、声音。另外可以隐含、显示、半透明等方式嵌入，形状、大小可以任意。

（4）安全性：非法用户不能获取或去除数字水印信息。

数字水印可应用于：数字作品的知识产权保护、商务交易中的票据防伪、声像数据的隐藏标识和篡改提示、隐蔽通信及其对抗。

数字水印通常根据它所加密的对象文档的类型分为两类：文本和图像水印^[51]。当对象文档为图像时，则在空间域中加数字水印。处理文档的数字水印有三种方法：文本行编码、字间空档编码和字符编码。

6.7 安全传输标准

Internet 蓬勃发展的同时，也表现出了大量安全性弱点。为了弥补这些安全漏洞，除了对 Internet 上传输的数据采取加密、签名和封装处理外，Internet 本身也提供了某些协议和相关机制，以保证数据的安全传输。下面对 Internet 的一些安全传输标准加以介绍。

6.7.1 IPSec

IPSec 是 IETF (Internet 工程任务组) 的 IP 安全性协议工作组目前正在定义的 IP 安全性附加协议，它为 IP 数据报这一层提供了认证、完整性和保密性服务的规范，在多个 RFC 文档中对它都有描述。虽然 IPSec 设计目的是 IPv6 专用，但它也可以用于 IPv4。它旨在为 Internet 通信提供安全基础。一些供应商和软件组织正在开发或提供了集成 IPSec 的产品。例如，芬兰的 SSH COMMUNICATIONS Security 公司有一种称为 IPSec Express 的产品，它是为方便符合 IPSec 的电子商业应用程序开发而设计的。此外，从 1999 年 6 月开始，NetBSD Foundation 已经将 IPSec 代码合并到了 NetBSD 发布版中。

虽然 IPSec 已经成为因特网安全性实现的一个事实上的标准，但却受到来自 Niels Ferguson 和 Bruce Schneier (后者是被广泛关注的 Blowfish 密码的设计者) 等人的批评。Ferguson 和 Schneier 认为 IPSec 已经变得过度复杂并趋于难以管理。他们认为，IPSec 比以前任何安全性协议都好得多，但是其设计中固有的复杂性已经导致了大量模糊、矛盾、低效和其它一些弱点，并最终变成了极其难懂的一套规范。他们怀疑 IPSec 是否能生成一个真正安全的可操作协议。但是，他们也承认：尽管不满意使用当前形式的 IPSec，但与当前其它协议相比，为了保护网络的安全，还是建议使用 IPSec。

6.7.2 Secure HTTP (SHTTP)

安全 HTTP (SHTTP) 是在应用层运行的 HTTP 安全性扩展。它旨在当支持不可抵赖性以及允许使用多种密码算法和密钥管理机制的同时，提供保密性和认证。虽然可以在会话之前获得经 Kerberos 服务器同意的初始密钥，或者可以在一个会话中生成下一个会话要使用的密钥，但通常将 RSA 用于初始密钥协商。

6.7.3 Secure Sockets Layer (SSL)

安全套接字层 (SSL) 是由 Netscape COMMUNICATIONS 开发的用来向因特网会话提供安全性和保密性的握手协议。它支持服务器和客户机认证，并且被设计成

协商加密密钥以及在交换任何数据之前认证服务器。它使用加密、认证和 MAC 来维护传输信道的完整性。

虽然 SSL 最适合用于 HTTP，但它也可以用于 FTP 或其它相关协议。它在传输层运行并且是独立于应用程序的，因此象 FTP 或 HTTP 之类的相关协议可以放在该层之上。使用初始握手来对服务器进行认证。在这一过程中，服务器把证书提交到客户机并指定要使用的首选密码。然后，客户机生成在即将进行的会话期间使用的密钥，然后将它提交给服务器，并相应地用服务器的公钥对它加密。服务器使用其私钥解密消息，恢复密钥，然后通过向客户机发送一条使用该密钥加密的消息来向客户机认证自己。使用这一达成协议的密钥对加密的数据进行进一步的交换。

可以用第二阶段（可选）来进一步增加安全性。这里，服务器发送一个质询，客户机对此作出响应，向服务器返回该质询的数字签名和客户机的公钥证书。

质询阶段通常是使用带有用于消息摘要的 MD5 的 RSA 执行的。也可以使用各种对称密码，包括 DES、三重 DES、IDEA、RC2 和 RC4。公钥证书符合 X.509 标准。SSL 目前的版本为 3.0。

6.7.4 Transport Layer Security (TLS)

传输层安全性 (TLS) 协议是 IETF 标准草案，它基于 SSL 并与之相似。它的主要目标是在两个正在通信的应用程序之间提供保密性和数据完整性。它由两层构成。较低的层称为 TLS Record 协议，且位于某个可靠的传输协议（例如，TCP）上面。这一层有两个基本特性，具体说该连接是专用的并且是可靠的。它用于封装各种更高级协议，但也可以不加密地使用。通常使用加密时，生成的用于这个加密的密钥专用于每个连接，这些密钥基于由另一个协议（例如，更高级别的 TLS Handshake 协议）协商的密钥。

TLS Handshake 协议提供了具有三个基本特性的连接安全性，即可以使用非对称密码术来认证对等方的身份，共享密钥的协商是安全的，以及协商是可靠的。

与 SSL 一样，TLS 是独立于应用程序协议的，其使用的加密算法的种类与 SSL 使用的相似。然而，TLS 标准把如何启动 TLS 握手和如何解释认证证书的决定权留给运行于其上层的协议的设计者和实现者来判断。

