

SAP企业信息化与最佳实践丛书

SAP公司
官方指南

数据流管理系统设计与应用

——SAP ESP 项目实战指南

Data Stream Management System
Design and Application

罗永强 付峰 马泽炯 / 著



清华大学出版社



SAP企业信息化与最佳实践丛书

数据流管理系统设计与应用—— SAP ESP 项目实战指南

罗永强 付峰 马泽炯 著

清华大学出版社

北 京

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

数据流管理系统设计与应用——SAP ESP 项目实战指南 / 罗永强 等著. —北京: 清华大学出版社, 2015

(SAP 企业信息化与最佳实践丛书)

ISBN 978-7-302-38728-2

I. ①数... II. ①罗... 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(2014)第 284242 号

责任编辑: 陈 莉 高 岫

封面设计: 周晓亮

版式设计: 思创景点

责任校对: 曹 阳

责任印制: 李红英

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: 清华大学印刷厂

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: 175mm×226mm 印 张: 21.5 字 数: 506 千字

版 次: 2015 年 2 月第 1 版 印 次: 2015 年 2 月第 1 次印刷

印 数: 1~4000

定 价: 45.00 元

产品编号: 055457-01

SAP企业信息化与最佳实践丛书

编 委 会

总 编 张志琦

执行主编 和轶东

编 委 王 强 黎文宪 彭俊松

鲁百年 Elke Simon-Keller(爱可)

刘 涛 毛 彦 曹乃刚

李国裕 陈 今 郭 磊

丛书总序 一

SAP是一家具有40年历史的全球最大的企业管理和协同化商务解决方案供应商。作为全球领先的企业应用软件解决方案提供商，SAP帮助各行业不同规模的公司提供商务软件解决方案和咨询服务，以帮助它们成为最佳运营企业。从企业后台到公司管理层、从工厂仓库到商铺店面、从电脑桌面到移动终端——SAP助力用户和企业更加高效地协作，同时更加有效地获取商业洞见，在竞争中保持领先地位。

目前，SAP的应用程序和服务帮助全球120多个国家的超过197 000家客户实现运营盈利、不断完善和稳步增长。《财富》500强80%以上的企业正在从SAP的管理方案中获益。10多年来，我们深耕中国，服务超过1 500家企业，其中超过80%是本地客户，70家以上的合作伙伴与我们维持稳定紧密的合作关系，2 000多名经过认证的顾问为我们的客户带来高品质、本地化的服务及商业价值。

SAP中国的发展价值与目标十分明确——建立互信、互助、互动、互惠的行业生态系统。我们借由与合作伙伴密切合作，将先进的管理理念与方法转变为切实帮助中国企业迈向成功、走入世界的现实。我们凭借与政府的携手配合，贴近国家经济与社会发展的步伐，积极参与行业相关标准与法规制定程序，并引入SAP几十年来在全球不同市场的最佳业务实践，从根本上协助政府注入创新技术与运营模式，从而推动国家信息化建设，完全秉持“在中国、为中国、与中国一起发展茁壮”的一贯信念。

SAP拥有多元且全面的产品与解决方案。在广度上，我们可以服务于不同行业、规模、类型的企业；在深度上，客户企业的各业务线与各类信息化需求都能通过缜密的部署实施得到满足。从后勤办公室到董事会、仓库到店面、桌面设备到移动设备，SAP支持人员和企业提高合作效率，更加高效地利用业务洞察力来保持竞争优势。我们可以通过企业预置、按需随选、移动应用等来扩展软件的可用性，从



而更加灵活地实现上述目标。

我们希望通过与清华大学出版社合作的《SAP企业信息化与最佳实践》丛书，帮助您更进一步地了解SAP先进的方案、产品、技术和SAP中国，同时期望您的许多关于信息化管理、技术应用以及企业运营的疑问能够获得解答。

我们的使命是帮助各种规模及行业的公司更卓越地运营。我们的愿景是帮助世界更卓越地运转。我们期待与您一起“蕴韬略，更卓越”！

萧洁云

SAP全球高级副总裁、中国区总裁
2012年12月于北京

丛书总序 二

长期以来，坚持不断创新和持续发展使SAP当之无愧地成为行业的领军者。SAP一直基于领先的产品技术、丰富的实践积淀和优质的保障服务，为中国企业提供企业预置、按需随选、移动应用等协调运营的企业管理解决方案，以满足不同行业、不同规模和不同类型企业的信息化需求，为客户提供高端业务洞察力与全方位管理企业商务智能和企业信息的技术平台。

