



数据集成与应用集成

宋晓宇 王永会等 编著



中国水利水电出版社
www.waterpub.com.cn



数据集成与应用集成

宋晓宇 王永会等 编著



中国水利水电出版社
www.waterpub.com.cn

内 容 提 要

本书主要内容分为数据集成和应用集成两部分。数据集成主要包括数据集成的方法与规范、数据仓库、异构数据集成、数据集成工具、元数据管理以及数据集成典型实例；应用集成部分主要包括应用集成的概念与标准、应用集成的主流技术（Microsoft 应用集成技术、OMG 应用集成技术、Java 应用集成技术、基于 Agent 的集成技术），并通过实例介绍了每一种应用集成技术如何在实际中使用以及不同集成技术间的互操作方法。

本书可作为信息系统开发、管理人员和计算机软件开发人员的培训用书和参考书，也可作为高等院校信息系统工程、计算机信息管理、管理工程、计算机科学与技术等专业本科生或研究生教材。

图书在版编目 (CIP) 数据

数据集成与应用集成/宋晓宇等编著. —北京: 中国水利水电出版社, 2008

ISBN 978 - 7 - 5084 - 5644 - 7

I. 数… II. 宋… III. 数据处理 IV. TP274

中国版本图书馆 CIP 数据核字 (2008) 第 085493 号

书 名	数据集成与应用集成
作 者	宋晓宇 王永会等 编著
出版发行	中国水利水电出版社 (北京市三里河路 6 号 100044) 网址: www. waterpub. com. cn E - mail: sales@waterpub. com. cn
经 售	电话: (010) 63202266 (总机)、68367658 (营销中心) 北京科水图书销售中心 (零售) 电话: (010) 88383994、63202643 全国各地新华书店和相关出版物销售网点
排 版	中国水利水电出版社微机排版中心
印 刷	北京纪元彩艺印刷有限公司
规 格	184mm×260mm 16 开本 17.75 印张 421 千字
版 次	2008 年 7 月第 1 版 2008 年 7 月第 1 次印刷
印 数	0001—4000 册
定 价	36.00 元

凡购买我社图书, 如有缺页、倒页、脱页的, 本社营销中心负责调换

版权所有·侵权必究

前 言

当代计算机信息系统是计算机硬件平台、网络、操作系统、数据库系统、应用组件等部分的集合体。各部分的协调配合是信息系统成功的前提，为适应需要产生了信息系统的集成技术。因此，信息系统集成技术是信息系统工程的重要技术基础。本书从综合型信息系统设计的角度出发，重点讲述了信息系统的体系结构，信息系统集成的概念、原则与目标以及信息系统集成的基础技术。从数据集成和应用集成两个层次深入描述了信息系统集成技术。本书具有以下特点：

1. 重视技术的可操作性。在介绍每一部分技术之后，均有相应用的实例，每个实例尽可能详细。读者可以参照书中的实例，结合具体应用背景来实现所需要的系统。

2. 突出系统集成的重点。集成技术中数据集成是基础，应用集成是核心。本书以数据集成与应用集为中心展开，同时涵盖了一些相关技术。读者只需详细了解这两种技术，即可成功构建相应的系统。

3. 综合介绍了多种应用技术及其互操作。不同的应用技术有不同的特点。

本书分为两大部分共9章。第1~3章为数据集成的内容。第1章介绍信息系统的基本概念、信息系统集成概述及存在的问题。第2章介绍数据集成的相关概念。第3章介绍一种数据集成典型技术，即数据仓库。第4~9章为应用集成的内容。第4章对应用集成进行了一个概述。第5~7章分别介绍 Microsoft 的应用集成技术、OMG 的应用集成技术和 Java 平台的应用集成技术。第8章讲述了各应用技术的互操作。第9章介绍了基于 Agent 系统集成。

