



这是一本SQL Server 2005数据库管理并结合
Microsoft Training and Certification (Microsoft培训及认证)
知识体系方面的应试教材

SQL Server 2005 初学者指南

北京希望电子出版社 总策划
吴豪 主编
宁义 郑兵 薛涛 编著

- 全面诠释SQL Server知识难点
- 细化数据库结构与范式设计
- 嵌入XML Web及分布查询技术
- 高效开发C/S、B/S应用程序

**直揭
数据库技术
核心**



科学出版社
www.sciencep.com



这是一本SQL Server 2005数据库管理并结合
Microsoft Training and Certification(Microsoft培训及认证)
知识体系方面的应试教材

SQL Server 2005 初学者指南

北京希望电子出版社 总策划
吴豪 主编
宁义 郑兵 薛涛 编著

- 全面诠释SQL Server知识难点
- 细化数据库结构与范式设计
- 嵌入XML Web及分布查询技术
- 高效开发C/S、B/S应用程序

**直揭
数据库技术
核心**



科学出版社
www.sciencep.com

内 容 简 介

这是一本作者亲历 SQL Server 2005 数据库管理及应用开发所有实践细节记录的图书。同时结合 Microsoft Training and Certification (Microsoft 培训及认证) 知识点进行逐项探讨, 对应试者及开发者都有很好的指导作用。

本书共分 12 章, 内容涵盖了 SQL Server 2005 的安装配置、数据管理器、数据库创建与管理、T-SQL 应用、联接表对象、数据库复制与备份、数据库优化及服务集成等内容。由浅入深、系统全面地介绍了 SQL Server 2005 数据库管理技术的所有细节。讲练结合, 实用性强。

本书非常适用于初级读者, 为他们提供最基本的 SQL Server 数据库管理知识; 同时也是中级用户、Web 数据库设计人员以及项目开发人员必须熟悉的后台数据库管理技能, 帮助他们快速完成大型系统开发与应用设计。另外, 也可作为社会培训班以及 Microsoft 培训及认证教材。

图书在版编目 (CIP) 数据

SQL Server 2005 初学者指南 / 吴豪 主编; 宁义、郑兵、
薛涛编著. —北京: 科学出版社, 2007
(数据库管理及应用超级技术)
ISBN 978-7-03-020492-9

I. S... II. ①吴...②宁...③郑...④薛... III. 关系数据库—数据库管理系统—SQL Server 2005 IV. TP311.138

中国版本图书馆 CIP 数据核字 (2007) 第 170040 号

责任编辑: 但明天 / 责任校对: 张月岭
责任印刷: 媛 明 / 封面设计: 刘孝琼

科学出版社 出版

北京东黄城根北街 16 号
邮政编码: 100717

<http://www.sciencep.com>

北京媛明印刷厂印刷

科学出版社发行 各地新华书店经销

*

2008 年 1 月第 一 版 开本: 787×1092 1/16
2008 年 1 月第一次印刷 印张: 30 1/4
印数: 1—3000 字数: 758 784

定价: 48.50 元

编委会名单

编委主任： 宁 义

副 主 任： 徐卸古 李瑞兴

主 编： 吴 豪 郑 兵 薛 涛 黄殿龙

副主编： 杨前凤 李卫东 马玉娟 彭东平

编 委： 阳 亮 周 雨 索向军 郭 歌

李清杰 蒲 卫 郭树森 刘希华

张宇辉 刘素刚 梁力武 胡湘林

薛万国 施建飞 李焕文 郭永平

王 宏 翁盛鑫 毕小龙 吕俊文

前 言

SQL Server 2005 在 SQL Server 2000 的基础上产生了飞跃发展。有了 SQL Server 2005, 组织内的用户和信息技术 (IT) 专家将从减少的应用程序停机时间, 提高的可伸缩性及性能, 更紧密而灵活的安全控制中获益。

SQL Server 2005 包括许多新的和改进的功能来帮助 IT 工作人员更有效率地工作, 并在企业数据管理中几个关键部分增加和增强了不少新功能。比如: 易管理性、可用性、可伸缩性、安全性等等。作为目前应用广泛的一种数据库, 它不仅采用了合理的体系结构, 而且利用了优越的可视化开发环境, 加上其优秀的编辑界面, 使得其很多性能又远远超过了 Oracle 及其他数据库, 从而为广大用户所青睐。虽然 SQL Server 2005 推出已有一段时间了, 但是为了使越来越多的用户逐步深入掌握 SQL Server 2005 中的博大功能, 并备战微软认证, 笔者编写了这本书, 并与大家分享。

笔者是具有多年数据库研发经验的资深系统分析员和软件高级工程师, 使用过 SQL Server 2005 和 Oracle 等大型数据库, 开发过医院信息管理系统和远程医疗视频会议等大型系统。在学习和研究这些数据库的过程中, 笔者深深地感受到学习过往的艰辛, 为把推敲再三的心得殷实详尽地记录下来, 为后学者铺出一条广阔的道路, 于是笔者与同事一起系统地编写了这本书。

