

SAP企业信息化与最佳实践丛书

SAP公司
官方指南

THE IN-MEMORY
REVOLUTION

HOW SAP HANA
ENABLES BUSINESS
OF THE FUTURE

内存革命

SAP HANA助力商业未来

[德] 哈索 贝恩·努克特 著 姚念民 李春晖 译



清华大学出版社

SAP 企业信息化与最佳实践丛书

THE IN-MEMORY
REVOLUTION

HOW SAP HANA
ENABLES BUSINESS
OF THE FUTURE

内存革命

SAP HANA助力商业未来

[德] 哈索 贝恩·努克特 著 姚念民 李春晖 译

清华大学出版社
北 京

Hasso Plattner, Bernd Leukert

The In-Memory Revolution: How SAP HANA Enables Business of the Future

EISBN: 978-3-319-16672-8

Translation from English language edition:

The In-Memory Revolution: How SAP HANA Enables Business of the Future by Hasso Plattner, Bernd Leukert

Copyright © Springer International Publishing Switzerland 2015

Springer Berlin Heidelberg is a part of Springer Science+Business Media

All Rights Reserved.

本书中文简体字翻译版由德国施普林格公司授权清华大学出版社在中华人民共和国境内(不包括中国香港、澳门特别行政区和中国台湾地区)独家出版发行。未经出版者预先书面许可,不得以任何方式复制或抄袭本书的任何部分。

北京市版权局著作权合同登记号 图字: 01-2015-8677

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话: 010-62782989 13701121933

图书在版编目(CIP)数据

内存革命: SAP HANA 助力商业未来/(德)哈索(Plattner, H.), (德)努克特(Leukert, B.) 著; 姚念民, 李春晖 译. —北京: 清华大学出版社, 2016

(SAP 企业信息化与最佳实践丛书)

书名原文: The In-Memory Revolution: How SAP HANA Enables Business of the Future

ISBN 978-7-302-42695-0

I. ①内… II. ①哈… ②努… ③姚… ④李… III. ①企业管理—应用软件 IV. ①F270.7

中国版本图书馆 CIP 数据核字(2016)第 018168 号

责任编辑: 陈 莉 高 岫

封面设计: 周晓亮

版式设计: 方加青

责任校对: 牛艳敏

责任印制: 宋 林

出版发行: 清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址: 北京清华大学学研大厦 A 座 邮 编: 100084

社 总 机: 010-62770175 邮 购: 010-62786544

投稿与读者服务: 010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈: 010-62772015, zhiliang@tup.tsinghua.edu.cn

印 刷 者: 北京鑫丰华彩印有限公司

装 订 者: 北京市密云县京文制本装订厂

经 销: 全国新华书店

开 本: 185mm×230mm 印 张: 17.5 字 数: 280 千字

版 次: 2016 年 2 月第 1 版 印 次: 2016 年 2 月第 1 次印刷

印 数: 1~3000

定 价: 88.00 元

产品编号: 066342-01

序 言

Clayton M. Christensen教授
哈佛商学院工商管理Kim B. Clark教授

有时，改变命运的事件似乎在你无法预测的时间和方向上发生。于我而言，这样的一个事件就在2014年发生。当时我正在哈佛商学院的办公室里想事，电话响了。来电话的是哈索(Hasso Plattner)，超级欧洲软件公司SAP的一位创始人。我确实知道他，但我们从没见过面。他直接跟我谈起来，而没有问我有没有时间。我又怎么能问呢？对于我来说，这是一次难得的机会——听一听这位世界软件天才想要告诉我什么。

哈索带来的消息是一个奇异的巧合。我是个基督徒，我通常在想上帝是否对我的生活满意。在最近一次我这么想的时候，我有了一个重要的领悟：上帝在天堂不需要会计。因为我们的脑力是有限的，我们需要把数据聚合为一个大点的数字来描述我们周边发生的事件。举个例子，我记不住每一个发给客户的发票。因此，感谢上帝，我们有会计可以把所有的发票额加起来得到一个数字，我们称这个数字为“销售额”。我们也经常得到其他像这样的数字，如个人预约量、花费、资产、负债等。没有这些数字或聚合，生活中的数据将变得太复杂而不能理解。无论好坏，聚合的需求经常让我们对芸芸众生有一种层次感：在这个世界，掌控更大数字的人比掌控较小数字的人更重要。比如公司里的梯形的职位高低。然而，我意识到因为上帝有无穷的脑力，他不需要做聚合来给每人一个级别，以对世界上发生的事有一个很好的理解。这意味着当他想要衡量我的生活时，他只需跟我讨论我做了什么以帮助他人，因为他不做聚合来评价个人的级别。

