

21

21 BEIJING HOPE IBM

IBM 核心技术内幕丛书 (1)

IBM 核心技术内幕

Data Warehouse System Development Guide

大型数据库系统 开发指南

IBM 核心技术内幕丛书编委会 编写



本书配套光盘内容包括：
与本书配套的电子书



01 CX-3045

IBM 核心技术内幕

Data Warehouse System Development Guide

大型数据库系统开发指南

北京

* BITI *
TP311.138

1



北京希望电子出版社
Beijing Hope Electronic Press
www.bhp.com.cn

IBM 核心技术内幕

Data Warehouse System
Development Guide

大型数据库系统 开发指南

IBM 核心技术内幕丛书编委会 编写



本书配套光盘内容包括
与本书配套的电子书



北京希望电子出版社
Beijing Hope Electronic Press
www.bhp.com.cn

内 容 简 介

IBM 是世界上最大的信息工业跨国公司之一, 拥有综合先进技术与结构的全系列产品, 包括新一代基于 CMOS 的并行企业服务器、采用 64 位 RISC 技术的 AS/400 高级系列、基于高性能 PowerPC604 微处理器的新 RS/6000 系列以及广泛的软件和网络产品等。在电子商务、复杂的网络管理、系统管理、密集型事务处理、庞大的数据库、强大的可伸缩服务器、系统集成等方面, IBM 具有很强的优势。

目前, IBM 技术和产品已被国内各行各业广泛应用, 占有相当的市场份额。为满足技术领域专业人员和用户开发、应用和学习的需要, 我社和美国 Austin 技术研究中心以及 IBM 有关专家合作, 共同组织出版了本丛书——**IBM 核心技术内幕丛书** (7 本)。

本书主要介绍大型数据库系统 AIX DB2 UDB/Oracle 的设计、实现和优化等技术, 由 3 部分, 13 章组成。第一部分介绍 RDBMS 的概念, 包括: 关系数据库系统概述, 工作负载类型, 特殊数据库, 并行数据库等; 第二部分介绍了最优性能的系统设计和估算尺寸, 包括: 估算一个数据库系统尺寸, 设计一个 RDBMS 系统, 设计磁盘子系统; 第三部分介绍了系统优化, 包括: 数据库实现, 监视 RDBMS 系统性能, 调整 RDBMS 系统, DB2 UDB 的调整, Oracle 调整, 问题解决等技术。

本书反映了 90 年代末、21 世纪初 IBM 最新技术的发展, 内容定位与国内外技术和产品市场同步, 技术内涵高, 指导性、实用性、操作性强, 特别针对 IBM 技术用户、应用与开发人员、技术支持和管理人员, 具有很强的技术参考价值, 是以上人员必备的重要技术参考书, 同时也是高等院校相关专业师生教学、自学参考书和国内各图书馆、科研机构重要的馆藏书籍。

本书配套光盘内容包括与本书配套的电子书。

- 系 列 书: IBM 核心技术内幕丛书 (1)
书 名: 大型数据库系统开发指南
文本著作者: IBM 核心技术内幕丛书编委会 编写
责 任 编 辑: 苏 静
CD 制 作 者: 希望多媒体创作中心
CD 测 试 者: 希望多媒体测试部
出版、发行者: 北京希望电子出版社
地 址: 北京海淀路 82 号, 100080
网 址: www.bhp.com.cn
E-mail: lwm@hope.com.cn
电 话: 010-62562329,62541992,62637101,62637102,62633308,62633309
(发行, 技术支持)
010-62613322-215 (门市) 010-62531267(编辑部)
- 经 销: 各地新华书店、软件连锁店
排 版: 希望图书输出中心
CD 生 产 者: 北京中新联光盘有限责任公司
文 本 印 刷 者: 北京双青印刷厂
开 本 / 规 格: 787 毫米×1092 毫米 1/16 开本 14.25 印张 316 千字
版 次 / 印 次: 2000 年 7 月第 1 版 2000 年 7 月第 1 次印刷
印 数: 0001-3000 册
本 版 号: ISBN7-900044-75-2/TP·75
定 价: 48.00 元 (1CD, 含配套书)

说明: 凡我社光盘配套图书若有缺页、倒页、脱页、自然破损, 本社负责调换。

IBM 核心技术内幕丛书

编委会名单

主 编：戈里高·海登博格

副主编：范加尔·茨格 沈 鸿

编 委：（按姓氏笔划排序）

马宏华 屈里奇·托马斯 刘晓融 陆卫民

张中民 米歇尔·李 李春葆 苏 静

奥列佛·帕登 斯蒂芬·高奇 汤米·贾维奇

莫里哀·琴妮

本书执笔人：李春葆 曾慧 曾平 刘斌 胡海生等

序

IBM——国际商业机器公司，1914年创建于美国，是世界上最大的信息工业跨国公司之一。

在世界经济不断发展，现代科学日新月异的年代，IBM以超前的技术和产品，优良的客户服务，成为中国金融、冶金、石化、交通、制造、商品流通等许多行业的信息技术提供和支持伙伴。在信息产业迅速发展，竞争日趋激烈的时代，将信息产业的最新技术转化为对用户有价值的解决方案，帮助用户更有效地开展业务，增强竞争优势，IBM在产品技术等方面作出了重大的调整与改革。

IBM拥有综合先进技术与结构的全系列产品，包括新一代基于CMOS的并行企业服务器、采用64位RISC技术的AS/400高级系列、基于高性能PowerPC604微处理器的新RS/6000系列以及广泛的软件和网络产品等。在复杂的网络管理、系统管理、密集型事务处理、庞大的数据库、强大的可伸缩服务器、系统集成等方面，IBM具有很强的优势。