这是一本 SQL Server 2005 进阶的从书, 更是一组开发经验和实际摸索的结晶, 可满足不同层次读者的切实需要。

本书目的在于使你阅读以后能在前进的道路上掌握更高级的技术。从一开始, 笔者就假设你已经是一位有经验的开发者 (并不一定在数据库方面)。为了彻底理解本书, 虽然需要你懂得编程的基础知识, 比如变量、数据类型、程序化设计等, 但并不要求你在以前就看过有关查询的内容 (尽管假设你可能有这方面的经验)。

只需从头到尾通读, 你就能很好地掌握本书所有知识。这套书虽然不是万能工具, 但它能够提供有关 SQL Server 开发的许多信息 (比如分析服务、报表服务和联机分析处理 [OLAP] 等)。如果你要掌握更深的有关开发的内容, 请参阅《SQL Server 2005 开发者指南》。

本书内容

为了使读者迅速掌握和利用 SQL Server 2005 进行数据库管理与开发, 设计出自己需要的数据库结构, 笔者从数据库设计的角度出发, 先介绍数据库简史, 然后介绍如何使用 SQL Server 2005 必备的工具, 再讨论如何进行数据库设计并开发应用的基础, 最后就高级数据库设计中可能遇到的各种问题做详细剖析。

在本书中, 笔者介绍了 SQL Server 2005 的版本、新增功能、常用工具、开发基础 (包括 T-SQL、联接、创建和修改表), 然后讨论了范式化设计与与设计相关的一些基本问题 (比如数学建模知识), 同时就数据库开发模式展开探讨, 再结合目前广为应用的 XML 技术, 探讨了如何将 SQL Server 2005 与 XML 结合起来, 然后介绍了有关复制的知识, 并就如何使用 XML 技术、本机 XML Web 查询和公共语言运行时进行了探讨, 从而对利用多种技术

进行高级查询做了深入研究。同时就如何设计合理的包展开描述,最后就综合服务和 Visual Studio 等内容做了详细讨论,再针对大家感兴趣的全文检索、分布式查询、服务分割器、报表服务和分析服务进行了探讨,并围绕性能优化和管理展开详细论述,最后以典型示例作为结束。

下面简要介绍一下本书各章节的内容:

在绪论中比较各种 SQL Server 版本和数据访问技术,并提供了 SQL Server 的一些历史背景信息。

第 1 章对 SQL Server 2005 的安装步骤进行探讨。在本章中,笔者详细讨论了数据库的版本和默认安装方法,包括存储过程、有效性规则和视图等其他常见的对象,并简要讨论了如何从 SQL Server 2000 升级到 SQL Server 2005 的方法。

第 2 章介绍了 SQL Server 2005 的组成,重点介绍了 SQL Server 2005 的管理工具、配置工具、性能工具和联机文档,同时简要地介绍了更新向导和开发工具,重点对 SQL Server 2005 的基本概念(比如基础语句、创建和修改表、约束、视图、脚本和批处理、存储过程、用户自定义函数、触发器、游标与快照等)进行了深入讨论。

第 3 章介绍初学者和开发者在 SQL Server 2005 中所使用的管理向导。通过本章,你将熟悉管理向导的使用方法和数据库的组成(比如数据库对象、数据类型和对象标识符等等),最后介绍了数据库引擎,还用一定篇幅介绍了数据库引擎的增强功能,比如公共语言运行时、用户定义聚合、用户定义类型和存储过程、触发器与用户自定义函数,包括 Web 访问。

第 4 章根据基本查询的需要,系统地介绍了 Transact-SQL,包括 SELECT、INSERT、UPDATE 和 DELETE 语句。通过本章的学习,你可以掌握最简单、常用的用语言操纵数据库数据的方法。

第 5 章基于前四章的内容,介绍了联接(joins: INNER、OUTER、FULL 和 CROSS)和联合(UNION)的概念。特别注意的是,其中介绍的替代语法,只是为了兼容性的缘故才介绍的,目前已不太常用。

第 6 章处理创建和修改表,实际介绍了创建数据库对象的方法。本章介绍的内容包括 SQL Server 2005 的对象名(比如架构、数据库名和服务器命名等),还介绍了默认值,最后重点介绍了 CREATE 语句,用该语句创建数据库、架构和表,并简要探讨了 ALTER 和 DROP 语句,最后介绍了如何使用 GUI 工具来创建数据库。

第 7 章介绍 SQL Server 2005 管理,在讨论安全性的基础上,介绍了安全性选项,重点讨论了用户权限、服务器和数据库角色、应用程序角色、XML 安全性、调度作业、自动响应目录和全文目录操纵等内容,并用一定篇幅探讨了备份与恢复的问题,包括索引重建、归档数据和元数据等概念。

第 8 章重点讨论了数据库的优化问题。实际上,优化是一个非常复杂的问题,如果你是初学者,可能本章是最难掌握的,因为实现优化涉及很多的知识,优化的方法也有多种。本章只是简要地探讨了其中的知识点,要想彻底阐述清楚有关优化的问题,可能还需要一写本书来进行介绍。

