

IBM信息管理开发部总经理吉燕勇推荐序

IBM DB2开发部资深经理干毅民推荐序

IBM信息管理首席架构师陈奇博士任技术顾问



DB2 设计与性能优化

—— 原理、方法与实践

王飞鹏 陈辉 张广舟 成孜论 编著



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>

· 舞动DB2系列 ·



DB2设计与性能优化

——原理、方法与实践

王飞鹏 陈辉 张广舟 成孜论 编著

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书原汁原味地展示了 DB2 设计和优化技术,深入剖析了 DB2 的工作原理。全书分为 4 部分,共 11 章。

第一部分解析数据库性能的本质问题。第二部分从设计高性能数据库的角度,详细介绍如何使用 DB2 数据库设计高质量的应用系统,内容包括 DB2 数据库逻辑设计、DB2 数据库物理设计、DB2 pureXML 数据库设计及 DB2 工作负载设计。第三部分介绍如何优化 DB2 数据库,内容包括性能监控及参数调整、SQL 语句优化策略、锁和日志优化等。第四部分阐述 DB2 高级性能加速技术,主要包括 DB2 pureScale 集群和 solidDB 高速缓存加速。

附录部分列出了常用监控命令用法和缩略语释义,还特别提供了 PAT 树的放大图,供读者进行数据库性能优化时使用。

本书适合 DB2 数据库设计人员、DBA、数据库性能分析人员、数据库开发人员、运维人员及应用开发工程师阅读和参考,也可用做高校相关专业或培训班的教材。

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

DB2 设计与性能优化:原理、方法与实践 / 王飞鹏等编著. —北京:电子工业出版社, 2011.4

(舞动 DB2 系列)

ISBN 978-7-121-13094-6

I. ①D… II. ①王… III. ①关系数据库—数据库管理系统, DB2 IV. ①TP311.138

中国版本图书馆 CIP 数据核字(2011)第 041981 号

责任编辑:高洪霞

印 刷: 三河市鑫金马印装有限公司

装 订:

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

开 本: 787×1092 1/16 印张: 26 字数: 627 千字 彩插: 2

印 次: 2011 年 4 月第 1 次印刷

印 数: 4 000 册 定价: 59.00 元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010) 88254888。

质量投诉请发邮件至 zltz@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线:(010) 88258888。

总 序

去国离家十六载后，2006 年，当我再次归来凝视上海这座不夜城时，不由得感慨世界真的变小了、变平了，中国作为地球村的一员，正以惊人的速度发展变化着。

世界正改变着中国，中国也不断影响着世界。在备感欣喜振奋之余，我亦常常思索，在一个越来越小、越来越平的世界里，中国会面对怎样的挑战和机遇？IBM 作为全球最大的信息技术和行业解决方案公司，又在哪些方面可以与中国分享我们的先进技术、经验与创新为要的精神呢？

中国自 2001 年入世以后，很多企业开始迅速发展，创造出了许多新的发展模式。但是，中国、中国的企业在做大做强并开始走向世界的过程中也面临着越来越多的挑战。如何拥有更多更强的创新能力？如何提升中国在全球一体化经济格局中的竞争力？如何增强企业的可持续发展能力并逐渐走向领袖地位？这应该是很多中国企业共同面对并思考的问题。

百年来，创新求变是 IBM 公司虽历经艰难仍屡铸辉煌的关键。因而，IBM 公司虽多次转型，却始终以超前的技术、出色的管理和独树一帜的产品领导着全球信息技术的发展，保证了世界范围内几乎所有行业用户对信息处理的全方位需求。关于创新，我们有很多的经验愿与中国业界分享。

中国开发中心（China Development Lab, CDL）是 IBM 在中国的窗口，是中国 IT 产业与世界交流的桥梁。CDL 不仅把全球先进技术引入中国，也将中国领先的技术与行业解决方案推向世界。经过十年的蓬勃发展，CDL 已经拥有 5000 名软件开发人员，成为 IBM 全球规模最大的软件开发基地之一，并成为跨国公司在中国最大的软件开发机构。十年来，一批又一批中国本土的人才，在 CDL 提供的平台上施展才华，并逐渐成长为技术与管理精英。

