

<<内存数据管理>>

图书基本信息

书名：<<内存数据管理>>

13位ISBN编号：9787302292562

10位ISBN编号：7302292566

出版时间：2012-8-2

出版时间：清华大学出版社

作者：(德)普拉特纳(Plattner, H.)

译者：(德)蔡尔(Zeier, A.)

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

## &lt;&lt;内存数据管理&gt;&gt;

## 前言

第二版前言本书第一版已经出版了一年有余。

在这一年间，内存计算技术对企业级计算与应用市场产生了重大影响，使其真正成为一个技术拐点。

由于这一发展，以及由此引发的新问题，有必要扩充本书至第二版。

第二版中的新内容主要侧重于数据密集型应用程序的开发和部署，数据密集型应用程序的设计需要尽可能利用内存数据库系统的功能。

在其他新增内容中，6.1.1 章节介绍了内存应用程序编程模型，涵盖了开发内存应用程序的重点和指导原则。

为减轻应用程序开发人员和数据库管理人员的工作负荷，我们在6.1.5 章节探讨了图形化创建数据库视图。

最后，对应用程序层上的新功能作了详细说明，例如6.2.4 章节中数据分析和文本搜索的组合以及9.2 章节中两个行业案例的展示。

当然，如果没有我们企业平台的学生们以及集成概念主席的帮助，我们不可能在这么短的时间内完成本书第二版的编纂工作。

因此，除了在第一版序言中对他们的付出和努力表示感谢之外，在这里我们还想再次感谢他们。

2012年3月1日，波茨坦哈索（Hass&middot;Plattner）亚力山大&middot;蔡尔（Alexander Zeier）第一版前言我们写这本书，是因为我们深信内存计算技术的运用将成为企业级应用的转折点。

在过去几年里，内存单价下降，内存容量大幅增长。

这就需要重新思考如何存储海量数据。

现在我们可以将数据库主要数据的副本存储在内存中而不是机械磁盘中，这样可以使性能得到几个数量级的提升，并且使得全新的应用成为可能。

这种数据存储方式的改变正在并将继续对企业级应用产生重大影响，进而最终影响企业的运营方式。

企业决策者能够以思想般的速度获得实时信息，他将具备前所未有的洞察力。

本书面向特定的读者群。

他们将了解数据管理方式的根本性转变如何影响企业级应用。

我们希望本书能对想充分利用内存计算来创建新业务流程的在校大学生、IT 专业人员、IT 经理和高级管理层有所启发。

本书分为三个部分：&middot;第一部分概述我们的愿景，即内存计算技术如何改变企业级应用。

这一部分适用于所有读者。

&middot;第二部分更深入地阐述我们打算如何实现愿景。

这一部分适用于那些希望深入了解内存数据管理的学生和开发人员。

&middot;第三部分阐述内存计算对企业级应用开发和功能所带来的深远意义。

这一部分适用于技术人员和业务型读者。

一本书往往凝聚着许多人的心血，绝非仅靠作者完成。

在此，我们要特别感谢德国波茨坦大学哈索&middot;普拉特纳学院的“企业平台和集成概念”小组的全体成员。

HANA 研究小组的成员包括：安佳&middot;伯格（Anja B&middot;g）、马丁&middot;格伦德（Martin Grund）、延森&middot;克鲁格（Jens Kr&uuml;ger）、史蒂芬&middot;穆勒（Stephan M&uuml;ller）、简恩&middot;夏弗纳（Jan Schaffner）和克里斯汀&middot;缙娜芬德（Christian Tinnefeld）。