TLS 协议的目标，按其优先级顺序来说，是密码安全性、互操作性和可扩展性。最后一个目标意味着 TLS 提供了一种框架，当新的和改进的非对称以及其它加密方法可用时，便可以将它们引入该框架。

6.7.5 Wireless Transport Layer Security (WTLS)

无线应用程序协议 (WAP) 体系结构中的安全性层协议称为无线传输层安全性 (Wireless Transport Layer Security, WTLS)。它在传输协议层之上操作，它是模

块化的，是否使用它取决于给定应用程序所需求的安全性级别。WTLS 为 WAP 的上一层提供了保护其下的传输服务接口的安全传输服务接口。另外，它为管理安全连接提供了一个接口。

WTLS 与 TLS 非常相似，但它最适合用于等待时间相对长的窄带传输网络。但是，它添加了一些新特性，例如，数据报支持、经过最优化的握手和密钥刷新。与使用 TLS 一样，它的主要目标是在两个正在通信的应用程序之间提供保密性、数据完整性和认证。

6.7.6 Secure Electronic Transaction (SET)

安全电子交易 (SET) 协议是由 VISA 和 MasterCard 国际财团开发的作为在开放网络上的进行安全银行卡交易的方法。它支持 DES 和三重 DES 以实现成批数据加密，并支持用 RSA 对密钥和银行卡号的公钥进行加密。

虽然 SET 被认为是非常安全的，但太安全使它的速度相对很慢。另外，用户需要正确发出的数字证书，因此不能象 SSL 或 TLS 那样以一种简单的特别方式使用它。由于这些原因，以及许多银行将风险和银行卡安全性漏洞的后果转嫁给其商业客户，所以 SET 的采用远没有开始设想的那么多。

6.7.7 Secure Wide Area Network (S/WAN)

安全广域网是由 RSA Data Security 推动的，它旨在促进基于 Internet 的虚拟专用网 (Internet-based Virtual Private Networks) 的广泛部署。S/WAN 支持 IP 级的加密，因此它比 SSL 或 TLS 提供了更基本的、更低级别的安全性。

VPN 提供了这样一种机制，当用户使用 Internet 时，可以在他们之间维护一个安全隧道。例如，可以连接远程办公室，而不需要专门租用线路，避免增加成本，或者避免专线点对点的不便。通过对通道传递的消息进行加密，来保证信息的安全性，可以避免第三方的有效截获。但是，由于以前不同标准的商业产品的开发，造成 VPN 技术的混乱。S/WAN 就是为了解决这种混乱而进行的一个尝试。

目前，最初的 S/WAN 设想已经不再进行了，但是虚拟专用网协会 (Virtual Private Network Consortium) 又提出了类似的设想——Linux FreeS/WAN，这是 IPsec 在 Red Hat Linux 中的实现，并遵照 GNU GPL 提供了可自制的 Linux 的 VPN 实现。

6.7.8 Secure Shell (SSH)

安全 Shell (SSH) 是目前正在由 IETF 的 SECSH 工作组标准化的协议。它允许网络上的安全远程访问。它可以使用多种方法来认证客户机和服务器，并能在支持 SSH 的系统之间建立加密的通信通道。该协议可以应用于许多领域，例如，建立 VPN 或者在服务器上创建安全远程登录以替换非安全的 telnet、rlogin 或 rsh。

6.7.9 Secure Mail (S/MIME)

S/MIME 是保证电子邮件安全的协议。发送明文电子邮件等于发送一张任何人都可读的明信片，因为电子邮件是在一条不确定的、分段路由的线路上传输，并且沿途的许多节点可以毫不费力地查看它。此外，电子邮件在一定阶段保存在网络邮件服务器的半公开区域或 ISP 的存储器中。电子邮件还可能被错误路由或错误地成批发送。许多产品提供安全电子邮件，安全电子邮件既可以作为使用 PGP (Pretty Good Privacy, PGP) 的补充产品提供，也可以使用如安全 MIME (S/MIME) 这样的协议将数字签名和加密工具添加到用户的应用程序（例如，Netscape）创建的邮件消息中。MIME (Multipurpose Internet Mail Extensions, 多用途因特网邮件扩展) 是一种因特网邮件标准化的格式，它允许以标准化的格式在电子邮件消息中包含增强文本、音频、图形、视频和类似的信息。然而，MIME 不提供任何安全性元素，S/MIME 添加了安全保护。目前，S/MIME 已经得到了许多厂商的认可，包括 Netscape、QualCOMm、Microsoft、Lotus、Novell 等。

然而，将加密的电子邮件引入大型组织会引起一些新问题。反病毒产品可能无法识别加密的危险附件。另外，扫描程序和防火墙也可能会遇到类似的问题。

6.8 XML 安全机制

6.8.1 XMLEnc

XMLEnc (XML Encryption) 是 W3C 的 XML Encryption Work Group 制定的对 XML 文档的内容进行加密/解密的规范，同时它还定义了相应的语法来表达加密的内容和接收者用来解密所需的信息^[58]。最新的 XML Encryption Requirements 发布于 2003 年 3 月 4 日，该规范介绍了 XMLEnc 的设计原则、使用范围，此外定义了加密语法、数据模型、格式、加密算法和其它外部要求。