带着这样的想法，在我向哈索解释我这个观点后，我非常惊异地听到他说在企业资源规划(ERP)软件中他已经看到了与此相同的模式。任何数据存储系统的架构核心都是聚

合。因为我们只有有限的脑力，我们需要会计和他们的计算机把细节数据整合起来作为聚合。我们使用算法定期地对这些聚合值做加、减、乘、除运算来概述在企业里发生的事件。

聚合帮助执行官来管理一个公司。然而，聚合阻碍执行官去试图改变企业。这是怎么回事？我观察到公司的发展有4个阶段。第一个阶段是公司开发第一个产品的时候，该产品由几个组件组成。在第二个阶段，公司有了一组组的雇员，雇员们为公司成功的产品或服务开发更好的组件。因此，雇员们的分组情况反映了产品的架构。第三阶段，当公司逐渐成功，组织结构图被开发出来，可以清楚地看出哪些人或哪些组负责什么和他们受谁领导。第四阶段，负责组织结构图上某一区域的负责人需要掌握本区域的情况数据——此处聚合即被定义为向执行官提供其所需的数据。

只要聚合数据忠实地反映了企业和市场的情况，基于聚合的会计系统才能给管理者一个有用的、但静态的现实视图。但是聚合的框架没有给我们对公司和市场的一个动态视图。唯一的得到动态数据的方法是等待下一个静态报告发表。ERP系统不能预示公司和市场的未来趋势，因为包含新方向的细节数据隐藏在设计用来反映旧情况的聚合中。搜寻出预示未来的分散数据非常困难，这是根基稳固的领导者在创新上比较慢的关键原因。

在SAP的经济成功上我没有任何经济利益。但从一种学术观点上看，SAP的新HANA系统给管理带来了即使现在看来都是不可想象的事物。原因是HANA没有实现聚合。这就像数据是大理石，被扔到一个非常平滑的地板上。没有哪一段数据被隐藏在另一段数据之下；每一个数据都有一个可见的、永久的位置。我把这种系统称为“瞬时结构化数据”。

下面是一些这种瞬时结构化数据能做的一些“不可思议”的事：

- 研究者还没有开发一个测量理论——选择错误的测量方式能够导致管理者做出不幸的选择。举个例子，毛利润比例是一个广泛使用的成功指标。它使得管理者跳出他们的低端市场，优先发展毛利率高的领域。这造成了公司的混乱。如果管理者选择了另一个指标(如每吨或每车的净利润)，则很有可能使公司呆在市场底部以

防止公司动荡。一个瞬时结构化数据的系统使执行官能够即时地开发出用多个不同方法来衡量成功的报告。

- 我可以想象这样的场景：使用瞬时结构化数据系统建立5个不同的组织结构图，同时使用这些图给雇员们讲解，告诉他们对于项目的不同部分，他们怎样对不同的领导负责；这使得执行官能够立即召集，然后解散不同成员的组，并且为了开发一个新产品，围绕这个系统来组织公司。
- 最后一点，我第一次能看到公司不再围绕经济情况来组织，也不必围绕客户。我可以想象到公司集中力量围绕客户想要完成的任务。因为要完成的任务是长时间稳定的，我可以想象有了瞬时结构化数据，执行官不需要再对公司进行重组。

我感谢哈索和SAP的雇员开发了这个技术，并且使它在市场上能被所有人获得。做个类比，当IBM在1964年开发出了计算机360时，我们那一代对这一技术成就感到敬畏，它使得机器像在变魔法。当IBM开发出个人计算机时，同时代的人同样感到震撼，普通人也能获得计算能力的技术真可以看作奇迹。然而，随着时间流逝，我越来越敬佩在360计算机和个人计算机中的管理成就。几万人的工作和数十亿美元必须在巨型集成公司与小创业公司内部或之间精心安排。我们为少数已知的历代计算机革命的应用感到激动，更为那些未知的应用而激动。