参加本书编著工作的有：宋晓宇、王永会、孙焕良、栾方军、孙丽梅、许景科、刘继飞老师及朱叶丽、刘峰、田春曼、周磊、李开等研究生。

限于作者的水平，错误和疏漏之处在所难免，敬请读者批评指正。

作者

2008年4月

目 录

前言

第 1 章 概述	1
1.1 信息系统的基本概念	1
1.1.1 信息	1
1.1.2 系统	1
1.1.3 信息系统	2
1.2 信息系统的体系结构	3
1.2.1 集中式计算模式	3
1.2.2 客户机/服务器计算模式	5
1.2.3 浏览器/服务器计算模式	7
1.3 信息系统的层次结构	8
1.3.1 网络与通信硬件	8
1.3.2 网络操作系统	10
1.3.3 信息系统开发工具	11
1.3.4 信息系统数据模型	14
1.3.5 信息系统功能模型	14
1.3.6 信息系统业务模型	15
1.3.7 用户界面	15
1.4 信息系统集成概述	16
1.4.1 信息系统集成的概念	16
1.4.2 网络集成	16
1.4.3 数据集成	17
1.4.4 应用集成	18
1.4.5 信息系统集成项目管理	20
1.5 信息系统集成存在的问题	25
1.5.1 体系结构混乱	25
1.5.2 技能欠缺	25
1.5.3 安全问题	26
第 2 章 数据集成	28
2.1 数据集成概述	28

2.2	常见数据集成方法	29
2.2.1	联邦数据库	29
2.2.2	中间件集成方法	30
2.2.3	数据仓库方法	31
2.3	常见数据集成技术与标准	31
2.3.1	ODBC (Open DataBase Connectivity)	31
2.3.2	JDBC (Java DataBase Connectivity)	33
2.3.3	OLE DB	35
2.3.4	ADO	36
2.3.5	XML	37
2.4	元数据管理	39
2.4.1	元数据的概念	39
2.4.2	元数据管理	40
2.4.3	元数据建模标准	40
2.5	数据集成实现实例——智能小区管理系统	42
2.5.1	智能化小区管理系统中异构数据库集成总体框架设计	42
2.5.2	中间件层的设计	43
第3章	数据仓库	46
3.1	数据仓库的基本概念	46
3.1.1	操作型数据库与数据仓库的区别	47
3.1.2	数据仓库的特点	48
3.2	数据仓库的系统结构	49
3.3	数据仓库中数据集成模型	50
3.3.1	多维模型的分类	51
3.3.2	多维模型上的 OLAP 操作	53
3.3.3	度量的计算	54
3.4	数据仓库保证数据质量方法	55
3.4.1	数据清洗定义	55
3.4.2	主要数据质量问题及解决方法	56
3.4.3	数据清洗框架	56
3.5	常见数据仓库解决方案	60
3.5.1	Oracle 数据仓库解决方案	60
3.5.2	Sybase 数据仓库解决方案	63
3.5.3	SQL Server 数据仓库解决方案	64
3.5.4	IBM DB2 数据仓库解决方案	68
3.5.5	SAS 数据仓库解决方案	71
3.6	数据仓库实现实例	73

3.6.1	需求建模	74
3.6.2	数据仓库设计及实现	75
3.6.3	分析应用	77
3.6.4	维护完善	79
第 4 章	应用集成	80
4.1	应用集成概述	80
4.1.1	应用集成的概念	80
4.1.2	应用集成与标准化	81
4.2	企业应用集成 (EAI)	82
4.2.1	EAI 概述	82
4.2.2	EAI 的类型	83
4.2.3	EAI 所需的支持和技术	84
4.3	EAI 的基本组成模块	84
4.3.1	通信模式	84
4.3.2	集成方法	86
4.3.3	中间件的选择	88
4.3.4	服务模块	91
4.4	基于中间件的集成方法	92
4.4.1	EAI 与中间件	92
4.4.2	面向消息的中间件	92
4.4.3	分布式对象中间件	94
4.4.4	事务处理监控器	103
第 5 章	Microsoft 应用集成技术	112
5.1	COM 技术	112
5.1.1	组件对象模型 (COM)	112
5.1.2	COM 应用模型	113
5.1.3	分布式 COM (DCOM)	117
5.2	COM+ 技术	118
5.2.1	从 COM 到 COM+	119
5.2.2	COM+ 结构	119
5.2.3	COM+ 系统服务	125
5.3	Microsoft .NET	132
5.3.1	.NET 概述	132
5.3.2	.NET 框架 (.NET Framework)	134
5.3.3	基于 .NET 应用的开发	137
5.3.4	Microsoft 新技术	140
第 6 章	OMG 应用集成技术	143