XMLEnc 必需实现以下规范：

- (1) 必需能够保证接收者从加密后的数据中识别原始的数据类型（如 XML 的 CDATA、image/gif）。
- (2) XML 中的二进制数据必需使用 Base64 编码。
- (3) 除了按 XMLEnc 定义的语法加密的信息之外，规范中不定以对非 XML 数据的表达格式。

此外，该规范还定义了对 XML 文档加密的步骤：对 XML 文档解析——验证文档——文档变换（压缩等）——加密、签名。

6.8.2 XMLSig

XMLSig (XML Signature) 是 W3C 的 XML Signature Work Group 和 IETF 联合制定的与 XML 兼容的 Web 资源和协议信息数字签名规则和语法^[59]。有关该规范的

描述目前已经有了正式的RFC3275、RFC3076 和RFC2807 文档。

该规范定义了如何生成和表示 XML 格式的数字签名。这种 XML 签名可以用于任何数字内容或数据对象上，其中当然也包括 XML 文档。

7. 软件工程

软件工程是指导计算机软件开发和维护的工程学科。采用工程的概念、原理、技术和方法来开发与维护软件，把经过时间考验而证明正确的管理技术和当前能够得到的最好的技术方法结合起来，这就是软件工程。

软件工程强调使用生存周期方法学和各种结构分析及结构设计技术。它们是在七十年代为了对付应用软件日益增长的复杂程度、漫长的开发周期以及用户对软件产品经常不满意的状况而发展起来的。随着计算机技术的不断发展，软件工程技术也逐步向系统化、集约化、智能化方向发展。在传统软件工程的基础上，又派生了许多新的理念和方法。

7.1 软件开发技术

软件工程的发展总是同编程语言的发展相辅相成的。第三代语言（3GL），如 FORTRAN、C，是于软件开发方法中的面向过程的方法相对应的。第四代语言（4GL），如 SQL，支持面向实体的风格。第五代语言，如 SmallTalk、C++，使用的是面向对象的软件的开发风格。而展望中的第六代语言（6GL），很有可能就是使用了面向主体的软件开发模型。下面比较了面向对象方法和面向主体方法。

7.1.1 面向对象（OO）

面向对象技术最初是从面向对象的程序设计开始的，它的出现以 60 年代 simula 语言为标志。80 年代中后期，面向对象程序设计逐渐成熟，被计算机界理解和接受，人们又开始进一步考虑面向对象的开发问题。这就是九十年代以 Microsoft Visual 系列 OOP 软件的流行的背景。面向对象的编程方法强调对象的“抽象”、“封装”、“继承”、“多态”。

传统的结构化分析与设计开发方法是一个线性过程，因此，传统的结构化分析与设计方法要求现实系统的业务管理规范，处理数据齐全，用户能全面完整地描述其业务需求。传统的软件结构和设计方法难以适应软件生产自动化的要求，因为它以过程为中心进行功能组合，软件的扩充和复用能力很差。

对象是对现实世界实体的模拟，因而能更容易地理解需求，即使用户和分析者之间具有不同的教育背景和工作特点，也可很好地沟通。区别面向对象的开发和传统过程的开发的要素有：对象识别和抽象、封装、多态性和继承。

对象（Object）是一个现实实体的抽象，由现实实体的过程或信息来定义。一个对象可被认为是一个把数据（属性）和程序（方法）封装在一起的实体，这个

程序产生该对象的动作或对它接受到的外界信号的反应。这些对象操作有时称为方法。对象是个动态的概念,其中的属性反映了对象当前的状态。类 (Class) 用来描述具有相同的属性和方法的对象的集合。它定义了该集合中每个对象所共有的属性和方法。对象是类的实例^[31]。

对象概念对软件解决方案具有莫大的好处,在设计优秀合理的情况下尤其如此。可以只编写一次代码而在今后反复重用,而在非 OOP 的情况下则要在应用程序内部各个部分反复多次编写同样的功能代码。所以说,由于面向对象编程减少了编写代码的总量,从而加快了开发的进度同时降低了软件中的错误量。

面向对象的另一优点是对代码结构的影响。像继承之类的面向对象概念通过简化变量和函数的方式而便利了软件的开发过程。面向对象可以更容易地在团队之间划分编码任务。同时,由于采用面向对象,辨别子类代码的依附关系也变得更简单了(比如说继承对象的代码)。此外,软件的测试和调试也得以大大简化。

7.1.2 面向主体 (AO)

20 世纪 90 年代,随着计算机网络、计算机通信等技术的发展,对于智能主体 (Intelligent Agent) 的研究不仅成为分布式人工智能研究的一个热点,而且也是信息技术关注的一个热点。智能主体是一种处于一定环境下包装的计算机系统,为了实现设计目的,它能在那种环境下灵活地、自主地活动。

当前对主体的研究主要集中在理论方面。Shoham 提出的面向主体编程 (AOP, Agent-Oriented Programming) 使用的主体定义为:“一个主体是这样—个实体,它的状态可以看作是由信念 (belief)、能力 (capability)、选择 (choice)、承诺 (COMmitment) 等心智构件 (mental COMponent) 组成。”

基于主体的软件开发方法的优点是便于软件的创建,有很强的交互操作能力,它代表了一种新的、功能强大的解决大规模软件工程问题的方法^[17]。在该方法中,应用程序编写为软件主体,即软件构件,这些构件之间通过主体通信语言可以进行比普通消息传递更规范、更明确的通信。主体通信语言的显著特征是其表达性,允许数据和逻辑信息、单个命令和程序的交换。由于主体通信语言具有平台无关性,在异构环境下更能体现出面向主体编程方法的优越性。

