

# 企业数据仓库设计与实现

张冬松

信阳学院  
大数据与人工智能学院

[dszhang@nudt.edu.cn](mailto:dszhang@nudt.edu.cn)



# 目录 content



## 第一节

## 企业数据仓库设计

## 第二节

## 企业数据仓库实现

## 第三节

## 企业数据可视化



# 第一节 企业数据仓库设计

- 企业数据仓库背景
- 企业数据仓库概述
- 企业数据仓库需求分析
- 企业数据仓库模型设计





## 1.1 企业数据仓库背景

### A.ERP

1. 随着信息技术的发展和企业管理决策支持的需要，数据仓库技术已经从纯理论研究转化为企业管理领域中的一种实用性技术。
2. 下面，举例说明，某集团型企业数据仓库是如何建设的。
3. ERP: Enterprise Resource Planning, 企业资源规划软件，是一种企业事务管理系统。
4. ERP系统是指建立在信息技术基础上，以系统化的管理思想，为企业决策层及员工提供决策及事务处理的管理平台。



## 1.1 企业数据仓库背景

### B.背景

- 1.ERP的成功实施和应用，使得企业流程规范化的同时，也使得企业管理系统能处理的数据量呈指数级增长。
- 2.因此，对企业管理者来说，企业管理某种程度上就是“信息的管理”，大量准确、及时的数据信息就像一座金矿，等待挖掘。
- 3.如何开采和利用这些金矿，就称为亟待解决的问题。



## 1.1 企业数据仓库背景

### C.举例说明

- 1.以某集团型企业数据仓库系统的构建为例，设计采用EDW（企业数据仓库）建模思想为企业构建相应的系统。
- 2.为保证新的企业数据仓库系统的顺利实施及与现有系统的无缝衔接，考虑基于现有SAP业务系统作为获取主要业务数据的数据源，将基于数据仓库分析的结果送给现有企业门户系统进行可视化展示，或通过接口等方式与其他业务系统进行数据交换。