6.1	CORBA 概述	143
6.1.1	ORB	143
6.1.2	IDL 语言和语言映射	144
6.1.3	存根和框架	145
6.1.4	动态调用	145
6.1.5	对象适配器	145
6.1.6	界面仓库和实现仓库	146
6.1.7	ORB 之间的互操作	147
6.2	CORBA 系统开发	147
6.2.1	开发过程	147
6.2.2	开发实例	148
6.3	CORBA 的服务	153
6.3.1	核心服务	154
6.3.2	数据库与事务处理	156
6.3.3	其他服务	158
6.3.4	CORBA 服务应用实例	159
6.4	CORBA 的 OTS	160
6.4.1	OTS 组成	161
6.4.2	OTS 中的接口	161
6.4.3	对象事务服务流程	166
6.4.4	CORBA OTS 服务与 XA 的集成	167
6.5	基于 CORBA 的应用集成实例	170
6.5.1	实例系统概述	170
6.5.2	实例系统实现	171
第 7 章	Java 应用集成技术	181
7.1	基于 Java 平台集成的优势	181
7.2	J2EE 应用体系结构	183
7.2.1	J2EE 简介	184
7.2.2	J2EE 的优势	184
7.2.3	J2EE 的结构	185
7.2.4	J2EE 的核心 API 与组件	186
7.2.5	J2EE 对应用集成的支持	189
7.3	JMS 技术	189
7.3.1	JMS 简介	189
7.3.2	JMS 接口	190
7.3.3	基于 JMS 的集成	191
7.4	EJB 技术	192

7.4.1	EJB 简介	192
7.4.2	EJB 容器	192
7.4.3	Enterprise Bean	193
7.4.4	会话 Bean	194
7.4.5	实体 Bean	195
7.4.6	消息驱动 Bean	196
7.4.7	远程和本地接口	196
7.4.8	会话 Bean 组件模型	197
7.4.9	实体 Bean 组件模型	199
7.4.10	基于 EJB 的集成	202
7.5	JCA 连接器架构技术	206
7.5.1	JCA 简介	206
7.5.2	JCA 体系协定	207
7.5.3	JCA 运行环境	209
7.5.4	JCA 连接管理	209
7.5.5	JCA 安全管理	209
7.5.6	JCA 事务管理	210
7.5.7	基于 JCA 的集成	210
7.6	集成应用开发实例	211
7.6.1	系统分析	211
7.6.2	系统设计	213
7.6.3	关键问题	216
7.6.4	技术优势	221
第 8 章	互操作技术	223
8.1	Java 和 COM 互操作技术	223
8.1.1	MS Java/COM 解决方案	223
8.1.2	Bridge2Java 解决方案	225
8.1.3	J-Integra 解决方案	226
8.2	CORBA 和 COM (DCOM) 互操作技术	227
8.2.1	CORBA 和 COM (DCOM) 的互操作模型	227
8.2.2	CORBA 和 COM 互操作的主要内容和实现策略	228
8.3	EJB 和 CORBA 互操作技术	229
8.3.1	EJB 和 CORBA 互操作的基础	230
8.3.2	基于 RMI-IIOP 的解决方案	231
8.3.3	基于 Web Service 的解决方案	233
8.3.4	基于 CCM 的解决方案	236
8.3.5	方案比较	236

8.4 互操作实例：EJB 与 CORBA 互操作的实现	237
8.4.1 EJB 会话 Bean 的实现	237
8.4.2 开发 CORBA 客户端	237
第 9 章 基于 Agent 系统集成	242
9.1 Agent 概述	242
9.1.1 Agent 的定义	242
9.1.2 Agent 的特点及基本特性	243
9.1.3 Agent 的分类	244
9.1.4 Agent 与相关领域的关系	245
9.1.5 Agent 系统的开放标准	245
9.2 基于 JADE 的 Agent 系统集成实例	248
9.2.1 JADE 平台特性及操作	249
9.2.2 JADE 的系统扩展关键技术	251
9.2.3 网络信息搜索与集成系统的分析与设计	255
9.2.4 Agent 系统集成的实现	260
9.3 多 Agent 系统的理论、技术及应用	268
9.3.1 多 Agent 系统的理论和结构体系	268
9.3.2 多 Agent 系统的通信	269
9.3.3 多 Agent 系统的学习	270
9.3.4 多 Agent 系统的协调与协作	271
9.3.5 多 Agent 系统的集成应用	272