在过去的5年里，他们一直专注于内存计算应用领域的研究。

他们的研究为本书的撰写奠定了坚实的基础。

此外，瓦迪姆&middot;波罗夫斯基（Vadym B&middot;r&middot;vskiy）、托马斯&middot;科沃科

（Th&middot;mas K&middot;wark）、拉尔夫&middot;屈内（Ralph K&uuml;hne）、马丁&middot;洛伦兹

（Martin L&middot;renz）、杰根&middot;穆勒（J&uuml;rgen M&uuml;ller）、奥克斯德&middot;帕奇

可（&middot;leksandr Panchenk&middot;）、马斯恩&middot;夏普拉娜（Matthieu Schapran&middot;w）

、克里斯汀&middot;施瓦兹（Christian Schwarz）、马提亚&middot;艾弗拉克（Matthias Uflacker）和约

## &lt;&lt;内存数据管理&gt;&gt;

翰尼斯·伍斯特 (Johannes Wust) 也为本书的编写做出了巨大贡献。

我们的助手安德里亚·兰格 (Andrea Lange) 也一直在帮助进行协调工作。

此外, 在编写本书的过程中, 我们还得到了许多 SAP 同事的热心帮助。

如果没有他们的支持, 本书将无法顺利面世。

卡夫尔·涂森 (Cafer Tunc) 在 HPI 和 SAP 之间起着桥梁作用, 他不仅负责协调我们与 SAP 的合作关系, 而且还积极为本书提供部分章节内容。

其团队成员安德鲁·麦考密克史密斯 (Andrew McCormick-Smith) 和克里斯汀·马蒂斯 (Christian Mathis) 为本书提供了许多重要文字段落。

借此书付梓之际, 我们衷心感谢乔斯·亨德里克·布斯 (Johannes Hendrik Bese)、伯恩哈德·菲舍尔 (Bernhard Fischer)、依娜·福克斯 (Ennke Fokerts)、安德烈亚斯·赫歇尔 (Andreas Herschel)、塞拉·卡佩斯 (Sarah Kappes)、克里斯汀·穆克尔 (Christian Munkel)、弗兰克·伦克思 (Frank Renkes)、弗雷德里克·全斯尔 (Frederik Transier) 所付出的辛勤劳动。

我们衷心感谢保罗·霍夫曼 (Paul Hofmann) 为本书付出的心血, 感谢他帮助我们与美国各大学共同合作的研究项目。

此外, 我们的研究成果也离不开其他 SAP 同事的热心帮助。

在此, 我们特别感谢弗朗兹·法波 (Franz Fieber) 及其团队所提供的反馈意见, 以及他们在过去几年里对我们的研究成果所做出的卓越贡献。

我们在整本书中提及的许多想法都源自弗兰兹。

此外, 他还负责 SAP 内部实施事务。

对此, 我们要特别赞扬。

最后, 我们要感谢 SAP 首席技术官维舍尔·斯卡 (Vishal Sikka)。

他一直鼎力支持我们的研究工作, 并亲身参与其中。

另外, 我们还要衷心感谢 SAP 首席运营官格哈德·奥斯瓦德 (Gerhard Oswald) 和 SAP 联合首席执行官吉姆·哈格曼·施杰翰 (Jim Hagemann Snabe) 与比尔·孟鼎铭 (Bill McDermott) 对我们项目的长期支持。

请您访问本书的官方网站。

网站中列出了本书的更新内容、评论、有关内存数据管理的相关博客文章以及学生试题。

2011年2月1日, 波茨坦哈索 (Hasso Plattner) 亚力山大·蔡尔 (Alexander Zeier)

## <<内存数据管理>>

### 内容概要

SAP创始人哈索教授在《内存数据管理（第2版）》中提出了创新的概念：高性能的内存计算将改变企业的工作方式，并最终真正实现“实时”计算处理的承诺。

内存计算技术将给以下三个相互关联的战略领域带来重大变革：降低总体拥有成本、创新应用、作出更快更好的决策。

## 作者简介

哈索教授，（Hasso Plattner），是SAP公司的创始人之一，并于2003年5月起担任SAP监事会主席。作为公司的监事会主席和首席软件顾问，他致力于制定SAP的中长期技术战略和发展方向。与此同时，哈索也负责领导SAP监事会技术委员会。

