

ERwin

数据建模

杨国强 路萍 张志军 等编著

CA 公司 监制

计算机专业人员书库

ERwin 数据建模

杨国强 路萍 张志军 等编著

CA公司 监制

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书结合数据库建模工具 ERwin, 以及笔者在这个领域的丰富经验, 全面介绍了数据库建模技术, 包括数据库理论、数据建模方法和工具操作。读者可以从本书中获得基本的数据库理论知识, 逐步熟悉并理解 ERwin 的使用方法, 掌握 ERwin 的高级功能, 充分发挥 ERwin 的作用, 以大大方便数据建模工作。全书分为 3 篇, 共 14 章。

本书主要的读者对象包括专业的数据库(包括数据仓库)分析和设计人员、ERwin 最终用户、相关专业的教师和学生等。

本书著作权归 CA 公司所有。未经许可, 不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有, 侵权必究。

图书在版编目(CIP)数据

ERWin 数据建模 / 杨国强等编著. —北京: 电子工业出版社, 2004.10

(计算机专业人员书库)

ISBN 7-121-00304-X

I. E… II. 杨… III. 数据库技术—建立模型 IV. TP311.13

中国版本图书馆 CIP 数据核字(2004)第 089552 号

责任编辑: 孙学瑛

印 刷: 北京智力达印刷有限公司

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

经 销: 各地新华书店

开 本: 787×980 1/16 印张: 28 字数: 509 千字

印 次: 2004 年 10 月第 1 次印刷

印 数: 4000 册 定价: 55.00 元

凡购买电子工业出版社的图书, 如有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系。联系电话: (010) 68279077。质量投诉请发邮件至 zlbs@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

序 一

在当今的市场中，企业要想保持竞争优势，就必须依据准确的数据来做出更快、更优的决策。妥善实施的建模环境不仅能够加快推出更加精确、及时的解决方案，而且为可选地捕捉、同步及使用企业的关键业务知识、信息资产和应用解决方案奠定了基础。因此，数据建模工具的选择，对于企业构建高性能的电子商务应用软件和数据库具有极为重要的意义。

在数据建模领域，CA 拥有世界级的产品——AllFusion ERwin Data Modeler。它可帮助企业轻而易举地设计、创建和维护数据库、数据仓库及企业数据资源模型。自推出以来，该产品已先后获得“最佳数据库工具”、“最佳设计和建模解决方案”等奖项，受到了用户的好评。我们相信，借助 ERwin 这一将管理与技术融为一体的先进建模手段，企业无疑将显著地增强自身的竞争能力。

当然，在 IT 应用高速发展的今天，仅仅有建模工具是不够的。因为不断创新的 IT 技术在为用户带来了新的、更好的产品的同时，也使 IT 应用变得更加复杂。这就促使客户需要以一种更加完整而综合的方式来管理其全部的技术基础设施，包括网络、存储和安全，从而更好地提高企业 IT 投资的效率，同时使 IT 投资更好地与企业的业务相配合。而这一切的实现，都要依赖于管理软件。

IT 计算正在进入一个崭新的时代，即管理软件时代。作为全球最大的管理软件公司，CA 为客户提供了广泛而全面的运营管理、安全管理、存储管理、应用生命周期管理和服务管理软件与服务。CA 的管理软件可以帮助企业发现尚未被有效利用的技术，并把它们充分利用起来，从而最大限度地发挥现有资源的能量，为企业带来更大的生产力。

多年来，CA 公司始终致力于不断提升和增强对管理软件的开发及应用，以满足客户不断增长的需求。CA 公司愿与各位客户和业界同仁一起，通过广泛的合作，推出更多、更优秀的管理软件，推动中国企业的信息化建设！

尹婉智

CA 公司中国区市场总监

序 二

对 ERwin（现在为 AllFusion ERwin Data Modeler）总有一种老朋友似的亲切感。记得最早在读书的时候，就在教科书上看到过 ERwin 的介绍，还依稀记得当时的任课老师用略带自豪的语气向我们讲述数据库设计工具——ERwin 的神奇之处（虽然许多功能现在已经是工业标准）。后来，进入软件行业，就开始真正地和 ERwin 面对面了。这些是许久之前的事了，但是对 ERwin 的感情却始终深厚。