第 1 章 概 述

1.1 信息系统的基本概念

1.1.1 信息

信息 (Information) 是近代科学的一个专门术语, 已广泛地应用于社会各个领域。关于信息有各种不同的定义, 狭义信息论是将信息定义为人们获得信息前后对事物认识的差别; 广义信息论认为, 信息是指主体 (人、生物和机器) 与外部客体 (环境、其他人、生物和机器) 之间相互联系的一种形式, 是主体和客体之间一切有用的消息和知识, 是表征事物特征的一种普遍形式。

这里我们可以理解为: 信息是向人或机器提供关于现实世界各种事实的知识, 是数据、消息中所包含的意义, 它不随载体物理形式的各种改变而改变。

信息具有如下的特点:

- 1) 信息的客观性。任何信息都是与客观事物紧密相联的。
- 2) 信息的适用性。信息对决策是十分重要的, 建立信息系统的目的就是为生产、管理和决策服务的, 因而信息具有适用性。
- 3) 信息的传输性。信息可以在信息发送者和接受者之间传输, 既包括系统把有用信息送至终端设备 (包括远程终端), 和以一定形式提供给有关用户, 也包括信息在系统内各子系统之间的传输和交换。
- 4) 信息的共享性。信息与实物不同, 它可以传输给多个用户, 为多个用户共享, 而其本身并无损失。

信息的这些特点, 使信息成为当代社会发展的一项重要资源。

数据 (Data) 是通过数字化或直接记录下来的可以被鉴别的符号, 不仅数字是数据, 而且文字、符号和图像也是数据。数据是可以载荷信息的物理符号, 它的格式往往和具体的计算机系统有关, 随载荷它的物理设备的形式而改变。数据只有对实体行为产生影响时才成为信息。例如同样的数据 1 和 0, 当用来表示某一种实体在某个地域内存在与否时, 它就提供了有 (1 表示)、无 (0 表示) 的信息。对同一数据每个人的解释可能不同, 因而获得信息量的多少与人的知识水平和经验有关。

信息与数据虽然有词义上的差别, 但信息与数据是不可分离的, 即信息是数据的内涵, 而数据是信息的表达。也就是说数据是信息的载体, 只有理解了数据的含义, 对数据作解释才能得到数据中所包含的信息。信息系统的建立和进行, 就是信息 (或数据) 按一定方式流动的过程, 在通常情况下, 并不严格区分地使用“信息”和“数据”两个术语。

1.1.2 系统

对于系统一词目前学术界尚无统一的定义。我们可以从以下描述开始讨论: 系统是若



干个相互作用、相互依存的部分所组成的整体。

系统与其说是一种具体对象，还不如说是一种研究事物的方法。任何事物都可以分为若干部分，所以可以说处处是系统，事事是系统。事实上，只有当我们把注意力集中于它们各部分及其相互关系时，才明确地指出系统性。我们常讲：把企业当作一个系统看待，把组织当作一个系统看待，就表明了这一点。

一个系统作为抽象模型来看，有其共同的基本组成部分。

1) 输入：由系统来处理的东西。

2) 处理过程（简称处理）：对输入按照一定的方式进行处理而产生输出，在社会经济系统中，更具体地是指做出决定与采取行动这两类活动。

3) 输出：由系统操作而得到的结果。

分析研究一个系统，除了要了解各个组成部分在系统内部的相互关系和相互作用外，还要看这些组成部分对于外界影响是如何反应的。用控制论的术语来讲：外界影响称为环境，环境对系统的影响称为刺激或冲击，而系统对环境的影响称为反响，刺激或冲击相当于系统的输入，而反响则相当于系统的输出。一个系统对于任何一个刺激是如何反响的，在很大程度上决定于各个组成部分在系统内部是如何组成的。系统与环境之间的相互作用表现为外界（系统的用户和设计者）提出的限制以及系统的反馈控制。

限制包括两种：一种是系统的目的，即系统应得出的结果，也就是系统的输出；另一种是约束，即对系统的目的（输出）所加的限定条件。例如，某一具体的计算机系统的目的是提供（输出）某种具体的报表，对这种报表所加的约束，即为对表的格式的规定，如共有多少栏，每栏的标题，共有多少行等；信息范围的规定，如起讫日期，表的内容等；使用规定，如怎样使用该表和由谁使用，使用该表得出的结论和由谁来采用这些结论等。