我们始终恪守着“中国人才、国际管理；中国创造、服务全球”的理念，希望将 IBM 的最新技术、创新和管理心得与大家共同分享，并鼓励 CDL 的专家走出实验室。于是，我们有幸选择了实力非凡、专业创新的电子工业出版社博文视点公司作为合作伙伴，推出这一系列丛书。该丛书的作者包括来自 CDL WebSphere、Information Management、Lotus、Rational、Tivoli 五大软件产品开发团队，以及来自中国开发中心新兴技术学院（Emerging Technology Institute）、行业解决方案中心（Industry Solution Labs）、软件服务团队（Software Service）等各个部门的优秀架构师、资深工程师、管理者，范围将涵盖从开发理论及平台、

方法论及实践、项目管理、最新软件技术与产品、先进行业解决方案架构，到最新技术标准和前沿技术发展趋势甚至创新的商业模式探讨等诸多方面。

我希望这套丛书能把我们一线专家宝贵的经验和我们的见解呈现给读者。在与业界分享经验和世界最新技术及趋势的同时，我们希望能为推动中国 IT 产业的加速发展奉献微薄之力。IBM 中国开发中心将一如既往地同业界同仁一起，共铸中国信息产业的辉煌明天！

王阳 博士

IBM 全球副总裁兼中国开发中心总经理

推荐序一

两千多年前的战国时代，信息记录在竹简上。从此信息文明更新换代。而最近 30 年发生的信息爆炸，使世界发展步伐加快，社会变化日新月异。在刚过去的经济危机中，IBM 逆势取得了骄人成绩，继续保持了优势地位，并紧紧把握了中国信息发展的时代脉搏。而信息管理（IM），作为 IBM 最重要的软件品牌之一，已成为 IBM 近年来投入最大的领域。在过去 5 年间，IBM 在这一领域仅在收购最新技术上的花费金额就超过 120 亿美元。

2010 年 10 月在拉斯维加斯举行的“2010 IBM IOD 全球大会”上，信息按需应变（IOD）的创新理念及丰富的“智慧”产品吸引了超过 9000 名来自世界各地、各行各业的代表参加。与往年相比，这次峰会众多来自中国的面孔增加了很多。在随后举行的客户座谈和用户研讨会上，我们更希望听到来自中国客户的声音。可喜的是，IBM 在中国的热情投入换来了中国客户的倾情赞誉和坚定支持；可忧的是，中国信息管理的声音不如期待中的那样响亮，这与中国近年来飞速发展的信息技术产业和市场规模是极不相称的。究其原因，中国的信息管理领域起步晚、经验还需积累是一方面，但最重要的因素就是中国本土缺乏优秀的技术人才储备，这导致在这一领域的高端人才十分匮乏，这必会成为中国信息产业走向辉煌的巨大障碍。

IBM 中国开发中心经过十年的积累和沉淀，已经培养建立起了一支具有精湛技术、熟悉中国市场和融合全球信息技术的优秀团队。这支团队通过坚持不懈、热情执着的努力，已成功伴随中国的客户应对信息社会的巨大挑战，并获得了飞跃式的发展。我们有能力，也深感有必要把先进的信息管理技术和理念进一步传播给中国广大的技术人员。这本由 DB2 资深工程师撰写的书，是信息管理领域的一朵奇葩，专注于数据库设计与性能优化方面，内容非常全面，定位非常准确，方法非常实用，相信她能帮助中国的高端信息管理技术人才更上一层楼。

