

资深SAP认证顾问倾力打造
为提升企业商务智能水平提供解决方案

谢爱华 编著

SAP

商务智能应用开发高级案例教程

SAP

Business Intelligence Functional Development & Enhancement



机械工业出版社
CHINA MACHINE PRESS

SAP 商务智能实用开发与高级功能详解

谢爱华 编著



机械工业出版社

本书主要介绍 SAP 商务智能和数据仓库建设管理,其内容不限于对 SAP BW 软件产品本身的介绍,更在于结合作者的实际项目实施经验,站在企业信息管理的角度对整个建设过程进行思考。

本书适用于企业 IT 经理、项目经理、商务智能工作者、数据仓库工作者以及与之相关的业务工作人员,同时也适合相关人员在实际项目过程中作为工具书使用,指导商务智能和数据仓库系统架构,加强系统管控,从一开始就为系统进行顶层设计,避免重复建设和无序管理,降低系统的总体拥有成本。

图书在版编目(CIP)数据

SAP 商务智能实用开发与高级功能详解/谢爱华编著. —北京:机械工业出版社, 2015.9

ISBN 978-7-111-50218-0

I. ①S… II. ①谢… 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(2015)第 100512 号

机械工业出版社(北京市百万庄大街 22 号 邮政编码 100037)

策划编辑:丁 诚 责任编辑:丁 诚

责任校对:张艳霞 责任印制:李 洋

三河市宏达印刷有限公司印刷

2015 年 6 月第 1 版·第 1 次

184mm×260mm·25.75 印张·636 千字

0001-3000 册

标准书号:ISBN 978-7-111-50218-0

定价:79.00 元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

电话服务

网络服务

服务咨询热线:(010) 88379833

机工官网:www.cmpbook.com

读者购书热线:(010) 88379649

机工官博:weibo.com/cmp1952

教育服务网:www.cmpedu.com

封面无防伪标均为盗版

金书网:www.golden-book.com

前 言

SAP BW (Business information Warehouse, BW) 业务信息仓库是 SAP 公司推出的数据仓库产品, 是实施企业商务智能系统的重要组成部分, 也是商务智能前端展现的数据基础。由于 BW 能够和 SAP 的相关解决方案, 如 SAP ERP (Enterprise Resources Planning, ERP) 系统、SAP CRM (Customer Relationship Management, CRM) 系统等, 进行天然无缝的数据集成, 同时也能很好地集成其他应用系统, 因此越来越多的企业选择实施和应用 BW, 以构建企业的数据仓库解决方案。

在 BW 数据仓库中, BW 本身提供了标准的“最佳业务实践”, 也就是通常所说的业务目录 (Business Content, BCT)。在标准的业务目录中, SAP 提供了除用户名以外的几乎涵盖所有与业务流程相关的数据仓库对象, 其中包括数据源、信息包、转换、数据传输进程、数据存储对象 DSO、信息块、多信息提供者、信息集、查询、权限对象、甚至是角色, 因此, 大大降低了构建和运行数据仓库的企业总体拥有成本。但是, 由于标准的业务目录不能满足用户所有的需求, 因此在项目实施过程中必须基于 BW 标准功能进行增强和开发, 以满足用户独特的需求。

由于 BW 系统本身是基于 SAP ABAP 应用服务器和 SAP JAVA 应用服务器进行设计和开发的, 因此, 项目中大部分的增强或个性化开发都需要使用 SAP ABAP (Advanced Business Application Programming, ABAP) 语言。但是, 在项目的实施和后续维护过程中, 很多顾问或 BW 的从业者已经在个人脑海中建立起了固定的、对于 BW 涉及的 ABAP 语言程序的喜好, 而对自己不熟悉的内容会说“这个需求 BW 系统满足不了”。这种喜好是因为这些人没有从 BW 的实施过程中获得足够多的 ABAP 语言知识, 因此, 作者建议所有的 BW 顾问或从业者都能参与 SAP 标准的 ABAP 培训课程, 以获得日常工作所需的所有 ABAP 语言知识, 因为具有丰富经验的 BW 顾问能够从大量的系统 BADI 对象增强接口和 User Exit 用户出口中对特定用户的特定需求进行完美的实施 (而不是通过大量的其他工作, 甚至是系统外的工具来实现既定的目标), 从而大大减少项目的工作量、降低实施难度, 以完美地满足用户需求, 这在项目实施中尤为重要。

本书的目的就是尽量将 BW 系统中的功能增强点列示出来, 并且说明这些接口的应用场景、数据接口规格、具体使用方法及案例, 使本书的读者在阅读完本书之后能够对后续的 BW 相关工作有所帮助, 也为 BW 顾问的项目实施提供一份技术参考。

