普通高等教育“十一五”国家级规划教材
计算机科学与技术系列教材 信息技术方向

应用集成原理与技术

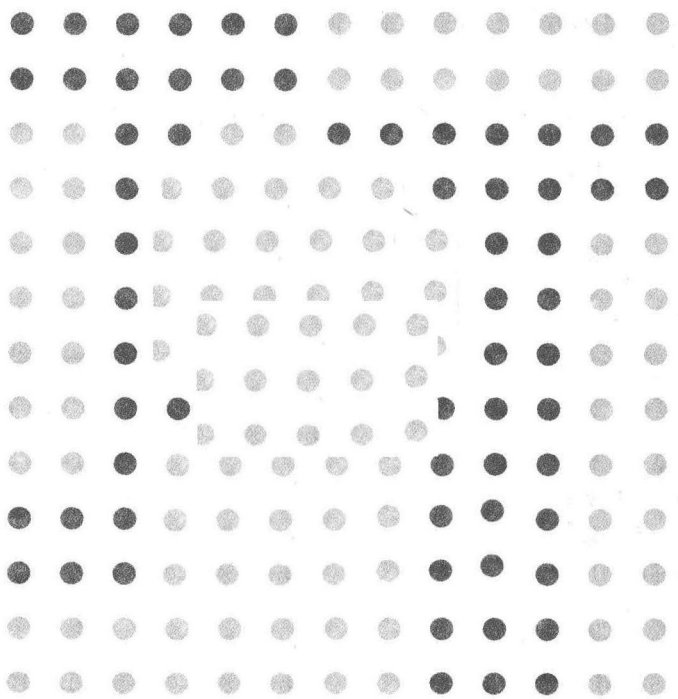
刘峰 郑滔 编著

清华大学出版社
http://www.tup.com.cn

普通高等教育“十一五”国家级规划教材
计算机科学与技术系列教材 信息技术方向

应用集成原理与技术

刘峰 郑滔 编著



清华大学出版社
北京

内 容 简 介

本书介绍了应用集成的原理、主要框架和主流技术，重点包含数据集成和应用集成两大部分内容。其具体内容包括 XML 技术、数据集成技术、不同程序设计语言的集成技术、应用集成中的软件复用技术、分布式对象技术、消息中间件技术以及 Web Service 技术，并通过相应示例讲解各主流技术的使用方法。

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、软件工程、计算机科学与技术等专业本科生教材。

图书在版编目（CIP）数据

应用集成原理与技术 / 刘峰，郑滔编著．—北京：清华大学出版社，2011.6
（计算机科学与技术系列教材·信息技术方向）
ISBN 978-7-302-25897-1

Ⅰ．①应…　Ⅱ．①刘…②郑…　Ⅲ．①信息系统集成－教材　Ⅳ．①G202

中国版本图书馆 CIP 数据核字(2011)第 110289 号

责任编辑：张瑞庆　薛　阳
责任校对：焦丽丽
责任印制：何　芊

出版发行：清华大学出版社　　**地　　址**：北京清华大学学研大厦 A 座
http://www.tup.com.cn　　**邮　　编**：100084
社　　总　　机：010-62770175　　**邮　　购**：010-62786544
投稿与读者服务：010-62795954，jsjjc@tup.tsinghua.edu.cn
质　量　反　馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
印 装 者：北京国马印刷厂
经　　销：全国新华书店
开　　本：185×260　　**印　　张**：15.25　　**字　　数**：369 千字
版　　次：2011 年 6 月第 1 版　　**印　　次**：2011 年 6 月第 1 次印刷
印　　数：1～3000
定　　价：26.00 元

产品编号：042630-01

计算机科学与技术系列教材　信息技术方向

编　委　会

序 言

随着高等教育规模的扩大以及信息化在社会经济各个领域的迅速普及，计算机类专业在校学生数量已在理工科各专业中遥遥领先。但是，计算机和信息化行业是一个高度多样化的行业，计算机从业人员从事的工作性质范围甚广。为了使得计算机专业能更好地适应社会发展的需求，从 2004 年开始，教育部高等学校计算机科学与技术教学指导委员会组织专家对国内计算机专业教育改革进行了深入的研究与探索，提出了以“培养规格分类”为核心思想的专业发展思路，将计算机科学与技术专业分成计算机科学(CS)、软件工程(SE)、计算机工程(CE)和信息技术(IT)四个方向，并且自 2008 年开始进入试点阶段。

以信息化技术的广泛应用为动力，实现信息化与工业化的融合，这是我们面临的重大战略任务。这一目标的实现依赖于培养出一支新一代劳动大军。除了计算机和网络等硬件、软件的研制开发生产人员外，必须要有更大量的专业人员从事信息化系统的建设并提供信息服务。