IBM 中国开发中心 信息管理产品开发部

总经理 吉燕勇

2010 年 11 月 25 日

推荐序二

自从埃德加·考特（Edgar F. Codd）在 1970 年发明关系数据库以来，IBM 在数据库技术的研究和开发方面一直处于世界领先地位。DB2 数据库是 IBM 公司数据管理产品线中最知名也是最成功的产品，除了可以运行于大型主机环境外，也适用于多种开放式平台。近年来，IBM 加大了 DB2 研发在中国开发中心的投资，目前已培养了一大批参与 DB2 产品核心开发、客户技术支持和服务的优秀工程师。

DB2 数据库在中国取得了非常广泛的应用，在金融、电信、政府、制造和零售等行业都有很多成功的客户。然而我们发现的一个问题是：在国内的广大用户中，对 DB2 精通的人还比较少，影响了 DB2 产品的使用。DB2 生态系统的繁荣依赖于用户、合作伙伴和 IBM 的共同努力，如果没有广大 DB2 用户技术水平的提升，DB2 生态系统的繁荣就很难实现。因此我们的一个重要任务就是要通过各种方式和途径来推广 DB2 产品技术。

随着企业信息应用系统越来越复杂，用户量和数据量正变得日益巨大，这对设计人员设计高质量的数据库及 DBA 优化数据库提出了严峻的挑战。本书就是为了解决这个问题而诞生的。在这本书中，不仅包括了 DB2 经典设计和优化方法，还涵盖了 DB2 最新设计和优化技术，例如表分区、工作负载设计、pureXML 设计、pureScale 集群、solidDB 高速缓存等。特别是为了帮助读者掌握性能优化技术，还提出了一套基于 PAT 树的性能优化方法系统。

本书由来自 DB2 的资深工程师编写，原汁原味地展示了 DB2 设计和优化技术，深入剖析了 DB2 内部工作原理。在为全国各地的 IBM 合作伙伴培训 DB2 数据库期间，作者更深入地了解到 DB2 用户在学习遇到的困难。另外，作者具有为政府、电信、金融等行业客户做数据库迁移方面的咨询和实施经验，充分了解生产环境中客户遇到的实际问题。

上面的一切使得本书能想读者之所想，写读者之所需。

IBM 中国开发中心 DB2 开发

资深经理 干毅民

2010 年 12 月 6 日

推荐序三

心有多大，舞台就有多大。当我看到这本书的初稿时，立刻被书名所吸引。“舞动 DB2”几个字，尽管还未经过美工处理，我已经意识到作者心怀天下，志存高远。而当我接连读完第 1 章和第 2 章，我又发现作者不但目标长远，下的功夫也极深。第 1 章总述数据库设计与性能优化，高屋建瓴般地提出了颠覆性的 PAT 方法学。第 2 章相对于传统性能分析方法，以电信行业的经营分析系统为例，全面展现了性能问题分析定位体系，其高度的实用性着实令人叹服。而案例分析部分环环紧扣，步步精彩。贯穿整本书的线索使我迫不及待地追逐一个又一个的疑问，一直读过数百页。我不得不惊叹，这本书本身就是一个完整的理论体系，扩能长龙飞舞，缩能稳如泰山。

我也不得不说起当今的一些数据库图书，更像在写教材，寻章摘句，呆板无趣。普遍浮躁的学术氛围造就了一批苍白无力的技术书籍，这样的平庸之作导致读者从一开始就迷惑，进而迷失，读到最后是一片迷茫。

在本书中，我们始终能读出作者清晰的思路，作者在这方面确实投入了巨大的精力。春节刚过，王飞鹏在为书制定目录时，严谨的学术风格就展现了出来——他邀请了多位技术专家进行评审，为其严格把关，为读者负责。从选题结构上就高人一筹，完全切中了当今 DB2 应用过程中设计与性能优化问题的要害。之后，他组织陈辉、张广舟、成孜论各显身手，充分展现团队威力。事实证明，他们把各自多年的积累与优势淋漓尽致地发挥出来，为读者带来一场数据库设计与性能优化的盛宴。在他们完成初稿后，我与其他两位数据库资深顾问逐字斟酌了内容，提出了建议。他们非常谦虚地接受了建议，仔细修改。当第二稿、第三稿、第四稿陆续发给我们时，我们所有人都为他们坚持不懈的努力和出众的才华赞叹不已。