本书结构和通常实施 BW 数据仓库的过程是紧密相关的, 也和 BW 数据仓库的数据流向紧密相关, 也就是从数据抽取、数据转换、数据加载、数据管理、数据展现等环节逐一展开进行讲解。

第 1 章主要讲解计算机技术的发展、数据处理的历史过程和数据仓库的由来。

第 2 章主要对 SAP 公司以及 SAP 公司的相关产品进行简要介绍, 同时, 重点对 BW 相关的基本概念做详细介绍, 如商务智能、数据仓库、星形模型、特性、关键指标、数据提供者、信息块和数据存储对象 (DSO) 等。

第3章在对 SAP ABAP 语言和 SAP 增强的概念进行简要介绍的同时,还对 BW 系统的增强以及在 BW 系统中使用 ABAP 语言的注意事项和性能提升建议进行了描述。

第4章主要讲解在 ERP 源系统中设置和增强数据源的各种方法。

第5章主要讲解在 BW 数据仓库中数据上载过程的各种增强,包括信息包的增强、转换的增强、数据传输进程的增强等。为了覆盖之前的版本,作者在最后也对 BW 3.x 的相关增强进行了说明。

第6章主要讲解数据仓库管理中的对象,包括对信息块的访问,以及在虚拟信息块和虚拟关键指标等数据管理过程中的增强。

第7章主要讲解 BW 数据展现过程中的增强。

第8章主要讲解使用 BW - BPS 和 BW - IP 进行企业全面预算管理过程中会涉及的 ABAP 例程增强的开发。

第9章主要讲解 BW 系统建成上线后,在后期维护过程中通常会使用的一些 ABAP 应用场景,包括处理链的调度和监控,以及数据的管理。

第10章主要将与 BW 数据仓库系统相关的、重要的 BADI、UserExit、函数、程序、程序类以及相关重要的表格等进行了罗列,这也是作者多年实施 BW 数据仓库的经验总结。

最后,附录列举了本书中涉及的相关技术术语的对照查询索引。

本书的 ABAP 语言逻辑和示例都是基于 BW 7.01 进行设计和说明的,但是这些增强点和实施方法同样适用于 BW 3.x 以及当前的最新版本 BW 7.4,只是部分代码段需要做一些调整。

为了使本书简单易懂,本书的相关代码段都尽量不使用面向对象语言 ABAP OO 的方式进行描述,因为根据作者的经验,很多顾问或用户对 ABAP OO 的知识了解甚少,同时也因为在 BW 中,使用 ABAP OO 的优点不是特别明显,所以作者更倾向于使用功能模块,也就是 Function Module 的方式进行代码的编写。当然,这对于那些已经熟悉掌握了 ABAP OO 程序编写方式的人来说,看懂本书和按照本书举例的代码去实施项目同样不是一件太难的事情。但是,在本书的第8章中,由于 BW - IP 预算和计划中 ABAP OO 新技术所带来的冲击,不可避免地涉及了一些 ABAP OO 的代码,所以相关代码也为初学者做了一步一步步地详细讲解。正是因为最近几年 ABAP OO 技术的突飞猛进,所以作者建议所有的 SAP 从业者都应该掌握或至少了解 ABAP OO 技术。

目前,商务智能和数据仓库都还在不断地飞速发展,新的概念和理念在不断地涌现,作者希望通过本书,能够给从业者一些建议和指引。限于作者的水平,书中难免存在疏漏,甚至是错误,欢迎广大读者批评指正。作者私人邮箱为 aihuaxie939@163.com,有任何意见或建议请通过邮件的方式联系作者。

编 者

目 录

前言

第 1 章 数据仓库和商务智能	1
1.1 计算机技术的发展	1
1.2 数据存储技术的发展	2
1.3 从数据库到数据仓库	3
1.4 数据仓库	4
1.4.1 数据仓库的定义	4
1.4.2 数据仓库的特性	4
1.5 商务智能	5
1.5.1 商务智能的定义	5
1.5.2 商务智能的特点	6
第 2 章 SAP 商务智能	7
2.1 SAP 公司简介	7
2.2 SAP 产品系列	7
2.2.1 业务应用	8
2.2.2 商务分析	8
2.2.3 数据库及其技术	9
2.2.4 移动应用	9
2.2.5 云服务	9
2.3 SAP BW 数据仓库	10
2.3.1 BW 出现的背景	10
2.3.2 BW 的优势	10
2.3.3 BW 的未来发展	12
2.4 BW 系统架构	13
2.5 BW 中的对象	17
2.5.1 数据仓库工作台	17
2.5.2 信息提供者	21
2.5.3 信息对象	23
2.5.4 信息块——一种扩展的星形结构	34
第 3 章 SAP BW 功能增强概览	43
3.1 ABAP/4 语言简介	43
3.1.1 报表程序	43
3.1.2 对话程序	44
3.1.3 功能模块	46