1972年，哈索和四位同事离开位于德国曼海姆的IBM公司，并创建了SAP（System，Applications，Products in Data Processing - 系统、应用与数据处理产品）。

如今，总部设在德国沃尔多夫的SAP公司已成为整合企业及业务部门间流程的企业管理软件供应商中的领导者。

1988年SAP上市的时候，他被任命为SAP执行董事会副主席。

1997年至2003年5月期间，哈索一直担任SAP执行董事会主席兼首席执行官。

2003年5月他接替另一位SAP的创始人DietmarHopp被选举为SAP监事会主席。

哈索在卡尔斯鲁厄大学获得通信工程硕士学位。

1990年，萨尔布吕肯大学授予他名誉博士学位，并于1994年授予他名誉教授头衔。

1997年，作为SAP美国公司的董事长，SAP的联席董事长和SAP R/3的总架构师，哈索在1997年被授予全球一体化信息技术领导奖，该奖是《计算机世界》史密森奖励计划的一部分。

1998年，他入选德国名人堂。

2002年哈索被波茨坦大学任命为名誉博士，并于2004年获得名誉教授头衔。

哈索于1998年在德国波茨坦大学内建立了以其名字命名的IT系统工程学院，同时创下了德国为单一大学私人捐款的最高纪录。

在他持续不断地财政支持下，该学院已经成为世界级的软件专业人才教育中心。

亚历山大·蔡尔（Alexander Zeier），毕业于维尔茨堡大学企业管理专业，并在开姆尼茨工业大学完成了信息技术专业的学习。

在获得埃朗根-纽伦堡大学供应链系统专业的博士学位之前，他做了多年的战略IT顾问。

他拥有19年SAP/IT系统方面的工作经验，2002年加入SAP后，他担任产品经理，全面负责SCM软件 - SAP首个大型内存应用程序。

自2006年开始，他在柏林/波茨坦的哈索·普拉特纳学院担任哈索的副教授，并致力于实时内存企业系统和RFID技术的研究。

他著作的出版物超过100多本，其中包括五本有关IT和SAP的书籍。

他不仅是麻省理工学院的客座教授，还担任麻省理工学院论坛欧洲地区的执行总监。

## 书籍目录

第一部分企业级应用的转折点第1章 可取性、适用性、可行性&mdash;&mdash;内存计算技术的影响1.1 实时信息--随时随地获取任何信息1.1.1 思想速度般的响应1.1.2 实时分析和动态计算1.2 最新硬件趋势的影响1.2.1 企业级应用的数据库管理系统1.2.2 主存是新磁盘1.2.3 从最大化 CPU 速度到多核处理器1.2.4 增加的 CPU 和主存之间的带宽1.3 通过内存数据管理降低成本1.3.1 总体拥有成本1.3.2 企业系统中的成本因素1.3.3 内存计算的性能促进成本降低1.4 结论第2章 企业级应用为何如此繁杂零乱2.1 当前的企业级应用2.2 企业级应用范例2.3 企业级应用架构2.4 企业级应用中的数据处理2.5 企业级应用中的数据访问模式2.6 结论第3章 SanssouciDB--企业内存数据库系统的未来蓝图3.1 重点关注多核和主存3.2 内存数据库系统设计3.3 SanssouciDB中数据的组织与访问3.4 结论第二部分 SanssouciDB：通过内存计算技术提供单一数据源第4章 SanssouciDB 的技术基础4.1 了解内存层次结构4.1.1 主存简介4.1.2 主存层次结构的组织结构4.1.3 内存层次结构的趋势4.1.4 从程序员的角度看待内存4.2 使用多核和跨服务器进行并行数据处理4.2.1 通过添加资源增加容量4.2.2 并行系统架构4.2.3 企业级应用数据库的并行化4.2.4 SanssouciDB 中的并行数据处理4.3 