SAP面向企业的行业需求，量身定制行业解决方案。融合了SAP及其合作伙伴40年的“最佳业务实践”，不仅仅为众多制造行业企业提供富有效率的工具和最佳业务实践经验，使得从行业领袖到不同规模大小的企业从中受益，同时更为包括电力、电信、物流等非制造业企业提供面向其特殊需求的解决方案，从而帮助他们提高管理水平，提升效益，增强竞争能力并制定正确的商业决策。

有别于市场上的其他管理软件供应商，SAP公司提供的SAP最佳业务实践，是一套整体的、可配置的、端到端的业务财务一体化解决方案。通过SAP最佳业务实践可以帮助企业管理财务、人力资源、采购、库存、制造、物流、新品开发，贯通从企业的内部服务到外部客户服务、连接销售和营销等各方面的商业事项。

SAP最佳业务实践为商务流程提供的支持和方法功能反映了SAP解决方案在超过125个国家，跨26个行业，超过40 000次实施中获得的经验教训，以利于企业实施和消费SAP系统软件的功能。借助SAP最佳业务实践，企业可以免去软件实施项目中最复杂的实施步骤，并且避免大量资源的投入。过去在启动项目时仅能开启很少的基础功能，现在企业可以通过SAP最佳业务实践获得丰富的行业知识，外加合作伙伴的行业知识与经验，支持标准的行业和跨行业流程，专为您的业务需求而预先配置。

一直以来，SAP通过产品和服务推动全球企业创新。我们相信，其将帮助全世



界的企业实现增长并创造新的价值源泉，最终惠及整个行业和经济。正是基于不断创新这样一个最基本的理念，SAP在2012年把创新分成五大市场，分别是移动解决方案、商务分析、企业应用、数据库与平台技术以及云计算。

移动应用无处不在，移动应用改变未来。SAP和Sybase(SAP旗下公司)一起致力打造无线企业，为客户提供广泛的移动应用和基础设施，将企业的人员、信息、移动设备整合起来，让工作者随时随地、在任何设备上安全地访问业务流程，实现企业移动化变革。SAP在其强大的移动应用平台基础上，不仅提供了移动销售、移动工作流等系统级移动应用以拓展企业CRM和ERP，还提供了专业的移动商务智能应用扩展企业决策支持的覆盖面。SAP在各应用和技术领域不断创新，为企业创造短、平、快的无线增值应用，以帮助企业能够快速实施、快速部署、快速实现应用价值。

SAP商务分析系列解决方案包含商务智能解决方案、信息管理解决方案、企业绩效管理解决方案，以及治理、风险管理及合规解决方案。其中，商务智能和信息管理解决方案均构建在一个得到市场公认的智能平台之上。SAP商务分析解决方案组合有助于各类机构有效地做到合规和应对业务风险，缩小战略与执行之间的差距，从而实现优化业务绩效的目标。采用SAP商务分析系列解决方案，企业将获得更为广泛的洞见，同时将各项战略有机地结合起来，优化决策流程，改善风险管理，而这些都不受底层特定事务系统的约束。SAP商务分析解决方案让企业把对业务的洞察转化为策略，再将业务策略转化为能够即刻执行的运营决策。

SAP企业应用的核心是SAP商务套件。SAP商务套件是一套真正完整集成的解决方案和应用平台。它提供开放的商务应用软件，通过集成人员、信息和流程来获得最大的回报。SAP商务解决方案使企业从容应对不断变化的客户和市场需求，使得企业成功地驾驭与客户、供应商、后勤伙伴、金融服务商以及员工的关系，并从这些关系中获得效率和收益。它包括功能强大的跨行业方案和26个行业解决方案。SAP商务套件可以优化企业的业务和IT战略，使其更有效率，更具灵活性，提高企业对业务的洞察力。其独特优势有：丰富的应用解决方案能针对每个行业的业务流程提供持续的管理创新；融合最佳行业实践的管理理念可以促进企业管理核心业务绩效的改善；其功能增强包可以帮助企业在不停机的情况下在线升级，灵活地应对业务需求与状况的改变，以更低的整体拥有成本实现企业创新；SAP平台能无限扩展端到端的流程，在确保流程的完整性的同时增强业务和IT效率；新套件基于企业级SOA可以全面提高IT对业务的敏捷性；其内置的分析功能能够实现对业务清晰透