第 9 章讨论了 SQL Server 2005 的服务,介绍了服务分割器,重点探讨了分析服务(包括终端用户的要求、维度数据库、OLAP 存储类型、数据仓库、数据转换、元数据、数据挖掘模型、数据挖掘算法和分析服务体系结构,并介绍了增强的商业智能功能),同时简要

探讨了报表服务和通知服务。

第 10 章重点介绍 SQL Server 2005 的灾备问题。这是任何一个数据库初学者必须掌握的内容。如何进行备份,才能保证自己的数据不丢失?才能使自己的数据具有完全可恢复性?如何进行恢复?本章将带你揭开其神秘的面纱。

第 11 章讨论 SQL Server 2005 中的复制问题。介绍了 BCP,然后讨论了架构复制、发布比喻、复制类型和复制模型和实例,并介绍了计划复制、启用发布和分发、事务/快照发布刊物、合并发布刊物以及从 Oracle 发布数据的问题,最后探讨了如何管理复制的内容。

第 12 章涵盖了综合服务的内容。本章不仅要介绍综合服务的体系结构,而且要讨论如何使用 Visio Studio 进行包设计、包配置和包移植等诸多内容。

本书适合 SQL Server 2005 数据库设计领域的初学者、开发者和系统管理员,但重点面向初学者。这意味着,本书内容覆盖面广,可以使初学者最大限度地掌握 SQL Server 2005 的主要内容,而且会讲透,但不会进行太过深入地讨论。为了简明扼要,本书的内容涵盖了初学者、管理员和开发者所要掌握的知识点,而且对开发者有影响或在开发过程中需要考虑的管理因素,会在部分章节里对相关内容进行阐述。

对于初学者或具有一定关系数据库经验而不熟悉 SQL Server 2005 的人来说,使用这套书可以迅速利用 SQL Server 2005 进行数据库开发,设计出自己需要的数据库结构;对于具有一定基础的中级水平的开发人员,可以利用本书提供的内容,融会贯通、灵活运用,并利用本书提供的最新特点和使用方法,帮助自己提高开发和运用工具的水平;对于高级开发人员,这也是一套不可多得的参考书,书中讨论的内容表明:新版软件包含了更多的内容,SQL Server 2005 将以一个全新的面孔展现出精彩的数据库系统的世界。

由于成书的时间紧、任务重,尽管作者做了不懈努力,但书中难免存在不妥和遗漏之处,敬请广大读者批评、指正。作者联系电子邮箱 jszczxwuhao@sina.com。

目 录

引论.....	1
数据库简史.....	1
数据库类型.....	1
关系数据库发展史.....	2
建立数据库连接系统.....	3
单层结构（主机系统）.....	4
二层结构（客户/服务器结构）.....	4
三层结构.....	6
N 层结构.....	7
数据存取模型.....	8
小结.....	9
第 1 章 SQL Server 2005 安装与配置.....	11
1.1 SQL Server 2005 简介.....	11
1.2 安装条件.....	14
1.2.1 版本说明.....	15
1.2.2 硬件和 OS 要求.....	16
1.3 安装.....	20
1.4 配置.....	24
1.4.1 SQL Server 实例.....	24
1.4.2 Licensing 选项.....	25
1.4.3 SQL Server 服务的安全性选项.....	25
1.4.4 校验考虑.....	26
1.4.5 升级.....	26
1.4.6 系统一致性检查器.....	27
1.4.7 智能化安装.....	27
1.4.8 服务器级选项.....	27
1.5 小结.....	28
第 2 章 SQL Server 2005 组成.....	29
2.1 SQL Server 管理工具.....	29
2.1.1 SQL Server Management Studio.....	29
2.1.2 综合服务（SSIS）.....	31
2.1.3 大容量拷贝程序.....	32
2.1.4 MS DTC 管理控制台.....	32

2.1.5 SQL Server Profiler.....	32
2.1.6 查询分析器.....	34
2.1.7 OSQL.....	34
2.1.8 sqlmaint.exe.....	34
2.2 配置工具（Configuration Tools）.....	35
2.2.1 通知服务命名提示符.....	35
2.2.2 报表服务配制器.....	36
2.2.3 SQL Server 配置管理器.....	36
2.2.4 向导错误和使用情况报告设置.....	43
2.2.5 SQL Server 表面区域配置器.....	44
2.3 性能工具（Performance Tools）.....	46
2.4 文档和指南（Documentation and Tutorials）.....	48
2.4.1 联机丛书.....	48
2.4.2 示例.....	49
2.4.3 指南.....	49
2.5 SQL Server 2005 Upgrade Advisor.....	49
2.6 开发工具.....	49
2.7 基本概念.....	50
2.7.1 T-SQL 基础语句.....	50
2.7.2 创建和修改表.....	52
2.7.3 约束.....	53
2.7.4 视图.....	54
2.7.5 脚本和批处理.....	55
2.7.6 存储过程.....	56
2.7.7 用户自定义函数.....	57
2.7.8 触发器.....	58
2.7.9 SQL 游标.....	59
2.7.10 数据库快照.....	61
2.8 小结.....	61
第 3 章 SQL Server Management Studio.....	62
3.1 SQL Server Management Studio 导航.....	62
3.1.1 Management Studio 窗口介绍.....	63