为了这些原因，我对来自SAP的瞬时结构化技术肃然起敬。我赞赏他们的工程实力和管理技巧。我喜欢思考这个技术会给管理者和他们的组织带来什么——一些我们想象不到的事。瞬时结构化数据不能彻底解决管理问题，但我确实相信该技术有潜力把速度、官僚和通用营销等问题变为次要问题。并且现在，考虑难以置信的事情是很愉快的。

前言

为什么我哈索与贝恩·努克特(Bernd Leukert)及其在SAP的几个同事一起写这本书呢？

自2006年以来，我一直在德国波兹坦大学的哈索·普拉特纳(Hasso Plattner)学院工作，致力于开发企业系统的新基础软件。同时，我的研究领域有很大扩展，从医疗保健和交通控制到基因组学和伊波拉病毒控制——但是，企业系统仍然是我的主要研究领域。新企业系统的前提是有一个非常快速的数据库——HANA。现在HANA已经成熟了，SAP可以构建一个新的企业商务系统套件了，它就是S/4HANA。

业界已出现很多相关的文章及著作，其中虽然不乏优秀之作，但大部分作品乏善可陈。因此，我觉得有必要讲讲我自己的故事，说一说什么是我认为在未来很重要的，还要讲一下SAP的实际开发经验，是SAP肩负着将现有客户无间断升级到S/4HANA的责任，并且做了建立新系统的所有困难工作。这需要新一级别的工程质量，我很荣幸有SAP作为同伴。

这本书应使管理人员认识到S/4HANA的潜力，并且在不久的将来，有信心去开始大大小小的项目。S/4HANA的潜力是巨大的，随着本书的分步讲解，我们将看到其在总体拥有成本、用户体验、新型商务过程以及与客户更好的交互等方面渐增的影响。你可能不相信我推测的所节省的总体拥有成本，但这里列出一些事实：数据输入快3~4倍，分析快10~1000倍，扩展模块的开发更快，数据管理工作显著减少，数据存储空间占原来的1/10，通过复制达到无限的负载能力。除此之外，我们有了新的用户体验、新的客户定制和全新的应用套件。

从R/2版到R/3版的改进是巨大的；升级到S/4HANA将有更多益处。

目 录

第 1 章	为了创新的反思	2
1.1	从头开始	6
1.2	从更理论化的角度看企业应用	8
1.3	关于本书	10

第1部分 内存计算技术的基础

13

第 2 章	内存中按列存储数据库的设计原则	16
2.1	核心原则	20
2.2	高级特性	26
第 3 章	HANA对企业应用程序设计的影响	42
3.1	对象和关系	44
3.2	使用零响应时间原则重新设计	50
3.3	在一个系统中进行联机事务和联机分析处理	52
3.4	时间尺度	55
3.5	实时扩展性	57
3.6	数据占用空间减少	58
3.7	大数据	60

第2部分 新型企业应用

63

第4章 新一代商务套件S/4HANA	68
4.1 具有颠覆性技术的无缝技术创新	69
4.2 再次强调，速度很重要	74
4.3 新的用户接口	77
4.4 总体拥有成本低的系统会更好	81
4.5 基于HANA的云端应用	97
4.6 组织变化的影响	101
4.7 硬件系统的考虑因素	102
4.8 可扩展性	107
4.9 解决方案软件包	110
第5章 重建和反思	112
5.1 未来会议室	114
5.2 财务	122
5.3 物流	128
5.4 销售和营销	130
5.5 分析及规划	137
5.6 经验学习	143

第3部分 HANA对于企业应用的关键优势

145

第6章 在合作创新项目中应用设计思考方法	150
6.1 SAP的合作创新项目	151
6.2 设计思考方法是合作创新的推动者	152
6.3 合作创新行动：房地产“驾驶舱”	154
6.4 合作创新行动：医学研究视角	156

第 7 章	降低复杂性	170
7.1	在企业系统中减少数据库占用空间	172
7.2	在云端进行财务规划	175
7.3	Celonis的过程挖掘——过程透明度的新水平	178
7.4	大市场中的信号检测	182
第 8 章	最佳粒度达成最高灵活性	188
8.1	企业模拟	190
8.2	挑战速度	194
8.3	欺诈管理	198
第 9 章	利用机会之窗	204
9.1	全渠道零售	206
9.2	康尼格拉公司用TMM管理从了解到行动的所有过程	209
9.3	自然灾害的风险分析	213
第 10 章	工作时的高质体验	218
10.1	决定消费者行为中的谁、什么和在哪里等因素	220
10.2	所有关键绩效指标一览	227
10.3	在企业数据中导航	231
10.4	通过更好地了解花费情况完美控制供货和采购	234
第 11 章	解决难解问题	238
11.1	智能销售	240
11.2	远程服务和预测性维护	242
11.3	使物流中心高效	247
11.4	在全球追踪货物	250
11.5	理解数字社会中人们的想法	253
第 12 章	结论	258
	缩略语	262