明的洞察和实时监控；SAP统一协调的用户界面能够有效地提升最终用户使用系统的效率。

SAP的数据库与平台技术的核心是业界领先的内存计算平台与解决方案——HANA。SAP HANA提供多用途的内存应用设备，企业可以利用它即时掌握业务运营情况，从而对所有可用的数据进行分析，并对快速变化的业务环境做出迅速响应。使用SAP HANA，企业可以即时访问相关信息，更快地做出更加可靠的决策，并降低获取洞察力时对IT部门的依赖。SAP HANA提供灵活、节约、高效、实时的方法管理海量数据。利用它，企业可以不必运行多个数据仓库、运营和分析系统，从而削减相关的硬件和维护成本。

云计算已经深远地改变了人们互动的方式。随着各企业以全新方式与其全球客户和合作伙伴互联和协作，这种影响力将更加难以估量。基于云的协作重新定义了业务网络创新的概念，而SAP在这一变革早期便已成为弄潮儿。随着SuccessFactors和Ariba的加入，SAP可以为企业创建业务网络，为客户实现立竿见影的价值。同时SAP在成长型企业信息化领域拥有领先的理念、技术、产品和实践经验，推动在云端部署的解决方案，为广大成长型企业用户带来符合自身需求的最佳实践。

SAP致力于通过创新来提高全球企业的运营效益。这五大市场的创新是我们的未来发展的指导愿景，是我们通过思考自己在全球市场中应该扮演角色后做出的。通过五大创新，我们希望创造一个更加高效的世界，所有企业都能在其中找到和把握新的机会，同时平衡风险，从而实现赢利性增长。

我们始终相信“立足中国，创新中国”！通过清华大学出版社《SAP企业信息化与最佳实践》丛书的出版，我们希望中国更多的企业管理者、IT从业人员以及院校师生能更好地了解SAP和SAP的解决方案。

欢迎您走近SAP！

张志琦

SAP中国区方案与架构设计部总经理
2012年12月于上海

前言

一提到 SAP 这个名字，大家首先想到的是 ERP 和 CRM，想到的是企业级应用软件、企业管理解决方案等。是的，SAP 一直基于领先的技术、各行各业丰富的实践经验以及优质的服务，帮助各行业不同规模的企业实现卓越的运营和管理。从企业后台到公司决策层、从工厂仓库到商铺店面、从桌面设备到移动终端——SAP 助力用户和企业高效协作，获取商业洞察力，并从竞争中脱颖而出。SAP 的软件和服务能够帮助客户实现盈利性的运营，不断提升应变能力，实现可持续的增长。

自从 1972 年成立以来，SAP 一直专注于企业应用和管理，有了自己独特的最佳实践和服务理念。但在最近的几年中，IT 技术不断发展，也促成企业应用需要更加开放，更具有多样性，更加适合当今最新的 IT 技术，例如大数据、云计算和实时分析等。从 2007 年开始，SAP 并购了一系列业界领先的公司，容纳更广阔的 IT 技术于企业应用之中，并加强在 IT 基础架构、数据分析、企业移动化等诸多领域。这就好比一棵笔直生长的大树，它一直在蹿个儿，换句话说，原来只做企业应用，而现在需要枝繁叶茂，需要在数据分析、企业移动化方面大展宏图。

从 2012 年开始，SAP 把创新分为了五大市场，分别是企业应用、商务分析、企业移动化、数据管理与技术平台和云计算。在这五大市场中，企业应用是 SAP 一贯坚持的方向，而其他四个方面则是适应当今 IT 技术的发展而发展起来的。例如：自从 2010 年 SAP 收购 Sybase 公司以后，SAP 采用 Sybase 移动化解决方案，使得移动化应用无处不在，我们可以在恰当的时间、适当的地点，通过移动应用随时随地发送和接收信息，完成工作上的要求。通过提供移动销售、移动工作流等系统级移动应用，方便地拓展了企业 ERP 和 CRM 的使用，延伸了企业应用的范围，提高了工作效率和水平。再譬如，随着 SuccessFactors 和 Ariba 加入 SAP 大家庭，以云计算为核心的企业应用不断涌现，