一直以来，虽然越来越多的人使用 ERwin 进行数据库、数据仓库设计，但是真正好的资料却寥寥无几。仅有的一些资料要么不具有权威性，要么只是偏重工具的操作。现在非常高兴地看到这本关于 ERwin 的专业书籍的出版，虽然相对于国内用户和专业人士的企盼来说有点晚，但是相信此书丰富的内容、专业化的阐述、富于实践性的案例，以及设计理念和工具集合的阐述思路都会让大家觉得等待还是值得的。

ERwin 的全称是 AllFusion ERwin Data Modeler，是 CA 公司 AllFusion 品牌下的建模套件之一，用于数据库建模。CA 的建模套件还包括业务过程建模工具 AllFusion Process Modeler（原 BPwin）、组件建模工具 AllFusion Component Modeler、模型管理工具 AllFusion Model Manager、数据模型审验工具 AllFusion Data Model Validator 和模型访问工具 AllFusion Model Navigator。这些工具在本书中也有介绍。

ERwin 最早是 Logic Works 公司推出的数据建模工具，由于其优秀的功能被广大数据库相关工作者所青睐，同时也被大型公司看好，所以逐步被越来越大的公司收购，先是被 Platinum 收购，后来又被 Computer Associates 公司收购，收编在 AllFusion 品牌下，该品牌专注于应用生命周期管理。ERwin 至今仍是数据建模领域的佼佼者，在许多世界级的评比中屡获大奖。

在和客户关于 ERwin 的交流中经常会被问到两个问题。第一个问题是：“为什么要使用工具？”问这类问题的不是对技术不懂的人，而恰恰是对技术非常“自信”的所谓骨灰级的专家。他们习惯于凡事自己动手，操作系统要用命令行，解决问题要用 Debug，所以数据库管理当然全部是 SQL。但是，我们今天所面对的 IT 业已经是一个商业感很强的行业，不再是为了满足某些“天才”的荣誉感而运作的。因此，采用自动化的工具，以及行业标准的标记语言和设计方法是提高效率的惟一办法，也是大势所趋。同时，数据仓库的复杂性决定了：在数据仓库模型设计时，不可能采用手工的办法进行设计和实施，必须要借助自动化的工具作为辅助手段。

另一个被经常问到的问题是：“为什么 ERwin 能确立在业界中的位置？”这可以说有两个方面因素决定。

首先要从数据库的设计角度入手。关系型数据库自从 20 世纪 70 年代诞生至今，不断发展成熟，其建模技术也随之发展，并陆续产生了多种建模方法，如 E-R 图、IE、IDEF 系列方法等。比较引人注目的是，近年来发展很快的 UML 也将数据建模纳入其中，但迄今为止还是讨论与研究居多。在这些数据建模技术中，比较成熟并广为应用的方法当属 IDEF1x 和 IE。IDEF1x 是 IDEF 家族中的一员。IDEF 是由美国军方的标准演化过来的，它主要的特点是层次化和全面性，其中 IDEF0、IDEF1、IDEF3 使用最为广泛（在本书中上述三种都有提及，以 IDEF1x 为主）。遵照 IDEF 方法进行系统的分析和设计，实际上就遵循了行业的一些最佳实践方法，为整个系统的设计和分析的水平、质量提供最佳保障。ERwin 和 BPwin 就是采用这样的标记语言作为基础的，因此，它们不仅给使用者一个画图工具，还可以直接提供设计和分析上的帮助。

另一方面，IDEF1x 和 IE 数据库建模方法的广泛应用，本质上是由于这些方法的优越性，但是从某种程度上还应该归功于这些方法的支撑工具。ERwin 就是这些支撑工具中的一个佼佼者。ERwin 的若干优秀的功能已经成为行业其他工具效仿的标准，另外，对最广泛数据库支持的中立特性、强大的完全比较功能、高效率的正反向工程，都为 ERwin 的使用者带来直接帮助。

实际上，ERwin 在 CA 公司（Computer Associates）的全面数据管理、数据仓库解决方案中，仍然扮演着重要角色，也在不断地发展和完善。现在 ERwin 的版本是 R4，在未来的一年内，CA 计划将推出 R6 版本，主要将提高界面友好程度，提供对“企业数据架构”支持，支持“构件”重用，提供更加完善的完全比较功能，支持影响性分析（Impact Analysis），并且对新版本的数据库提供支持（包括 Oracle 9i/10g）。相信上述的新特性可以更加巩固 ERwin 市场占有率第一的位置。

本书是一本结合数据库建模工具 ERwin，全面介绍数据库建模的书籍，包括数据库理论、数据建模方法、工具操作，以及笔者在这个领域的经验。主要的读者对象包括专业的数据库（包括数据仓库）分析和设计人员、最终用户、教师、学生和研究人员等。