# 第一节 企业数据仓库设计

- 企业数据仓库背景
- 企业数据仓库概述
- 企业数据仓库需求分析
- 企业数据仓库模型设计

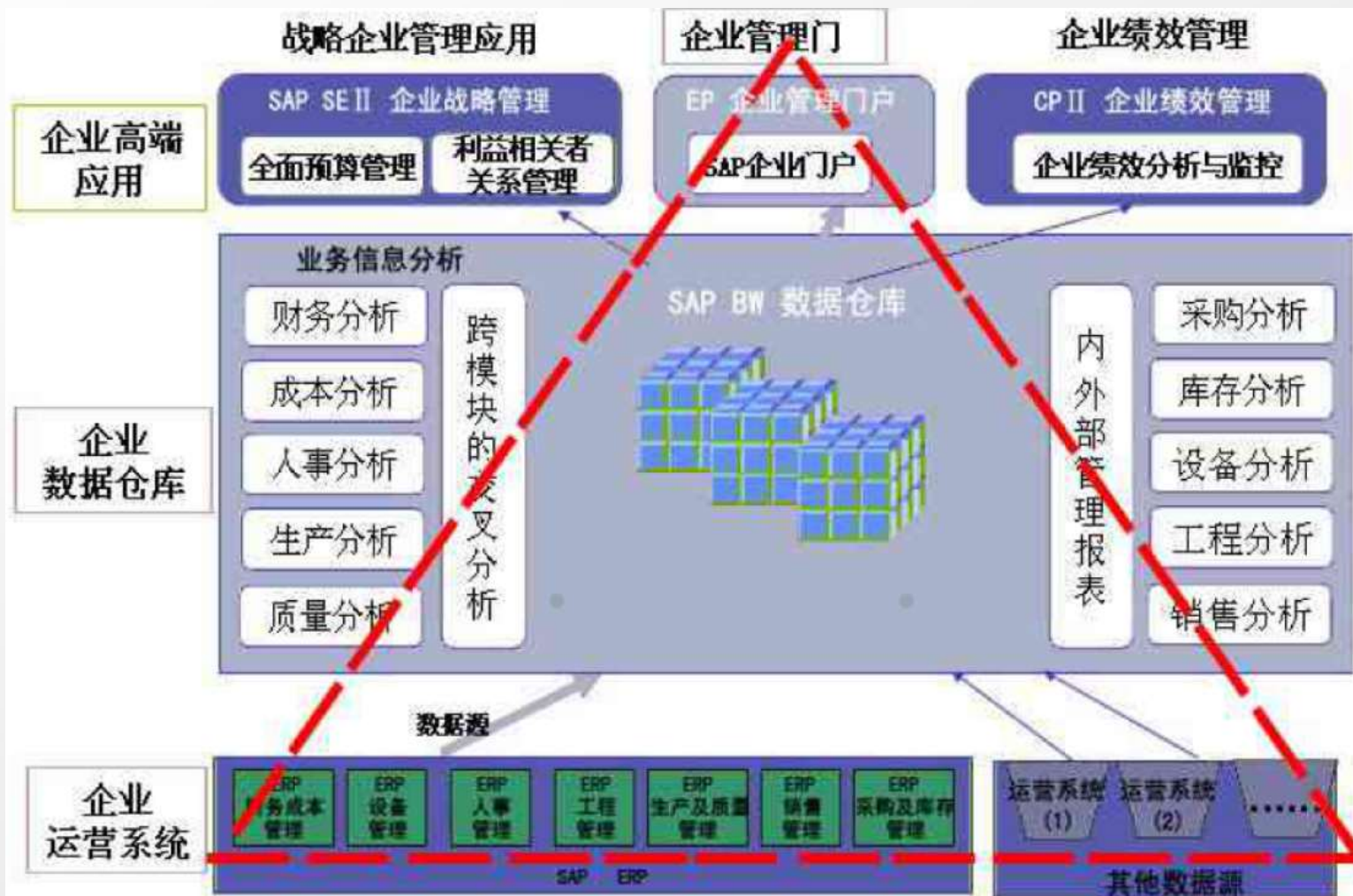




## 1.2 企业数据仓库概述

### A. 总体架构

以企业中运行的SAP系统和ERP业务系统作为数据仓库的数据源提取数据，数据在数据仓库中根据规则层上载并根据需要加工、处理、合并，将数据存放在立方体中，基于立方体设计报表，在企业门户中展示。

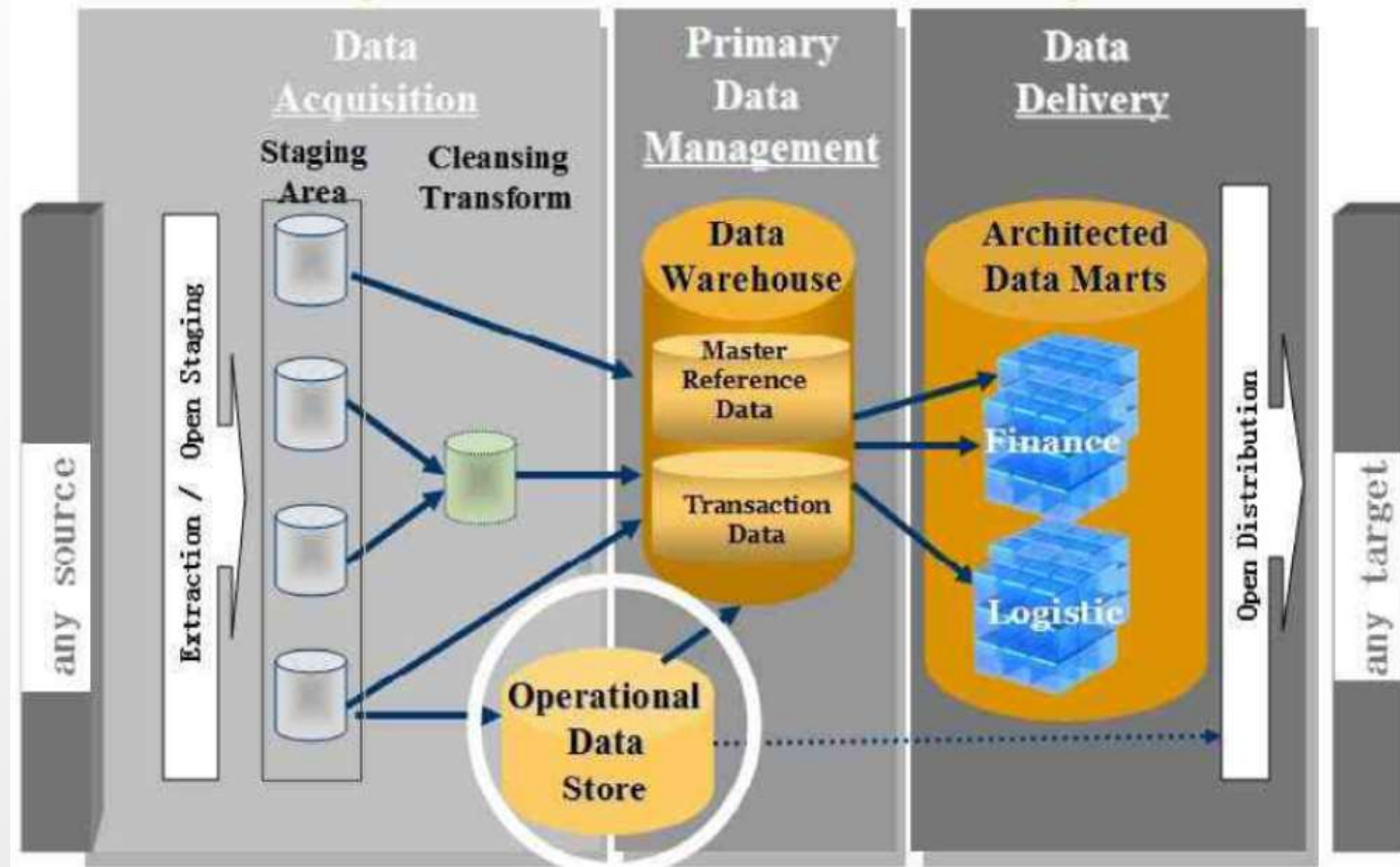




## 1.2 企业数据仓库概述

### B.EDW模型及特点

- 1.数据不直接进数据分析区，先进操作数据存储区（ODS）。
- 2.不同业务含义的数据、不同数据源的数据、不同粒度级别的数据分别存储在不同的ODS中。
- 3.ODS中数据进数据分析区前，需要数据转换和清洗。