THE IN-MEMORY
REVOLUTION

HOW SAP HANA
ENABLES BUSINESS
OF THE FUTURE

内存革命

SAP HANA助力商业未来

第 1 章

为了创新的反思

在

2006年的秋天，我是德国波兹坦市哈索·普拉特纳研究所(HPI)的一名计算机科学教授。我主要研究企业系统架构，当时我必须为我的博士研究生找一个新的研究领域。当然，他们必须自己寻找论文的题目，但我想要引导他们研究一个我熟悉的东西，即一个新的企业资源规划(ERP)的概念。我所有的职业生涯都在这些系统上工作，我问自己：如果我们可以重新开始，这些系统将是什么样子？

一沓空白纸，一堆铅笔，一个橡皮擦，只有这些工具供我开始，我希望新想法会在我脑海中闪现。但是新想法没有出现，然后我借助谷歌和维基百科在互联网进行了小小的搜索，这个网上调查的目的是看看5年后世界上的计算机将会变成什么样。

在波士顿的麻省理工学院，他们说我们将会CPU速度仅快一点的计算机，但计算机会有很多的CPU。多核结构正在兴起和到来，已经有人预测会有超过1 000个核的系统。直接访问存储，DRAM(Dynamic Random Access Memory，动态随机存取存储器)或者SSD，(容量或速度)都会提高100倍，并且价格越来越便宜。这些已经成为令人鼓舞的事实。在SAP，我们已经有了一个完全在内存中运行的数据库，它作为SAP的一个商业数据仓库的一个加速器来运行，即商业数据仓库加速器(BWA)。这只是证明新型ERP系统必须运行在一个内存数据库上的一小步。但这一小步对于ERP系统的架构意味着什么呢？

自从我们建立这样的ERP系统，刚开始在IBM，现在在SAP，它们都基于这样的想法：我们确切地知道用户想要知道什么。为了在合理的时间内回答用户的问题，我们实时地维护聚合的数据，这意味着每当我们记录一条商业事务，就必须更新所有影响到的相关事务。因此，这种系统随时准备好回答任何可预见的问题，所以称之为实时系统。然后我有了一个新想法，即完全抛弃涉及全部数据的思路，仅仅把每天的事务数据压缩在一起，保持其他数据不变。到此为止，新想法的内容还不多，仅仅一页纸，但这是一个开始。

这个学生完全没有动容，他说：“从学术观点来看，这个压缩率并不可观。”我的新想法，也就是我的所有，被粉碎了。我拿起板擦，擦去白板上的与聚合相关的内容，回答说：“好吧，那么不再聚合。”