当然，书中提到的内容还有很多不完善的方面，希望读者提出宝贵意见。

牛晓强

CA 公司高级技术顾问

前 言

数据库建模技术对于应用软件开发来说具有重要的意义。大多数应用软件都依赖于数据库，因此，数据库设计、开发技术对于应用软件开发工作来说是不可或缺的。同时，应用软件规模越来越大，复杂程度越来越高，人们不得不依靠各种模型来对应用系统进行分解、分析，以求开发工作能够顺利进行。作为开发工作的一部分，数据库的分析设计工作也同样离不开建模技术。

本书结合数据库建模工具 ERwin，全面介绍了数据库建模，包括数据库理论、数据建模方法、工具操作，以及本书编者在这个领域中的经验。

本书的编者多年从事软件工程领域中的工作，曾经研究过多种 CASE (Computer Aided Software Engineering) 工具，尤其是在数据建模方面。在对这些工具研究的过程中，编者深深体会到了一些优秀工具对于软件开发组织和相关工作者的重要性，非常希望能够向大家推荐这些工具，并分享自己的经验，于是经过精心准备，推出这本书，希望对业内人士有所帮助。

本书不是讲述数据库原理的教材，但你可以从中获得基本的数据库理论知识；它也不是 ERwin 的使用手册，但你可以按照本书的指导逐步熟悉并掌握 ERwin 的使用方法，掌握 ERwin 的高级功能，充分发挥 ERwin 的作用，大大方便数据建模工作；更重要的是，本书融合了编者数据建模及工具使用的经验，能够让新用户快速掌握工具、少走弯路，对于老用户来说，本书也是一本很有价值的参考用书。

本书第 1、7、9、10、11、13 章由杨国强编写，第 2、3、4、5、6、8、12、14 章由路萍、张志军编写。全书由杨国强统稿。由于编者水平和时间有限，难免会有疏漏之处，欢迎读者批评指正。

本书的约定

为了便于读者的理解，这里简单介绍一下本书的格式约定。

- 菜单描述：用“→”表示菜单的级联关系。例如，主菜单→子菜单→下一级子菜单。
- 术语解释：除了附录中的术语表之外，对于一些不太常用的英文术语或者缩写，在第一次出现或者上下文中第一次提到的地方，直接将解释附在后面的括号中。比如，Domain (类型域)。

- 对于插图中特别需要注意的地方，用椭圆形或者线条进行标注，这些线条并非工具中所显示的。比如，图 0-1 中的椭圆就是为了标注“Create Entity”这个按钮。

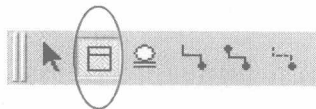


图 0-1 插图标注举例

- 所有操作均按操作步骤编号，其中的插图都是按顺序操作产生的，所以如果你对工具不是太熟悉的话，请从第一步开始操作。

目 录

入 门 篇

第 1 章 数据建模与 ERwin	2
1.1 ERwin 历史和现状	3
1.2 ERwin 入门	4
第 2 章 数据建模方法详解	9
2.1 数据建模的背景	9
2.1.1 三层模式的概念	9
2.1.2 数据建模的益处	12
2.2 IDEF1x 数据建模方法论	13
2.2.1 逻辑模型	14
2.2.2 物理模型	15
2.3 构建逻辑模型	16
2.3.1 定义实体和属性	16
2.3.2 关系	17
2.3.3 多对多关系	18
2.3.4 确认逻辑模型的设计	19
2.4 设计基于键的模型	21
2.4.1 键	21
2.4.2 关系和外键	24
2.5 命名和定义实体属性	28
2.5.1 命名实体和属性	28
2.5.2 同义词、同名异物和别名	29
2.5.3 实体定义	30
2.5.4 属性定义	30
2.5.5 角色名	31
2.5.6 定义业务规则	33

2.6	精化数据模型	33
2.6.1	关系基数	33
2.6.2	特殊关系	36
2.7	案例介绍	51
2.7.1	背景信息	51
2.7.2	技术支持业务问题陈述	52