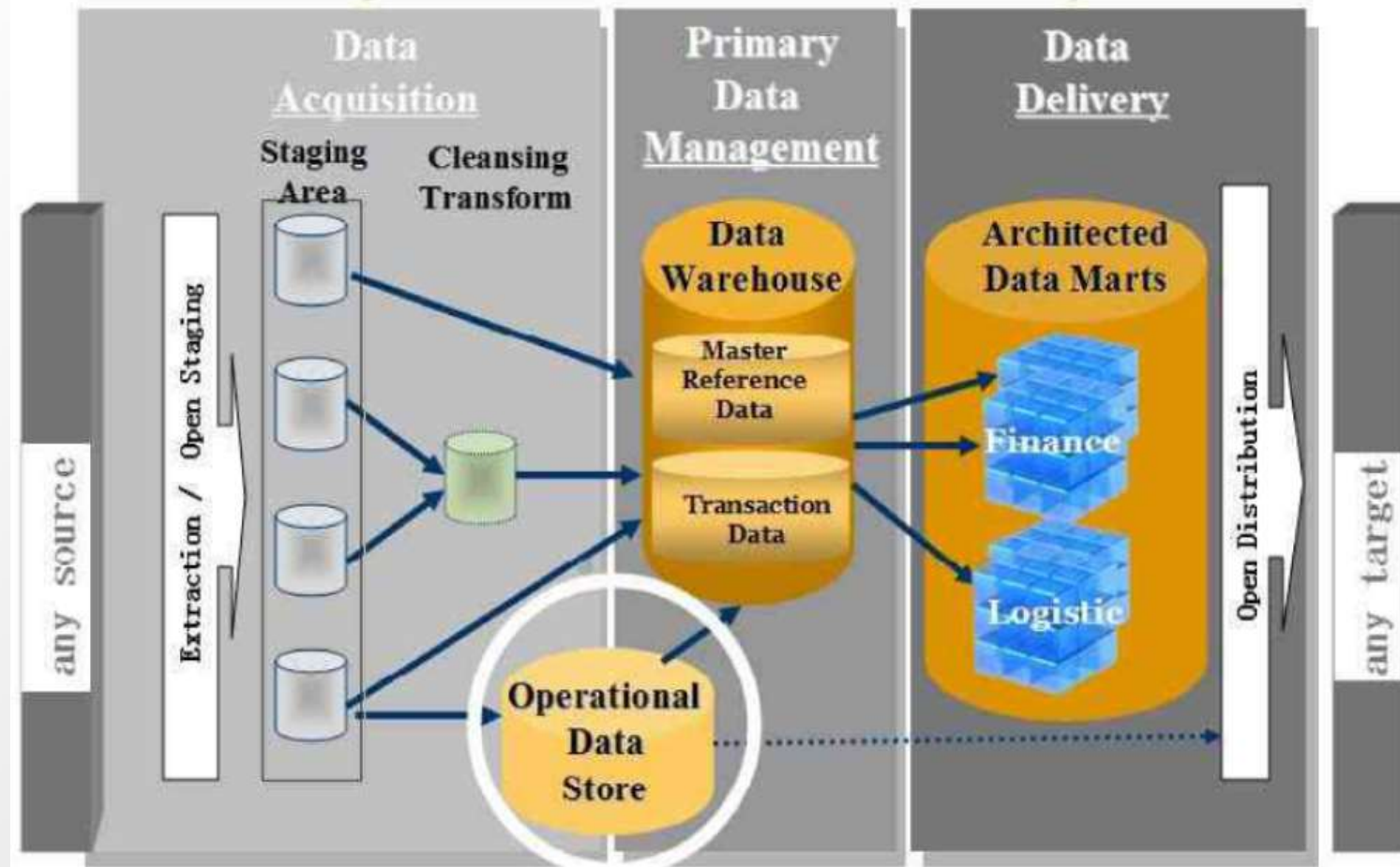


## 1.2 企业数据仓库概述

### C.EDW建模思想

将数据仓库划分为3个层次：

- (1) 数据提取层
- (2) 数据合并层
- (3) 数据集市层





# 第一节 企业数据仓库设计

- 企业数据仓库背景
- 企业数据仓库概述
- 企业数据仓库需求分析
- 企业数据仓库模型设计





## 1.3 企业数据仓库需求分析

数据仓库构建时通常采用“**自顶向下、逐步求精**”的方法，因此，企业数据仓库的系统分析与设计过程可以分为：**系统需求分析、模型设计和ETL设计**，这里，ETL设计可以包含进模型设计中。



## 1.3 企业数据仓库需求分析

### A.需求分析

- 1.主要是确定数据仓库的主要主题及相互关系，即对需求范围内的业务及其间关系进行高度概括性描述，并把业务对象进行归类，划分主题域，对每个主题域进行较为明确的描述。
- 2.具体可分为：概要分析、详细分析。
- 3.举例说明。



## 1.3 企业数据仓库需求分析

### B.概要分析

- 1.抽调各业务模块人员组成数据仓库项目组，根据与业务部门的相关人员进行交流和调研，对每个主题模块的需求，从关键指标、分析维度、分析方法进行分析，最终形成《概要分析报告》，并对今后工作列出详细的时间计划。
- 2.举例，以生产模块为例，得到关键分析指标：
  - (1) 量：产量，计划产量，交库数，入库数量，出库数量，库存数量等
  - (2) 率：计划完成率、成材率等