服务器形成了网络，服务器开启了电子商务。服务器正在人们的生活、工作中发挥越来越大的作用。IBM服务器家族目前包括四大主力系列产品：Netfinity，AS/400，RS/6000，S/390。它们都是各自领域中的顶尖服务器，在中国乃至世界各地的大小企业中担当着推动电子商务运行的核心任务。

目前，IBM技术和产品已被国内广泛应用，为满足技术领域专业人员和用户工作、学习的需要，我社和美国Austin技术研究中心以及IBM有关专家合作，共同组织出版了本丛书——IBM核心技术内幕丛书。本丛书由以下7种图书组成，全面介绍了IBM产品和技术的应用、开发、管理。

1. 大型数据库系统开发指南。本书介绍在DB2 UDB和Oracle环境中如何优化AIX的数据库性能。

本书由3部分13章组成，第一部分介绍RDBMS的概念，包括：关系数据库系统概述，工作负载类型，特殊数据库，并行数据库等；第二部分阐述最优性能的系统设计和估算尺寸，包括：估算一个数据库系统尺寸，设计一个RDBMS系统，设计磁盘子系统；第三部分讨论系统优化，包括：数据库实现，监视RDBMS系统性能，调整RDBMS系统，DB2 UDB的调整，Oracle调整，问题解决。

本书系统地论述了在AIX操作系统下的DB2 UDB和Oracle两种数据库管理系统中性能调整的技术和方法。帮助读者理解如何规划、运行和调整关系数据库管理系统的性能，在此基础上提供了很多经验规则，以指导读者进行数据库系统的优化。本书具有内容全面、新颖和权威的特点。

2. 小型机数据库应用完全解决方案。本书是IBM公司的IMS数据库产品使用的入门书籍。

全书共分为五大部分24章，第一部分是IMS的综述，讲述了IMS数据库产品的基本情况及系统组成，IMS数据库与OS/390操作系统之间的关系，以及IMS基本功能操作；第二部分介绍了IMS事务管理器，详细说明了各种事务的处理；第三部分关于IMS数据库管理器，讲述了IMS数据库的基础、IMS层次数据库模型，以及如何使用IMS数据库；第四部分具体介绍了IMS数据库的应用开发，涉及应用编码、信息格式化服务和应用代码等内容；第五部分说明了IMS数据库系统的管理，详细介绍了数据库恢复控制技术、记录类型、系统生成过程和系统安全功能。

3. 大型网络管理完全解决方案。本书主要讲述怎样用Tivoli管理PeopleSoft。

本书共分3部分，由12章组成。第一部分“PeopleSoft、BEA TUXEDO和Tivoli概述”包括：PeopleSoft简介、BEA TUXEDO简介；第二部分“Tivoli Manager For BEA TUXEDO”包括：安装和配置、流线型操作、保证BEA TUXEDO可用性；第三部分“管理PeopleSoft”包括：Tivoli Managers for Oracle、Manager for PeopleSoft安装和配置、流线型操作、保证PeopleSoft可用性、测量PeopleSoft环境中的性能、在PeopleSoft中管理批处理作业等等。

本书是 Tivoli、PeopleSoft 与 BEA TUXEDO 的宝贵参考资料，其图配文的方式别具一格。

4. **网络应用程序开发指南。**本书主要讲述 IBM 流行的 Tivoli Service Desk 6.0 应用程序软件的应用、开发。

IBM Tivoli Service Desk 6.0 是一个流行的应用程序软件，其广泛应用于企业、商业的管理系统中，本书共分 9 章、两个附录，详细介绍了 Tivoli Service Desk 6.0 的组成、工作原理和使用方法。其主要内容包括：Tivoli Service Desk 概要、Tivoli Service Desk 结构、Tivoli Service Desk 服务器的要求、分布式 TSD 的实现、Tivoli Service Desk 桥、TSD5.0.2 版本到 6.0 版本的数据移植用户界面/商务示例、分布式数据管理器、提示和技巧等。在附录 A 中包含了建立 Tivoli Service Desk 环境的小组所需使用的重要的脚本和命令，附录 B 中描述了与 Tivoli Service Desk 6.0 一起使用的一些数据库引擎的安装和准备方法。

5. **网络动态主机配置解决方案。**本书讨论了一种 TCP/IP 协议——动态主机配置协议 (DHCP)。

全书由十一章组成，内容包括：TCP/IP 概述，使用 DHCP 启动和运行网络，DHCP 服务器和客户的交互，服务名，集成文件和打印设备，扩展网络，移动用户，DHCP 和动态 DNS 的安全性，可靠性，性能，管理网络。本书从基础开始讲解，并逐步涉及到一些比较高深的理论和一些技术性的知识，由于 DHCP 提供的动态寻址能使网络更好地运行，所以本书重点介绍了它在网络上的应用，同时对涉及到的一些相关的内容也进行了介绍，例如：网络的安全性和可靠性。

6. **大型网络系统安全完全解决方案。**本书主要介绍 IBM Secure Way Host On-Demand 技术及其应用。

全书由 17 章组成，主要内容包括：IBM SecureWay Host On-Demand、系统规划、系统安装、在 System/390 上安装 Host On-Demand 服务器、系统管理、使用 LDAP 目录、Host On-Demand 客户程序、使用 Host On-Demand 的会话、3270 主机打印、5250 主机打印、系统安全、IBM SecureWay 屏幕定制器、主机访问类库、Host Access JavaBeans、Database On-Demand、双字节字符集 (DBCS) 因素以及问题确定等等。