3.1.4 面向对象的 Class 程序	50
3.2 SAP 系统的增强概念简介	54
3.2.1 第一代 User Exit	55
3.2.2 第二代 User Exit	55
3.2.3 字典增强 Append Structure	55
3.2.4 BAdI	55
3.2.5 业务交易事件 BTE	56
3.3 BW 中的系统增强点概览	56
3.4 针对 BW 增强的 ABAP 指引	57
3.4.1 ABAP 语言中的表类型	57
3.4.2 对内表的循环访问处理	59
3.4.3 指针的应用	60
3.4.4 数据存取和缓存	60
3.4.5 ABAP 小技巧	61
3.4.6 ABAP 运行时间分析工具 SE30	62
第4章 SAP 源系统数据抽取过程的功能增强	70
4.1 一般数据源的制作	70
4.1.1 应用场景描述	70
4.1.2 一般数据源创建准备工作	70
4.1.3 通过功能模块进行数据源的开发	72
4.1.4 通过表和视图进行数据源的开发	79
4.1.5 单元小结	81
4.2 对数据源的增强处理	81
4.2.1 应用场景描述	81
4.2.2 使用用户出口 RSAP0001	82
4.2.3 使用 BADI RSU5_SAPI_BADI	97
4.2.4 单元小结	103
4.3 使用 BTE 制作带增量的一般数据源	103
4.3.1 应用场景描述	103
4.3.2 实施步骤详解	103
4.3.3 单元小结	121
4.4 实战经验分享——如何在项目中更好地组织增强接口开发	121
第5章 数据上载过程中的功能增强	126
5.1 信息包中的增强实施	126
5.1.1 应用场景描述	126
5.1.2 文件名的动态选择	126
5.1.3 特性选择条件的动态实施	128
5.1.4 单元小结	132
5.2 转换中的 ABAP 增强实施	132

5.2.1 应用场景描述	132
5.2.2 转换中的增强实现详解	132
5.2.3 单元小结	163
5.3 数据传输进程中的增强实施	163
5.3.1 应用场景描述	163
5.3.2 在数据传输进程中选择文件名称	164
5.3.3 在数据传输进程中指定特性的选择条件	165
5.3.4 在数据传输进程中使用报表变量	169
5.3.5 单元小结	182
5.4 层次结构的处理	182
5.4.1 层次机构的同步(下载+上载)	182
5.4.2 从数据表生成层次结构	199
5.5 BW 3. X 中的传输规则	207
5.5.1 传输规则中的开始例程	207
5.5.2 传输规则中的对象例程	211
5.6 BW 3. X 中的更新规则	220
5.6.1 应用场景描述	220
5.6.2 实施步骤详解	220
5.6.3 单元小结	234
5.7 实战经验分享——如何在项目中更好地组织例程的编写	235
5.7.1 建立集中的转换管控配置表	235
5.7.2 读取配置表的 INCLUDE 程序的编写	236
5.7.3 开始例程函数的编写	239
5.7.4 结束例程函数的编写	242
5.7.5 转换中的开始例程和结束例程	245
5.7.6 大型项目例程管控小结	252
第6章 数据仓库对象管理中的功能增强	254
6.1 对信息块的数据访问	254
6.1.1 应用场景描述	254
6.1.2 向信息块中随机增加测试数据	254
6.1.3 读取信息块的数据	260
6.1.4 单元小结	268
6.2 虚拟信息块	268
6.2.1 应用场景描述	268
6.2.2 虚拟信息提供者的实施	268
6.2.3 实施虚拟信息块	274
6.2.4 单元小结	281
6.3 虚拟关键指标	281
6.3.1 应用场景描述	281

