

Oracle Design



Oracle设计

O'REILLY®
中国电力出版社

Dave Ensor & Ian Stevenson 著

林琪 王宇 译

Oracle 设计

Dave Ensor & Ian Stevenson 著

林琪 王宇 译

O'REILLY®

Beijing • Cambridge • Farnham • Köln • Paris • Sebastopol • Taipei • Tokyo

O'Reilly & Associates, Inc. 授权中国电力出版社出版

中国电力出版社

图书在版编目 (CIP) 数据

Oracle 设计 / (英) 恩索尔 (Ensor, D.), 史蒂文森 (Stevenson, I.) 著; 林琪, 王宇译. - 北京: 中国电力出版社, 2002.9

书名原文: Oracle Design

ISBN 7-5083-1112-4

I. O... II. ①恩... ②史... ③林... ④王... III. 关系数据库-数据库管理系统, Oracle - 设计 IV. TP311.138

中国版本图书馆 CIP 数据核字 (2002) 第 053058 号

北京市版权局著作权合同登记

图字: 01-2002-1203 号

©1997 by O'Reilly & Associates, Inc.

Simplified Chinese Edition, jointly published by O'Reilly & Associates, Inc. and China Electric Power Press, 2002. Authorized translation of the English edition, 1997 O'Reilly & Associates, Inc., the owner of all rights to publish and sell the same.

All rights reserved including the rights of reproduction in whole or in part in any form.

英文原版由 O'Reilly & Associates, Inc. 出版 1997。

简体中文版由中国电力出版社出版 2002。英文原版的翻译得到 O'Reilly & Associates, Inc. 的授权。此简体中文版的出版和销售得到出版权和销售权的所有者——O'Reilly & Associates, Inc. 的许可。

版权所有。未得书面许可, 本书的任何部分和全部不得以任何形式重制。

书 名 / Oracle 设计

书 号 / ISBN 7-5083-1112-4

责任编辑 / 宋宏

封面设计 / Edie Freedman、张健

出版发行 / 中国电力出版社 (www.infopower.com.cn)

地 址 / 北京三里河路 6 号 (邮政编码 100044)

经 销 / 全国新华书店

印 刷 / 北京市地矿印刷厂

开 本 / 787 毫米 × 1092 毫米 16 开本 33.5 印张 493 千字

版 次 / 2002 年 9 月第一版 2002 年 9 月第一次印刷

印 数 / 0001-5000 册

定 价 / 59.00 元 (册)

O'Reilly & Associates 公司介绍

为了满足读者对网络和软件技术知识的迫切需求,世界著名计算机图书出版机构 O'Reilly & Associates 公司授权中国电力出版社,翻译出版一批该公司久负盛名的英文经典技术专著。

O'Reilly & Associates 公司是世界上在 UNIX、X、Internet 和其他开放系统图书领域具有领导地位的出版公司,同时是联机出版的先锋。

从最畅销的《The Whole Internet User's Guide & Catalog》(被纽约公共图书馆评为二十世纪最重要的 50 本书之一)到 GNN(最早的 Internet 门户和商业网站),再到 WebSite(第一个桌面 PC 的 Web 服务器软件),O'Reilly & Associates 一直处于 Internet 发展的最前沿。

许多书店的反馈表明,O'Reilly & Associates 是最稳定的计算机图书出版商——每一本书都一版再版。与大多数计算机图书出版商相比,O'Reilly & Associates 公司具有深厚的计算机专业背景,这使得 O'Reilly & Associates 形成了一个非常不同于其他出版商的出版方针。O'Reilly & Associates 所有的编辑人员以前都是程序员,或者是顶尖级的技术专家。O'Reilly & Associates 还有许多固定的作者群体——他们本身是相关领域的技术专家、咨询专家,而现在编写著作,O'Reilly & Associates 依靠他们及时地推出图书。因为 O'Reilly & Associates 紧密地与计算机业界联系着,所以 O'Reilly & Associates 知道市场上真正需要什么图书。