通过压缩提高速度和减少内存消耗4.3.1 轻量级压缩4.3.2 重量级压缩4.3.3 数据相关的优化4.3.4 压缩感知的查询执行4.3.5 真实数据的压缩分析4.4 列优先、行优先、混合方式--优化数据布局4.4.1 垂直分区4.4.2 寻找最佳布局4.4.3 混合型数据库面临的挑战4.4.4 应用情景4.5 虚拟化的影响4.5.1 分析型工作负载的虚拟化4.5.2 数据模型和基准测试环境4.5.3 虚拟执行与本地执行4.5.4 使用并行虚拟机减少响应时间4.6 技术概念汇总4.7 结论第5章 SanssouciDB 中数据的组织与访问5.1 用于访问内存数据的 SQL5.1.1 SQL 的角色5.1.2 查询的生命周期5.1.3 存储过程5.1.4 数据组织和索引5.1.5 任何属性均可作为索引5.2 凭借数据老化提高性能5.2.1 主动和被动数据5.2.2 老化过程在实现上的考虑5.2.3 销售线索水平分区的用例5.3 高效检索业务对象5.3.1 从数据库中检索业务数据5.3.2 对象数据指南5.4 高效执行业务函数5.4.1 区分业务函数与应用程序函数5.4.2 比较业务函数5.5 处理读优化数据库中的数据更改5.5.1 对 SanssouciDB 的影响5.5.2 合并过程5.5.3 通过单列合并提高性能5.6 只添加、不删除，保持历史记录完整性5.6.1 “只插入”实施策略5.6.2 通过“只插入”操作最小化锁定5.6.3 对企业级应用的影响5.6.4 “只插入”方法的可行性5.7 支持事务数据分析5.7.1 动态聚集5.7.2 无星型模式的分析查询5.8 不停机扩展数据布局5.8.1 行存储中的重组5.8.2 列存储中的动态附加5.9 利用高级日志技术提高业务恢复能力5.9.1 列存储中的恢复5.9.2 行优先数据库的差分日志记录5.9.3 提供高可用性5.10 对混合负载进行优化调度的重要性5.10.1 调度简介5.10.2 混合负载的特征5.10.3 运行时间较短与较长任务的调度5.11 结论第三部分 内存计算技术所带来的改变第6章 应用程序开发6.1 优化 SanssouciDB 的应用程序开发6.1.1 内存应用程序的编程模式6.1.2 应用程序架构6.1.3 将业务逻辑移到数据库中6.1.4 最佳实践6.1.5 视图的图形创建6.2 创新的企业级应用6.2.1 全新分析应用程序6.2.2 运营处理帮助简化日常业务6.2.3 创新用户界面让信息触手可及6.2.4 合并分析与文本搜索6.2.5 基本搜索类型6.2.6 企业搜索功能6.3 结论第7章 即将呈现的真正的商务智能系统7.1 运营数据分析7.1.1 过去的商务智能7.1.2 如今的商务智能7.1.3 将分析从日常运营中分离出来的弊端7.1.4 为分析系统设计的专用数据库7.1.5 分析和查询语言7.1.6 促进商务智能变化的驱动因素7.1.7 未来的商务智能7.2 改变之后如何评估数据库7.2 企业计算基准测试7.2.2 为混合负载量身定制的新基准测试要求7.2.3 日常运营和分析的新基准测试7.3 结论第8章 在云计算中扩展 SanssouciDB8.1 什么是云计算8.2 云应用程序的类型8.3 从提供商的角度看云计算8.3.1 多租户8.3.2 低端硬件与高端硬件8.3.3 复制8.3.4 凭借内存计算技术提高能源效率8.4 结论第9章 内存计算技术革命已拉开序幕9.1 无风险过渡到内存数据管理9.1.1 内存系统和传统系统并肩工作9.1.2 系统整合和可扩展性9.2 客户验证点9.2.1 柏林夏洛特医科大学9.2.2 Hilti9.3 结论关于作者参考文献术语表索引