## 1.3 企业数据仓库需求分析

### B.概要分析

#### 1.主要分析维度:

- (1) 组织维度: 公司, 工厂, 库存地, 车间
- (2) 物料维度: 大类, 中类, 小类, 细分类, 产地, 牌号, 物料组, 生产调度员
- (3) 时间维度: 年月, 年月日, 年, 季, 周

#### 2.主要分析方法:

- (1) 比较分析
- (2) 比例分析



## 1.3 企业数据仓库需求分析

### C.详细分析

类型	分析对象/指标	说明	数据来源说明
报表名称：产成品产量月报			
分析对象	公司	XX 股份公司	
	工厂	股份公司工厂 (8000、8800)	
	库存地	所有生产成品库存地	8301-8320;
	物料类型	物料类型为产成品的物料	物料主数据的物料类型
	物料分类		按物料统一分类
	时间	年/月	
	车间	各生产车间	物料主数据的生产调度员 生产调度员 (MARC-FEVOR)
指标	产量	生产单位实际生产下线量	
	...		
说明	1、产成品涉及的库存地:8301; 8302; 8303; 8304; 8305; 8306; 8310; 8311; 8316; 8317; 8318; 8319; 8320; 2、钢种说明：棒线材取物料编码2-7 位，次品取3-8 位；板材取物料编码3-8 位，次品取 4-9 位；镀锌彩涂板卷取物料编码2-7 位，次品物料无钢种； 3、烧结作业区产量：生产收货 I01/I02 -改判 (309/310) 精整、镀锌彩涂车间产量：移出现场库的量X01/X12 其他作业区产量：生产收货 I01/I02 -质量改判 (Z09/Z10)		



## 1.3 企业数据仓库需求分析

### D.用户确认需求分析

- 1.梳理完成所有的主题和指标，以及确定主题数据的来源，最终提交《详细分析报告》，经过业务部门相关人员和领导、内部顾问、外部顾问**签字确认**后，进入实际建模阶段，也就是实际设计阶段。



# 第一节 企业数据仓库设计

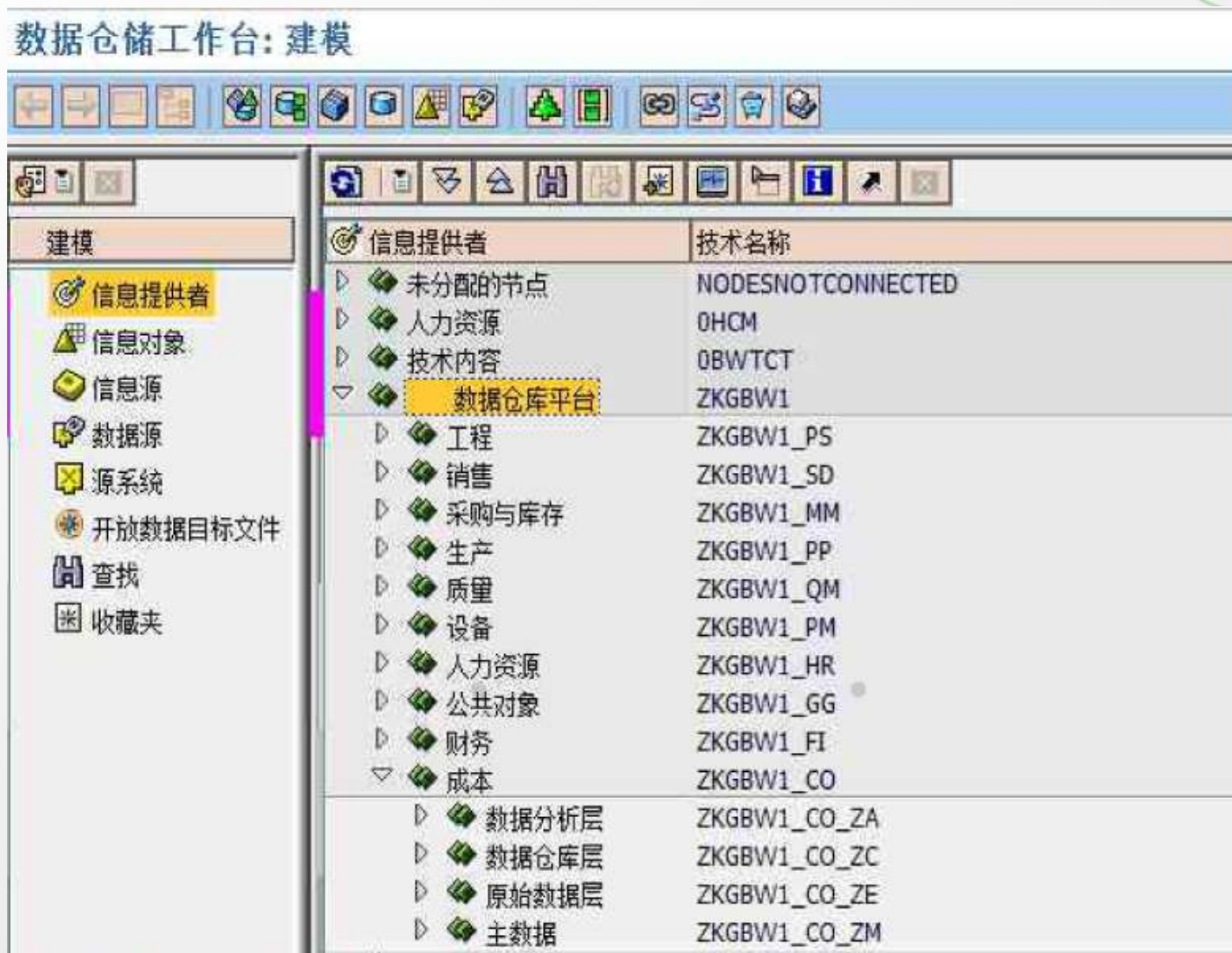
- 企业数据仓库背景
- 企业数据仓库概述
- 企业数据仓库需求分析
- 企业数据仓库模型设计