本书具有结构清晰，内容新颖、丰富，技术含量高的特点，可供 IBM 公司的大型机和小型机用户，以及为结合传统方式的大型机和小型机系统的工作方式提供解决方案的高级软件开发人员使用，也可作为从事计算机工作、需要了解 IBM 技术的社会相关人士阅读。

7. **网络结构体系设计完全解决方案。**本书由 21 章组成，主要包括：SNA 在多网络时代，SNA 的发展，数据链路层教程，SNA 和 APPN 教程，非 SNA 协议——TCP/IP、IPX 和 NetBIOS 教程，应用程序设计接口教程，网络计划简介，SNA, APPN 与 TCP/IP——比较和对照，穿过数据链路的 SNA，通过 TCP/IP 传输 SNA，LAN 网络与 SNA 互连，在终端系统中，使 SNA 网络成为 Internet 的一部分，多协议路由器将 IP 与 APPN/HPR 聚集一起，管理 SNA 网络，与传统网络共存；3270 访问 TCP/IP 应用，使用 SNA 网关从 PC 到主机，集成 Web 技术和企业系统，SNA 应用和 Java 承诺，将几个协议永远共存，传输协议变革。

本书的每章都是一个关于 SNA 网络的专题，涉及当前许多有争议的和技术性很强的问题，通过仔细的分析和具有说服力的讲解，给出了权威的见解。

本丛书反映了 IBM 最新技术的发展，内容定位与国内外技术和产品市场同步，技术内涵高、指导性强，特别针对 IBM 技术用户、应用与开发人员、技术支持和管理人员，具有很强的技术参考价值，是以上人员必备的重要技术参考书，也是高等院校相关专业师生教学、自学参考书和国内各图书馆、科研机构重要的馆藏书籍。

藉本丛书出版之际，特别感谢美国 Austin 技术研究中心主任戈里高·海登博格教授，IBM 全球支持中心副主任范加尔·茨格博士，本丛书就是在他们的大力帮助和协调下才得以完成。感谢美国 Austin 技术研究中心米歇尔·李博士、IBM 系统开发专家奥列佛·帕登博士、IBM 加拿大公司技术总监屈里奇·托马斯教授、IBM 数据仓库专家莫里哀·琴妮女士，以及 MIT 计算机科学系斯蒂芬·高奇教授和 Tivoli 公

司专家汤米·贾维奇先生，由于他们的技术指导和全力参与，本丛书才得以及时完稿。还要感谢马宏华、张中民、李春葆、陆卫民、苏静等，是他们夜以继日的辛勤劳动，使本丛书及时面市。真诚感谢参与本丛书编写的全体专家和技术人员，以及编辑、美工设计人员和录排人员、光盘制作人员等，是他们的加班、加点、忘我的工作，才使本丛书如期付梓出版。

因出版时间紧迫，书中错误在所难免，敬请读者谅解，并请拨冗指正，以期再版时修订。

IBM 核心技术内幕丛书编委会

2000年6月



目 录

引言.....	1	3.1.4 进程.....	37
第一部分 RDBMS 概念		3.1.5 SQL 扩展——存储过程.....	38
1 关系数据库系统概述.....	3	3.1.6 管理工具.....	39
1.1 RDBMS.....	3	3.2 Oracle 数据库体系结构.....	41
1.2 一个 RDBMS 提供的功能.....	6	3.2.1 内存结构.....	41
1.3 数据库性能策略.....	8	3.2.2 逻辑存储结构.....	42
1.4 RDBMS 包含的组成部件.....	9	3.2.3 物理存储结构.....	43
1.5 RDBMS 术语和概念.....	13	3.2.4 进程.....	44
1.5.1 RDBMS 术语.....	13	3.2.5 SQL 扩展——存储过程.....	46
1.6 结构化查询语言.....	18	3.2.6 管理工具.....	46
1.7 如何安全地维护数据.....	18	4 并行数据库.....	48
1.8 备份和性能.....	21	4.1 数据库环境中的并行概念.....	48
1.8.1 备份介质.....	21	4.1.1 共享内存.....	48
1.8.2 完整或部分备份.....	22	4.1.2 共享磁盘.....	49
1.8.3 物理和逻辑备份.....	22	4.1.3 共享分布式.....	49
1.8.4 联机 and 脱线备份.....	23	4.2 DB2 UDB 企业扩展版本 (EEE).....	50
1.8.5 备份建议.....	24	4.2.1 概念和功能.....	50
2 工作负载类型.....	25	4.2.2 优化程序.....	51
2.1 联机事务处理 (OLTP).....	25	4.2.3 分区间和分区内并行.....	51
2.2 联机分析处理 (OLAP).....	26	4.2.4 硬件实现.....	52
2.3 决策支持系统 (DSS).....	26	4.3 Oracle 并行服务器.....	53
2.3.1 数据仓库.....	27	4.3.1 并行 Oracle 体系结构.....	53
2.3.2 数据中心.....	27	4.3.2 虚拟共享磁盘 (VSD).....	55
2.3.3 商务智能 (BI).....	28	4.3.3 分布式封锁管理员 (DLM).....	56
2.3.4 数据挖掘.....	28	4.4 并行数据库的优点和缺点.....	57
2.4 企业资源规划 (ERP).....	28	第二部分 最优性能的系统设计	
2.5 电子商务.....	29	和估算尺寸	
2.6 报表.....	30	5 估算一个数据库系统尺寸.....	59
3 特殊数据库.....	32	5.1 估算尺寸约束.....	59
3.1 DB2 UDB 数据库体系结构.....	32	5.2 估算尺寸技术.....	60
3.1.1 内存结构.....	32	5.2.1 从数据尺寸来估算尺寸.....	61
3.1.2 逻辑存储结构.....	33	5.2.2 从事务速率估算尺寸.....	61
3.1.3 物理存储结构.....	34	5.2.3 从用户个数估算尺寸.....	61