信息技术方向作为计算机科学与技术专业中分规格培养的一个方向，其目标就是培养在各类组织机构中承担信息化建设任务的专业人员。对他们的能力、素质与知识结构的要求尽管与计算机科学、软件工程、计算机工程等方向有交叉，但其特点也很清楚。信息技术方向培养能够熟练地应用各种软、硬件系统知识构建优化的信息系统，实施有效技术管理与维护。他们应该更了解各种计算机软、硬件系统的功能和性能，更善于系统的集成与配置，更有能力管理和维护复杂信息系统的运行。在信息技术应用广泛深入拓展的今天，这样的要求已远远超出了传统意义上人们对信息中心等机构技术人员组成和能力的理解。

信息技术在国外也是近年来才发展起来的新方向。其专业建设刚刚开始起步。本系列教材是国内第一套遵照教育部高等学校计算机科学与技术教学指导委员会编制的《高等学校计算机科学与技术专业发展战略研究报告暨专业规范(试行)》(以下简称专业规范)，针对信息技术方向需要组织编写的教材，编委会成员主要是教育部高等学校计算机科学与技术教学指导委员会制定专业规范信息技术方向研究组的核心成员。本系列教材的着重点是信息技术方向特色课程，即与计算机专业其他方向差别明显的课程的教材建设，力图通过这些教材，全面准确地体现专业规范的要求，为当前的试点工作以及今后信息技术方向更好的发展奠定良好的基础。

参与本系列教材编写的作者均为多年从事计算机教育的专家，其中多数人直接参与了计算机专业教育改革研究与专业规范的起草，对于以分规格培养为核心的改革理念有着深刻的理解。

当然，信息技术方向是全新的方向，这套教材的实用性还需要在教学实践中检验。本系列教材编委和作者按照信息技术方向的规范在这一新方向的教材建设方面做了很好的尝试，特别是把重点放在与其他方向不同的地方，为教材的编写提出了很高的要求，也有很大的难度，但对这一新方向的建设具有重要的意义。我希望通过本系列教材的出版，使得有更多的教育界的同仁参与到信息技术方向的建设中，更好地促进计算机教育为国家社会经济发展服务。

2009.6.1

中国科学院院士

教育部高等学校计算机科学与技术教学指导委员会主任

前 言

随着计算机技术和企业信息化的发展，为提高企业综合竞争力，企业往往会根据自身规模采用少则几个、多则几百的各种应用软件系统以及遗留的应用系统来支撑其业务的发展。而复杂多变的市场环境要求企业都必须能够高效灵活地管理和运作。这需要管理者不断改进其信息化系统以保证商业流程更加高效而准确，并且还要求其信息化系统易于扩展以适应新的需求。因此，实现企业内部、企业间、扩展企业网络间的业务流程自动化，将不同的应用系统进行集成，达到面向业务的企业流程的无缝连接，成为企业的迫切需求。

另一方面，随着网络、Internet 的发展及分布式系统的日益流行，大量异构网络及计算机厂商推出的软硬件产品广泛存在互操作的问题。要实现异构环境下的信息交互，实现系统在应用层的集成，需要集成技术。应用集成是将基于各种不同平台的、用不同方案建立的异构应用集成的一种方法和技术。应用集成关注的是多个系统之间互操作一致性的问题。

应用集成涉及应用系统分析、设计、实现、可靠性、安全以及重构等诸多方面的内容，技术面相当广泛。本书从应用集成的基本原理和实现技术的角度出发，讲述应用集成的概念、模型和部分技术。本书着重介绍集成系统的相关实现技术，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，都有该技术的应用示例和分析。特别对于基于 XML 的数据集成和基于分布式对象技术的应用集成，本书详细描述了两个具体应用场景，给出一般的解决方案以及主要的实现细节。

(2) 围绕异构系统集成介绍了多种实现技术，并分析了各种应用技术的特点。

本书共有 8 章内容。第 1 章概述了应用集成的概念、模型和主要技术组成。第 2 章描述了 XML 的相关技术，因为 XML 可以作为信息表示的规范而在应用集成中广泛使用。第 3 章介绍数据集成的概念、相关技术及其一般解决方案。

第4～8章介绍了应用集成相关的5个技术主题。第4章综述了程序设计层面的集成技术，第5章概述了软件复用技术，第6章介绍分布式对象技术，第7章介绍消息中间件技术，第8章描述Web Service的技术原理与实现。

本书由刘峰、郑滔老师编著，许林、苏赛杰、钱凯、周新等研究生参与部分整理工作。

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

2011年4月

目 录

第 1 章　应用集成概述

1.1　应用集成的概念

近二十多年来，信息技术行业中最富于戏剧性的变化，莫过于大型计算机在时代舞台上的逐渐隐去，而让各种网络工作站唱上了主角。在这个变化中，终端用户获得了比以前更为强大的处理能力，分布于整个企业各处的硬件资源也拥有了比以前更强大的功能。这种变化最初是从硬件上开始的，而现在则越来越多地体现到软件方面上来。