这正是我一直在寻找的突破。没有人曾建立一个不进行实体聚合的财务系统，无论是实时更新或是批处理更新。

回到刚才面对的挑战，我问道：“如果我们假设一个数据库总是有零反应时间会怎

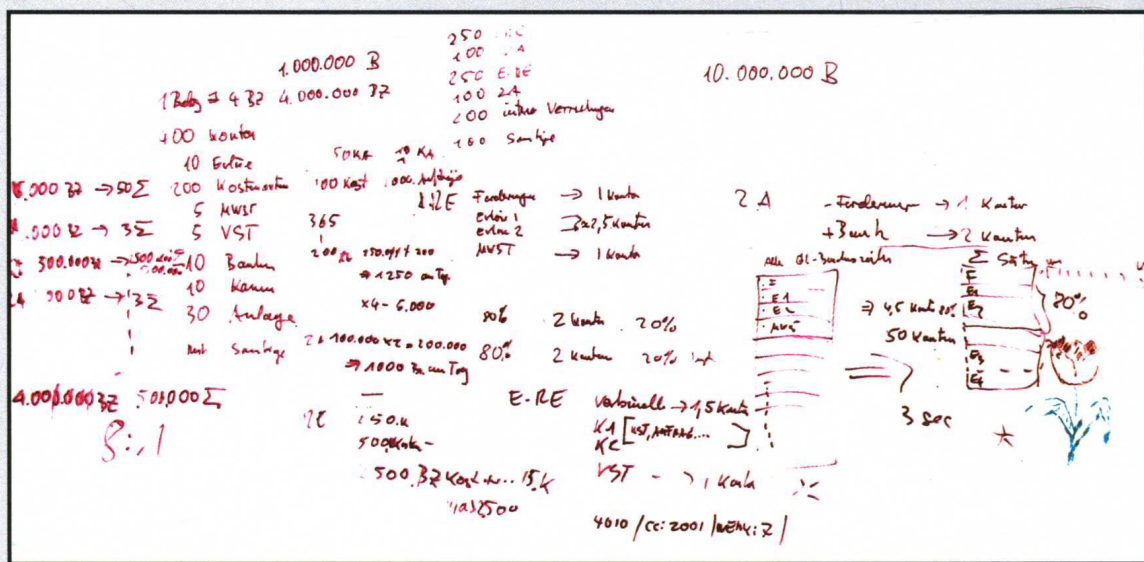


图1.1
HPI的原始白板，HANA的故事从这里开始

样？这样的ERP系统会是什么样的？”这是一个恰当的研究式提问，学术上的工作可以由此开始。但能由此开启一些实验工作吗？我们不是应该检查一下是否有数据库可以接近这个理想状态吗？最终，这个零反应时间是可能的吗？这是在大学里做研究的美妙所在。通常，零反应时间这个想法在SAP就不会被广泛接受。在大学里，你可以做梦，至少做一段时间的梦。只要你能产出有意义的论文，基本上就没有问题。

我们请求SAP允许我们可以使用他们三个数据库的背后技术：TREX，一个按列存储的内存数据库；P*Time，一个按行存储的内存数据库；MaxDB，SAP的关系数据库。我的博士研究生们开始研究这些系统，很快发现建立一个新的ERP系统远没有建立一个新的数据库有趣。经过这个阶段，这些研究生们本来是计算机科学家，但学了一大堆商学院学生的知识，例如会计、销售、采购和人力资源管理。最后的折中是我们要从头建立一个数据库原型，所有我们需要验证超快数据库概念的应用场景，都要与真实的企业系统一致。

原因很简单，有了零反应时间的数据库，我们不再需要在所有企业系统中确保合理反应时间的任何结构。

聚合机制已经去掉了，接下来就要去掉冗余表了，它们显示了聚合机制是如何建立的，还要去掉所有应用中的缓存和所有处理数据库索引的技巧的代码。我们只留下事务数据，其他都将在每次需要的时候再计算。异步插入/更新任务去掉了，这是我1970年在ICI(帝国化学研究所)引入的，它仍然在为IBM工作。在数据库对象更新中不必要的锁操作也去掉了。新的应用程序代码看起来干净、简短。数据库占用内存将会小得多，估计能减小到1/10，这对于我们想把所有数据都放到内存帮助很大。我们的数据库原型仅能做一些必须的功能，但从第一天我们就明确：联机事务处理(OLTP)和联机分析处理(OLAP)应该运行在单个系统上。我们将看到这是建立企业应用的新方法的基础。

为什么没有其他人曾考虑过接近零反应时间的事务和分析处理呢？大概我们太专注于通过索引直接访问，以至于不能想象在大规模事务处理中使用按列存储，或者不能想到一个全部数据不在一个多维的数据立方体中的信息管理系统。

我们获得了匿名化的真实客户数据来进行实验，数据来自一个德国啤酒小公司和一个非常大的包装消费品(CPG)公司。处理啤酒公司的损益表中的所有匿名化记录，仅用了三秒钟。刚开始，大CPG公司的处理用了三秒钟，后来仅用了少于一秒钟的时间。我们更加