6.3.2	使用系统提供的增强进行处理	282
6.3.3	使用系统提供的 BADI 进行处理	287
6.3.4	单元小结	291
第 7 章	报表中的功能增强	293
7.1	报表变量出口的 ABAP 增强实施	293
7.1.1	应用场景描述	293
7.1.2	函数 EXIT_SAPLRSR0_001 的接口说明	293
7.1.3	过程号 I_STEP = 0 的处理	294
7.1.4	过程号 I_STEP = 1 的处理	295
7.1.5	过程号 I_STEP = 2 的处理	297
7.1.6	过程号 I_STEP = 3 的处理	300
7.1.7	单元小结	301
7.2	报表跳转的 ABAP 增强实施	301
7.2.1	应用场景描述	301
7.2.2	实施步骤详解	302
7.2.3	单元小结	304
7.3	通过邮件发送查询结果	305
7.3.1	应用场景描述	305
7.3.2	实施步骤详解	305
7.3.3	单元小结	310
7.4	报表 MDX 的编写	310
7.4.1	应用场景描述	310
7.4.2	实施步骤详解	310
7.4.3	单元小结	315
7.5	在 APD 中使用 ABAP 例程实施	315
7.5.1	应用场景描述	315
7.5.2	实施步骤详解	316
7.5.3	单元小结	332
第 8 章	计划中的功能增强	333
8.1	应用场景描述	333
8.2	变量的 ABAP 应用实施	333
8.2.1	BW - BPS 中的变量实施	333
8.2.2	BW - IP 中的变量实施	339
8.3	对特性派生的 ABAP 应用实施	341
8.3.1	BW - BPS 中的特性派生	341
8.3.2	BW - IP 中的特性派生	349
8.4	退出功能的 ABAP 应用实施	358
8.4.1	BW - BPS 中的退出功能	358
8.4.2	BW - IP 中的退出功能	366

第 9 章 系统运行中的功能增强	372
9.1 通过 ABAP 程序启动处理链	372
9.1.1 应用场景描述	372
9.1.2 详细实施步骤	372
9.1.3 单元小结	373
9.2 当天数据上载结果监控	373
9.2.1 应用场景描述	373
9.2.2 详细实施步骤	374
9.2.3 单元小结	376
9.3 在处理链中安排数据选择性删除	376
9.3.1 应用场景描述	376
9.3.2 详细实施步骤	376
9.3.3 单元小结	380
第 10 章 涉及 BW 系统的程序和数据表汇总	381
10.1 BW 系统涉及的数据表汇总	381
10.1.1 ERP 中涉及 BW 系统的数据表汇总	381
10.1.2 BW 系统中的常用事务码	381
10.2 BW 系统中其他相关的 BAdl 和 User Exit	383
10.2.1 BAdls	383
10.2.2 User Exit	384
10.3 BW 系统标准表	384
10.3.1 ERP 中与 BW 系统相关的重要表格概览	384
10.3.2 BW 系统中的重要表格概览	385
10.4 BW 对象的生成表	393
10.4.1 信息块可能产生的表	393
10.4.2 DSO 可能产生的表	393
10.4.3 信息对象可能产生的表	394
10.4.4 定义数据流可能产生的表	394
10.5 BW 系统中有用的程序索引	394
10.5.1 BW 系统中常用的程序列表	394
10.5.2 BW 系统中常用的功能函数列表	395
10.5.3 BW 系统中处理转换错误的类列表	396
10.5.4 BW 系统中常用的程序列表	397
附录 术语对照	398
参考文献	401

第1章 数据仓库和商务智能

1.1 计算机技术的发展

1946年,美国宾夕法尼亚大学的埃克特和莫希里研制出世界上第一台电子多用途计算机“埃尼阿克(ENIAC)”,从而开启了人类社会信息技术革命的序幕。从第一台多用途计算机诞生开始,截至目前,信息技术经历了大型主机、小型计算机、微型计算机、客户/服务器、互联网、云计算六大阶段。

第一阶段即大型主机阶段。此阶段主要是20世纪40~50年代,这也是第一台多用途计算机诞生的年代,此时的计算机大都体型庞大,如第一台计算机就重约27 t,占地150 m²,由18 800个电子管构成。在经历了电子管数字计算机、晶体管数字计算机、集成电路数字计算机和大规模集成电路数字计算机等发展历程后,计算机技术逐渐走向成熟。虽然改良之后的大型主机在中小企业已经难觅踪迹,但是在银行、保险、民航等公用领域还继续发挥着其作用。

第二阶段即小型计算机阶段。此阶段主要是20世纪60~70年代,由于半导体和集成电路技术的不断升级发展,人类已经能将身形巨大的大型机“缩小”到一个可以接受的范围,并且在使用成本上也降低到中小企业能够接受的范围之内。现在很多企业使用的服务器都属于小型计算机的范畴,其体型要小于大型主机,大于常见的个人计算机。

第三阶段即微型计算机阶段。此阶段主要是20世纪70~80年代,是对大型主机的第二次“缩小”,此时的计算机已经小到能放置在用户的桌面,因此又被称为“微机”或个人计算机。美国苹果公司于1977年在原有Apple基础上推出了第二代计算机Apple II,大获成功。而在1981年,国际商业机器(IBM)公司推出了IBM-PC,通过若干年的演进,其功能得到不断增强,并迅速占领了计算机市场,使得个人计算机得到了很大程度的普及。