5.3 从一个特殊应用估算尺寸	62	6.7.4 处理大量的意外的增长	83
5.4 CPU 目标和估算尺寸	62	6.7.5 预期的增长领域	83
5.4.1 单处理器 (UP) 系统	62	6.7.6 加载大量数据	84
5.4.2 对称多处理器 (SMP) 系统	62	6.8 性能和可用性	84
5.4.3 CPU 利用率	63	6.9 在同一台机器上生产开发和测试	85
5.5 内存目标和估算尺寸	63	6.9.1 生产	85
5.5.1 AIX 操作系统	64	6.9.2 开发	86
5.5.2 AIX 文件系统高速缓存 (AIX 缓冲区高速缓存)	64	6.9.3 测试	86
5.5.3 RDBMS 高速缓存和结构	64	6.9.4 综合机器	86
5.5.4 用户应用和数据库连接	64	6.10 AIX 和 RDBMS 升级	86
5.6 磁盘目标和估算尺寸	65	7 设计磁盘子系统	88
5.6.1 通用数据库估算尺寸——高级	66	7.1 磁盘子系统的设计方法	88
5.6.2 特殊的一个表一个表地估算尺寸 ——细目级	67	7.2 与性能有关的带宽的考虑	88
5.6.3 选择何种磁盘尺寸	68	7.3 物理数据库布局的考虑	89
5.6.4 磁盘保护	68	7.3.1 数据库数据文件的分布	89
5.7 由组成部件代价平衡一个系统	70	7.4 逻辑卷管理的概念	90
6 设计一个 RDBMS 系统	71	7.4.1 物理分区带和 LVM 微带	91
6.1 工作空间	71	7.4.2 LVM 策略的使用	92
6.1.1 基本的和将来的 AIX 资源	71	7.5 原始逻辑卷和相对的日志文件 系统(JFS)	94
6.1.2 基本的和将来的应用程序资源	71	7.6 各级 RAID 概述和性能考虑	95
6.1.3 基本的 RDBMS 资源	72	7.6.1 0 级 RAID	95
6.1.4 未来 RDBMS 资源	74	7.6.2 1 级 RAID	95
6.2 考虑工作负载	75	7.6.3 2 级和 3 级 RAID	96
6.3 考虑网络	75	7.6.4 4 级 RAID	96
6.4 考虑内存和数据库	76	7.6.5 5 级 RAID	96
6.4.1 DBW UDB 内存需求	76	7.6.6 RAID 0+1	97
6.4.2 Oracle 内存需求	76	7.6.7 各种 RAID 级别之间的比较	97
6.5 系统资源利用	77	7.6.8 RAID 和 AIX LVM 镜像	97
6.6 数据库的备份和恢复	78	7.7 使用镜像写一致性(MWC)	98
6.6.1 DB2 UDB 备份/恢复方案	79	7.8 串行存储结构(SSA)	100
6.6.2 Oracle 备份/恢复方案	79	7.8.1 技术概述	100
6.6.3 通用备份考虑	79	7.8.2 SSA 特殊性能考虑	100
6.7 处理增长性	81	7.9 集成的磁盘存储系统	103
6.7.1 DB2 UDB 重组方法	81	7.9.1 IBM 企业存储服务器 (ESS)	103
6.7.2 Oracle 重新组织方法	82	7.10 磁盘性能测试和观察	104
6.7.3 何时需要并如何避免数据库重组	82	7.11 选择你的磁盘子系统	106