目前国际上比较通用的主体通信语言是 KQML (Knowledge Query and Manipulation Language) 和 ACL (Agent COMmunication Language)。而本体论则是主体之间通信的基本保证。

KQML 标准

KQML (Knowledge Query and Manipulation Language) 标准是美国 ARPA 的知识共享计划 (knowledge sharing effort) 提出的。KQML 定义了一种主体之间传递信息的标准语法以及一些“动作表达式”。这些“动作”主要是从言语行为理论中

演化来的。

KQML是一种最通用的主体通信语言。1997年Finin和Labrou建议一种KQML新的规范^[5]。这种新规范在KQML消息的句法和保留的执行参数方面差别很小，但是在保留的消息类型集、它们的含意和使用方面有重大的改变。

KQML规定了消息格式和消息传送系统，为多主体系统通信和协作提供了一种通用框架。特别是提供了一种识别、连接建立和消息交换的协议。消息的语义内容在KQML中规定并不详细。因为标准是开放的，各种语言可以用来传送知识，也可以集成在KQML中。

KQML分为三个层次：通信、消息和内容。全部技术通信参数协议都在通信层规定。消息层规定与消息有关的言语行为的类型。内容层规定消息内容。

ACL 标准

ACL (Agent COMmunication Language) 标准是由智能物理主体FIPA (Foundation for Intelligent Physical Agents) 制定的。FIPA是一个在瑞士日内瓦注册的非赢利协会^[6]。它的目的是为了促进基于主体的应用、服务和设备的成功实现。FIPA通过制定能及时获得的国际承认的规范来达到这一目标，这种规范可以最大限度地增大基于主体的应用、服务和设备的互操作性。截至1999年4月，FIPA已经有50多个合作成员，他们来自全世界12个国家。

FIPA97规范定义了一种语言和支持工具，如协议，用于智能软件主体相互通信。软件主体技术对这类主体采用了一种高层的观点，它的许多思想都来源于采用其他方式的社会交互作用，例如人与人之间的通信。本规范并非试图定义通常与分布式软件系统之间通信相关的底层和中间层服务，例如网络协议、传输服务等。事实上，用于物理传输组成交互主体通信动作的比特序列的这类服务都是假设存在的。

对于软件主体没有单一的、通用的定义，但是主体行为的一些特性是广泛接受的。FIPA定义的通信语言用于支持和促进这些行为。这些特性包括但不限于这些：

- (1) 目标驱动行为
- (2) 动作过程的自主决定
- (3) 通过协商和委托进行交互
- (4) 心智状态模型，例如信念、意图、愿望、规划和承诺
- (5) 对于环境和需求的适应性

主体技术对于复杂系统行为和交互操作的贡献尤其表现在高层交互上，FIPA提出的ACL是以这一观点为基础。目前，ACL在主体传输服务上定义了如下的一组最小要求：

- (1) 消息服务能够传递一个与比特序列一样的消息到目的地，通过它的界面

消息服务能发现它是否能可靠地处理高阶层被设置的 8 位比特流。

(2) 正常情况下消息服务是可靠的 (包装好的消息能够达到目的地)、准确的 (接收的消息在形式上和发送的一样)、有序的 (从主体 a 发送到主体 b 的消息到达 b 时和它从 a 发送的顺序一样)。

(3) 如果消息传递服务不能保证以上一个或者所有的特性, 它将通过消息传递服务的界面以某种方式表示出来。

(4) 主体能选择是否暂停和等待消息结果或者在等待信息回复时继续其他无关任务。这种行为的有效性是执行细节, 但是这种行为是否支持必须明确。

(5) 传递消息动作的参数。

(6) 消息传递服务将发现和报告出错情况并返回给发送主体。

(7) 一个主体将有一个名字使得消息传递服务能将消息传到正确的目的地。消息传递服务能够决定正确的传输机制 (TCP/IP, SMTP, HTTP 等), 允许主体位置的变化。

本体论 (ontology)

在网络技术快速发展的信息高速公路时代, 组织和个人之间, 以及软件系统之间的交流与协作越来越重要。但是这往往受到彼此之间不同的背景、语言、协议和技术的制约。本体论 (ontology) 是概念化 (conceptualization) 的明确的表示和描述。对某一领域中的概念有共同理解, 可以提高交流和协作的效率, 从而提高了软件的重用性、互操作性和可靠性。在本体论中概念化 C 可以表示为 $C = \langle D, W, R \rangle$, 其中 D 是一个领域, W 是领域中相关的事务状态 (state of affairs) 的集合, R 是领域空间 $\langle D, W \rangle$ 上概念关系 (conceptual relation) 的集合。本体论的研究是多领域研究的交叉, 涉及到哲学、知识获取和表示、规划、过程管理、数据库模式集成、自然语言处理和企业建模等领域, 同时它的研究又为上述各项研究服务。

对本体论的研究和应用近年来发展很快。自 90 年代以来, 举办了很多这方面的研讨会, 同时高水平的论文和成果不断涌现。在 1998 年 6 月, 第一届“信息系统中的形式化本体论国际会议”的召开标志着这一领域在逐渐走向成熟。对本体论的研究主要集中于一下几个方面:

- (1) 建立本体的方法和过程
- (2) 本体设计和评估的形式方法
- (3) 本体论的应用研究

国际上, 已经完成和正在进行的本体论研究计划主要有:

(1) 过程交换格式 (process interchange format)。目标是支持采用不同过程表示的商业过程模型的交换, 并通过在本地格式和 PIF 之间的交换实现互操作。

(2) KRSL 规划本体。知识表示规范语言 (Knowledge Representation Specification Language, KRSL) 可以用来表示规划和规划信息。

(3) 已建立的标准有：STEP、EXPRESS、CORBA、KIF 和概念图等。

(4) 已经实现的本体有：CYC、TOVE、Enterprise、KACTUS、Plinius。

(5) 计算机辅助工具：KSL ontology server 是一个 WWW 的本体编辑和浏览器，并且可以把本体翻译成目标语言。

本体论的应用主要集中在以下几个方面：

(1) 组织内部和组织间的交流。通过提供统一的框架，本体论减少了组织内部在概念上和术语上的混淆，从而使共同的理解和交流成为可能。

(2) 软件系统的互操作性。解决不同软件系统之间数据的交换和协作问题，主要研究集中于企业建模和多主体系统上。

(3) 软件工程。辅助建软件系统的规范，而且以自然语言写成的非形式化的本体可以用于对规范的手工检查，从而提高软件的重用性和可靠性。

本体论采用某种语言对概念化进行描述，因此本体论依赖于所采用的语言。按照表示和描述的形式化的程度不同，可以分为完全非形式化的、半非形式化的、半形式化的和严格形式化的本体论。形式化程度越高，越有利于计算机进行自动处理。目前人们从不同的形式化程度对各个领域设计了本体论。

从概念化对象的定义来看，一个领域中的术语、术语的定义以及各个术语之间的语义网络应是任一个领域本体论所必需包含的基本信息，同时本体论中还应包含关于同义词的描述。

7.2 过程管理方法

随着软件规模日趋庞大，软件的开发过程也成为影响软件质量的关键所在。标准化、严格的软件开发过程才能保证高质量的软件按期交付，下面比较了国际上通用的软件开发过程管理标准。

7.2.1 ISO/IEC 12207 (IEEE/EIA 12207)

ISO/IEC 12207 (IEEE/EIA 12207) 是一个侧重描述软件生命周期过程内容的模型。ISO/IEC 12207是国际标准化组织 (ISO) 和国际电工委员会 (IEC) 于1995年8月1日发布的。IEEE/EIA 12207是IEEE和EIA的联合工作组在ISO/IEC 12207上的修订，使之囊括了美国的实际工业标准。1997年发布了IEEE/EIA 12207草案第二版。

IEEE/EIA 12207的目的是建立为软件行业建立一个软件生命周期的通用框架，说明开发中各种必要的活动和相关的产品，并为使用这个框架提供了指导性建议。IEEE/EIA 12207将软件生命周期过程分为3类17种过程，每一类过程可以分解为一组活动，每个活动又可以分解为一组任务^[32]。

IEEE/EIA 12207不仅规定了各个过程中需要有哪些活动，活动中包含哪些任务，还规定了过程的输出产品，承担的角色等内容。提供过程内容的可见性。与CMM不同的是在IEEE/EIA 12207没有区分哪些活动（任务）是属于哪个成熟级别的。其

任务是“平铺”的；而CMM中则分成了5个等级的成熟度。意味其侧重点是不同的。

IEEE/EIA 12207规定了应该作什么，但并没有规定如何来实施过程。为了起到指导的作用，给出了一个裁剪过程，包括如下4个活动：

- (1) 标志项目环境；
- (2) 征求输入，考虑受剪裁影响的个组织的意见；
- (3) 选择过程活动和任务；
- (4) 将剪裁决策和原则写成文档。

上述活动标准给出了一个普通的框架，基本表示出了软件开发过程中必须遵循的步骤，需要输出的产品，实现的量化规定。

IEEE/EIA 12207采用非形式化的自然语言来描述。

7.2.2 能力成熟度模型（CMM）

软件开发能力的成熟度模型（Capability Maturity Model for Software, CMM）是软件工程协会SEI（Software Engineering Institution）在卡内基·梅隆大学开发完成的对一个组织软件开发能力进行评价的标准,它侧重于对软件开发过程和开发方法论的考察^[33]。CMM包括五个成熟等级,开发的能力越强,开发组织的成熟度越高,等级越高。目前,大多数公司处于第一级和第二级,只有很少的公司可以达到第五级。五级的具体定义如下:

(1) 初始级（Initial）

初始级的软件过程是未加定义的随意过程，项目的执行是随意甚至是混乱的。也许，有些企业制定了一些软件工程规范，但若这些规范未能覆盖基本的关键过程要求，且执行没有政策、资源等方面的保证时，那么它仍然被视为初始级。

(2) 可重复级（Repeatable）

根据多年的经验和教训，人们总结出软件开发的首要问题不是技术问题而是管理问题。因此，第二级的焦点集中在软件管理过程上。一个可管理的过程则是一个可重复的过程，一个可重复的过程则能逐渐进化和成熟。第二级的管理过程包括了需求管理、项目管理、质量管理、配置管理和子合同管理五个方面。其中项目管理分为计划过程和跟踪与监控过程两个过程。通过实施这些过程，从管理角度可以看到一个按计划执行的且阶段可控的软件开发过程。