实 践 篇

第3章	ERwin 的工作空间	56
3.1	概述	56
3.2	ERwin 绘图区	57
3.2.1	模型的类型	57
3.2.2	选择建模语言符号	57
3.2.3	改变符号	58
3.3	ERwin 的菜单和工具条	58
	建模语言符号工具条	60
3.4	模型导航器	60
3.5	使用帮助	61
第4章	构建逻辑模型	63
4.1	创建实体-关系图	63
4.1.1	新建模型	64
4.1.2	用 ERwin 创建实体	66
4.1.3	编辑实体定义	68
4.1.4	用 ERwin 创建实体间的关系	70
4.2	基于键的模型	73
4.2.1	创建主键属性	73
4.2.2	创建非主键属性	75
4.2.3	键组对话框	80
4.2.4	创建候选键	80
4.2.5	创建查询项	83

4.2.6	分析关系类型	84
4.2.7	关系的描述	86
4.2.8	外键	88
4.2.9	使用关系角色名	89
4.2.10	解决多对多关系	93
4.3	全属性模型	97
4.3.1	创建其他非键属性	97
4.3.2	编辑属性定义	99
4.3.3	创建分类关系	106
4.3.4	精化模型	111
第 5 章	构建物理模型	113
5.1	从逻辑模型到物理模型	113
5.1.1	构建物理模型	114
5.1.2	选择目标 DBMS	120
5.2	物理表设计	122
5.2.1	反规范化	122
5.2.2	物理字段设计	145
5.2.3	表的其他属性	157
5.3	键和索引	162
5.3.1	键	162
5.3.2	索引	163
5.4	规划数据库容量	168
5.5	数据库视图	170
5.6	其他数据库对象	180
5.6.1	物理存储对象	180
5.6.2	存储过程	181
5.6.3	触发器	185
第 6 章	构建维模型	191
6.1	数据仓库简介	191
6.1.1	数据仓库发展历史	191

6.1.2	什么是数据仓库	192
6.1.3	OLTP 与 OLAP	193
6.1.4	从 OLTP 到 OLAP	194
6.2	维建模	196
6.2.1	维建模简介	196
6.2.2	维建模与 E-R 建模	197
6.2.3	维建模的基本组成	199
6.2.4	维建模的优点	202
6.3	维建模示例	203
6.3.1	新建模型	203
6.3.2	创建矩阵	206
6.3.3	设计事实表	207
6.3.4	在 ERwin 中创建维和事实表	208
6.3.5	在 ERwin 中定义数据仓库规则	213
第 7 章	模型报告	219
7.1	使用 Data Browser	219
7.1.1	基本操作	219
7.1.2	共享报告	224
7.1.3	利用 Data Browser 审核模型	226
7.2	使用 Report Templates Builder	229
7.2.1	使用模板生成报告	229
7.2.2	创建报告模板	230
7.3	利用宏生成简单文本报告	235
第 8 章	双向工程	237
8.1	正向工程	237
8.1.1	什么是正向工程	237
8.1.2	执行正向工程	238
8.1.3	小结	243
8.2	反向工程	244
8.2.1	什么是反向工程	244

8.2.2	执行脚本文件的反向工程	245
8.2.3	执行数据库反向工程	248
8.3	完全比较	249
8.3.1	什么是完全比较	249
8.3.2	执行完全比较	250
8.4	创建 Quill 技术支持系统数据库	255
8.4.1	生成数据库脚本	255
8.4.2	对脚本执行反向工程	258
8.4.3	比较模型和数据库脚本	261

提 高 篇

第 9 章	规范化建模	266
9.1	规范化的几个方面	266
9.1.1	模型模板	267
9.1.2	术语表与命名标准文件	271
9.1.3	命名语法	275
9.1.4	命名选项和命名映射	279
9.1.5	类型域	281
9.1.6	DBMS 之间的数据类型映射	284
9.2	规范方法的综合利用	286
第 10 章	模型的审验	288
10.1	审验应考虑的因素	288
10.2	审验工具 Data Model Validator	289
10.2.1	Data Model Validator 入门初体验	290
10.2.2	典型问题	294
第 11 章	深入了解 ERwin: 使用技巧和高级功能	318
11.1	操作技巧	318
11.1.1	模型图形的美化	318
11.1.2	可存储的显示方式	318
11.1.3	存储显示方式的操作	319