第三部分 系统优化

8 数据库实现	108
8.1 硬件和 AIX 就绪检查列表.....	109
8.2 启动前的检查列表.....	111
8.3 数据库数据.....	111
8.4 硬件测试.....	113
8.5 安装 RDBMS 代码.....	114
8.6 数据库物理布局.....	115
8.7 建立脚本.....	115
8.8 建立小型系统.....	116
8.9 安装后.....	117
8.10 备份和恢复测试.....	117
9 监视 RDBMS 系统性能	118
9.1 RDBMS 工具.....	118
9.1.1 DBW UDB 监视工具.....	118
9.1.2 Oracle 监视工具.....	128
9.2 常规、特定或报警监视方法的使用.....	132
9.2.1 常规 (regular) 监视方法.....	132
9.2.2 特定 (ad-hoc) 监视方法.....	133
9.2.3 报警 (alert) 监视方法.....	133
9.3 性能监视脚本.....	133
9.4 监视和调谐责任者.....	133
9.5 性能问题的报告.....	134
9.5.1 检查.....	134
10 调整 RDBMS 系统	136
10.1 调整技能.....	137
10.2 参考手册和参考书.....	137
10.2.1 有关 RDBMS 调整和 RDMBS 性能调整参考书.....	138
10.3 调整策略.....	138
10.4 合理的好的调整方法.....	139
10.4.1 弄清楚成功准则的定义.....	139
10.4.2 限制活动.....	140
10.4.3 重复.....	140
10.4.4 每次一个改变.....	140
10.4.5 决定优先权.....	140
10.4.6 热点.....	141
10.4.7 很好地了解重要领域.....	141
10.4.8 可再现的工作负荷.....	141
10.4.9 如何测定响应时间.....	142
10.4.10 注意测试设备和测量.....	142
10.4.11 文档.....	142
10.4.12 安排测试.....	143
10.4.13 检验改善.....	143
10.4.14 调整小组.....	144
10.5 同时改变的方法.....	144
10.5.1 忽略谣传.....	145
10.5.2 收集信息.....	145
10.5.3 检查错误.....	147
10.5.4 升级到最后确定的级别.....	147
10.5.5 研究系统.....	147
10.5.6 检查和设置顶级性能参数.....	147
10.6 瓶颈、利用和资源.....	147
10.6.1 利用目标.....	149
10.6.2 不足的 CPU 和潜在请求.....	149
10.6.3 不足的内存.....	150
10.6.4 不足的磁盘 I/O.....	151
10.6.5 不足的网络资源.....	153
10.6.6 不足的逻辑资源访问.....	153
10.7 调整.....	153
10.7.1 调整窗口.....	154
10.8 典型错误列表.....	155
11 DB2 UDB 的调整	156
11.1 性能提高过程.....	156
11.2 通常调整的元素.....	156
11.2.1 操作性能要考虑的事情.....	156
11.2.2 环境要考虑的事情.....	157
11.2.3 应用要考虑的事情.....	157
11.2.4 系统目录统计.....	157
11.2.5 SQL 编译程序.....	157
11.2.6 SQL 解释工具.....	157
11.2.7 使用 DB2 UDB 管理员.....	157
11.2.8 配置可扩充性.....	157
11.2.9 DB2 DUB 的内存应用.....	158
11.3 可执行的改变.....	159

11.4 选项.....	159	12.3.5 分页空间和监视分页.....	178
11.4.1 数据库管理员配置参数.....	159	12.3.6 对 Oracle 没有分配足够的内存....	179
11.4.2 数据库参数.....	161	12.3.7 较差的使用 AIX 磁盘特征.....	179
11.4.3 DB2 DUB 注册表变量.....	162	12.3.8 忙碌的磁盘.....	180
11.5 选项的作用.....	162	12.4 检查最常见的 Oracle 错误.....	180
11.5.1 缓冲池尺寸 (buffpage).....	162	12.4.1 索引.....	180
11.5.2 I/O 服务器数目 (num_ioservers)	164	12.4.2 分析.....	180
11.5.3 异步分页消除器数		12.4.3 基本的 Oracle 参数.....	180
(num_iocleaners).....	165	12.4.4 分析数据库的表和索引.....	180
11.5.4 改变的页阈值 (chnpggs_thresh)	165	12.5 AIX 和 Oracle 使用的调整提示分类....	182
11.5.5 排序堆尺寸 (sortheap).....	166	12.6 估计最重要的 10 个 Oracle 参数.....	183
11.5.6 排序堆阈值 (sheapthres).....	166	12.6.1 db_block_size.....	183
11.5.7 语句堆尺寸 (stmtheap).....	167	12.6.2 db_block_buffers.....	183
11.5.8 程序包高速缓存尺寸 (pckcachesz)	167	12.6.3 use_async_io 或 disk_asynch_io..	185
11.5.9 数据库堆尺寸 (dbheap).....	168	12.6.4 db_writers、db_writer_processes	
11.5.10 目录高速缓存尺寸		和 dbwr_io_slaves.....	185
(catalogcache_sz).....	168	12.6.5 shared_pool_size.....	185
11.5.11 日志缓冲区尺寸 (logbufsz).....	169	12.6.6 sort_area_size.....	186
11.5.12 最大主体数 (maxagents).....	169	12.6.7 sql_trace.....	186
11.5.13 锁列表的最大存储 (locklist) ...	170	12.6.8 timed_statistics.....	186
11.5.14 升级前最大锁列表百分比		12.6.9 optimizer_mode.....	186
(maxlocks).....	171	12.6.10 log_buffer.....	186
11.5.15 并行的最大查询度		12.6.11 rollback_segments.....	186
(max_querydegree).....	171	12.7 其他关键的 Oracle 参数.....	187
11.5.16 DB2MEMDISCLAIM 和		12.8 可重复的好的调整步骤.....	188
DB2MEMMAXFREE.....	171	12.8.1 访问方法调整.....	188
11.5.17 DB2_PARALLEL_IODB2_STRIP-		12.8.2 内存调整.....	189
ED_CONTAINERS.....	172	12.8.3 磁盘 I/O 调整.....	189
11.5.18 重新组织表.....	173	12.8.4 CPU 调整.....	190
11.6 通过 SYSSTAT 视图模拟.....	173	12.8.5 竞争调整.....	191
12 Oracle 调整.....	175	12.9 按 Oracle 提示调整 AIX.....	192
12.1 可以进行的改变.....	175	12.9.1 AIX 异步 I/O.....	192
12.2 Oracle 调整顺序.....	175	12.9.2 AIX 逻辑卷管理程序或 Oracle	
12.3 检查最常见的 AIX 配置错误.....	178	文件.....	192
12.3.1 改变控制.....	178	12.9.3 按标准尺寸建立逻辑卷.....	193
12.3.2 使用异步 I/O 故障.....	178	12.9.4 AIX JFS 或原始设备.....	194
12.3.3 差的磁盘子系统的安装.....	178	12.9.5 AIX 磁盘布局考虑.....	195
12.3.4 重做日志磁盘.....	178	12.9.6 命名约定.....	195