(3) 定义级（Defined）

在第二级仅定义了管理的基本过程，而没有定义执行的步骤标准。在第三级则要求制定企业范围的工程化标准，而且无论是管理还是工程开发都需要一套文档化的标准，并将这些标准集成到企业软件开发标准过程中去。所有开发的项目需根据这个标准过程，剪裁出与项目适宜的过程，并执行这些过程。过程的剪裁不是随意的，在使用前需经过企业有关人员的批准。

(4) 管理级（Managed）

第四级的管理是量化的管理。所有过程需建立相应的度量方式，所有产品的质量（包括工作产品和提交给用户的产品）需有明确的度量指标。这些度量应是详尽的，且可用于理解和控制软件过程和产品。量化控制将使软件开发真正成为一种工业生产活动。

（5）优化级（Optimizing）

第五级的目标是达到一个持续改善的境界。所谓持续改善是指可根据过程执行的反馈信息来改善下一步的执行过程，即优化执行步骤。如果一个企业达到了这一级，那么表明该企业能够根据实际的项目性质、技术等因素，不断调整软件生产过程以求达到最佳。

除了第一级,其它每一级都有几个特别值得注意的关键过程。第二级的关键之处是建立基本的项目管理控制。他们是需求管理、软件项目计划、软件项目的跟踪和监督、软件转包管理、软件质量保证和软件组态管理。

CMM 和 ISO9001 的出发点都是通过对生产过程进行管理,来确保产品的质量。虽然它们之间有很多区别,但也有相似之处。比如,通过 ISO9001 认证的组织,可以基本满足 CMM 二级的标准和很多 CMM 三级的要求。因为 CMM 中的很多要求并没有列入 ISO9000 标准之中,所以,CMM 一级的组织也可能获得 ISO9001 的登记。同样,有些 ISO9001 规定的内容并没有列入 CMM 标准。一个 CMM 三级组织获得 ISO9001 认证几乎没有困难,CMM 二级组织申请 ISO9001 认证也有明显优势。

7.2.3 ISO9000

国际标准化组织（ISO）自1987年推出了ISO9000系列标准以来，很快得到了工业界的广泛承认，并被各国标准化机构所采用，在世界范围内形成"ISO9000现象"。一套国际标准，在如此短的时间内被众多国家采用，且影响如此广泛，是国际标准化史上从未有过的。

对提供产品的企业而言，ISO9000的作用有两个方面,体现为企业内部和企业外部。在企业内部（管理职能），能够提高质量意识,建立保证质量的机制——质量体系;实现文件化管理，改善工作机制;改进效率，提高生产率。

在企业外部（保证职能），能够提高顾客的满意程度，赢得市场；便于与用户、中间商等合作、交流；能提高企业的宣传效果，扩大声誉。

ISO9000起源于硬件产品，但它的应用并不限于硬件领域。ISO在制定ISO9000时就充分考虑了各行业及其产品的特点，在给出了通用工业领域应用标准之外，又分别针对不同行业制定了一些补充性标准和指南，其中ISO9000-3就是ISO9000在软件领域的补充性指南。在ISO9000-3的前言中对此是这样描述的：“软件的开发和维护过程不同于大多数其它工业产品，由于这一技术领域的迅速发展需要考虑其技术现状，有必要对涉及软件产品的质量体系提供补充性指南。”由此可见ISO9000已经为软件企业的应用提供了标准。

ISO9000国际标准族最初发布在1987年，第二个版本于1994年发布，当前正在使用的为ISO9000-2000年版。

ISO9000-2000年版与以前的版本相比，是一次全面彻底的“重规划”。整个被称之为“ISO9000家族”的二十七个标准，已被完全重新规划，主要内容集中在4个重要的标准中，即：

(1) ISO 9000: Quality management systems – Fundamentals and vocabulary (品质管理体系 - 基本原理和词汇)

(2) ISO 9001: Quality management systems – Requirements (品质管理体系 - 要求)

(3) ISO 9004: Quality management systems – Guidance for Performance Improvement (品质管理体系 - 绩效改进指南)

(4) ISO 9011: Guidelines on Quality and Environmental Auditing (品质与环境审核指导纲要)

基本内容遵循“品质管理八项基本原则”，采用面向过程的体系原理，对以前的要求有补充或加强，此外还特别考虑了与环境管理体系标准ISO14000的协调、配合。

8. 结束语

本报告讨论了数字图书馆支撑领域的一些相关标准。随着数字图书馆技术与应用的迅速普及，教育、科研、出版发行、博物馆、档案馆、图书馆、企业等各个应用领域的多元化的建设主体广泛投入到数字资源和服务建设之中，构成了分布式的数字图书馆应用环境。这个环境的支撑领域几乎涉及了所有的信息、计算机、网络技术。因此，支撑领域相关研究将是数字图书馆发展的基础课题之一，今后还必须不断分析总结相关的各项标准及技术。

参考文献

- [1] PVM version 3.URL: <http://www.netlib.org/pvm3> (检索日期: 2002-10-24)
- [2] MPI General Information.URL: <http://www.lam-mpi.org/mpi> (检索日期: 2002-10-25)
- [3] Java Document.URL: <http://java.sun.COM> (检索日期: 2002-10-28)
- [4] RFC 2460.URL: <http://www.ietf.org> (检索日期: 2002-10-28)
- [5] Labrou,Y.,Finin,T.A Proposal for a New KQML Specification. In: TR CS-97-03, UMBC. URL: <http://www.cs.umbc.edu> (检索日期: 2002-10-28)
- [6] FIPA.FIPA Application Types.URL: <http://www.csel.stet.it/fipa> (检索日期: 2002-10-28)
- [7] William Y.Arms.Digital Libraries.施伯乐等译.电子工业出版社, 2001.
- [8] About The Object Management Group.URL: <http://www.omg.org> (检索日期: 2002-10-28)
- [9] COM.URL: <http://www.microsoft.COM/COM/default.asp> (检索日期: 2002-10-28)
- [10] 林宁. 数字图书馆相关标准探讨.中国数字图书馆工程标准规范论坛会议文件,2002
- [11] The MPEG Homepage.URL: <http://mpeg.teleCOMitalialab.COM/> (检索日期: 2002-10-28)