目录

前言	1
----------	---

第一部分 开始 Oracle 设计

第一章 简介	13
--------------	----

什么是设计?	13
一个案例研究	14
寻找替代的方法和途径	25
制定设计阶段的计划	29
设计的任务	30

第二章 Oracle 设计的重要性	46
-------------------------	----

特殊体系结构的设计	46
性能设计	51
其他设计考虑	53
Oracle 的版本演变及其对设计的影响	55
小结	72

第三章 数据建模	73
模型的类型	74
什么是数据建模?	76
实体关系图	78
对概念信息模型的质量保证	82
实体生命历史和数据流图	106
数据驱动的设计和元模型	110

第二部分 数据库设计

第四章 决定何时反向规范化	115
反向规范化: 是什么, 为什么, 什么时候?	115
实现反向规范化的方法	119
其他类型的反向规范化	121

第五章 选择数据类型和 null	127
可用的数据类型	128
数字数据	129
日期和时间(时间)数据	132
字符串数据	134
非结构化的数据和 BLOB	137
其他数据类型	138
null 值	140

第六章 选择键和索引	146
主键	147
其他键	156
索引的概述	158
索引类型和建立索引的技术	162

第七章 时间数据的处理	179
时间数据的问题	179
时间数据库的设计问题	183
时间数据：小结	204
 第八章 加载和卸载数据	 205
与外部系统的交互	205
数据兼容性问题	208
数据迁移步骤	215
数据转换	217
文件格式	221
排序、恢复和提交的频率	223
使用 SQL*Loader	223
从 Oracle 到 Oracle：一个特殊的案例	226
引出数据	227
 第九章 如何放置和存储对象	 230
指定对象的放置位置	230
规划大小	233
指定存储参数	235
创建脚本	238
实现计划	238
 第十章 保护数据	 240
存档	241
审核	248
安全	259
备份	270

第三部分 特殊体系结构的设计

第十一章 客户 / 服务器的设计	283
为什么使用客户 / 服务器?	284
什么是客户 / 服务器?	285
客户 / 服务器采用哪种类型的硬件?	286
客户 / 服务器设计的基本问题	288
客户 / 服务器的设计	289
SQL*Net 基础	292
一个 SQL 语句的剖析	293
从中间划分	296
特殊的设计考虑	298
 第十二章 分布式数据库的设计	 314
什么时候应该使用分布式数据库?	314
Oracle 对数据分布支持的演变过程	316
选择一个数据分布策略	328
实例分析	343
使用分布式数据库以便回退	347
其他设计考虑	347
对数据分布的小结	348
 第十三章 数据仓库的设计	 350
为什么要使用数据仓库?	350
什么是数据仓库?	352
数据仓库的设计问题	362
提取和加载数据	373
元数据	378
转换类型和转换方法	379
通过数据挖掘和 OLAP 工具检索数据	381

第十四章 并行处理的设计	383
为什么需要并行机制?	384
并行机制的设计实现	389
并行查询选项	395
Oracle 并行服务器	398

第四部分 代码模块的设计

第十五章 代码设计介绍	407
分析阶段提交的内容	408
将功能映射为模块	411
不要忘记系统模块	413
源代码与版本控制	416
模板代码	419
测试的设计	420
运行单元设计中的 CASE 产品	422
第十六章 如何划分处理逻辑	424
数据规则、处理规则和接口规则	424
逻辑的放置	430
锁定问题	432
处理逻辑的划分与三层体系结构有何关系?	434
第十七章 度量、原型和规范	437
开发模块的设计和建立度量	437
消除很大的模块	440
我们要设计原型吗?	442
如何建立模块规范? 关于建立模块规范的指导方针	447
详细说明屏幕和报表	449
详细说明批处理	450

第十八章 锁定	451
锁定策略	451
死锁	454
可串行性	455
其他设计问题	456
 第十九章 选择工具集	 460
工具的类型	460
哪些选择标准是重要的?	461
客户/服务器工具	463
WWW 的设计	466
 第二十章 屏幕、报表和批处理程序等的设计	 467
屏幕设计	467
报表设计	470
批处理程序的设计	472
错误处理	481
导航	483
在线帮助	485
 第五部分 附录	
 附录一 第三方提供的开发包	 489
 附录二 设计的窍门	 499
 词汇表	 515

前言

为什么要写这本书呢？因为我们强烈地感到大家低估了Oracle设计的重要性，况且目前也迫切需要一本深入探讨这方面内容的书。除了这些想法，写这本书时我们还有其他一些考虑。