几位作者经过数月的奋斗，终于完成了这部巨著，这个过程真正的辛苦只有他们自己最清楚，但我相信他们的意志力一定经受住了极大的考验。这本真材实料的技术专著，也一定会为广大读者带去最真切的帮助，并创造出新的奇迹。

IBM 信息管理产品首席架构师

美国加州大学

陈奇 博士

2010 年 11 月 30 日

推荐序四 用之道

许多 DBA 对一则定律都深有感触：

If there are two or more ways to do something, and one of those ways can result in a catastrophe, then someone will do it.

这句话被称作“墨菲定律”，是由毕业于西点军校的空气动力专家 Edward A. Murphy 发现的。无独有偶，早在五百年前，冯梦龙就写道：屋漏偏逢连夜雨，船迟又遇打头风。两位学者身处年代不同，所受文化熏陶迥异，但对事件所反映出来的结果有着相同的描述：在某个时刻，问题会集中爆发，原本丰富的资源在那个时刻突然枯竭，原本充沛的体力在那个时刻立刻透支。在数据库领域，有人把这句话解释为：如果存在隐患，哪怕只是一点点，它一定会在某个时刻爆发，并且后果严重。责骂、混乱、恐慌、歇斯底里，在那个时刻会统统地冒出来。这对于 DBA 来讲，绝对是个恶梦。太多的教训应验了无论是什么数据库，一旦“墨菲定律”发作，会让 DBA 们痛彻心扉。

我想起郎咸平先生这几年一直非常火爆的一本书《我们的日子为什么这么难》。书的视角开阔而又着眼于民生，社会生活的方方面面都涵盖于其中。尽管郎先生的观点众说纷纭，但确实说出了老百姓的无奈和难处。而 DBA 的无奈和难处，又由谁来代表大家倾诉？又由谁来帮助大家解决？我看完本书后找到了答案，而这要从作者说起。

IBM 是一家令人尊敬的公司。在西方世界，IBM 是“阿波罗登月计划”的英雄，被 Bill Gates 称为世界上最大的对手；在东方国度，IBM“蓝色巨人”是前沿科技的代名词。本书的作者均来自 IBM 中国研发中心，他们有着年轻人特有的激情和活力，在自己的工作岗位中表现出众，王飞鹏就是其中一位。

飞鹏生长于三晋大地，毕业后，他进入 IBM 成为一名研发工程师，大家都称呼他 William。2008 年端午节，我与 William 在成都不期而遇，那时正值震后救援与震后重建的交替期，特殊时期的救灾经历让我们的友谊迅速升温。再见已是一年以后的西湖边上了，那天 William 作诗一首，时过境迁，我仍不时念诵：

西子水光竞涟漪，	南屏山色叠峦峰。
栖霞烟霞入云霄，	苏堤白堤成双龙。
意气勃发思将飞，	壮怀九天欲乘风。
雷锋塔前青衫湿，	愿为此地田舍翁。

今年我们见面的次数少了，William 整日飞来飞去，忙于各种数据库项目的咨询、实施和讲座。有一次我们碰巧都在杭州，我去听了一次他给浙江电信做的数据库性能优化培训。来的人很多，位子不够，最后面也站了一排人（后来才知道，一些数据库爱好者也慕名而来）。整个培训期间，William 文采飞扬，激情澎湃，与他私下里稍偏内向的性格截然不同。他的演讲诙谐轻松，简练易懂，时而高屋建瓴犹水瀑，时而烟枝细雨落风檐。我想 William 对数据库技术的精湛理解和讲授，跟他丰富的经历是分不开的。William 曾为中央级政府机关、电信企业、金融机构、高铁系统等制定数据库优化方案，指导项目实施以及解决性能问题。这些高强度的工作使他本人的组织能力、协调能力和项目控制能力达到了新的高度。此外，他对事物不同寻常的分析和领悟能力，还有坚毅执着的信念，在他所获得的多项专利中也都有深刻的体现。