## 章节摘录

4.1.3内存层次结构的趋势 在过去几年里,存储介质与计算系统其余部分之间在延迟方面的性能缺口继续迅速扩大。

为消除缺1:3并保持性能的高速增长,人们进行了大量的研究工作,开发新的存储技术,以突破传统磁盘存储的限制。

一组称为存储级内存的颇有潜力的新存储技术正在开发过程中。

这种新型存储介质的一般特性包括非易失性、极低延迟、不含旋转机械部件的固态实施、价格低廉以及高能效。

闪存是试图突破传统硬盘限制的存储级内存技术的一个早期版本。

在闪存中,数据存储由浮栅晶体管组成的单元中,这些晶体管即使在切断电源之后仍能保存电流。

由于无需补充电流,因此与易失性的DRAM相比可以降低能耗。

此外,闪存还可以提供大大低于硬盘的快速随机存取时间[87]。

不过,闪存也有一些不足之处。

比如,闪存单元无法直接更新,这意味着必须要先清除,然后才能在单元中存储新值。

另外,向单元中注入电荷的时间(写入时间)要长于读取当前单元状态的时间。

闪存的写入性能取决于存取模式,并且与顺序写入相比,随机写入的性能显著降低。

要利用这种存储技术的潜力,在设计闪存算法时必须要考虑这种不对称的读取,写入性能[104]。

与磁盘存储相比,闪存的主要缺点是其有限的物理耐用性。

与主存(约10<sup>15</sup>次写入)相比,每次写入存取都会轻微损害存储单元,大大缩短闪存的生命周期(约10<sup>4</sup>~10<sup>5</sup>次写入)。

为提升闪存有限的耐用性,人们采用一种称为耗损均衡的技术,尝试将写入访问均匀分布在现有物理地址空间中,以避免出现写入热点。

在传统数据库管理系统环境中,经检验,闪存可以有两种不同使用方案。

首先,闪存技术可像内存一样用做附加缓存层次结构级别。

该方案的问题在于闪存有限的耐用性与缓存频繁的读写存取之间存在矛盾。

人们经常建议将闪存用做持久性存储介质。

在这种情况下,为了获得闪存的全部性能潜力,必须要调整DBMS的数据结构和存取方式,特别是要考虑读取和写入之间的不对称性(参考文献第[104]条描述了一种利用闪存存储特性的记录方法)。

在内存数据库管理系统环境中,闪存及其他存储级内存计算技术会扮演双重角色。

首先,将闪存卷用做主要的持久存储设备,而将磁盘用做备份和存档设备。

内存数据库的只插入模式与闪存的优势相匹配。

在只插入式数据库中,随机写入的次数可能会减少,甚至完全没有。

另外,由于不会发生更新和删除数据,耐用性有限的缺陷也会得到弥补。

其次,闪存存储的低读取延迟可在系统停机,甚至出现故障时确保系统快速恢复。

闪存的另一种使用方案是用做内存映射存储,以保留不常使用的数据或主要用于读取的大型二进制对象。

DBMS可以基于简单的试探方法或用户手工移动的方法,将不常使用的列转移到一个特殊内存区域——闪存上。

这样一来,其可以利用低能耗的闪存来减少主存的数据量,从而在保持性能的同时降低系统的整体能耗。

.....

## <<内存数据管理>>

### 编辑推荐

《内存数据管理（第2版）》通过对技术发展趋势的分析，举出了有关主内存将在数据库系统中占主导地位的可信例证。

业界第一本有关内存数据库系统的权威书籍，可为学生和数据库设计人员等相关人士提供完整的参考信息。

《内存数据管理（第2版）》不仅描述了这项以市场为导向，并极有可能改变整个企业软件市场的德国创新技术，而且还展示了一个工作原型。

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>