- [12] About JPEG.URL: http://www.jpeg.org/jpeg_about.html (检索日期: 2002-10-28)
- [13] About JBIG.URL: <http://www.jpeg.org/jbighomepage.html> (检索日期: 2002-10-28)
- [14] What is MHEG.URL: <http://www.mheg.org/> (检索日期: 2002-10-28)
- [15] 绍佩英.分布式数据库系统及其应用.科学出版社,2000
- [16] 网络存储系统.URL: <http://www.ict.ac.cn/kexue/xm7.htm> (检索日期: 2002-10-28)
- [17] 史忠植.智能主体及其应用.科学出版社,2001
- [18] 并行处理的理论、算法与结构.URL: <http://www.ict.ac.cn/kexue/xm11.htm> (检索日期: 2002-10-28)
- [19] 试验床概况. URL: <http://www.ipv6.net.edu.cn/> (检索日期: 2002-10-28)
- [20] 蔡忠善. 移动计算和移动数据库在我国的应用前景. URL: <http://www.swm.COM.cn> (检索日期: 2002-10-28)
- [21] RFC 1932.URL: <http://www.ietf.org> (检索日期: 2002-10-28)
- [22] 雷震洲,杨然.新一代的因特网——光因特网.URL: <http://economy.zz.ha.cn> (检索日期: 2002-10-28)
- [23] 吴亦川等.IP over WDM技术概述.URL: <http://61.185.224.156> (检索日期: 2002-10-28)
- [24] IEEE P802.11.URL: <http://grouper.ieee.org/groups/802/11> (检索日期: 2002-10-28)
- [25] 陈新,张成昱.信息检索协议标准 Z39.50 和 ISO 10162、ISO 10163 简介.中关村文献信息网络建设论文专集,《现代图书情报技术》增刊, 1997
- [26] 陈海燕等.Z39.50 信息检索标准机器应用.图书馆建设, 2002 (2), 66-68
- [27] 蔡吉人.信息安全与密码学.URL: <http://www.netfront.COM.cn/attention/expert> (检索日期: 2002-10-28)
- [28] DES加密算法介绍.URL: <http://www.cnpgp.org/step/des.html> (检索日期: 2002-10-28)
- [29] 加密: RSA加密算法介绍.URL: <http://www.cx66.COM/cxgz/program/delphi/920.htm> (检索日期: 2002-10-28)
- [30] Internet上的安全认证——加密与数字签名.URL: <http://it.rising.COM.cn/safety> (检索日期: 2002-10-28)
- [31] 刘强. 也谈面向对象.URL: <http://www.zdnet.COM.cn> (检索日期: 2002-10-28)
- [32] IEEE/EIA 12207.URL: <http://www.software.org/quagmire/descriptions/ieee-eia12207.asp> (检索日期: 2002-10-28)
- [33] 李华北.软件能力成熟度模型及其应用.URL: <http://www.ceprei.org/chinese/artice> (检索日期: 2002-10-28)
- [34] WebSphere:Getting Started with WebSphere Application Server.
URL:
http://www-3.ibm.COM/software/webservers/appserv/doc/v40/aes/infocenter-zh/infocenter/wass_content/pdf/atw00.pdf (检索日期: 2002-12-19)
- [35] Product Brief.URL:http://www.bea.COM/products/weblogic/platform/wl_platform7_pb.pdf (检索日期: 2002-12-19)
- [36] 史忠植.知识发现.清华大学出版社,2002
- [37] 萨师宣,王珊.数据库系统概论.高等教育出版社,1995
- [38] 钟玉琢等.多媒体计算机技术基础及应用.高等教育出版社,2001
- [39] W.H.Inmon.Building the Data Warehouse.New York:John Wiley&Sons,1996
- [40] Jiawei Han,Micheline Kamber.Data Minings:Concepts and Techiques.机械工业出版社,2002
- [41] 毛军,张晓林等.数字图书馆数字对象唯一标识符问题.国家科学数字图书馆资助项目, 2002,7