今年夏天，William 本人亲自来上海郑重邀请我作序。原来，他正在准备写一本有关 DB2 数据库设计和性能优化方面的书。他讲这本书的目的不是头痛医头，脚痛医脚，而是从数据库设计阶段就开始考虑性能问题，从系统整体层面优化数据库，一劳永逸地解决本质问题。更为重要的是，他打算在这本书里面提供一套经过验证、简单实用的 PAT 方法学，让数据库技术人员思路清晰、有条不紊地看清问题、解决问题，从而达到终极目标——让庞大的数据库轻盈地舞动起来。这本书另外的几位作者分别是：陈辉，DB2 技术知名专家，拥有 DB2 内核开发的丰富经验；张广舟，号称“DB2 诊断医生”，具有丰富的数据库优化实战经验；成孜论，多年往返于国内外，视角开阔，技术功底深厚，一定会为读者带来别样的惊喜。

秋天到了，窗外瓜果飘香，屋内我品读着他们的硕果。在这辛勤耕耘的几十万字中，我读出了 DB2 的力量，读出了中国数据库工作者的希望。书里既有惊心动魄的危机处理纪实，又有摸爬滚打积累的实践心得。有时你会享用到最新技术的说明资料，有时你会感同身受实战技巧，有时你也会对作者独特的分析视角感叹不已。我很欣慰，一部 DBA 案头必备，可以传世的经典之作诞生了！

本书作者其实代表了一部分热血青年，他们本可以养尊处优、高谈阔论地过着小资的生活，但是却选择了一条荆棘密布、困难重重的长征之途。而这条路途，最终通向的是梦想之乡，通向的是雪山之巅。无论是这条路的开辟者，还是有共同志向的开拓者，都会为中国数据库生态系统的繁荣与昌盛贡献出自己的力量。

IT 诸葛网总裁 焦新港

2010 年秋 写于上海

前言 关系论

提到数据库的历史，就不得不说起一位传奇人物：E.F.Codd。作为数据库理论的奠基人，E.F.Codd 首次提出了“关系型数据库模型”。这位从牛津走出来的绅士，“二战”期间曾是英国皇家空军的飞行员。后来他来到纽约加入了 IBM。1981 年，凭借对数据库发展的划时代贡献，Codd 荣获被誉为“计算机科学界诺贝尔奖”的“ACM 图灵奖”。这位通透数据关系的先驱，在讲述“关系”时有过这样的描述：

我的家乡在波特兰岛，这个孤独的小岛几百万年来访客不多。这使得岛上的物种与外界有很大差别。这里有翅膀已经退化的海鸥，这里的蟋蟀长得像大型工蚁，既不会跳也不会叫，还有一种我喜爱的叫做黄衣草的植物。的确，波特兰岛是个比较封闭的生物圈，但是植物群与动物群的关系，甚至与人的关系却是如此紧密，这令我非常惊讶，也非常着迷。

在这里人们总是诅咒兔子，以至于“dead rabbit”总是挂在嘴上。原来，住在土洞里的先民，如果发现洞口有兔子出现，经常会有土块从洞顶掉落下来，砸到人的头上。这种离奇的关联，竟然可以从达尔文的《物种起源》找到缘由。达尔文说：“任何生物之间都有关系，只是往往你察觉不到。”后来发现岛上的兔子喜爱吃一种荠菜。这种荠菜的根茎被兔子吃光后留出的空间，是兔子良好的休憩场所。进而兔子把洞越挖越深，有的可达十英尺。而人们恰恰钟爱这里荠菜的叶子，配上沙丁鱼是最美味的佳肴。荠菜被种在洞顶的土坡上，那里是最合适的菜地。就这样，害怕人类的兔子为了荠菜，越来越多地涌入人类住的洞顶。当过于拥挤时有的兔子就会游逛到下面的洞口活动。而这个举动证明，洞顶已经被兔子挖得到处是安乐窝了，自然会有土块松动而坠落下来。