反馈是对系统的一种控制，它把输出与预定的标准相比较，看这个输出是否符合标准；若有任何差异，即采取纠正措施来进行控制。一个库存控制系统就可以是一个带反馈的系统：预先规定最低库存量，当库存达到或低于它时，就立即发出通知，要求再订货。

系统有简单和复杂之分。这种区别主要是取决于系统内各个组成部分之间关系的多少以及在一定程度上决定于系统所包含的组成部分的个数。例如，一台电动洗衣机就是一个简单系统，因为其中部件不多，且部件之间的关系也不复杂。一个企业就是一个复杂系统，而一个大企业则是一个很复杂的系统。

分析研究一个大型复杂的系统时，要详尽而具体地了解系统的运行状态，就有必要把这种系统分解为若干部分，即子系统，而且可逐层分解下去；并且还能反过来，逐级集成，最后回到原来系统的整体。

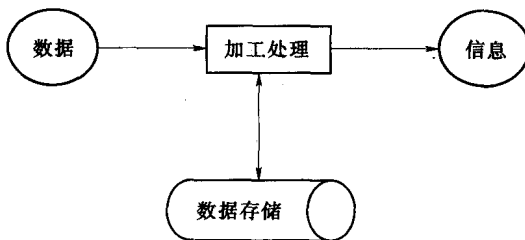


图 1-1 信息系统的层次模型

1.1.3 信息系统

信息系统，简单地说就是输入数据，经过加工处理后输出各种信息的系统。它的主要功能是对信息进行采集、处理、存储、管理、检索和传输。信息系统的基本



模式如图1-1所示。它是一种由人、计算机（包括网络）和管理规则组成的集成化系统，是复杂系统中的小系统，作用是沟通各个子系统使整体协调一致。社会组织的功能和类型各不相同，但它们都有自己一定形式的信息系统，而且信息系统功能的好坏直接影响全局组织的效益，因此可以说信息系统是整个系统的中枢。信息系统的功能包括以下几个方面：

(1) 信息的采集

这是信息系统最基本的功能，它将分散在各部门、各地方的相关信息收集起来，记录其数据，然后整理成为信息系统所需要的格式。

(2) 信息的加工

信息系统对信息进行加工处理的具体含义是：排序、分类、归纳、检索、统计、模拟、预测及各种数学运算。所有这些工作，都依靠规模大小不同的计算机来完成。

(3) 信息的存储

信息系统经过上述两个步骤后，形成了对管理决策有价值的信息，由信息系统完成对这些有用信息的存储保管。若信息过于庞大，就必须有良好的组织结构和先进的存储技术。通常，信息的存储分为逻辑组织和物理存储两个方面，前者依赖合理的数据结构，后者把信息存储在合适的介质上。

(4) 信息的检索

检索查询是信息系统重要的功能。存储在系统内的信息要便于查询，满足各方面用户的需求。通常信息检索需要用到数据库的技术与方法。

(5) 信息的传输

信息系统中的信息必须具有相互交换的功能，各部门之间要完成通信工作，一定要解决好信息的传输问题，以便使信息准确迅速地送往使用者手中。

1.2 信息系统的体系结构

1.2.1 集中式计算模式

(1) 概念

在计算机诞生和应用的初期，计算所需要的数据和程序都是集中在一台计算机上进行的，称为集中式计算。

(2) 优点和缺点

优点：随着网络的发展，这种集中式计算往往形成一种由大型机和多个与之相连的终端组成的网络结构。当支持大量用户时，大型机自顶向下的维护和管理方式显示出集中式处理的优越性。它具有安全性好、可靠性高、计算能力和数据存储能力强以及系统维护和管理费用较低等优点。

缺点：该模式存在着一些明显的缺点，如大型机的初始投资较大、可移植性差、资源利用率低以及网络负载大等缺点。

(3) 发展

1) 主机/终端系统。主机/终端系统又简称主机系统，是20世纪60年代后期形成的



以一台计算机（大、中、小型机）为中心的多用户系统。在该系统中，用户通过终端与主机相连。在主机操作系统的管理下共享主机的硬件资源，包括中央处理器、内外存、输入/输出设备等。其特点是可同时为多个用户服务，如图 1-2 所示。