## 1.4 企业数据仓库模型设计

### A.模型设计

- 1.根据数据用途不同将数据仓库分为不同的业务模块，即分析主题。如右图。
- 2.每个分析主题又可分为4个层次，如右图。





## 1.4 企业数据仓库模型设计

### A.模型设计

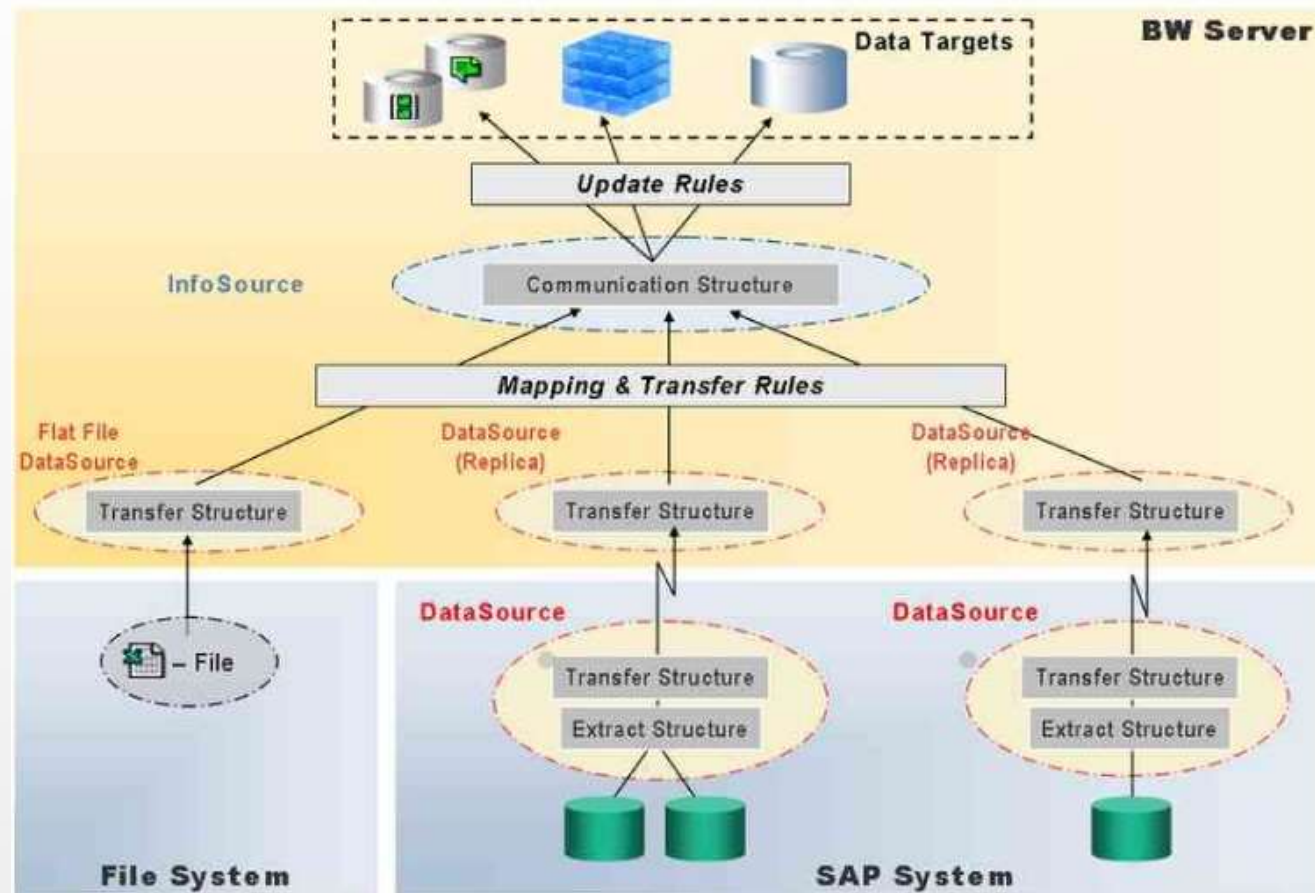
- 1.主数据：包含该主题模块建立数据仓库时需要的信息对象。
- 2.原始数据层：保存从外部数据源中获取的数据于ODS中，不加以任务修改计算。
- 3.数据仓库层：将原始数据层中存储的原始数据进行加工后保存于ODS。
- 4.数据分析层：以数据立方体（Cube）形式存储数据仓库层加工后的数据。