3.1.2 编写 T-SQL 语句.....	68	5.6.3 CROSS JOIN 的替代语法	174
3.1.3 使用 Management Studio	71	5.7 联合	175
3.2 SQL Server 数据库组成.....	77	5.8 小结	178
3.2.1 数据库对象.....	77	第 6 章 创建和修改表	179
3.2.2 SQL Server 数据类型.....	93	6.1 SQL Server 的对象名.....	179
3.2.3 SQL Server 对象标识符.....	97	6.1.1 架构	179
3.3 SQL Server 数据库引擎.....	98	6.1.2 数据库名	181
3.3.1 连接数据库引擎.....	98	6.1.3 用服务器命名	181
3.3.2 数据库引擎的增强功能.....	100	6.1.4 回顾默认值	182
3.4 小结.....	105	6.2 CREATE 语句	182
第 4 章 T-SQL 基础语句.....	106	6.2.1 创建数据库	182
4.1 最基本的 SELECT 语句.....	106	6.2.2 创建架构	190
4.1.1 SELECT 和 FROM 从句.....	107	6.2.3 创建表	191
4.1.2 WHERE 从句	110	6.3 ALTER 语句.....	200
4.1.3 ORDER BY 从句	112	6.3.1 ALTER DATABASE	200
4.1.4 利用 GROUP BY 从句聚集数据.....	116	6.3.2 更改表	202
4.1.5 利用 HAVING 从句在分组上 放置条件.....	125	6.4 DROP 语句.....	205
4.1.6 用 FOR XML 从句输出 XML	127	6.5 使用 GUI 工具	206
4.1.7 使用 OPTION 从句, 利用提示.....	128	利用数据库引擎创建数据库	206
4.1.8 DISTINCT 和 ALL 谓词.....	128	6.6 小结	210
4.1.9 LIKE 比较	130	第 7 章 SQL Server 2005 管理	211
4.2 INSERT 语句.....	135	7.1 安全性基础	212
4.3 用 UPDATE 语句更新取得的数据	140	7.1.1 固定登录 ID 和口令	212
4.4 DELETE 语句	142	7.1.2 口令有效期	213
4.5 小结.....	143	7.1.3 长度与组成	214
第 5 章 联接表.....	144	7.1.4 登录次数	215
5.1 联接.....	144	7.1.5 存储用户和口令	216
5.2 内部联接.....	148	7.1.6 信任	216
INNER JOIN 类似于 WHERE 从句	151	7.2 安全性选项	217
5.3 外部联接.....	156	7.2.1 安全性范围	217
5.3.1 简单的 OUTER JOIN.....	157	7.2.2 SQL Server 2005 安全性.....	218
5.3.2 处理更复杂的 OUTER JOIN.....	162	7.2.3 操作系统集成安全性	222
5.4 完全联接.....	169	7.3 用户权限	224
5.5 交叉联接.....	170	7.3.1 对特定数据库授予访问权	224
5.6 探讨联接的替代语法.....	171	7.3.2 在数据库中授予对象许可权	225
5.6.1 INNER JOIN 的替代语法.....	172	7.3.3 用户权限和语句级的许可权	228
5.6.2 OUTER JOIN 的替代语法.....	172	7.4 服务器和数据库角色	229
		7.4.1 服务器角色	229

7.4.2 数据库角色.....	230	8.3 数据处理.....	269
7.5 应用程序角色.....	233	8.4 策略要求.....	270
7.5.1 创建应用程序角色.....	234	8.5 程序维护.....	271
7.5.2 给应用程序角色添加许可权.....	234	8.6 优化存储过程.....	271
7.5.3 使用应用程序角色.....	234	8.6.1 尽可能缩短事务.....	271
7.5.4 删除应用程序角色.....	235	8.6.2 尽可能使用限制最少的 事务隔离层.....	271
7.6 XML 的安全性.....	235	8.6.3 必要时可以执行多种解决方案.....	272
7.7 更高级的安全性.....	236	8.6.4 尽可能避免使用游标.....	272
7.7.1 处理 Guest 账户.....	236	8.7 使用临时表.....	272
7.7.2 TCP/IP 端口设置.....	236	8.8 硬件考虑.....	273
7.7.3 禁用 sa 账户.....	237	8.8.1 服务器的独占使用.....	273
7.7.4 xp_cmdshell.....	237	8.8.2 I/O 与 CPU 速度.....	274
7.7.5 其他工具.....	237	8.8.3 OLTP 与 OLAP.....	276
7.8 调度表作业.....	237	8.8.4 在线或离线.....	276
7.8.1 创建操作员.....	239	8.8.5 宕机.....	277
7.8.2 创建作业和任务.....	241	8.8.6 丢失数据.....	277
7.8.3 SQL Server 代理.....	252	8.8.7 性价比.....	277
7.9 自动响应警报.....	253	8.8.8 服务器.....	278
7.9.1 在数据库引擎中创建自动响应 警报.....	254	8.8.9 理想系统.....	278
7.9.2 在 T-SQL 中创建自动响应警报.....	254	8.9 解决方法.....	278
7.10 全文目录操纵.....	256	8.9.1 各种显示计划和统计信息.....	279
7.10.1 备份与恢复.....	256	8.9.2 数据库一致性检验程序.....	283
7.10.2 安排注入的调度.....	256	8.9.3 查询控制器 (Query Governor).....	286
7.11 异地复制.....	256	8.9.4 SQL Server 事件探查器.....	287
7.11.1 复制数据库向导.....	256	8.9.5 性能监视器 (Perfmon).....	289
7.11.2 备份和恢复.....	257	8.10 小结.....	291
7.11.3 连接/分离.....	257	第 9 章 SQL Server 2005 服务.....	292
7.12 索引重建.....	257	9.1 服务分割器.....	292
7.13 归档数据.....	258	9.1.1 Service Broker 概念.....	293
7.14 数据库选项.....	258	9.1.2 Service Broker 体系结构.....	295
7.15 数据库元数据.....	260	9.2 分析服务.....	301
7.16 小结.....	263	9.2.1 终端用户的要求.....	302
第 8 章 优化.....	264	9.2.2 维度数据库.....	305
8.1 索引选择.....	265	9.2.3 OLAP 存储类型.....	310
8.2 数据库引擎优化顾问.....	266	9.2.4 数据仓库.....	311
8.2.1 工作负荷.....	267	9.2.5 数据转换.....	313
8.2.2 注意事项.....	268	9.2.6 元数据.....	314