Ian在参与了Oracle6和7的几个项目的设计阶段之后，萌生了写这本书的念头。他注意到许多同样的问题和困难经常会一次次地出现。尽管这些问题很类似，但是其解决方法却会因为项目的不同而有所差异。显然，对于具体的解决方法不太可能简单地断言孰是孰非。事实上，就特定的项目而言，每个解决方法都可能是上策。所以他力图建立一个信息库，借此将做出特定设计决策的理由记录下来。这样，如果将来他（或者其他人）碰到类似的问题时，就能够从信息库中获得帮助。不难相信，这本书将成为Oracle设计领域的一个绝好媒介，大家可以由此在Oracle社区共享一些难得的设计经验。

Dave多年以来一直任Oracle全球性能研究组（Oracle's Worldwide Performance Studies Group）的负责人。在实践中，他不只一次地发现Oracle性能问题总是源自相似的根本原因。它们往往并非是由于构造技术上存在着缺陷，也不是由Oracle性能调整造成的。真正的原因通常是项目小组的设计上存在着错误。不过也不能否认“传统”因素确实也会对性能造成影响。但是很明显，正是由于设计决策有问题，许多项目根本无法提供较好的性能。基于这样的背景，Dave在一个研讨会上为所有开发人员准备了一篇题为“Designing for Oracle7”的发言稿，尽管它很长，但是仍

不足以提供完全可靠且面面俱到的设计方案。和Ian一样，Dave在撰写发言稿的时候，也发现如果要评判设计方案，就必须根据项目的实际需求和其所处的具体技术环境来决定。

假如没有固定的规则可依，又为什么要写一本这方面的书呢？还有为什么目前市面上关于设计方面的书这么少呢？首先回答第二个问题，我们认为，由于无法给问题提供直截了当的答案（即使是最简单的问题，比如“每个表都必须要有有一个主键吗？”），便使得许多作者和出版商退缩了，因为他们认为读者只愿意接受对所有问题都能给予绝对答案的技术资料。其实事实并非如此。

我们写这本书有两个绝好的理由：

1. 在参与的所有项目中，对那些解决起来希望很渺茫的问题，我们都曾付出大量时间探索其解决方案，但是无论我们如何千方百计地努力，也永远无法解决这些问题。如果我们知道这些，就可以在一开始时避免这些问题，以节省宝贵的时间。
2. 此外我们发现，随着经验的逐步积累，即使能从以前的项目设计技术中汲取一些影响系统性能的粗略方法，也是难能可贵的。那些经验丰富的设计人员，有时也难免会忽略某个特殊问题的解决方法。因此，倘若有些其他类型的、即使并非是绝对的参考资料，对于探讨不同的设计方案也可能会是一种价值不菲的工具。

写此书的初衷就是要给大家（那些和我们一样的人）提供一个日常使用的工具。我们非常乐意与你分享以前的经验，虽然有时很困难，但它们确实是有益的经验。

关系数据库的设计问题非常棘手。即使看起来相对简单的业务需求，说不定也会引出复杂的设计问题。当你仔细斟酌各种满足业务需求的设计方案时，常常转了一圈又会回到最初的选择，此时你可能已经记不清这个选择有何缺陷。但却切实感到问题没有取得任何进展，于是有点儿灰心丧气。在这里我们要告诉你，即使有本书作为参考，上述情况也在所难免。从事设计工作的人都曾有过同样的经历，你一定要相信自己最终能够冲破黑暗，走向光明。

本书的目的是帮助你工作得更潇洒、更轻松。若能把这些经验真正应用到你的设计中，必将收获丰硕的果实。

为什么针对 Oracle?