## 1.4 企业数据仓库模型设计

### B.ETL流程设计

Data Flow in BW





## 第二节

# 企业数据仓库实现

- 数据上传
- 数据多立方体
- 数据处理流程



## 2.1 数据上传

### A. 整体流程

信息提供者	技术名称
数据分析层	ZKGBW1_PP_ZA
报工数量	ZAPP_C02
生产产量	ZAPP_C04
ODSO ZCMM_003 -> CUBE ZAPP_C04	02GNETYAT7V3UYSXARZDOU86ZQWRY...
物料移动凭证	ZCMM_003
ODSO ZEMM_003 -> ODSO ZCMM_003	0KV3ZAZVB8UKE6BLNYCOA4LAAZD1694J
物料移动凭证	ZEMM_003
ZEMM_003 ZMM_2LIS_03_BF	48J0TYTG04IWXOKDHYU0GZXTU
物料移动凭证	ZMM_2LIS_03_BF
从物料移动为 PRD 集团 800	2LIS_03_BF      PROD800
库存管理的货物移动	2LIS_03_BF

## 2.1 数据上传

### B.数据传输

信息源: ZMM\_2LIS\_03\_BF 物料移动凭证

通讯结构

传输结构/传输规则

源系统: PROD800 - PRD 集团 800 (X)

数据源: 2LIS\_03\_BF - 库存管理的货物移动

状态: 活动的(已保存) 分配 DS

传输方法: PSA, IDoc

数据源/业务结构 传输规则

☐ 开始例程

信息对象	描述	Tp	规则
OSTORNO	冲销标识	←	OSTORNO
ORT_PROMO	促销	←	ORT_PROMO
OVAL_CLASS	估价类	←	OVAL_CLASS
ODOC_DATE	凭证日期	←	ODOC_DATE
OSTOCKTYPE	库存类型	←	OSTOCKTYPE
OSTOCKCAT	库存类别	←	OSTOCKCAT
OPSTNG_DATE	记帐日期	←	OPSTNG_DATE
OCOMP_CODE	公司代码	←	OCOMP_CODE
OBWAPPLNM	应用公司	←	OBWAPPLNM
OMOVETYPE	移动类型	←	OMOVETYPE
OSTOCKRELEV	BW: 适当库存	←	OSTOCKRELEV
OCPPVLC	BW: 采购价值	←	OCPPVLC

信息对象	描述	字段
OSTORNO	标识符: 取消数	ROCANC
ORT_PROMO	促销	AKTNR
OCOORDER	订单编码	AUFNR
	订单中工序的工	AUFPL
OVAL_CLASS	评估类	BKLAS
ODOC_DATE	凭证中的凭证日	BLDAT
ODLV_COSTS	用本市计算的交	BNBTR
OSTOCKTYPE	后勤信息系统库	BSTAUS
OSTOCKCAT	后勤信息系统中	BSTTYP
OPSTNG_DATE	凭证中的记帐日	BUDAT
OCOMP_CODE	公司代码	BUKRS
OBWAPPLNM	应用程序组件	BWAPPL

## 2.1 数据上传

### C.传输监控

数据传输处理监视器

请求标识: 55,387  
开始时间: 2008.10.07 03:35:20  
结束时间: 2008.10.07 03:36:54

关键日期/时间: 2008.10.07 09:39:44 0 运行 当前运行

请求处理	消息	数据	时间	持续时间
请求 55387			2008.10.07 03:35:20	
生成请求			2008.10.07 03:35:20	37 辅助
状态:可执行的(用户 BWREMOTE)			2008.10.07 03:35:38	
处理请求			2008.10.07 03:36:03	51 辅助
数据包 1 (16088 数据记录)			2008.10.07 03:36:04	50 辅助
从数据存储器中提取 ZCOMM_003 : 16088 数据记录			2008.10.07 03:36:03	7 辅助
开始提取			2008.10.07 03:36:04	
结束提取			2008.10.07 03:36:05	1 辅助
准备错误处理: 16088 -> 16088 数据记录			2008.10.07 03:36:11	
开始处理...			2008.10.07 03:36:11	
转发所有的记录			2008.10.07 03:36:11	
DDSO ZCOMM_003 -> CUBE ZAPP_C04 : 16088 -> 100			2008.10.07 03:36:11	37 辅助
更新至信息块 ZAPP_C04 : 1009 -> 1009 数据记录			2008.10.07 03:36:48	6 辅助
主处理结束			2008.10.07 03:36:10	44 辅助
将技术状态设置为绿色			2008.10.07 03:36:54	
将整体状态设置为绿色			2008.10.07 03:36:54	
启动了进一步处理			2008.10.07 03:39:17	
将状态设置为'进一步处理的'			2008.10.07 03:40:18	