第四阶段即客户/服务器阶段。这一阶段出现的标志是1964年IBM与美国航空公司建立的第一个全球联机订票系统,将全美2 000多个订票终端连在了一起。客户/服务器的结构到今天仍然被相当多的应用系统所采用,即使后来风靡全球的三系统架构,也只不过是客户/服务器结构的优化,而且随着时间的推移和互联网的出现,部分客户端进一步被移植到浏览器上。在这一结构中,服务器是网络的核心,客户端是网络的基础,客户端依靠服务器来获得所需的网络资源,其优点在于能够充分发挥客户端的计算能力。

第五阶段即互联网阶段。互联网即广域网、局域网及单机按照一定的通信协议组成的国际计算机网络。互联网起源于1969年,当年美国国防部研究计划署制定了协定,将加利福尼亚大学洛杉矶分校、斯坦福大学研究学院、加利福尼亚大学和犹他州大学的4台主要计算机联接起来。此后,互联网经历了从文本到图片、再到现在的语音和视频阶段,网络带宽也越来越大,互联网功能也越来越强,以至于1990年之后出生的人会被贴上“互联网时代原住民”或“互联网一代”的标签。

第六阶段即云计算阶段。从2008年开始,“云计算”“大数据”“内存计算”和“移动应用”变成了炙手可热的计算机概念,而这样的专业术语正在成为一个通俗且大众化的词语。其中,“云计算”被视为“革命性的计算模型”,因为在互联网的支撑下,它使超级计算能力通过互联网自由流通成为了可能。企业与个人用户无须投入昂贵的硬件购置成本,只需通过互联网来购买计算能力即可,用户只需为自己使用的计算能力付钱。同时,还省去了传统软件在硬件、软件和专业技能等方面的花费。云计算囊括了开发、架构、负载平衡和商业模式等,是软件业的未来模式。

1.2 数据存储技术的发展

计算机硬件经历了从大型机到微型机的进化过程,但是对于数据处理不断增长的需求却贯穿了整个计算机的发展历程。数据处理主要就是对数据的管理,包括对数据的分类、组织、编码、存储、检索和维护等,这些也是数据处理的中心问题。

近半个世纪以来,数据管理经历了3个发展阶段,分别是人工管理阶段、文件系统阶段和数据库系统阶段。

20世纪50年代以前,计算机(主要是大型机)主要用于科学计算,其本身并不带有存储设备,也没有操作系统的概念,数据处理只有批处理方式。所有数据都不在计算机上保存,而是由应用程序员管理。应用程序员不仅要规定数据的逻辑结构,而且要在应用程序中设计物理结构,包括规定数据的存储结构和读取方式等。

20世纪50年代至60年代中期,数据管理处于文件系统阶段。计算机上出现了磁盘、磁鼓等随机存取设备,文件系统成为了专门的数据管理软件。这样,数据可以长期保存,并且使用文件系统来进行管理,使得程序和数据之间有了一定的独立性,尽管这种独立性还不是很强。此时的应用程序设计者必须对所有文件的逻辑结构和物理结构有清楚的了解,同时由于文件系统只提供了最为基础的读、写等文件操作命令,因此对文件的查询、修改、插入和删除等操作必须在应用程序内编写相应的程序代码来解决,这使得文件系统只能专门为某一特定的应用程序服务,故应用程序的应用效率不高,数据冗余度大。

20世纪60年代后期以后,计算机进入了小型机时代,同时数据管理进入了数据库系统阶段。这一时期,计算机具有了庞大容量的存储设备和高速的信息处理能力。计算机软件 and 硬件都飞速发展,信息急剧膨胀,数据库技术也日新月异,蓬勃发展。数据库按照某种数据模型组织数据,不仅文件内部数据彼此关联,而且文件之间在结构上也有有机地联系在一起,当描述数据时,不仅描述数据本身,也描述数据之间的联系。在数据库中,数据不再分属于各个应用程序,而是集中存放在一起,实行统一管理。这一时期,也出现了专门的数据库系统。

1969年,IBM公司开发了第一个数据库系统IMS(Information Management System),这是一个层次数据库系统,在数据库系统发展史上意义重大。同年,美国的数据系统语言委员会下属的数据库任务组提出了著名的DBTG报告,并在1970年提出了该报告的修订版。这份报告定义了数据库操作语言、模式定义语言和子模式定义语言的概念。数据库操作语言用于编写操作概念视图的应用程序,模式定义语言用来编写概念视图和内部视图相结合的模式程序。在20世纪70年代,有许多遵循DBTG报告的网状数据库系统出现,如IDMS、IDS、