9.2.7 数据挖掘模型.....	317	11.3.4 架构一致性.....	386
9.2.8 数据挖掘算法.....	317	11.3.5 其他考虑.....	387
9.2.9 分析服务体系结构.....	320	11.4 发布比喻.....	388
9.2.10 商业智能增强功能.....	324	11.4.1 订阅刊物.....	390
9.3 报表服务.....	326	11.4.2 订阅服务器类型.....	390
9.3.1 创建管理和报表.....	328	11.4.3 过滤数据.....	390
9.3.2 功能与概念.....	329	11.5 复制类型.....	391
9.3.3 报表服务体系结构.....	333	11.5.1 快照复制.....	391
9.4 通知服务.....	333	11.5.2 合并复制.....	394
9.4.1 通知服务定义.....	333	11.5.3 事务复制.....	397
9.4.2 关键术语.....	334	11.5.4 立即更新的订阅服务器.....	400
9.4.3 通知服务体系结构.....	335	11.5.5 混合复制类型.....	400
9.5 小结.....	337	11.6 复制模型假设.....	401
第 10 章 SQL Server 2005 灾备.....	338	11.6.1 标准模型.....	401
10.1 重要性及目标.....	338	11.6.2 混合模型.....	403
10.2 备份介质.....	340	11.6.3 附加说明.....	404
10.3 备份策略.....	341	11.7 实现示例.....	404
10.4 恢复数据.....	345	11.7.1 承包人供应公司.....	404
10.5 恢复模式.....	346	11.7.2 清除服务.....	405
10.6 恢复优化.....	348	11.8 计划复制.....	406
10.7 备份方法举例.....	351	11.8.1 数据考虑.....	406
10.8 恢复方法.....	358	11.8.2 复制类型.....	407
10.9 小结.....	365	11.8.3 复制向导.....	407
第 11 章 复制.....	366	11.9 启用发布和分发.....	407
11.1 bcp.....	367	11.9.1 启动向导.....	408
11.1.1 bcp 语法.....	367	11.9.2 默认配置.....	408
11.1.2 bcp 导出.....	370	11.9.3 禁用发布.....	410
11.1.3 格式化文件.....	371	11.9.4 T-SQL 过程.....	410
11.1.4 bcp 导入.....	374	11.10 事务/快照发布刊物.....	410
11.1.5 BULK INSERT.....	376	11.10.1 创建和管理发布向导.....	410
11.1.6 OPENROWSET.....	377	11.10.2 配置之后.....	413
11.1.7 XML 大容量加载.....	379	11.10.3 T-SQL 过程.....	415
11.1.8 优化大容量导入的方法.....	384	11.11 合并发布刊物.....	415
11.2 支持分布式数据.....	385	11.12 异类数据库复制.....	415
11.3 架构复制时的考虑.....	386	11.12.1 从 Oracle 发布数据.....	416
11.3.1 独立性.....	386	11.12.2 将数据发布到非 SQL Server 订阅服务器.....	423
11.3.2 延迟性.....	386	11.13 管理复制.....	424
11.3.3 数据一致性.....	386	11.13.1 复制脚本.....	424

11.13.2	支持各种复制.....	425
11.13.3	发布到 Internet	425
11.13.4	复制和架构变化.....	425
11.13.5	复制监视器.....	427
11.13.6	后续.....	428
11.14	企业范围的自动化管理.....	428
11.14.1	创建多服务器环境.....	429
11.14.2	管理作业.....	430
11.14.3	轮询服务器.....	430
11.14.4	管理事件.....	431
11.15	高可用性解决方案.....	432
11.16	小结.....	435
第 12 章	综合服务.....	436