日期: 00:00:00

BWP (1) B88 KGBWPRD INS



## 第二节

# 企业数据仓库实现

- 数据上传
- 数据多立方体
- 数据处理流程





## 2.2 数据多立方体

信息提供者	技术名称	M = A 版本	执行功能	显示树	对...
未分配的节点	NODESNOTCONNE...		更改		
数据仓库平台	ZKGBW1		更改		
公共对象	ZKGBW1_GG		更改		
财务	ZKGBW1_FI		更改		
成本	ZKGBW1_CO		更改		
工程	ZKGBW1_PS		更改		
销售	ZKGBW1_SD		更改		
采购与库存	ZKGBW1_MM		更改		
生产	ZKGBW1_PP		更改		
数据分析层	ZKGBW1_PP_ZA		更改		
报工数量	ZAPP_C02	=	管理		
生产产量	ZAPP_C04	=	管理		
生产年计划	ZAPP_C05	=	管理		
生产订单行项目	ZAPP_C01	=	管理		
生产订单	ZAPP_M01	=	更改		
报工数量	ZAPP_C02	=	管理	信息提供者	
生产产量	ZAPP_C04	=	管理	信息提供者	
生产年计划	ZAPP_C05	=	管理	信息提供者	
生产订单行项目	ZAPP_C01	=	管理	信息提供者	