DMSII 等。20 世纪 70 年代初, IBM 公司下属的研究所发表了题为《大型共享数据库数据的关系模型》一文。在此论文中提到了关系数据模型的概念, 而关系代数和关系演算的提出, 使关系型数据库从理论到实践都取得了辉煌的成就, 许多著名的关系型数据库也应运而生, 如 System R、Ingres、Oracle 等。1986 年, 美国国家标准协会 (ANSI) 通过了关系数据库查询语言 SQL 的文本标准。

20 世纪 80 年代以后, 随着计算机硬件技术的发展, 计算机应用得以不断深入, 从而产生了许多新的应用领域, 如计算机辅助设计、计算机辅助制造、计算机集成制造、办公自动化和地理信息处理等。但是由于没能设计出统一的数据模型来表示这些新型数据及其相互关系, 因此出现了百家争鸣的局面, 产生了面向对象数据库、工程数据库、时态数据库、地理数据库、模糊数据库和积极数据库等, 但是真正统一的新一代数据库系统未能出现。

进入到 21 世纪, 计算机软、硬件技术飞速发展, 不仅在数据处理速度上突飞猛进, 在数据的存储上也是日新月异, 同时不断增加的企业数据容量和对数据实时获取的需求, 迫切需要新一代的数据库系统的出现, 德国的 SAP 公司顺势推出了自己的内存数据库产品。内存数据库, 顾名思义就是将数据放在内存中直接进行读、写操作的数据库。相对于传统的磁盘, 内存的读写速度要高出好几个数量级, 将数据保存在内存中相比对磁盘进行读写能极大地提高应用的性能, 从而解决实际的应用需求。目前, 内存数据库技术还在不断地发展、进化过程中。

1.3 从数据库到数据仓库

目前的商业、企业管理的数据处理大致可以分成两类, 即在线交易处理 (On - Line Transactional Process, OLTP) 和在线分析处理 (On - Line Analysis Process, OLAP)。交易处理也称为事务性处理, 一般针对非常具体的业务, 是对数据库联机的日常操作, 通常是对一个或一组记录的查询和修改, 主要为特定的应用场景服务。人们关心的是响应时间、数据的完整性和安全性, 在线分析处理一般是针对某个主题, 在时间上有比较大的跨度, 它操作的是大量, 甚至是海量的数据, 这些数据是操作性数据的一种积累和遴选, 它主要面向业务分析和决策支持。

数据库系统在相当长的时间内作为数据管理的主要手段, 从其诞生的第一天起, 就被主要用于事务处理, 经过数十年的发展, 在这些数据库中已经保存了大量的日常业务数据。随着技术的进步, 人们也试图让计算机执行更多的任务, 而数据库技术也一直力图能使自己胜任从事务处理、批处理到分析处理等不同类型的信息处理任务。后来人们逐渐意识到, 在目前计算机处理的能力上, 根本无法完全实现这些功能, 而在另一方面, 在线事务处理和在线分析处理具有极不相同的性质, 直接用事务处理环境来支持分析和决策在很多方面具有先天的缺陷。

在线交易处理不适宜决策应用的原因如下:

首先, 在在线交易处理环境中, 用户一般是具体的操作人员, 他们的行为特点是数据的存取操作, 并不用去理会数据对决策的用处, 操作频率高且每次操作的时间短, 但对系统的响应速度有极高的要求; 而对在线分析处理而言, 用户是企业的管理人员或数据分析师, 属于信息的探索者, 他们的目的是要将产生的数据抽象为信息, 以便于决策。其行为模式和在

线交易处理完全不同，他们可能会使一个决策支持系统（Decision - making Support System, DSS）应用程序联机工作几小时，从而消耗大量的系统资源，进而影响到业务操作的高效运行。

其次，DSS 过程需要集成的、历史的、综合的数据，全面而正确的数据是有效分析和决策的首要前提。相关数据收集得越完整、反映的信息面越广、清洗得越干净，最后得到的结果就越可靠。目前，绝大多数企业数据的实际分布情况都是以信息孤岛的形式存在，在具体的在线交易环境中，部分数据记录甚至被“束之高阁”，在降低交易系统性能的同时，也使宝贵的信息资源被白白浪费。

近些年来，随着数据库技术的应用和发展，人们尝试对数据库中的数据进行再加工，构建一个综合的、面向分析的环境，以更好地建设商务智能支持企业决策，从而形成了数据仓库技术。数据仓库系统包括多维数据仓库、联机分析处理 OLAP 或商务智能（Business Intelligence, BI）、数据挖掘（Data Mining, DM）等多个方面。