12.1	体系结构	436
12.2	综合服务实用工具	438
12.2.1	综合服务与 Studio 环境	438
12.2.2	SSIS 设计器及概念	440
12.2.3	SSIS 向导	454
12.2.4	执行包实用工具	457
12.3	SSIS 设计	457
12.3.1	生成解决方案	457
12.3.2	设计和创建 SSIS 包	459
12.3.3	部署包	469
12.4	典型用途	473
12.5	小结	476

引 论

在学习后面的章节之前，有必要先介绍一下有关数据库方面的知识，使读者能够了解数据库的发展历程，并熟悉数据库管理与开发的流程。

简要说来，数据库是存储数据的工具，主要用来采集、管理、分析各种信息数据，为使用者进行决策提供依据。

笔者将在本章简要地讨论数据库的种类，比较它们的优缺点。其次，笔者还要用点时间来探讨数据库开发流程，并简要讨论系统结构方面的一些入门问题，这些问题将对数据库设计产生重大影响。

本书侧重于为备战微软认证考试的人所用，笔者会重点介绍与初学者相关的维护与管理知识。在《SQL Server 2005 开发者指南》中，笔者会重点探讨 SQL Server 2005 数据库的实现。

数据库简史

SQL Server 是一个关系数据库管理系统 (RDBMS, Relational Database Management System)。RDBMS 是用于存储和检索数据的系统，在该系统中，数据被组织到相互关联的表中 (本章不深入讨论有关关系数据库管理系统的定义，因为它涉及许多相关的理论，在随后的章节中，笔者会逐步介绍一些相关知识)。

本节将带你回顾一下数据库的发展史，并了解过去使用的各种数据库，从而使你更好地理解数据库技术的来源及发展现状，熟悉数据库风靡世界并走向辉煌的原因。

数据库类型

不要将数据库的概念局限于计算机术语，其意义远不止这些。数据库实际上是有组织的数据集合。

办公室里有序的文件抽屉在某种程度上就是一个数据库。实际上，数据库几乎贯穿了“文明”世界的所有历史阶段，甚至可以追溯到奴隶社会。

数据库与计算机的关系如此紧密，其原因关键在于计算机对大多数数据库而言 (事实上，并非所有数据库) 不仅速度快而且存储数据效率高。实际上，大约在 1962 年，数据库这一术语才正式发源于计算领域。随后，数据库被分成以下几类：

- ❑ 纸张形式 (Paper-based) 尽管目前人们不认为这种形式是数据库的一种，但它是当今数据库的最大组成部分。我们可以认为，当今世界上精心组织保存的成千上万的纸张实际就是数据库的一种形式。
- ❑ 遗留主机 (Legacy Mainframe) 通常是虚拟存储存取方法 (VSAM, Virtual Storage Access Method, 它在 IBM 的主机上最普遍) 数据库。不要低估了遗留主机的存在数量及其重要性，本书很多章节采用了这种解决方案，而没有采用目前广为流行的客户/服务器或基于网页的模型。

- ❑ **Dbase 和其他基于文件的数据库** 通常, 这种数据库包括早期的索引顺序存取方法 (ISAM, Indexed Sequential Access Method) 数据库。虽然 ISAM 通常使每张表使用不同的文件, 但是其概念来源于数据存储和物理访问途径, 而不是文件。目前, ISAM 数据库仍然具有相当广泛的应用范围, 有时甚至用于一些新开发, 包括一些小型数据库等。在 RDBMS 系统出现之前, 这些小型数据库曾经相当火爆, 它们在独立的小数据库上应用很广, 其缺点是同时访问这些数据库数据的用户不能超过一定数量。
- ❑ **RDBMS 系统** 该系统是大量数据的集合, 具有更好的数据完整性。该系统所做的工作不仅仅完成存储和提取数据, 而且特别关心数据的完整性。典型的, 虽然 VSAM 和 ISAM 可以用于存储数据, 但其本身并不控制数据的输入和输出, 而是使用这些数据库的应用程序负责实现数据完整性规则。如果多个程序同时访问存储在这些数据库中的数据, 那么必须保证这些程序相互之间在正常运行时不违背数据完整性规则。虽然 RDBMS 系统不存在 VSAM 和 ISAM 在数据完整性上的缺陷, 但在编程过程中仍然需要了解数据完整性规则, 以尽可能避免数据库产生出错消息 (现有数据库会提示一部分出错消息, 以保证数据更加安全)。
- ❑ **面向对象数据库 (OOD, Object Oriented Database)** 这种数据库已经存在一段时间了, 但直到最近才得到飞速发展。这种数据库以完全不同的方式管理数据, 而且到目前为止只找到了一些特殊用途, 这一点与文件管理系统相似。但是, 它不是将文件存储在几张表中, 而是存储在单一对象中, 该对象具有可以维护其状态的属性。面向对象的数据库管理系统 (OODBMS, Object Oriented Database Management System) 通常提供了面向对象的概念, 如继承和封装。

本章不对以上几种类型的数据库历史进行一一介绍, 而是简要介绍一下目前处于数据库领域主宰地位的 RDBMS 系统的发展史。因为 RDBMS 系统设计的场合不是基于单表工作的概念, 而是在与数据相关的截然不同的表上进行操作, 并以多种方式存储组合数据, 从而尽可能避免重复存储, 加快了事务处理速度。