出版商所问的第一个问题就是：“为什么是一本关于 Oracle 设计的手册？为什么不针对所有的关系数据库？Oracle 有什么不同吗？”这的确是一个很重要的问题，我们也很愿意回答。

包含计算机应用程序在内的所有设计，都要充分利用所提供的技术和资源。当选择某个设计方案时，必须充分考虑到 Oracle7 所具有的许多特殊的、惟一的技术特点。本书用很大的篇幅来讨论那些特点。在多数情况下，我们建议你应该首先在自己的环境中进行测试，然后再选择某个特殊的设计方案。

我们不敢妄断这本书会使你一夜之间成为一名成功的 Oracle 设计人员。因为没有什么可以代替设计经验。那些出色的设计人员总是在不断地学习进取。在本书中，我们只给出了一个更加结构化、更容易控制的设计环境框架。针对主要的设计决策，每一章都给出了相应的替代方案，而且还强调了该设计的优点和缺陷。当然，最终的决定权在于你，我们只希望你能够三思而后行。

目前 Oracle 已经是一个成熟的产品。从最初发布这个产品开始，Oracle 公司就开始从应用程序（或用户界面）中提取业务规则的智能因素，将其纳入数据库（或服务器）中。这一变革意味着在同样的情况下，采用 Oracle7 的项目比采用 Oracle6 或 3GL（第三代程序语言）开发的项目需要花更多的时间进行数据库设计。如果是一个优秀的设计方案，创建应用程序的任务会更加简单。一旦应用程序的建立阶段已经开始，就很难再修改设计阶段做出的任何错误决定或者疏漏，而且会为此付出高昂的代价。

一个差的设计方案会令开发周期在长时间内困难重重，最终系统的质量也很低劣。事实上，如果方案设计得太差，就很难形成一个最终的系统，即使能够成系统，也是一个相当拙劣的产品。设计一个系统时，总要考虑它的性能和可用性。如果你被告知响应时间无关紧要或者“以后再修正响应时间”，那我们建议此时一定要进行争辩。在产品系统中，修正一个设计缺陷或错误所付出的代价非常大。应当抱着“极其苛刻”的态度解决性能设计问题。在我们的印象里，即便是一个低端系统，其修正也无法得到准确无误的保证。

还要记住的是，最终用户对任何应用程序的感觉，通常决定于他们第一周内使用该系统的印象。新的计算机系统的性能、用户应答时间、易用性和健壮性，直接决定了用户最终能否接受这个系统。

总而言之，我们在任何设计过程中都要达到以下目标，这些内容也是本书力求帮你做到的。贯穿系统运行的整个过程中，要求的性能指标包括：

- 功能性和适应性
- 吞吐量
- 响应时间
- 可用性
- 易用性
- 安全性

本书的组织结构

本书所述内容共分五部分。虽然我们希望你能通篇逐字地阅读，但我们也知道，有些读者不想从前往后一字不差地翻一遍；即使要看大部分，也不愿意按照顺序进行。为了满足大家的需要，我们尽量使每章内容相对独立，减少交叉引用的次数。那些经验丰富的设计人员、具有良好的项目生命周期管理经验的管理人员，可以跳过前面几章（主要是第一章和第二章）。使用过 Oracle 最新版本的人也可以跳过第二章。许多描述特殊技术的章节，也只是在具体设计时才需要细细研读。分布式数据库（第十二章）和数据仓库（第十三章）专门介绍了这些特殊的技术。

第一部分：开始 Oracle 设计

- 第一章，“简介”，介绍项目开发的各个阶段，特别是策略制定阶段、分析阶段、设计阶段和建立阶段；本章采用了一个真实项目的简化例子，来解释各个项目阶段及其需要交付的内容。
- 第二章，“Oracle 设计的重要性”，针对 Oracle，描述了特殊的设计问题，并逐条说明 Oracle7 的重要功能；对 Oracle8 及以上版本做了一个简单介绍。

- 第三章,“数据建模”,详细考查数据建模;定义典型的关系数据库术语,如实体、关系和第三范式(3NF);逐条说明项目分析人员应当向设计人员提交哪些内容,才能将概念数据模型转化为逻辑数据模型。