12.9.7	AIX 预先顺序读	195	12.11.14	Oracle db_file_multiblock_read_	
12.9.8	AIX 分页空间	196		count 参数	205
12.9.9	AIX 分页率	196	12.11.15	Oracle 重做日志锁存	205
12.9.10	消除热磁盘	196	12.11.16	Oracle 重做缓冲区尺寸	205
12.9.11	避免热磁盘的磁盘集	197	12.11.17	Oracle 共享缓冲池尺寸	205
12.9.12	在 SMP 机上平衡 CPU 利用	197	12.11.18	Oracle 表空间和表建立	206
12.10	高级的 AIX 调整提示	197	12.11.19	Oracle 回滚段数	206
12.10.1	AIX readv()特征	197	12.11.20	Oracle 并行处理	206
12.10.2	AIX 直接 I/O	198	12.11.21	Oracle 存档程序缓冲区	207
12.10.3	AIX 写滞后	198	12.11.22	Oracle 使用 TRUNCATE 而	
12.10.4	AIX 磁盘 I/O 定步	198		不是 DELETE 删除所有行	207
12.10.5	在 SMP 机上绑定 AIX 处理器	198	12.11.23	Oracle 标记和批量删除行	207
12.10.6	SMP 机上 AIX 旋转计数	199	12.11.24	Oracle SQL*Loader I/O 缓冲区	207
12.10.7	AIX 进程优先级	199	12.12	其他调整提示	207
12.10.8	AIX 进程时间片	199	12.12.1	网络 TCP/IP	207
12.10.9	AIX 空闲内存	199	12.12.2	使用嵌入的 Oracle SQL 编译	
12.10.10	AIX 缓冲区高速缓存尺寸	200		程序	208
12.11	Oracle 调整提示	201	13	解决问题	209
12.11.1	按照 Oracle 灵活的体系		13.1	Perfpmr——性能数据集工具	209
	结构安装 Oracle	202	13.1.1	获得 perfpmr 最新版本	209
12.11.2	Oracle ARCHIVEMODE	202	13.1.2	AIX 媒体提供的版本	210
12.11.3	Oracle 控制文件	202	13.2	出现问题前	210
12.11.4	AIX 的 Oracle post-wait 核心扩充	202	13.3	提升问题管理记录 (PMR)	210
12.11.5	Oracle 块尺寸	202	13.3.1	PMR 信息	211
12.11.6	Oracle SGA 尺寸	203	13.4	数据库性能 PMR 的大多数公共资源	212
12.11.7	Oracle 数据库写程序	203	13.5	避免下一个性能危机	213
12.11.8	Oracle 缓冲区高速缓存命中				
	率调整	203			
12.11.9	从 AIX 磁盘中分离数据库磁盘	204			
12.11.10	Oracle 重做日志应该有一个				
	专用磁盘	204			
12.11.11	镜像重做日志或使用 RAID 5 fast-				
	write cache 选项	204			
12.11.12	Oracle 重做日志组或 AIX 镜像	204			
12.11.13	Oracle 并发恢复	205			

引 言

关系数据库管理系统(RDBMS)越来越成为公司和大型企业的核心 IT 系统。由于 RDBMS 系统用于保存诸如销售额、库存品、资金和定单收入的数据，它们对于公司的盈利方法是至关重要的。这些数据可以被很多人访问，通常是每天 24 小时，每周 7 天，特别是公司扩展业务到全球市场时，人们需要从所有地区访问该系统。在公司的数据库系统中，不仅要保存每天商务生产数据，而且还要保存历史数据。这些数据对于研究和提供重要的管理决策是十分有用的。

对于一个公司的成功来说，这些系统的高性能(high performance)是至关重要的。快速处理公司定单，根据定单在仓库中找到可用的库存品并交赋给相应的定单，特别对于那些时间紧迫的公司，例如航空公司，必须要有好的性能。

因为系统每年越来越大，数据库越来越复杂，并且 RDBMS 系统意味着一个大的资源（包括人力物力）的大投资，每一部分都需要花钱。所以性能变成一个日益突出的问题。

数据库性能是一个具有很多不同方面的广泛领域，它取决于很多因素，例如，硬件、应用程序、工作负载、磁盘子系统的布局和无数的系统与数据库参数。

在本书中，提供给我们的知识和经验，可以帮助你理解在规划、运行和调整关系数据库管理系统的一些性能。我们发现在多数情况下，由于改变很少的参数就会产生很大的影响，所以我们很难控制整个局面，但根据我们的一些经验规则，为你提供了大量的信息以便确定你的数据库系统所需的性能和效果。

我们也希望消除有关调整性能选项的很多神话和传说，并且给出 AIX 最近特性的修改版能够用于产生最佳数据库性能。

你可能发现本书在数据库的生命周期——在规划和估算尺寸阶段、在实现中和在运行一个生产数据库系统中的很多阶段中都是有用的。我们使本书的结构适应于这个生命周期并划分为三个主要部分：

- RDBMS 概念——覆盖关系数据库管理系统的概念，不同的工作负载特性和 DB2 UDB 和 Oracle 数据库概述，包括并行数据库系统的简要介绍。
- 为了优化性能的系统设计和估算尺寸——覆盖 RDBMS 的前期阶段，估算尺寸以满足声明的工作负载的需要，以及为了优化性能的系统设计和布局。
- 系统优化——集中于一个 RDBMS 的实现以及在数据库安装后的监视与调整任务。