1.4 数据仓库

1.4.1 数据仓库的定义

数据仓库（Data Warehouse, DW）是决策支持系统（Decision Support System, DSS）和联机分析应用（On - Line Analysis Processor, OLAP）数据源的结构化数据环境。数据仓库研究和解决从数据库中获取信息的问题。数据仓库的特征在于其面向主题、集成性、稳定性和历史积累性。

数据仓库之父 William H. Inmon 在 1991 年出版的《Building the Data Warehouse》一书中所提出的定义被广泛接受，即数据仓库是一个面向主题的（Subject Oriented）、集成的（Integrated）、相对稳定的（Non - Volatile）、反映历史变化（Time Variant）的数据集合，用于支持管理决策（Decision Making Support）的信息系统。

1.4.2 数据仓库的特性

数据仓库不同于传统的数据库，其特点主要体现在“目标任务”“数据覆盖面”“数据粒度”“数据稳定性”“存取数据量”和“响应时间”等方面。

1. 面向主题

操作型数据库的数据组织面向事务处理任务，各个业务系统之间各自分离，而数据仓库中的数据是按照一定的主题域进行组织的。主题是与传统数据库的面向应用相对应的，是一个抽象概念，是在较高层次上将企业信息系统中的数据综合、归类并进行分析和利用的对象。每一个主题对应一个宏观的分析领域，通常的主题包括“财务”“成本”“采购”“生产”“库存”“销售”“人力资源”“质量”和“市场信息”等。数据仓库排除了对于决策无用的数据，提供了特定主题的全部视图。

2. 集成的

数据仓库中的数据是在对原有分散的数据库数据进行抽取和清洗的基础上，经过系统加工、汇总和整理得到的，必须消除原始数据中的不一致性，以保证数据仓库内的信息是关于

整个企业的、一致的、全局的信息，而不是一个一个的信息孤岛。

3. 高度汇总的

一般来说，存储于数据仓库系统中用于做出管理决策的数据，大都是经过提炼之后的数据，而不是像在线交易系统那样，存储了作为明细级别的数据，以供业务操作使用，因此其数据的粒度相对数据库系统而言比较粗，也就是属于高度汇总的数据。

4. 相对稳定的

数据仓库中的数据主要供企业决策分析之用，所涉及的数据操作主要是数据查询，一旦某个数据进入数据仓库以后，一般情况下将被长期保留，也就是数据仓库中一般有大量的查询操作，但修改和删除操作很少，通常只需要定期的加载和刷新。

5. 反映历史变化

数据仓库中的数据通常包含历史信息，系统记录了企业从过去某一时间点（如开始应用数据仓库的时间点）到目前的各个阶段的信息，通过这些信息，可以对企业的发展历程和未来趋势做出定量分析和预测。而在应用这些历史数据时，系统读取和运算的数据往往是巨大的，有时候甚至可达几十亿条。

6. 时效性要求低

假设读者去医院排队挂号，但是却被告知由于系统本身性能太差，而导致了挂号的延误，可想而知后果会是怎样的，因此对于数据库应用系统的时效性要求是非常高的。但是对于数据仓库而言，由于使用对象是数据探索者，因此对于时效性的要求不是那么高，只需要在可以接受的时间段内能够得到想要的结果即可，甚至有些时候被期望的时间可能长达一天或几天。

1.5 商务智能

1.5.1 商务智能的定义

商务智能（Business Intelligence，BI）的概念最早于1996年由高德纳咨询公司（Gartner Group）提出，其将商务智能定义为：商务智能描述了一系列的概念和方法，通过应用基于业务实际的支持系统来辅助商业决策的制定。商务智能技术提供使企业迅速分析数据的技术和方法，包括收集、管理和分析数据，将这些数据转化为有用的信息，然后分发到企业的各个部门。

可以认为，商务智能是对商业信息的搜集、管理和分析，目的是使企业的各级决策者能够获得足够的知识或洞察力，促使他们做出对企业更有利的决策。商务智能一般由数据仓库、联机分析处理、数据挖掘、数据备份和恢复等部分组成。商务智能的实现涉及软件、硬件、咨询服务及应用，其基本体系结构包括数据仓库、联机分析处理和数据挖掘3部分。

商务智能通常被理解为将企业中现有的数据提炼为信息，进而转化为知识，帮助企业做出明智的业务经营决策的工具。这里所说的数据包括来自企业业务系统的订单、库存、交易账目、客户和供应商等，来自企业所处行业和竞争对手的数据，以及来自企业所处的其他外部环境中的各种数据，如国际原油价格和大宗商品期货价格等。而商务智能能够辅助的业务经营决策，既可以是操作层的，也可以是战术层和战略层的。为了将数据转化为知识，需要