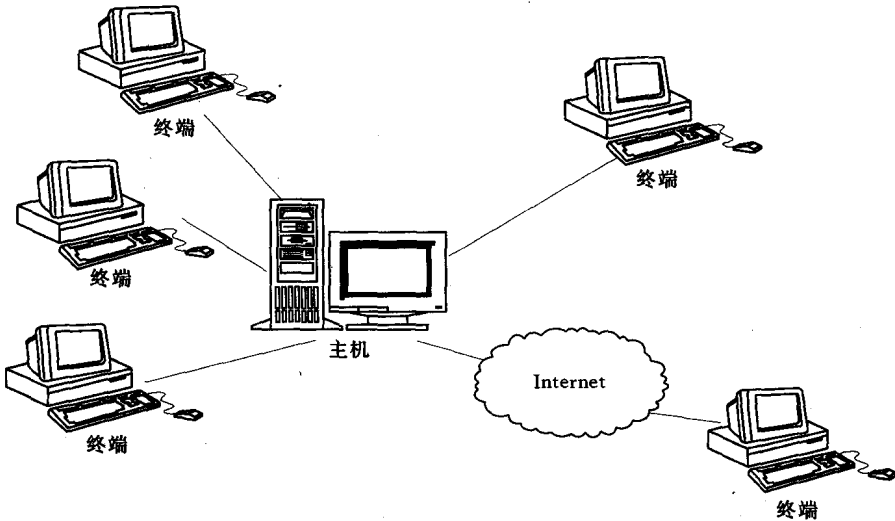


图 1-2 主机/终端系统

在主机/终端系统中，所有数据和程序都在主机上，进行集中管理，各终端只相当于一个显示器加键盘的功能。这种系统便于集中处理大量的信息，如大型科学计算、人口普查等。

该系统有以下不足：

- 主机负担过重。所有的计算、存储都集中在主机上，一旦主机出故障，系统将全面瘫痪。
- 扩充不易。当用户量不断增加时，必须更换主机，否则服务质量就要受到影响。
- 系统的购置、安装、维护费用较高，不易普及。

2) 工作站/文件服务器系统。20 世纪 80 年代初，随着局域网的兴起，联网的微机被分为两类：一类称为服务器，专门为网络上的其他用户提供共享文件（或数据），因此也称为文件服务器。它是网络的核心，管理网络通信，网络操作系统也安装在文件服务器中；另一类称为工作站，它可访问文件服务器中的数据和文件，而本工作站的资源不被其他工作站或服务器共享。每个局域网中可有一台或多台文件服务器。工作站用户通过磁盘映象，使用文件服务器上的硬盘。网络上传递的只是文件，计算任务在工作站上完成，如图 1-3 所示。

在工作站/文件服务器中，数据库和全部应用程序都存储在文件服务器上，但应用程序的执行却在微机工作站上进行，从而增加了网络传输负荷。同时当服务器负荷超过某个限度后，其效率会明显下降。这种系统由于数据的安全性较差，无法发挥服务器的信息处理能力，现在已很少采用。

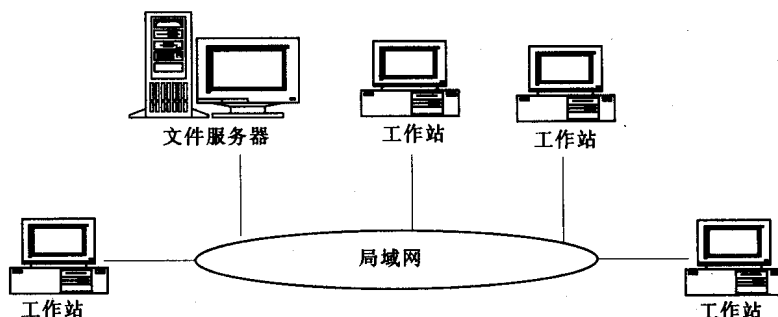


图 1-3 工作站/文件服务器系统

1.2.2 客户机/服务器计算模式

(1) 概念

客户机/服务器计算结构（简称 C/S 结构）的定义为：客户机是一种单用户工作站，它提供与业务应用有关的表现、计算、联网、访问数据库和各类接口服务；服务器是一种存储器共享型的多用户处理机，它提供业务所需的计算、联网、数据库管理和各类接口服务。客户机/服务器计算结构是由客户机、服务器构成的一种网络计算环境，它把应用程序所要完成的任务分派到客户机和服务器上共同完成。