自信，实现一个不预先把所有部分聚合的允交(ATP)系统是可行的，并且它在解决潜在冲突中具有很大的优势。

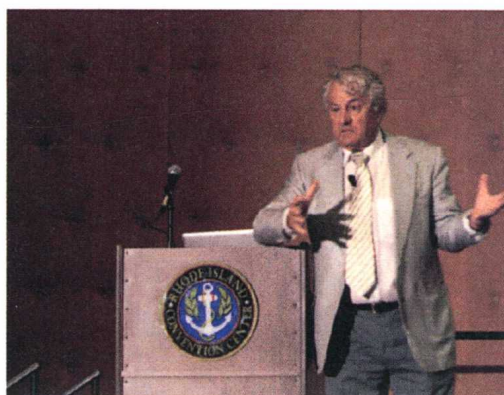
两年后，我们把研究所得发表在2009年在普洛威顿斯罗德岛举行的SIGMOD数据库会议上(见图1.2)[Pla09]。当时已有的数据库生产商的代表们真是很好的听众，会后不久，他们都在自己的主要产品上开始应用内存计算技术。麻省理工学院的一个教授迈克尔·斯通布雷克问我，SAP怎么能够重写4亿行代码。我回答，是的，那是SAP商务套件的代码量，SAP必须找到方法来全部重写。

1.1 从头开始

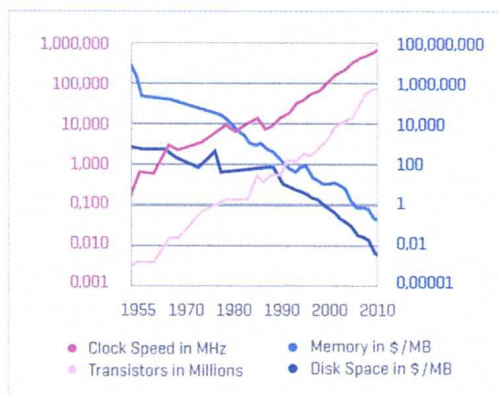
SAP当时听从HPI的建议，开始开发专业的针对OLTP和OLAP负载的按列存储的内存数据库，即HANA。从一个研究型系统到一个产品，中间隔着巨大的鸿沟。理解这个你只要想想需要实现的关系数据库的特点ACID(原子性、一致性、隔离性、持久性)，高可用性，与先锋产品的兼容性，应用开发的架构，等等，而这只列出了一小部分。一些世界范围的数据库专家批评我们的研究和SAP公司的大胆行动。但这只给了我们继续该计划的信心。现在看来，当时SAP找来几个Sybase公

司(一个老牌数据库公司)的数据库专家具有非常重要的意义。

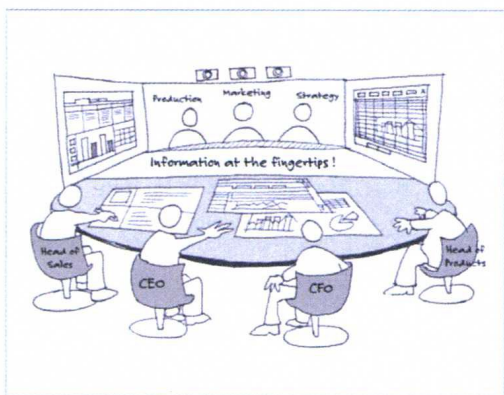
在我的ERP系统梦成真之前，数据库必须先成熟起来，SAP使用了数据库的早期版本作为它的数据仓库产品，即SAP商业数据仓库(BW)的基础。自2012年中开始，SAP一直重写它的企业应用(ERP企业资源规划、CRM客户关系管理、SCM供应链管理、SRM供货商关系管理和PLM产品生命周期管理)，现在，第一版已经发售了，其有着极精简的数据模型，极高的性能提升以及拥有新的应用功能。人力资源管理(HCM)与SuccessFactors的应用合并了，因此这是另一个项目。从SAP ERP Central Component 6.0(ECC 6.0)开始的所有ERP系统都可以无缝升级，这本身是一个奇迹，但如果你能充分理解SQL(结构查询



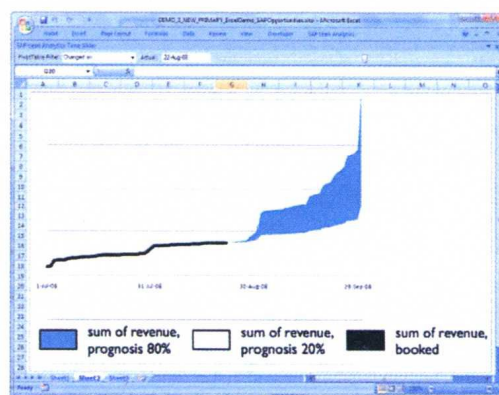
演讲



硬件趋势



未来的董事会会议室



现金预测原型

图1.2
SIGMOD数据库会议上的图片，显示了HANA背后的研究结果和第一次展示的事务系统的分析处理示意图