企业可以以全新的方式来运营，而不必遵循传统的管理方式。SAP 完全可以通过托管、租赁等不同的方式，把优质和创新的应用交付给不同规模的企业，尤其是成长型企业，为其部署在云端的解决方案，为广大成长型企业用户带来适合自己需求的最佳实践。

2007 年 SAP 融入 BusinessObject 公司，使得其在商务智能、信息管理、企业绩效以及治理、风险管理与合规等方面有了自己完整的解决方案。采用 SAP 商务分析系列解决方案，企业可获得更为广泛的洞察力，可以优化决策流程，改善风险管理等，同时让企业把对业务的洞察力转化为策略，再转化为运营决策。另外，SAP 商务分析解决方案既能支持传统 ERP 系统中的数据分析，也能支持大数据的智慧分析。

在 SAP 这棵大树上，有一个数据管理方面的重要分支，那就是流式数据管理。说到流式数据管理，就不得不提 CEP(Complex Event Processing)复杂事件处理。在 CEP 的领域，国外已经有了很多研究成果和相当成熟的产品，而在中国才刚刚起步。流式数据处理是 CEP 的一个核心技术，流计算来自一个信念：数据的价值随着时间的流逝而降低，所以事件出现后必须尽快地对它们进行处理。那么什么是复杂事件呢？复杂事件是一种事件，通常多个简单事件发生时预示着这种事件的发生。一个人发球，一个人接球，另一个人扣球，每个都是彼此孤立的事件，在一起组成了排球比赛这个复杂事件。

在复杂事件处理中，数据是不断变化的，而其“操作”是“静态”的。这里说的“操作”是用户预定义或订阅的规则，根据预定义或订阅的规则来发现可能原本不被注意的模式和趋势(即复杂事件)，从而使用户有能力建模、确认、预计及迅速响应看似无关的事件所带来的机遇和威胁。换句话说，复杂事件处理具备了分析高速数据流并鉴别重要事件的能力，而这些重要事件通常是按照一定顺序或在特定时间窗口内发生在多个数据点的组合，虽然对这些事件的鉴别过程是复杂的，但结果却是无价的。复杂事件处理能够帮助企业及时全面地洞察市场变化，降低风险和提高决策效率。

我们需要一个复杂事件处理(CEP)平台提供专业开发应用的功能，它使得企业能够以超过每秒 100 万条消息的处理速度分析数据流，其延迟时间为毫秒级或更少，从而提供一个快速应用开发和实施的平台。就如同其他平台一样，一个 CEP 平台也能够加快开发，降低成本，消除对专业开发技巧的依赖，可以说，CEP 平台使得我们摆脱既不能自己开发又无处购买的二难境地。

SAP ESP 就是一个 CEP 的开发和运行平台，基于这个平台能够创建和运行自己的复杂事件处理应用程序，以便于从实时的流数据中获得连续的智能分析结果。早在 2008 年 3 月，Sybase 基于 Coral8 内核发布了 Sybase CEP 产品，但 Coral8 在 2008 年 5 月被 Aleri 公司收购。异军突起的 Aleri 在 2009 年，被 Forrester 列为 CEP 领域杰出的领

军企业，同时，Aleri 还是首家获准作为安全技术评估中心基准的 CEP 厂商。随后在 2010 年 2 月，Sybase 收购了 Aleri，成功跻身成为 CEP 领头羊并强化实时分析平台。在 2011 年 9 月，也就是 SAP 收购 Sybase 之后，SAP 发布了基于“Aleri 架构升级+CCL(Coral8)描述模型”的 SAP ESP 产品(其过渡产品曾经命名为 SAP Sybase ESP)。

在实时分析能力上，SAP Sybase ESP 解决方案在高速度、高容量的环境下极具价值。原因就是 SAP ESP 使用与传统商务智能系统完全不同的方法来进行数据分析。传统的方法是——捕获数据，将其结构化为一种固定的格式，将数据存储在于基于磁盘的数据仓库中，然后定期在这些数据之上运行报表，进行数据分析。SAP ESP 与之不同的是它在数据到达时通过规则检查数据，所有的操作都不落地，因为一旦落地就失去时效性。SAP ESP 鉴别重要事件的能力几乎是实时的，并能从最新的数据中抽取出商务智能，形成多种用途。