Codd 博士也许是从家乡关于兔子的传说中找到了“关系型数据库”的灵感吧，当今无论是 DB2 还是 Oracle，都要感谢当年在英吉利海峡辛苦挖洞的兔子。很显然：关系，无所不在。从厄尔尼诺现象到多米诺骨牌效应，从热核武器裂变到蝴蝶效应，事物之间的关系如此丰富，有时引起的后果却又很严重。面对事物关系引发的连锁反应，如果我们未雨绸缪，积极主动地做好防范，准备好应对措施，就可以最大程度地降低风险。这些理论在数据库领域，同样适用。

随着各领域业务的迅速发展，支撑业务发展的大型数据库性能面临的压力与日俱增。我们在数据库设计与性能优化的工作中注意到如下一些不良现象：

1. 不重视数据库设计，后期出现大量性能问题，改造和维护代价高昂。
2. DBA 往往依靠临时判断处理问题，缺乏一套行之有效的系统化的优化思路和科学可靠的方法来指导。

3. 数据库优化存在技术门槛，工具繁多，学习周期长，而时间的压力往往导致进度过于仓促，造成工作混乱。

我们发现，这些问题表面上是工作内容混乱造成的，而实际上是工作方法和工作思路混乱造成的。我们撰写本书的目的正是为了帮助读者梳理流程，提供解决问题的思路和方法。本书列出了数据库设计与性能优化工作中频繁遇到的问题、难以解决的问题、措手不及的问题和容易忽视的关键问题，通过两条线索：缜密的技术阐述和贯穿始终的案例说明，帮助读者更好地吸收和理解。

需要特别指出的是，本书是由来自 IBM 的资深研发工程师执笔撰写的，他们都具备真实生产环境中大型数据库设计与性能优化的丰富实践经验，这是区别于目前市场上所有同类书籍的显著特点。

本书结构

本书分为四大部分，共 11 章。

第一部分，解析数据库性能的本质问题。

决定数据库性能的两个阶段：构建数据库之前的设计阶段，数据库上线运行后的优化阶段。第一阶段完成高质量的数据库设计，是高性能系统的基础；第二阶段基于 PAT 树的方法系统，是掌控数据库性能的秘密所在。我们通过金融、电信等行业经营分析系统的实际案例，来一层层地揭开这个秘密，领略 PAT 树调整设计和解决性能问题的威力。

第二部分，从设计高性能数据库的角度，详细介绍如何使用 DB2 数据库设计高质量的应用系统。内容主要包括 DB2 数据库逻辑设计、DB2 数据库物理设计、DB2 pureXML 数据库设计及 DB2 工作负载设计。

第三部分，介绍如何优化 DB2 数据库。内容主要包括性能监控及参数调整、SQL 语句优化策略、锁和日志优化等。

第四部分，阐述 DB2 高级性能加速技术。本部分内容是对 DB2 数据库设计和优化基础之上的有益补充，主要包括 DB2 pureScale 集群和 solidDB 高速缓存加速。

附录部分列出了常用监控命令用法和缩略语释义，还特别提供了 PAT 树的放大图，供读者进行数据库性能优化时使用。

本书特点

1. “用”字当先

很多人喜欢读成功人士的自传，慢慢就会按照书中所说去做。但其实，每个人的生活

和工作轨迹都不同，一种成功模板不一定适用于所有人。数据库技术系列丛书中，相当一部分书籍都是教科书般地讲解数据库知识，或许中间会穿插一些例子，但缺乏真实案例的有力支撑，也没有形成一套系统化的工作方法。因此我们要引导读者走上更有意义的应用大道，为读者提供自我分析问题的思维方式和实现方法。特别值得一提的是，贯穿本书的 PAT 方法学，是已被大量实践证明的解决应用性能问题的高效方法学。该方法学立足于实践，对设计人员提高 DB2 数据库设计水平，对 DBA 提高性能优化水平有很高的实用价值。