本书是从 AIX 和 RS/6000 的观点编写的，集中于在 RDBMS 上如何使用这些产品的高级特征。

尽管我们介绍的多数内容是覆盖 DB2 UDB 和 Oracle 数据库的，但本书中的大部分内容也适合于任何其他关系数据库管理系统。我们之所以选择 DB2 UDB 和 Oracle 数据库，是因为它们代表了 RS/6000 上安装的所有数据库的 80%，RS/6000 Enterprise 和 SP Systems 中也有 65% 运行这些数据库的版本。

数据库设计或应用程序设计是所有平台的公共的大型主题。有很多讨论这些主题的书籍，因此我们在本书中不讨论这些主题，也不详细讨论结构化查询语言(SQL)。

简单地提及了并行数据库，以便你知道在什么时候考虑它们。但讨论并行数据库设计

和性能超出了本书的范围。

设计、估算尺寸和调整一个 RDBMS 与其说成是科学倒不说成是艺术，需要具备一些技巧和个人经验。因此本书是你变成一个富有 RDBMS 性能经验的专家的有价值的信息资源。

第 1 部分 RDBMS 概念

1 关系数据库系统概述

本章讨论现代关系数据库管理系统(RDBMS)的基本原理以及其如何在 AIX 的概要层上实现。还详细讨论了 RDBMS 获取数据的手段以及如何保持数据安全。如果你是数据库的初学者或者忘记了基本知识并需要提示,那么最好从本章开始阅读。

1.1 RDBMS

数据就是信息。数据库是保存数据的地方,数据库可以存储三种不同的数据类型:

1. 普通数据——包括数值、日期和字符串,例如姓名和地址。
2. 复杂和大型对象——一个数据库可以存储和管理很多秘密数据类型(例如声音),地理数据(例如地图、图片、图形和视频)。
3. 新的用户数据——大多数现代数据库系统还允许用户存储他们自定义和操作的数据类型。

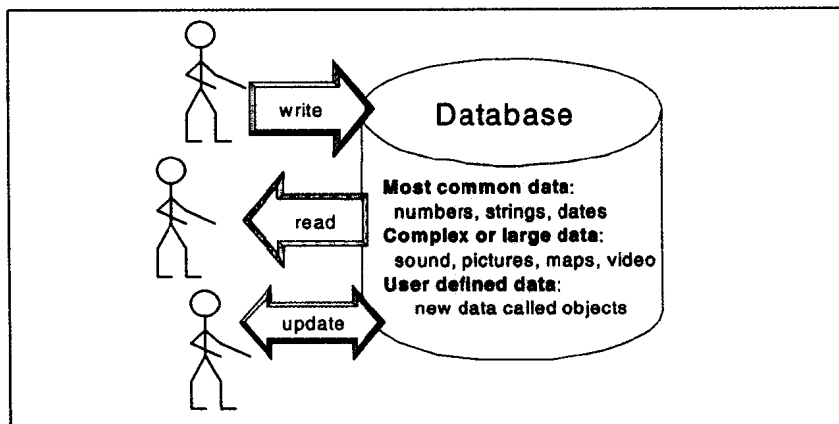


图 1.1 数据库可以写、读和修改很多数据类型

数据库必须允许用户:

- 写信息到数据库
- 检索以后的信息
- 修改数据

数据库必须能够以一种非常可靠的方式(否则不能将重要的数据存放在该数据库中)和高速地执行所有操作,它们还必须能够同时向很多人提供这些功能。作为另一种作用,一个数据库可以当作存放信息的公共位置,将一个公共视图提供给很多人。这样意味着,假如该数据库中包含零件存货,那么每个人都能看到可用的零件个数(只有一个正确的答案)。

虽然在一个现代数据库中存储各种数据类型，但大量的产品数据库还是用于存储简单的数值、字符串和日期记录。在本书中，我们主要考虑这些数据类型。

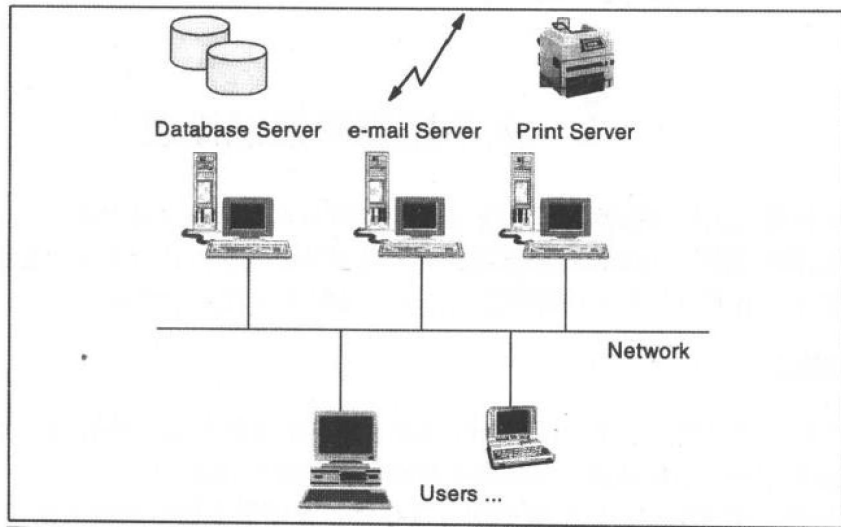


图 1.2 传统的使用多服务器的 UNIX 方法

在 UNIX 系统环境中，传统上每台 UNIX 机器有一种目的，例如一台机器作为 NFS 服务器，另一台机器作为打印机，再一台机器作为 E-mail 服务器(参见图 1.2)。这样做的一些理由如下：