客户机/服务器计算模式将应用一分为二：前端是客户机，一般使用微型计算机，几乎所有的应用逻辑都在客户端进行和表达，客户机完成与用户的交互任务，具有强壮的数据操纵和事务处理能力；后端是服务器，可以使用各种类型的主机，服务器负责数据管理，提供数据库的查询和管理、大规模的计算等服务。

客户机/服务器计算结构就是结合主机系统擅长数据处理和工作站/文件服务器系统便于使用、成本低廉这两种结构的优点，在局域网中，增加服务器的处理能力，即网络上的服务器不仅提供文件服务功能而且也提供数据处理功能。同时工作站不是以仿真终端的方式与服务器相连，而是在整个系统运行中也保持完善的信息处理功能。从而就可以把一个应用问题的解决分为两部分：一部分与用户业务有关的处理交给工作站完成，如建立用户接口菜单、数据输入的前期检查工作等，在这里，把工作站称为客户机（Client）；另一部分成批的、集中的大量信息处理任务交给服务器完成，如大型数据库查询、数据分析等，又把这种服务器称为数据库服务器（Data Server）。数据库服务器对局域网上的所有客户机提供数据处理服务。

(2) 优点及缺点

优点：通过异种平台集成，能够协调现有的各种 IT 基础结构；分布式管理；能充分发挥客户端 PC 的处理能力，安全、稳定、速度快，且可脱机操作。

缺点：它必须在客户端安装大量的应用程序（客户端软件）、开发成本较高，移植困难，用户界面风格不统一、使用繁杂、不利于推广使用，维护复杂、升级麻烦，信息内容和形式单一，新技术不能轻易应用等。

(3) 比较

C/S 结构与工作站/文件服务器系统在硬件组成、网络拓扑结构、通信连接等方面基



本相同。即与图 1-3 相似，但文件服务器变为了数据库服务器，工作站改称为客户机。它们的最大区别在于在 C/S 结构中，服务器控制管理数据的能力由文件管理方式上升为数据库管理方式，原先在工作站/文件服务器中由工作站所承担的数据加工任务（即应用的一部分）现改由服务器来承担，从而使整个系统的性能有了质的飞跃：一方面由高档微机组成的数据库服务器的效率得到了充分的发挥；另一方面数据经服务器加工后在局域网中传输的仅仅是客户机所需的那一小部分而不是整个文件，从而大大降低了网络流量。

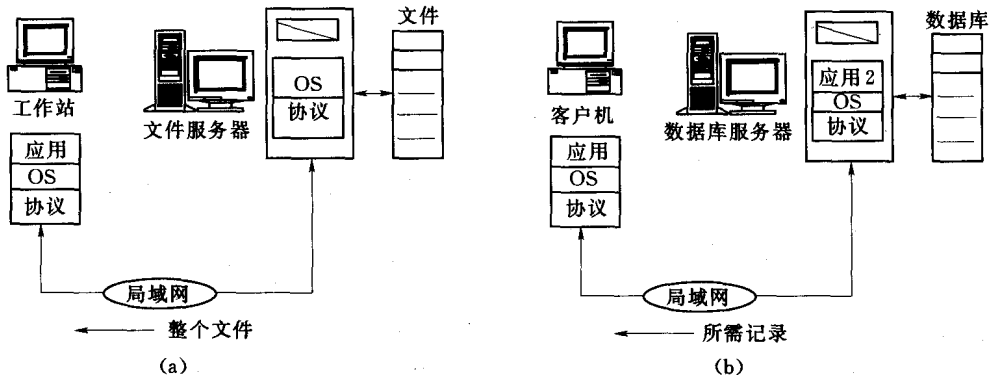


图 1-4 两种系统的工作方式的比较
(a) 工作站/文件服务器结构；(b) C/S 结构

如图 1-4 (a) 所示，表示工作站/文件服务器系统的工作方式。如果某台工作站需要共享某数据库文件中的一个特定记录，它得送一个“读整个文件”的请求给文件服务器，服务器将整个文件通过网络传送给该工作站，然后再运行工作站上的应用程序从整个文件中搜寻出所需的记录。而在图 1-4 (b) 所示的 C/S 结构中，一个应用程序分为客户机部分（应用 1）和数据库服务器部分（应用 2）。当客户机执行应用 1 后，需要共享数据库中的某个记录时，客户机只要送一个“读记录”的请求发送给数据库服务器，在数据库服务器上启动应用 2 对数据库文件进行所需的搜索、索引和排序，只选出所需的记录通过网络传送给客户机，从而节约了资源，减轻了对网络传输的压力。