利用数据仓库、联机分析处理工具和数据挖掘等技术。因此,从技术层面上讲,商务智能不是什么新技术,它只是“数据仓库”“联机分析处理”和“数据挖掘”等技术的综合运用。

所以,把商务智能看成是一种解决方案应该比较恰当。商务智能应用的关键是从企业许多不同业务系统中提取出有用的数据并进行清洗,以保证数据的正确性,然后经过抽取(Extraction)、转换(Transformation)和装载(Load),即 ETL 过程,合并到一个企业级的数据仓库里,从而得到企业数据的唯一全局视图,再在此基础上利用合适的查询和分析工具,以及数据挖掘工具对其进行分析和处理(这时信息变为辅助决策的知识),最后将知识呈现给管理者,为其决策提供支持。

1.5.2 商务智能的特点

1. 成熟的数据仓库管理能力

目前,商务智能市场上有多种数据仓库管理软件,它们都可以提供全面的数据仓库建设和管理工具。这些工具涵盖了从源系统链接、数据来源获取、数据清洗和转换、数据装载调度、明细数据存储、汇总数据处理、元数据管理以及用户管理和权限管理等数据仓库建设的基本过程。

2. 便捷的报表功能

许多商务智能系统都提供了友好的用户操作界面,使用户可以轻松地创建报表,并支持用户根据实际情况制订特殊要求的查询,同时还可以利用报表工具自动生成报表,而不用像以前那样要通过 IT 专业人员来实现业务人员的具体报表需求。例如, SAP 公司的报表工具 Webi 即席查询和 Dashboard 水晶易表就提供了一个便捷易懂的图形化操作界面,解决了用户写 SQL 语句的不便,用户借助鼠标的拖曳和单击即可进行报表或查询界面的生成。

3. 强大的数据挖掘能力

鉴于数据挖掘和分析功能是商务智能系统的基本功能,各商务智能厂商都非常注重这方面的研究,并开发了各种功能强大的分析挖掘工具。例如, Cognos 公司开发的 OLAP 工具 Powerplay, Oracle 公司基于数据仓库的数据挖掘工具 Oracle Darwin、Oracle Express 产品系列和 Oracle Discoverer,都具有探索数据和 OLAP 功能,可以帮助企业从数据中挖掘出更有价值的信息。

第2章 SAP 商务智能

上一章介绍了数据仓库和商务智能的定义和特点,下面将对 SAP 的数据仓库和商务智能解决方案(SAP BW)进行简要介绍,以便在后续相关的实际开发使用过程中对这些概念不至于太过陌生。

2.1 SAP 公司简介

SAP 公司成立于 1972 年,总部位于德国沃尔多夫市,是全球最大的企业管理软件及协同商务解决方案供应商,也是全球第三大独立软件供应商。通过 mySAP 商务套件,全球各类企业可以改善与客户、伙伴的关系,精细企业运作,并在其供应链管理中获得显著效益。从航空航天到公共设施行业,各个行业独特的核心业务流程都由 SAP 的 28 个行业解决方案支持着进行有效运行。目前,全球有 120 多个国家、超过 29 800 家企业正在运行着 100 600 多个 SAP 软件。SAP 在全球 50 多个国家拥有分支机构,并在多家证券交易所上市,包括法兰克福证券交易所和纽约证券交易所。

SAP 早在 20 世纪 80 年代就同中国的国营企业合作并取得了成功经验。1994 年底, SAP 在北京建立了代表机构,1995 年正式成立 SAP 中国公司,1996 年、1997 年陆续设立上海和广州分公司。作为中国 ERP 市场的绝对领导者, SAP 的市场份额已经达到 30%, 年度业绩以 50% 以上的速度递增。

SAP 荣获的主要赞誉如下:

1996 年:欧洲年度最佳公司。

1997 年:《计算机世界》Smithsonian 奖。

1998 年:最佳管理公司。

1999 年:《财富》500 强、《商业周刊》全球科技百强、最佳 ERP 应用奖。

2000 年:英国《金融时报》“全球最有声誉公司”前 30 强、mySAP.com 荣获“最佳电子商务系统”。

2001 年:《商业周刊》“全球最有价值品牌”、美国《环球财经》“全球最佳公司”。

2002 年:《福布斯》“全球计算机软件最佳企业”、英国《金融时报》“500 强进步最大的公司”。

2003 年:《商业周刊》“全球 1000 强”排名第 98 位,被评为全球 100 家最有价值公司之一。

2.2 SAP 产品系列

SAP 成立之初只有 SAP ERP 一个产品,经过 40 年的发展和近几年一系列的商业收购,目前产品线已经覆盖“业务应用”“商务分析”“数据库及其技术”“移动应用”和“云服