为了满足企业对能支持实时分析、高速交易的应用程序的需求，SAP ESP 提供了强大的复杂事件处理解决方案。它能够从多个数据源中吸收新的数据，并在数据之间创建关系从而能鉴别出代表机会和问题的模式。在事件存储之前，能实时持续分析多个事件流。它通过使用大量原始而重要的数据流，将跨多个流的事件关联起来，从而鉴别出新的关联业务条件，并在几毫秒(或更短)的时间内对这些事件做出响应。正因为具备了访问流式数据、实时数据和历史数据的能力，SAP ESP 产品能迅速地让智能决策生效。另外，SAP ESP 同样可用来生成简洁、信息丰富的数据流，并将它们分发至下游的系统、应用程序和历史库中。

总之，希望大家在读完本书后，能够对 CEP 领域和 SAP ESP 解决方案有一个清晰的认识，结合自己的工作和生活，灵活运用这些 CEP 方法和技术，为企业的实时分析、实时决策以及业务运营带来应用的新理念、新方式。

本书由 SAP(北京)软件系统有限公司罗永强带领中国移动(深圳)有限公司付峰、SAP(北京)软件系统有限公司马泽炯共同编写完成。中国移动(深圳)有限公司敖锦蓉也参与了编写工作，在此表示感谢。

宋一平

SAP 中国数据管理技术首席架构师

2014 年 4 月

目 录

 第一章 数据流管理系统	1
第一节 数据库管理系统	1
第二节 数据处理新问题	4
第三节 数据流管理系统	7
第四节 SAP 事件流处理器	10
第五节 SAP 实时数据平台	20
 第二章 SAP 事件流处理器安装	23
第一节 规划安装	23
第二节 安装场景	27
 第三章 SAP 事件流处理器快速入门	39
第一节 事件流处理器工作台	39
第二节 可视化编辑	43
第三节 CCL 编辑器	49
第四节 加载项目	53
第五节 运行测试项目	55
第六节 工作环境配置	59
 第四章 SPLASH 程序开发	61
第一节 SPLASH 程序语言	61
第二节 SPLASH 语句	66



第三节	SPLASH 数据结构	70
	第五章 CCL 程序设计	81
第一节	CCL 项目基础	81
第二节	CCL 语法	87
第三节	CCL 查询语句	94
第四节	高级 CCL 程序设计	101
第五节	声明	107
第六节	从句	118
第七节	查询	123
第八节	函数	131
	第六章 SAP 事件流处理器实用程序	197
第一节	概述	197
第二节	事件流处理器服务器实用程序	197
第三节	命令和控制实用程序	201
第四节	发布和订阅实用程序	203
第五节	开发实用程序	212
	第七章 高级程序设计	215
第一节	实体生命周期和访问模式	215
第二节	发布	219
第三节	订阅	224
第四节	故障恢复处理	225
第五节	应用程序接口(API)参考	226
	第八章 SAP 事件流处理器安全配置	229
第一节	身份认证	230
第二节	访问控制	246
第三节	SSL 安全链接	256
第四节	密钥及密钥库	256

第五节 数据存储安全	259
第六节 其他安全性相关配置	267
 第九章 SAP 事件流处理器集群	273
第一节 集群架构	273
第二节 多项目支持	276
第三节 高可用性	276
第四节 项目部署选项	277
第五节 集群持久化(Persistence)和缓存(Caching)	280
第六节 集中安全	281
第七节 集群启动	281
第八节 集群配置	282
第九节 项目配置文件示例	290
 第十章 SAP 事件流处理器应用开发	293
第一节 数据流计算框架	293
第二节 应用开发流程	297
第三节 开发调试环境	317

第一章 数据流管理系统

第一节 数据库管理系统

从 20 世纪 40 年代第一台计算机 ENIAC 出现, 计算机主要用于科学计算, 早期计算机处理数据包含在程序中, 数据通过输入设备输入到计算机, 利用中央处理单元进行计算后, 数据再输出到输出设备。由于受到技术发展限制, 早期的计算机外部存储设备主要是穿孔卡片和穿孔纸带, 在数据处理期间频繁访问外设数据存在困难。这也限制了计算机在数据处理方面的用途, 直到 1951 年雷明顿兰德公司 (Remington Rand Inc.) 研发的一种叫做 Univac I 的计算机推出了磁带驱动器, 可以快速存储一定量的数据记录, 从而引发了数据管理的革命。