随着计算机技术和企业信息化的发展，为提高企业综合竞争力，企业往往会根据自身规模采用少则几个、多则几百的各种应用软件系统以及遗留的应用系统来支撑其业务的发展。如办公自动化(OA)、财务管理、客户关系管理(CRM)、企业资源计划(ERP)、供应链管理(SCM)、企业资产管理(EAM)以及其他商业应用和管理系统等。这些应用系统虽然相互独立运行，包含各自的应用、流程以及数据，但是它们之间往往有很多交叉点，包含重复的信息与数据。由于这些应用系统可能是在不同时间、不同条件、由不同的开发者在不同平台上用不同计算机程序语言开发出来的，相互之间无法进行通畅的信息交流与共享，因而在企业内部形成一个个“信息孤岛”，给企业信息化带来严重的问题，具体有如下体现：

(1) 由于缺乏不同系统之间的互操作性，很难实现跨系统的信息共享，造成信息资源共享的困难。

(2) 各系统之间交叉公共数据，不仅给使用者带来负担(如多次录入)，还造成系统中的数据冗余，数据的一致性很难得到保障。

(3) 由于各系统的异构性，难以用各系统数据进行企业的全局信息分析和处理。

(4) 各系统的数据和数据处理紧密绑定，缺乏柔性，不能快速适应企业业务流程的变化和灵活的信息服务需求。

(5) 分散的子系统为信息管理、系统维护等工作带来诸多不便。

另一方面，复杂多变的市场环境要求企业都必须能够高效灵活地管理和运作。这需要管理者不断改进其信息化系统以保证商业流程更加高效而准确，并且还要求其信息化系统易于扩展以适应新的需求。企业往往通过建立新的应用系统来满足新的需求，但当企业面对以下变动时，建立新应用也未必能解决问题。

(1) 流程管理：当企业迈向信息化后，内部的工作流程以电子化的方式进行。举例来说，当 ERP 系统进行存货查询时，就要透过库存系统的信息来进行，应用程序之间仍存有系统整合及流程问题；当商业环境转变时，工作流程也必须保持一定的弹性，可以随企业最新的需求变更。

(2) 电子商务：整合企业内的应用程序，达到营运流程自动化及跨应用程序间交易的整合，打破原本存在于各应用软件间的鸿沟，对于电子商务而言是重要的竞争优势。此外，对于企业内部部分封闭的作业环境，需要通过企业应用整合将电子商务与这些封闭的系统结合在一起。

(3) 协同商务：企业透过因特网与供货商整合进行协同商务时，为能实时反应、防止错误数据产生、避免重复输入并达到营运流程整合，需要将内部的应用程序与供货商的应用程序整合起来。面对不同的协同商务需求，企业需要一个简单且单一的解决方案，透过内部应用程序整合，提供对外单一的窗口会是企业最佳的选择。

(4) 企业并购：当两家企业进行合并时，将会有分属在两个企业中大量的数据需要重整，而应用程序间的连接则是关键。

因此，实现企业内部、企业间、扩展企业网络间的业务流程自动化，将不同的应用系统进行集成，达到面向业务的企业流程的无缝连接，成为企业的迫切需求。而随着网络、Internet的发展及分布式系统的日益流行，大量异构网络及计算机厂商推出的软硬件产品广泛存在互操作的问题。要实现异构环境下的信息交互，实现系统在应用层的集成，需要研究新的技术。企业应用集成(Enterprise Application Integration，EAI)是完成在组织内、外的各种异构系统，应用和数据源之间共享和交换信息和协作的途径、方法学、标准和技术。

集成是一种把多个系统的数据和功能组合或连接成一个具有凝聚力的集合的方法。集成关注的是多个系统之间各种互操作的一致性。应用集成是将基于各种不同平台、用不同方案建立的异构应用集成的一种方法和技术。从应用集成的一致表示、行为和功能的角度来看，应用集成为应用提供了一种一致无缝的认知和操作模型。

实施应用集成需要理解并保持应用系统之间相互关联和相互操作。对这些相对独立的应用系统之间的数据共享和通信的需求就是应用整合的最本质的内容。应用集成把多个应用系统正确地组合起来建立一个由多个模块组成的复合应用，其中的这些模块可以相互调用或进行合作。在这种情况下，需要获取并记录现存模块之间的关联关系和依赖性，并在一定程度上保持这种关系。

应用集成有三个目标，首先是对存储于不同应用系统中的数据提供访问接口，这些信息由不同的应用系统来管理，并且可能存在着格式差异；其次是在各个应用系统中的信息发生变化时保证相关信息之间的依赖关系和约束条件；最后是使得不同系统中管理数据的程序可以互操作，互相调用获取数据或者实施服务。