2. “厚”的积累

我很幸运，能够站在蓝色巨人 IBM 的肩膀上看世界，更幸运的是，我周围聚集着一群极其优秀的专才。这使我的经历和见识都急速提升。本书记录了在移动通信系统优化项目中，我与伙伴们如何排除系统中的“定时炸弹”；在金融系统优化项目中，我的团队如何应对挑战，又如何通力协作；在中央机关数据库优化项目中，我们如何面对棘手状况，又如何有针对性地制定优化方案，最终突出重围取得成功；在各地高校，看到诸多热衷 DB2 的年轻脸庞，那是让人感动和激动的时刻！这些“厚”的积累，一定能帮助读者增加成功的可能性。

3. “新”的技术

技术变革会带来生产力的飞跃。MRI（核磁共振成像）横空出世后，这种革命性的冲击让医生们对 X 光、CT 技术再也无法产生热情。而在 DB2 设计和性能优化过程中，如果能合理使用 DB2 提供的最新技术，你一定会惊呼：事半功倍原来如此简单！本书中所阐述的 pureXML 设计、工作负载设计、表分区、数据压缩、新监控框架以及游标稳定性隔离级的当前提交特性等都是 DB2 的最新特性。另外，为了提高应用程序事务处理能力，本书还介绍了 DB2 pureScale 和 solidDB 高速缓存 DB2 数据库。

读者对象

如果您学习数据库知识仅一个月，本书可能不是最适合您的。三年前一位来自 IBM 多伦多实验室的数据库大师曾告诫我：你不要指望一本书可以兼顾入门者和专业者，更不要做这样的事情。我记住了这句话，因而这本书专注的对象是 DB2 数据库设计人员、DBA、数据库性能分析人员、数据库开发人员、运维人员及应用开发工程师。

还请注意的，在阅读本书前，读者无须具备性能优化的专业知识，但希望读者接触过 DB2 数据库并且了解 SQL 语言。同时，本书还提供了三条学习路径：

第一，从解决问题的角度来看，读者可以在学习了第 1 章介绍的 PAT 方法学和第 2 章介绍的金融电信行业经营分析系统的真实案例后，直接使用 PAT 方法学并结合后续章节来

解决实际工作中遇到的性能问题。

第二，从掌握 DB2 知识的角度看，读者可以直接从第 3 章开始学习 DB2 数据库设计、监控和优化等知识。

第三，从解决问题和学习结合的角度来看，PAT 树最重要的特点就是叶子节点给出的优化方法准确对应到本书的相应章节，所以在解决性能问题的同时，可以利用这个特点有针对性地学习 DB2 数据库相关知识。

另外，考虑到本书可以作为数据库设计以及性能优化的教学用书，所以对数据库相关专业的高校学生以及从事相关课程教学的教师具有良好的参考价值。

本书涵盖哪些版本

本书中绝大多数概念适用于包括 DB2 8 for Linux, UNIX and Windows 以来的所有版本。但不可避免的是，我们讨论到的某些特性只是针对特定版本的，具体如下。

- 表分区和 MDC 特性（第 3 章）：适用于 DB2 9 及其以后版本。
- 自动数据库管理、数据压缩和 STMM（第 4 章）：适用于 DB2 9 及其以后版本。
- pureXML 数据库设计（第 5 章）：适用于 DB2 9 及其以后版本。
- DB2 工作负载设计（第 6 章）：适用于 DB2 9.5 及其以后版本。
- 新监控框架（第 7 章）：适用于 DB2 9 及其以后版本。
- 当前提交特性（第 9 章）：适用于 DB2 9.7 及其以后版本。
- DB2 pureScale 海量事务处理技术（第 10 章）：适用于 DB2 9.8 及其以后版本。
- solidDB Cache 高速缓存技术（第 11 章）：适用于 DB2 9.1 及其以后版本。