第二部分: 数据库设计

- 第四章,“决定何时反向规范化”,为提高性能,针对数据库数据,介绍一些特殊的反向规范化技术。
- 第五章,“选择数据类型和null”,介绍各种Oracle数据类型,描述如何为数据库中的对象选择最恰当的数据类型,并探讨那些最令人头痛的问题,包括什么是null,如何对待null。
- 第六章,“选择键和索引”,讨论如何为特定的数据库选择最合适的键。
- 第七章,“时间数据的处理”,探讨一个与Oracle以及其他关系数据库都有关系的特殊问题——关系数据库没有充分支持时间序列(时间)数据。本章提供了大量特定于Oracle的技术,可以使用它们来克服时间数据所造成的限制。
- 第八章,“加载和卸载数据”,研究了从外部数据源生成Oracle数据库的各种方法。还讨论了如何从Oracle7中提取数据。
- 第九章,“如何放置和存储对象”,了解一些重要的数据库物理设计方法,比如数据库大小和文件放置。
- 第十章,“保护数据”,涉及备份、存档、审核和安全等主题。

第三部分: 特殊体系结构的设计

- 第十一章,“客户/服务器的设计”,采用Oracle设计方法设计客户/服务器模型。分析了各种优化性能和可用性的分布式处理技术。
- 第十二章,“分布式数据库的设计”,进一步探讨数据分布模型,介绍为什么以及如何使用分布式数据库。分析了Oracle提供的各种可用选项,以及如何在不同的方案中使用它们。
- 第十三章,“数据仓库的设计”,掌握如何建立和维护数据仓库。如何进行空间建模,如何将数据保存到数据仓库中,以及如何再将它们提取出来。

- 第十四章,“并行处理的设计”,描述并行处理的概念,Oracle的Parallel Query Option (并行查询选项)和Oracle Parallel Server (Oracle 并行服务器)在实际应用程序中的应用,探讨磁盘分段处理和RAID 等技术。

第四部分: 代码模块的设计

- 第十五章,“代码设计介绍”,描述代码模块设计的基本概念。
- 第十六章,“如何划分处理逻辑”,讲述一种划分应用程序逻辑的新方法。
- 第十七章,“度量、原型和规范”,描述代码设计的正规方法;特别要集中探讨如何确保模块满足需求。
- 第十八章,“锁定”,包含一些有用信息,能够帮助你最小化应用程序中的争议。
- 第十九章,“选择工具集”,比较各种支持 Oracle 后端的前端产品的优缺点。
- 第二十章,“屏幕、报表和批处理程序等的设计”,解决有关屏幕、报表、批处理程序、错误处理、导航和在线帮助等特殊的设计问题。

第五部分: 附录

- 附录一,“第三方提供的开发包”,它比较了购买一个打包的解决方案与完全从头开发应用程序的各自优点。
- 附录二,“设计的窍门”,集中讨论两个特殊的设计技术。我们提出一个方法来避免恼人的 Oracle 触发器“变异表”问题;简要介绍了扩展的 SQL。

本书的读者

本书认为你已经具备 Oracle 方面的知识,但是对于那些需要转到 Oracle,并具有其他关系数据库管理系统(RDBMS)使用经验的读者,本书也会给予他们很多帮助。无论你目前具备的是开发经验还是数据库管理经验都没关系,应该先熟悉本书中提到的相关术语和概念。我们还假设你具有项目开发生命期的基本知识。

本书怎样帮你成为一名设计人员呢?它并没有回答所有方面的问题,只是提供一个设计框架和一张常用的设计问题检查表。希望你在进行某项特殊设计任务时,能求

助于本书。它会帮助你澄清大量的、无法考虑周到的问题；我们还希望本书能指导你解决设计问题，激发更深入的思考和更热烈的讨论，为你解决重要问题提供参考，让你的设计更加完美。

针对哪种 Oracle?

本书涉及的问题适用于 Oracle 的所有发行版本。有些极少数的例外情况，我们将明确指出所适用的版本。希望 Oracle 的新版本能为本书中所讨论的一些问题提供解决方法。无论如何，我们相信书中所提供的大部分内容会继续适用于后继的 Oracle 版本。

排版约定

下面是本书英文所使用的约定：

斜体字 (*Italic*)

用来表示文件、变量和函数名。有时也用来强调或突出第一次提及的关键术语。

黑体字 (**Bold**)

用来表示标题。

等宽字符 (Constant width)

用于代码示例。

大写

在代码示例中，表示 Oracle 关键字。

小写

在代码示例中，表示用户定义的条目，如变量和参数。

标点符号

在代码示例中，实际输入的与显示的相同。

缩排

在代码示例中，帮助显示代码的结构，但不是要求的。