应用集成通过建立底层结构，来联系横贯整个企业的异构系统、应用、数据源等，完成在企业内部的 ERP、CRM、SCM 数据库、数据仓库，以及其他重要的内部系统之间无缝地共享和交换数据的需要。基于应用集成，企业就可以将企业核心应用和新的 Internet 解决方案结合在一起。应用集成保证底层子系统和数据的异构性对用户来说是透明的，将进程、软件、标准和硬件联合起来，在两个或更多的企业系统之间实现无缝集成，使它们就像一个整体一样。

对应用集成的技术要求大致有以下三项内容：

(1) 能提供应用间的互操作性。应用的互操作性不仅提供不同系统间信息的有意义交换，而且还提供系统间功能服务的方便使用，特别是资源的动态发现和动态类型检查。

(2) 能提供分布式环境中应用的可移植性。提供应用程序在系统中迁移的潜力并且不会破坏应用所提供的或正在使用的服务。这种迁移包括静态的系统重新部署以及动态的系统重构。

(3) 能提供系统中应用分布的透明性：分布的透明性屏蔽了由系统分布所带来的复杂性。分布的透明性使应用实现者不必过于关心系统的实现细节，从而大大减少应用集成编程的复杂性。

1.2　应用集成的历史

应用集成的发展演变经历了十多年的时间，产生了几代从不成熟到逐渐成熟的应用集成技术。

在 20 世纪 60—70 年代期间，企业应用大多是用来替代重复性劳动的一些简单设计。当时并没有考虑到企业数据的集成，唯一的目标就是用计算机代替一些孤立的、体力性质的工作环节。

20 世纪 80 年代，企业规模开始扩大，企业业务和数据日趋复杂，一些公司开始意识到应用集成的价值和必要性，很多公司的技术人员试图在企业系统整体概念的指导下对已经存在的应用进行重新设计，以便将它们集成在一起。此时点到点(point-to-point)的集成技术开始出现，在各个应用系统之间通过各自不同的接口进行点到点的简单连接，实现信息和数据的共享。点对点的应用集成也被称为第 0 代应用集成技术。

20 世纪 80 年代末和 90 年代初，随着企业规模的进一步扩大，应用系统不断增加，信息共享变得越来越重要，简单的点到点连接已经很难满足不断增长的应用集成要求，企业迫切需要新的集成方法：可以少写代码，无须巨额花费，就可以将各种旧的应用系统和新的系统集成起来。这时候以 VITRIA 为代表的公司开拓了第一代集成技术。第一代应用集成技术的出现在一定程度上解决了这些问题，它采用 CORBA、DCOM、MOM(消息中间件)等技术实现了对企业信息的集成，促进了企业的进一步发展。

20 世纪 90 年代中后期，ERP 开始流行起来，企业快速发展与电子商务结合对应用集成解决方案提出了更高的标准，仅仅局限于信息共享的集成技术无法实现企业业务流程自动处理、管理与监控。20 世纪 90 年代中期在 EAI 技术的基础上，发展了以业务流程管理为核心的第二代集成技术，即业务流程集成技术(Business Process Integration，BPI)。基于业务流程管理/集成(BPM/BPI)的第二代应用集成技术成为更加合适的集成选择方案。第二代应用集成技术通过实现对企业业务流程的全面分析管理，可以满足企业与客户、合作伙伴之间的业务需求，实现端到端的业务流程，顺畅企业内外的数据流、信息流和业务流。

21 世纪以来，B2B 与 B2C 的应用逐渐流行，企业间建立合作联盟逐渐成为发展趋势。跨企业边界的业务流程像企业内部的流程整合一样，实现跨组织的流程整合，将整合的对象延伸至整个供应链的上游和下游。商业社群中的企业个体间的流程整合需要将不同组织间的商业应用程序与商业流程进行有效整合，达到组织内部系统与外部系统统一的层次。其主要的整合范围包括社群内交易流程、信息分享流程及合作流程等。企业内部与外部集成方案由于管理与业务上的需求不同，决定了采用的技术不同，适用内部业务支撑的系统的紧密结合方式不再适应于外部系统的互连，因此出现了面向服务的第三代应用集成。第三代应用集成与第二代的差别在于基于服务架构(SOA)。把业务流程管理与基于服务架构结合起来，它在实现流程集成的同时，以基于服务架构提供流程重组与发布。第三代应用集成将企业的应用集成基础应用进一步扩展，通常的网络通信将由 Internet 完成，数据格式是标准 XML，集成平台的功能包括：安全、智能代理、数据映射、应用描述等，更加适合 B2B 与 B2C 等应用。