关系数据库发展史

关系数据库的发展史可以追溯到 1960 年, IBM 公司的 E.F.Codd 第一次引入了关系数据库结构和结构化英文查询语言 (SQL, Structured Query Language) 的基本原理。这个概念实际上相当简单, 就是改进并增加当时其他数据库中相当普遍的数据完整性问题, 以减少重复数据, 从而降低成本。

关系数据库在很长时间内没有发生实质性变化, 直到 20 世纪 70 年代中后期, 世界上才出现了真正的关系数据库产品, 如 Oracle 和 Sybase。这些系统不仅提供了一种新的管理数据库结构的方法, 而且在多平台上运行, 可以实现在多系统之间共享数据。

➤ 20 世纪 80 年代, 美国国家标准化组织 (ANSI, American National Standards Institute) 制订了 SQL 语言规范, ANSI-SQL 诞生了, 这是 RDBMS 计算领域的关键时刻。ANSI 规范定义了各种系统之间的兼容性, 这意味着不同厂家的产品将具有更好的兼容性, 也意味着相互竞争的系统可以在一个 RDBMS 中嵌入多种专业技术; 这

对增加 SQL 领域的开发者数量产生了很大的推进作用。

微软 SQL Server (本书中简称 SQL Server) 最初起源于 Sybase 的 SQL Server (简称 Sybase)。1989 年, 微软与 Sybase 联合开发了一个用于 OS/2 版的 SQL Server。1993 年, 他们将 4.2 版的 SQL Server 移植到 Windows NT 上。6.0 版推出后, 这种合作关系解除了。6.5 及以后版本的 SQL Server 已经成为微软的专有产品, 并首次包含了“复制”的内容, 微软通常将这一版本称为 SQL Server 变革历程中的第一代产品。在 7.0 版中, 微软对该产品进行了彻底重写, 并取得了高度成功, 从而使该产品成为第一个可以用于 Windows 9x 的产品 (目前, SQL Server 中几乎没有保留 Sybase 的原有代码), 而且重新架构了关系型服务器, 实现了广泛的自动资源管理。在 2000 版中, 微软修补了 7.0 版存在的漏洞, 并增加了许多功能, 改进了系统性能并提高了系统可扩展性, 支持 XML 语言编程功能等, 并支持联机分析处理 (OLAP)、数据压缩、转换和加载 (ETL, Extract Transformation Load), 这是 SQL Server 变革历程中的第二代产品。为充分适应大规模联机事务处理 (OLTP)、数据仓库和电子商务应用程序的需要, 微软历时多年, 最终于 2005 年 11 月推出了 SQL Server 变革历程中的第三代产品, 这就是今天的 SQL Server 2005, 该产品更能适应大规模协同开发并充分使用了基于 .NET 的技术。在这一代产品中, 微软在开发时特别注重其高可用性、安全性, 并致力于提高开发人员的生产力。

从外观界面看, 2005 版与 4.2 版 (实际能在某处找到安装的最老版本) 已面目全非, 而且它们在实现和性能支持上存在本质区别。6.0 版增加了如游标支持的细节, 6.5 版增加了分布式事务、复制和 ANSI 兼容性。重写后的 7.0 版使得像数据存储定义的设备方面的问题消失了。2000 版支持 XML 语言。而 2005 版内含的 XML 功能则深受广大开发者的欢迎, 其服务分割功能深受使用者的喜爱, 这一版本的企业级 ETL 和深度数据挖掘则为数据共享和重复利用提供了可靠途径。SQL Server 2005 在企业数据管理中聚焦于自管理与优化功能, 针对企业应用具有高可用性, 而且在安全和性能特性上进行了更高的设计。SQL Server 2005 的开发产品具有内含的 XML 技术, 与 Visual Studio 及 .NET 集成在一起, 具有与 Web Service 的互操作性, 在业务综合性能上, 具有整合与转换数据的功能, 可以分析、存储及挖掘数据, 在报告与数据的交互功能上具有独到的特点。

建立数据库连接系统

近几年产生了一些系统模型, 通常可以将这些模型根据所服务的对象分为 3 种:

- 服务于用户 通常包括描绘用户接口 (UI, User Interface) 和基本格式及字段规则等方面的内容。例如日期格式, 包括使你知道给定字段是日期字段, 而且在输入数字到该字段时必须预先校验并确定该数字为日期型。该模型是一种表达式, 在每个字段中至少有一种预定的数据类型。
- 服务于业务 这部分就是众所周知的业务规则。在 3-层结构或 n-层结构中, 业务服务对象可以驻留在自身服务器上, 也可以分为多个服务器, 或者以最小安装的方式与数据服务共享服务器。
- 服务于数据 该服务完全用于数据存储和提取。数据服务通常以数据完整性规则

著称（例如出库数量不能为负数）。这就是 SQL Server 的优势。
下面介绍实现这些模型的一些传统的结构。

单层结构（主机系统）

这是早期的主机和小型机模型。这种系统实质上没有桌面逻辑，而只有一个无声的终端。所有通过电缆发送到终端的内容就是显示器所显示的信息（当然包括显示的数据）。其优缺点如表 1 所示。