- [42] 张晓林,曾蕾等.数字图书馆的开放元数据机制.国家科学数字图书馆资助项目,2002,5
- [43] 张晓林,梁娜等.数字图书馆元数据登记系统.国家科学数字图书馆资助项目,2002,6
- [44] 张晓林,曾蕾等.数字图书馆管理元数据的发展与应用.国家科学数字图书馆资助项目,2002,9
- [45] 张晓林,曾蕾等.数字图书馆建设的标准与规范体系.国家科学数字图书馆资助项目,2002,4
- [46] 刘炜,张晓林等.信息资源集合开放性描述研究.国家科学数字图书馆资助项目,2002,7
- [47] 赵继海.DRM 技术的发展及其对数字图书馆的影响.大学图书馆学报,2002,1 (14-17)
- [48] EBX Working Group.The electronic book exchange system (EBX) Version 0.8. 2000,7.<http://www.ebxwg.org> (检索日期: 2003-1-17)
- [49] 谭建龙,庄超等.一种实用 Internet 内容版权保护系统的设计与实现.计算机研究与发展, Vol.38,No.10, 2001,10
- [50] Digital Watermarking World. <http://www.watermarkingworld.org/> (检索日期: 2003-1-17)
- [51] Hal Berghel, Lawrence O'Gorman. Digital Watermarking. http://www.acm.org/~hlb/publications/dig_wtr/dig_watr.html (检索日期: 2003-1-17)
- [52] Apache Http Server Version 2.<http://www.apache.org> (检索日期: 2003-1-17)
- [53] Apache Tomcat.<http://jakarta.apache.org/tomcat/index.html> (检索日期: 2003-1-17)
- [54] Websphere MQ family. <http://www-3.ibm.COM/software/ts/mqseries/> (检索日期: 2003-1-17)
- [55] IIS Document.<http://www.microsoft.COM/windows2000/en/server/iis/> (检索日期: 2003-1-17)
- [56] Raisin Core.<http://www.caucho.COM/resin/index.xtp> (检索日期: 2003-1-17)
- [57] The VLDB.<http://www.vldb.org/dblp/db/journals/vldb/vldb10.html> (检索日期: 2003-1-17)
- [58] XML Encyption Requirement.<http://www.w3.org/TR/xml-encryption-req> (检索日期: 2003-3-17)
- [59] D. Eastlake et.RFC3275.<http://www.ietf.org/rfc/rfc3275.txt> (检索日期: 2003-3-17)
- [60] 张云帆.Linux OpenSSL服务器.<http://www.swm.COM.cn/lwm/2001-3/6.html> (检索日期: 2003-3-17)
- [61] Scott Cantor, Marlana Erdos.NSF Middleware Initiative Draft: Shibboleth-Architecture Draft v50.<http://shibboleth.internet2.edu/draft-internet2-shibboleth-arch-v05.html> (检索日期: 2003-3-17)
- [62] Microsoft and SAML.<http://www.pconline.COM.cn/news/cn/10207/77757.html> (检索日期: 2003-3-17)
- [63] Reg Quinton.Certificates, HTTP, X509.<http://ist.uwaterloo.ca/security/ssl-pki/index.html> (检索日期: 2003-3-17)
- [64] S. Santesson, et al. Internet X.509 Public Key Infrastructure Qualified Certificates Profile. http://www.itu.int/newsarchive/press_releases/2000/05.html (检索日期: 2003-3-17)
- [65] Kerberos WG.<http://www.ietf.org/html.charters/krb-wg-charter.html> (检索日期: 2003-3-17)
- [66] IBM网格计算概述.<http://www-900.ibm.COM/cn/grid/index.shtml> (检索日期: 2003-7-27)
- [67] 高端计算(网格)环境.<http://www.ict.ac.cn/kexue/xm2.htm> (检索日期: 2003-7-27)
- [68] 网格计算理论及其应用.<http://www.chinagrid.COM/gridstudy/gridstudyD.htm> (检索日期: 2003-7-27)

附录

英文缩略表（按字母排序）

- ACL Agent COMmunication Language, 一种主体通信语言
- AS Authentication Server, 认证服务器
- CA Certifactae Authority, 认证中心
- COM COMpONENT Object Model, 组件对象模型
- CORBA COMmon Object Request Broker Architecture, 公共对象请求代理体系结构
- DCOM Distributed COMpONENT Object Model, 分布式组件对象模型
- DDBS Distributed Database System, 分布式数据库系统
- DES Data Encyption Standard, 数据加密标准
- DRM Data Right Management, 数据版权管理
- EJB Enterprise JavaBean, 企业 JavaBean
- FIPA Foundation for Intelligent Physical Agents, 智能物理主体联盟
- GGF Global Grid Forum, 全球网格论坛
- IETF Internet Engineering Task Force, Internet 工程任务组
- ITU International TeleCOMmunications Union, 国际电信同盟,
- JBIG Joint Bi-level Image Experts Group, 黑白二值文件的无损压缩标准
- JPEG Joint Photographic Experts Group, 彩色和单色灰度级的连续色调静止图像的压缩标准
- KQML Knowledge Query and Manipulation Language, 一种主体通信语言
- MHEG Multimedia and Hypermedia information encoding Experts Group, 多媒体与超媒体信息编码专家小组, 该标准是用于实时交换的多媒体和超媒体信息对象的表示法
- MPEG Moving Picture Experts Group, 活动视频的压缩标准
- MPI Message Passing Interface, 消息传递接口
- NAS Network Access Server, 网络存储服务器
- OMG Object Management Group, 对象管理组
- PKI Public Key Infrastructure, 公钥系统
- POA IP over ATM, ATM, 网络上实现的 IP
- POS IP over SONET/SDH, 同步光纤网络上实现的 IP
- PVM Parallel Virtual Machine, 并行虚拟机
- SAN Storage Area Network, 存储区域网
- SET Secure Electronic Transaction, 安全电子交易
- SGML Standard Global Makeup Language, 标准通用标记语言
- SSL Secure Socktes Layer, 安全套接字层
- TLS Transport Layer Security, 传输层安全性
- WDM Wavelength Division Multiplex, 波分复用
- WTLS Wireless Transport Layer Security, 无线传输层安全性
- XML eXtensible Markup language, 可扩展标记语言