随着硬件技术的不断发展，客户机同样可提供服务，即在某一时刻，客户机向服务器提出申请，而在另一时间，它们可以颠倒过来。客户机和服务器可以共存于一个多任务系统中，既可以运行某个应用的客户程序，也可以运行其他应用的服务程序。

(4) 分布式计算结构

分布性主要指数据分布和计算分布，数据分布是指数据可分散存储在网络上的不同计算机中；计算分布则是把操作计算分散由不同的机器处理。在物理上，分布式数据库的资源结构与分布式的组织结构（企业单位或社会团体）相吻合；而逻辑上的相关性，又使它们可以利用计算机网络集成为功能强大的系统。分布式计算结构在系统可靠性、可扩展性、可用性和并行处理等方面具有明显的优势。

分布式计算结构就是综合集中计算结构和 C/S 结构的优点，在 C/S 结构的环境中通过软件技术创造出类似于集中式结构的操作环境。分布式计算结构可分为物理层、分布式计算平台和分布式应用层三个方面。在物理层，采用 C/S 结构的服务器仍散置在网络各



处,并保持传统的 C/S 结构的特性。在分布式计算平台,其网络上的所有服务器均被抽象成一个大型的计算单元,使用户不感觉有网络和服务器。在应用层面上,程序开发人员和用户不必考虑其低层有哪些服务器,自己是在一个统一的、一致的操作环境下工作,如图 1-5 所示。

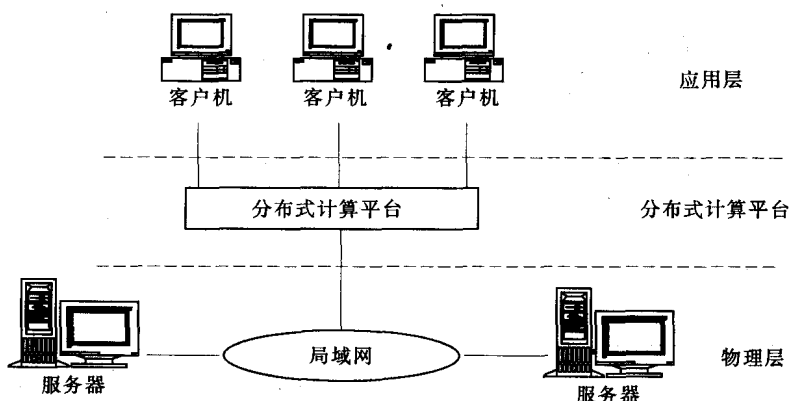


图 1-5 分布式计算结构

换句话说,分布式计算平台隐藏了网络的物理层面,它使传统的网络操作环境内的所有机器间的操作,能如同同一部机器内的不同元件之间的操作一样,创造出所谓的无缝的分布式计算环境,这种网络计算环境的产生为 Web 计算模式的推出创造了良好的条件。

1.2.3 浏览器/服务器计算模式

(1) 概念

浏览器/服务器 (B/S) 计算模式是一种基于 Web 的协同计算,是一种三层架构客户机/服务器的计算模式。其中,第一层为客户端表示层;第二层是应用服务器层;第三层是数据中心层,主要由数据库系统组成。

早期的 Internet 的计算结构是以 Web 技术为核心的浏览器/Web 服务器体系结构,如图 1-6 所示。

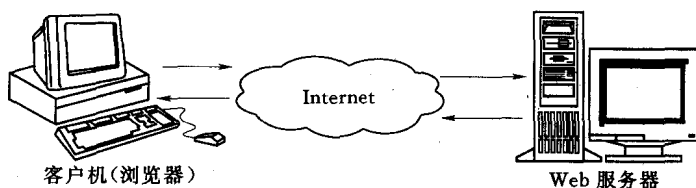


图 1-6 浏览器/Web 服务器结构

Web 服务的最大特色在于为用户提供良好的信息查询界面。Web 服务把各种形式的信息,如文本、图像、声音和视频等无缝隙地集成在一起,用户只需要通过浏览器提出自己的查询要求,Web 服务器自动完成查询结果,并将与查询条件相关的文件取回并显示在屏幕上,用户无需关心这些文件存放在 Internet 上的哪台计算机中。

实现 Web 服务的通信协议是 HTTP (Hypertext Transfer Protocol: 超文本传输协