随着人们对数据处理和数据共享的要求越来越高, 如何对数据进行存储管理出现了大量的理论研究, 也就出现了网状数据模型、层次模型和关系模型。而进行统一管理和共享数据的数据管理软件也就应运而生, 这就是我们说的数据库管理系统 (DBMS)。根据其在进行数据储存管理时采用的不同模型, 被分成网状数据库、层次数据库和关系数据库三类。

数据库发展到今天, 使用最为广泛, 也是最重要的数据库是关系型数据库, 1970 年埃德加·科德首先提出关系模型理论, 并配合“科德十二定律”。关系模型概念清晰, 结构简单, 实体、实体联系和查询结果都采用关系表示, 用户比较容易理解。另外, 关系模型的存取路径对用户是透明的, 程序员不用关心具体的存取过程, 减轻了程序员的工作负担, 具有较好的数据独立性和安全保密性。关系数据库是建立在关系模型基础上的数据库, 借助于集合代数等数学概念和方法来处理数据库中的数据。现实世界中的各种实体以及实体之间的各种联系均用关系模型来表示。标准数据查询语言 SQL 就是一种基于关系数据库的语言, 这种语言执行对关系数据库中数据的检索和操作。

1970 年埃德加·科德关系模型建立之后, IBM 公司推出了著名的 System R。该系统全面验证了关系型数据库管理系统的可行性, 是第一个实现了结构化查询语言 SQL 的数据库管理系统。1973 年加州大学伯克利分校的 Michael Stonebraker 和 Eugene Wong 利用 System R 开发开源的关系数据库系统 Ingres。而 IBM 公司为保护



其层次数据库 IMS, 直到 1980 年才投产 System R, 因此一般把 Ingres 称为第一个关系型数据库。

一、联机事务处理和联机分析处理

自从关系数据库出现以后, 大量的信息系统利用关系型数据库来存储和管理业务数据, 并建立相应的应用系统来支持企业日常业务运作。这种应用以支持企业日常业务处理为主要目的, 被称为联机事务处理(OLTP, On-line Transaction Processing)应用, 它所存储的数据被称为操作数据或者业务数据。

随着数据库技术的广泛应用, 企业信息系统产生了大量的数据, 如何从这些海量数据中提取对企业决策分析有用的信息成为企业决策管理人员所面临的重要难题, 但面向联机事务处理的应用系统对分析处理的支持一直不能令人满意。因此, 人们逐渐尝试对 OLTP 数据库中的数据进行再加工, 形成一个综合的、面向分析的、更好地支持决策制定的决策支持系统(Decision Support System, 简称 DSS)。联机分析处理(On-Line Analytical Processing, OLAP)的概念最早是由关系数据库之父 E. F. Codd 于 1993 年提出的, 他同时提出了关于 OLAP 的 12 条准则。OLAP 的提出引起了很大反响, OLAP 作为一类产品同联机事务处理 (OLTP) 明显区分开来。表 1.1 列出联机事务处理和联机分析处理的区别。

表 1.1

比较项目	联机事务处理	联机分析处理
使用用户	操作人员, 低层管理人员	操作层人员、管理层人员、决策人员、高级管理 人员
支撑功能	日常业务操作	日常操作分析、短期战术分析、长期战略分析
应用设计	面向应用	面向主题
存储数据	当前的, 最新的, 细节的, 二维的, 分 立的	历史的、聚集的、多维的、集成的、全面的
处理数据	读/写数十条记录	读百万条到上亿条记录
应用形式	简单的事务处理	复杂的查询统计、多维分析、数据挖掘
数据库大小	1GB~1TB	100GB~1PB

二、数据仓库

当出现联机分析处理相关理论后, 数据分析应用于日常内部管理和外部市场竞争, 让数据发挥其价值以外的价值。人们初步设想从数据中提炼知识, 却发现难以实现, 而原因往往是找不到完整的数据, 因为业务数据存储分散在分散的、异构的数据库中。所以人们开始思考把数据格式统一并集中到一起的方法, 这就是数据仓库