11.1.4	用主题域划分模型	322
11.1.5	主题域操作	323
11.2	用户自定义特性	325
11.2.1	利用 UDP 为接口表自动授权	326
11.3	宏	328
11.3.1	可以使用宏的地方	328
11.3.2	宏的基本元素及示例	330
11.3.3	宏的综合示例	337
11.4	使用 API	339
11.5	使用 Pre-Script 和 Post Script	342
11.6	触发器的定义与覆盖关系	345
第 12 章	业务过程建模	355
12.1	业务建模简介	355
12.1.1	活动和过程模型	355
12.1.2	业务建模和商业系统工程	357
12.2	业务过程建模方法	358
12.2.1	业务流程建模 (IDEF0)	358
12.2.2	数据流图 (DFD)	361
12.2.3	工作流建模 (IDEF3)	362
12.2.4	三种业务建模方法的结合使用	363
12.3	业务建模工具 Process Modeler	364
12.3.1	Process Modeler 工作空间	365
12.3.2	创建 IDEF0 模型	366
12.3.3	揭示业务信息流	375
12.3.4	绘制业务的地图	381
第 13 章	使用模型管理器进行团队建模	386
13.1	团队化建模中的协作问题	386
13.2	ModelMart ABC: 模型的打开、保存	388
13.3	ModelMart 库结构的组织	391
13.4	安全权限管理	392

13.4.1	安全相关对象	392
13.4.2	安全设置的认识和操作	395
13.4.3	Library 结构与权限设置实例分析	396
13.5	处理模型的更改	398
13.5.1	审查更改	399
13.5.2	冲突解决	400
13.5.3	刷新	402
13.5.4	锁定模型	403
13.6	模型版本维护	404
13.6.1	Model Manager 的模型版本管理功能	404
13.6.2	Model Manager 中的几种版本	404
13.6.3	常用的几种版本操作	406
13.6.4	版本策略	409
13.7	用 Model Manager 管理模型标准	409
第 14 章	建模最佳实践	411
14.1	模型驱动开发	411
14.2	建模实践	413
14.2.1	选择合适的模型	413
14.2.2	团队协作建模	413
14.2.3	选择合适的工具	414
14.3	常用建模技术	415
附录 A	整合数据资产, 促进业务开展	417
附录 B	应用案例	420
ERwin	软件试用申请表	424

入门篇

ERwin 数据建模

- 数据建模与 ERwin
- 数据建模方法详解

第 1 章 数据建模与 ERwin

信息技术几乎已经渗透到现代社会的每一个角落，现代人生活的每一天几乎都要直接或者间接地使用各种计算机应用：打电话，使用了电话机和交换机中的软件；刷卡消费，使用了 POS 机及银行的交易系统；更不用说日常工作，会使用到各种各样的办公系统。作为这些软件的普通使用者，自然不关心软件是如何构建的；而作为一个软件开发人员，应该知道所有软件都有一个不可或缺的基础——数据，而大部分软件的数据存储、访问等操作都是依赖于目前主流的数据库技术——关系型数据库。因此，关系型数据库的开发技术对于大多数应用软件来说都是至关重要的。

有一种说法是：软件是迄今为止人类能够构造的最为复杂的产品。姑且不论这种说法合适与否，单从有人提出这种说法这一点来看，软件的复杂程度也可见一斑。构造如此复杂的系统显然是非常困难的，人们使用了很多方法来降低这种复杂度，建模技术便是其中的一种。

建模技术就是借助模型来分析、设计应用系统的技术。模型是现实世界中的某些事物的一种抽象表示。抽象是指抽取事物的本质特性，而忽略事物的其他次要因素。模型是理解、分析、开发或改造事物原型的一种常用手段。比如说，建筑物的蓝图就是高楼大厦的模型，建筑师利用蓝图对建筑物分析和设计，以便在动工前就能有一个完整而清晰的掌握，并可以据此对建筑物进行评审，及早发现问题并改进，而不必等到建成之后再补救。

建模技术还可以把系统的复杂程度降低到人脑可以处理的程度，并用可视化的方式来展示复杂的系统，给人以形象的思维，帮助人们更好地认知和分析。同样，对于数据库的设计而言，建模技术也有同样的好处。

关系型数据库自从占据主导地位以来，似乎还没有其他类型的数据库能够撼动它的位置。近年来比较热门的多媒体数据库、多维数据库等，现在看来都只是关系型数据库的一种补充，而绝不是替代。由于处在主导地位，对关系型数据库建模技术的研究也最为充分、最为成熟，应用也最为广泛，所以相关的建模支撑工具也比较多，比如 Oracle 的 Designer，Sybase 的 Power Designer，Embarcadero 的 ERStudio 等。这些工具各有特色，而本书将要介绍的是业界非常著名的数据库建模工具——ERwin。