如果没有特别指明某特性只针对某个版本，则默认对所有版本都适用。

在线资源

读者可以从 www.db2ring.net 下载本书案例中用到的文件，也可以从该网站上找到本书最新的补遗及勘误。同时也欢迎大家到网站相互交流和学习。

致谢

首先感谢电子工业出版社的高洪霞编辑，她在本书出版过程中给予了我们大力支持和帮助！感谢 IBM 信息管理产品架构师陈奇博士及 Cognos 行业解决方案架构师王庆法，他们在本书的选题方面提供了重要帮助。同时，本书在审稿中还得到了 IBM 软件开发实验室经理刘慎锋的热心帮助，在此表示衷心感谢。此外，感谢赵文峰博士，他审核了本书的

PAT 方法论及大部分章节并提出了宝贵建议；感谢秦怡，她在繁忙的工作之余审核了所有章节的引言和案例部分，并提供了非常重要的建议；感谢张昱，他审核了书中部分章节并制作了 PAT 树的插图；感谢刘畅，她审核了本书的前两章并提出了有益的修改意见；感谢所有参与本书审核的同事们，孙岳、李玉明、李睿明、侯占友和陈诗然。

王飞鹏 2010 年感恩节写于圣淘沙岛

目 录

第一部分 DB2 基础

第 1 章 我看 DB2 设计与优化	1
1.1 数据库设计与性能优化	2
1.1.1 收集需求	3
1.1.2 设计概念模型	4
1.1.3 设计逻辑结构	5
1.1.4 设计物理结构	6
1.1.5 实施、运行和维护阶段	8
1.2 性能问题	8
1.2.1 什么是性能问题	8
1.2.2 为系统做性能基准测试	9
1.2.3 制订性能优化目标	10
1.2.4 把问题分类	10
1.3 使用 PAT 方法学解决问题	11
1.3.1 什么是 PAT 方法学	11
1.3.2 PAT 方法学优化策略	11
1.3.3 使用 PAT 方法学	15
1.3.4 使用 PAT 方法学的步骤	25
1.3.5 PAT 树使用建议	26
1.4 小结	26
第 2 章 性能优化利器——PAT 方法	28
2.1 优化步骤	29
2.2 业务分析	29
2.2.1 经营分析系统的架构	29
2.2.2 性能问题的提出	31
2.3 系统分析	31
2.3.1 硬件和操作系统	32

2.3.2	数据库物理设计	33
2.3.3	分析结果	34
2.4	制订优化计划	35
2.4.1	优化计划安排	35
2.4.2	实践优化策略	35
2.5	使用 PAT 方法	36
2.5.1	优化“每天下午系统响应慢”问题 (Perf_SlowAfterNoon_Sys)	36
2.5.2	优化“大数据转入”问题 (Perf_Load_App)	37
2.5.3	优化“客户流失分析”问题 (Perf_Customer_App)	38
2.5.4	优化“账户资费”问题 (Perf_ACC_App)	39
2.5.5	优化“数据质量管理”问题 (Perf_Data_App)	41
2.5.6	优化“系统逐渐变慢”问题 (Perf_SlowDown_Sys)	42
2.5.7	优化总结	45
2.6	小结	45

第二部分 DB2 数据库设计

第 3 章	高质量逻辑设计	48
3.1	常规表设计	49
3.1.1	表设计的原则	49
3.1.2	列定义原则	52
3.1.3	数据完整性设计	53
3.1.4	其他考虑因素	54
3.2	索引设计	56
3.2.1	索引的概念	56
3.2.2	深入解析索引原理	60
3.2.3	创建索引	67
3.2.4	索引设计原则	67
3.2.5	索引维护	70
3.3	MDC 表设计	71
3.3.1	多维群集 (MDC)	71
3.3.2	块索引	72
3.3.3	MDC 表的优势	73
3.3.4	MDC 表设计原则	74
3.4	表分区设计	76