- 在早期，UNIX 机器的功能不像其他系统(例如，IBM 主机)那样强大，因此，给工作负载指定一个服务器最大化了该工作负载的计算机功能。
- UNIX 在网络上强大的，使得客户服务器系统更容易实现，因此，采用这种方式分离工作负载对于 UNIX 来说是自然的。
- 在不同的机器上分担工作负载避免了应用之间的干扰。例如，如果两个计算集中(或 I/O 集中式)应用在一台机器资源上，则这两个应用的性能都会受到影响。
- 单台 UNIX 机器上支持的磁盘个数是有限的。

传统地，RDBMS 运行在指定的 UNIX 服务器上。本书假设大多数数据库运行在指定的机器上。这意味着该机器可以调整为最大数据库性能而不影响其他工作负载类型。在最近两年内，UNIX 机器的功能已经变得十分强大(采纳了很多主机设计特性)。这导致在大型 SMP UNIX 机器上可以运行多个工作负载，但通过逻辑或物理分区减少了工作负载之间的 CPU、内存或 I/O 资源之间的竞争。这意味着这些机器的运行把很多较小的数据库服务器连接在一起一样。

注意：在本书中，我们假设 RDBMS 运行在指定的机器上。因此，系统的性能调整不必考虑其他应用或工作负载。

本书是关于关系数据库管理系统(RDBMS)的。“关系”一词的含义是什么呢?在简单的数据库中存在很多记录，应用可以添加更多的记录到数据库中。这种操作称为 write, put, save 或 add a record，它们都指的是同一事情。这些记录也可以在后面检索，这种操作称为

retrieve, read, fetch 或 get a record。图 1.3 说明了如何从非关系数据库中读取数据的。为了指出检索哪些记录, 该应用通知该数据库如何查看记录(访问方法, 例如, 以 last name 排序的顾客记录)和记录标识(一个编号或姓名)。如果这个记录包含其他数据(例如, 第一个记录是一个雇员, 它包含部门编号), 那么该应用包含用于访问这些其他记录的代码。这要建立访问方法, 然后读该部门的记录以查找该部门名称、地址、主管等等。这有时称为 one row at a time(一次一行)应用, 这样做的效率很低。

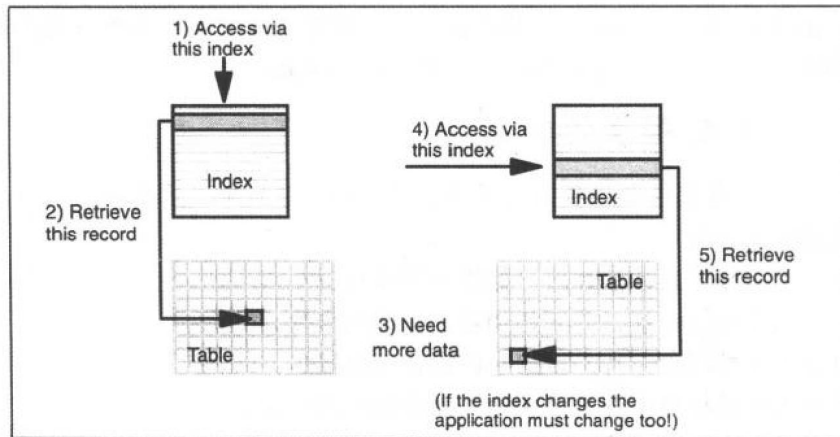


图 1.3 非关系的数据库方法

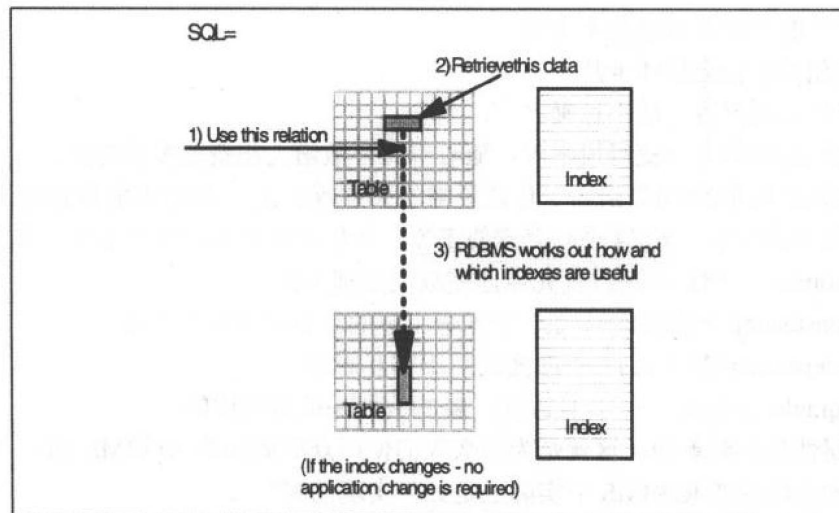


图 1.4 关系数据库方法

另一种方法是 RDBMS 使用的关系方法。数据库表示数据之间的关系, 在单个操作中使用这种关系提取所需要的数据。图 1.4 说明了如何访问 RDBMS 数据, 它显示了该应用同时从两个表中读取两个记录并只取回需要的相关信息。RDBMS 使用该关系:

- 存储产生两个请求的应用。