表 1 单层结构的优点和缺点

优 点	缺 点
为了在网络上得到快速响应，所需带宽较少，在广域网或国际互联网带宽相当昂贵的情况下非常有用	硬件非常昂贵
相当可靠，主机可以持续工作多年而不停机	本质上是典型的独占，因此很难与其他系统共享信息
开发新软件十分轻松。只要在主机系统上安装，就可以使每个用户拥有新版软件，不需对逐个机器进行更新	可用的“复活”软件包非常有限。由于潜在客户数量较少，所以这些软件包的价格非常昂贵

二层结构（客户/服务器结构）

二层结构，或者叫客户/服务器系统，于 20 世纪 90 年代早期开始流行。实际上这种结构有两种类型：客户中心和服务器中心两种模式。

1. 客户中心模式

基于 PC 机便宜的角度（这是大多数人使用客户/服务器开发的原因所在）和在分配计算任务时尽可能得到最大效率的观念，应用客户/服务器的客户中心形式。因此，这种模式下通常只有数据业务部分在服务器上处理，正常事务和 UI 端的事情都在客户端执行，从而确保各个计算机公平分担所有任务（至少，这就是客户中心形式的思想）。其原理如图 1 所示。

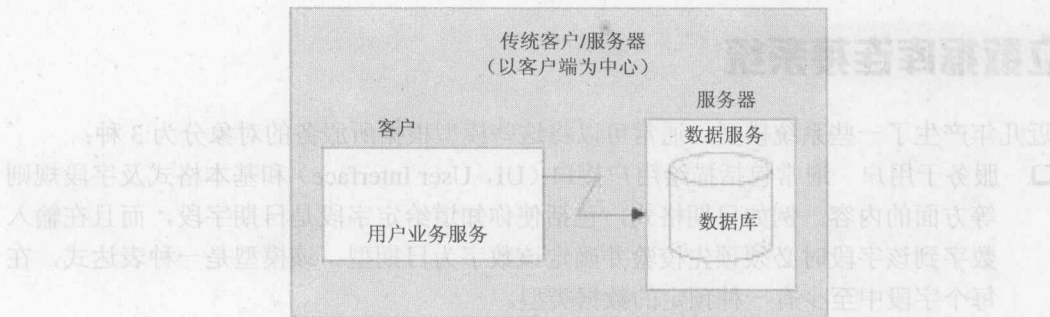


图 1 客户中心模式原理图

以客户端为中心的客户—服务器结构的最大问题是带宽（目前仍然存在）。如果所有业务逻辑都发生在客户端，则客户端与服务器端将存在大量交互（网络发射和接收）。大块大

块的数据经常被发送到客户端，最终导致网络阻塞，从而降低个人读取和写入的速度。表 2 显示了这种形式的优缺点。

表 2 客户中心结构的优缺点

优点	缺点
将工作分配到大量相关的相对便宜的客户机上执行	存在可怕的抢占带宽现象，而且很快会导致网络阻塞
如果个别用户需要更快速度，只要为其购买一台更快的系统即可，不需要购买一台每人都占用一块资源的极其昂贵的主机系统	安装需要时间，而且很难整理。新软件或新版本必须安装在多台机器上。版本更新时，问题尤其严重，因为老客户端与新服务器组件总是不兼容，反之亦然。所有客户端不得不一次更新，因而软件升级问题非常严重
在客户端花费相同的钱购买计算能力，也购买了其他产品应用程序的效率，如字处理和电子表格应用程序等	根据厂家而定，每一个客户端软件可能需要一个独立许可证，这样会增加成本

2. 服务器中心模式

这是从 PC 的计算成本比主机系统的计算成本要便宜的观点出发，希望尽量充分利用中心化系统优点的一种模式。客户端只分配用户业务，正常业务和数据业务放在服务器中。所占网络带宽小，其原理如图 2 所示。表 3 显示了这种形式的优缺点。

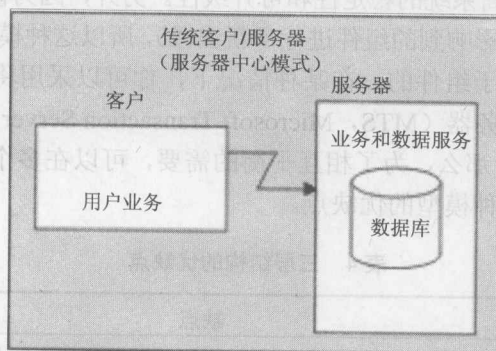


图 2 服务器中心模式原理图

表 3 服务器中心模式的优缺点

优点	缺点
一些升级工作可以完全在服务器端实现	其他的升级仍然需要与每个客户端接触，升级和新软件的安装是一个非常枯燥乏味的工作
大量相似产品都可重复使用，而且定制软件比较便宜	一个用户的长时间运行和过载工作会影响其他用户
所占网络带宽比客户中心模式小	服务器价格以指数幂增长，有一些服务器与主机系统一样昂贵
	尽管这种模型开始看上去与主机系统模型有点相像，但是通常停机很频繁