## 第二节

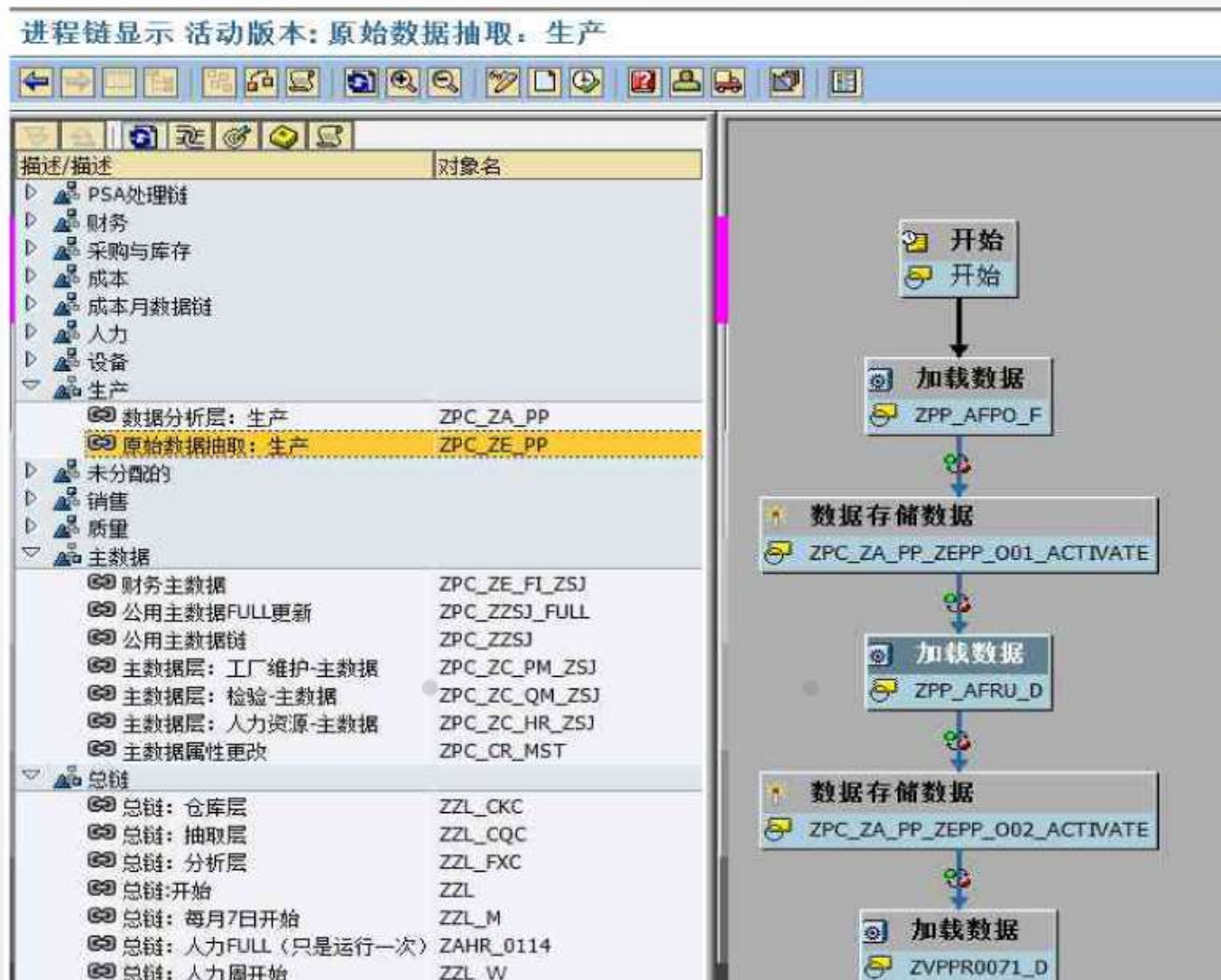
# 企业数据仓库实现

- 数据上传
- 数据多立方体
- 数据处理流程



## 2.3 数据处理流程

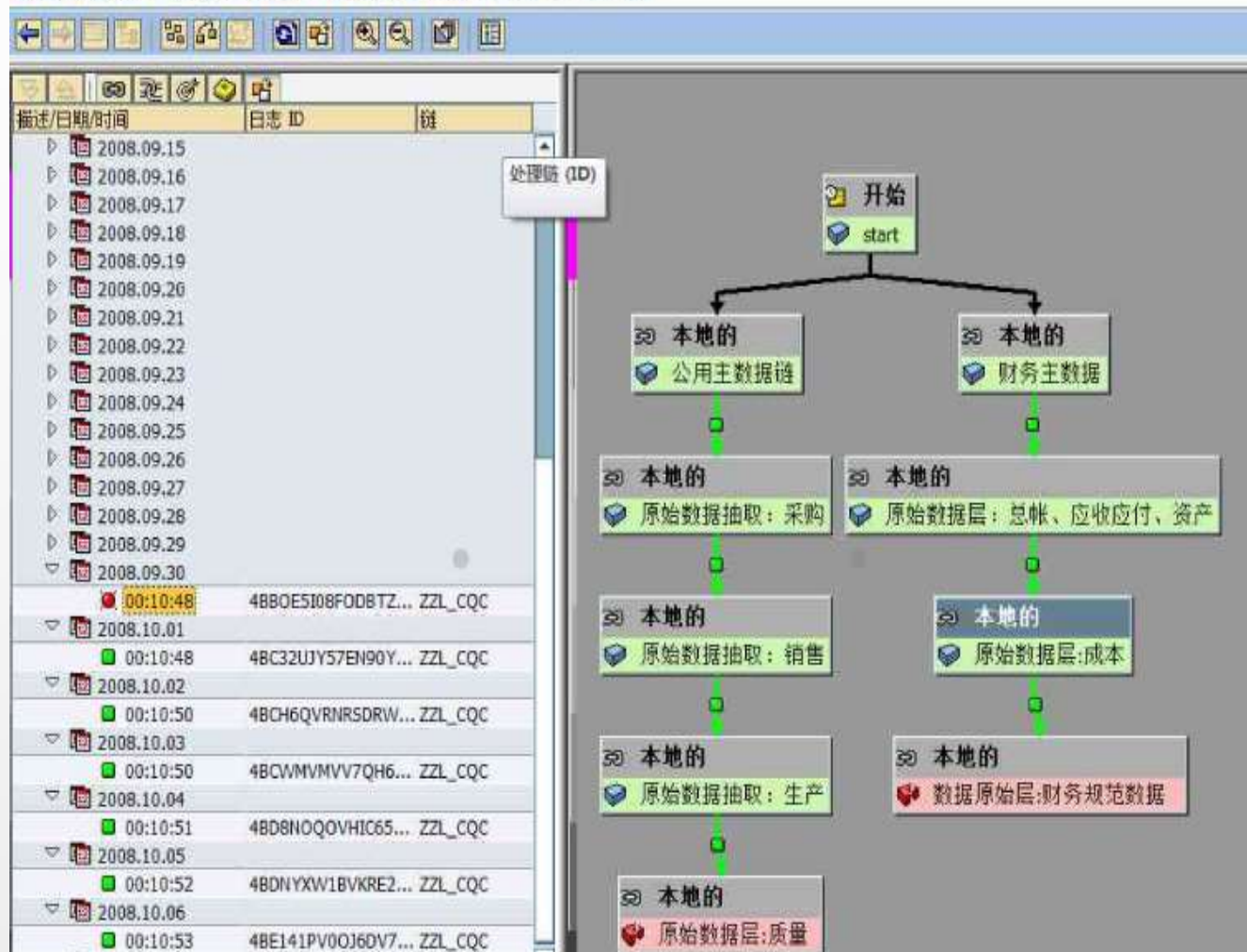
### A.示例



## 2.3 数据处理流程

### A.示例

进程链显示 日志视图 总链：抽取层 2008.09.30 00:10:48





## 第三节 企业数据可视化

- 图形展示
- 报表展示
- OLAP展示







## 3.1 图形展示







## 第三节 企业数据可视化

- 图形展示
- 报表展示
- OLAP展示





## 3.2 报表展示

生产产量对比分析 - KISC Portal

向上一级 字体大小 编码 内容控制 微游网址 智能报表 屏幕截图 清除记录

欢迎 eptest, 企业门户 (EP) 帮助 注册

搜索 系统管理

主页 管理报表 办公系统 安全信息 商务秘书 常用网站

日常管理报表 财务管理报表 成本管理报表 物资管理报表 生产管理报表 销售管理报表 设备管理报表 质量管理报表 人力资源报表 项目管理报表

生产产量对比分析 | 历史记录 | 返回 | 转发

详细导航

- 财务
- 成本
- 采购
- 生产
  - 股份生产经营日报
  - 生产产量对比分析
- 销售
- 库存
- 设备
- 质量
- 人事
- 项目

门户收藏夹

主要工序产量分析

工厂	月计划数量(吨)	当日毛产量(吨)	当日产量(吨)	月计划完成率	月进度比	月品种兑现率
总计结果	1,825,280.000	58,070.583	26,096.382	17.477 %	17.477	-34,274.3%
股份公司工厂	1,224,380.000	42,470.000	17,445.705	19.335 %	19.335	-247.0%
炼铁厂烧结一烧	19,000.000			1.170 %	1.170	-3,455.1%
炼铁厂烧结二烧	5,000.000				0.000	-967.7%
炼铁厂烧结三烧	175,000.000	6,645.000		20.051 %	20.051	1,218.0%
炼铁厂烧结球团	120,000.000	7,530.000		32.078 %	32.078	15,267.1%
炼铁厂烧结四烧	120,000.000	5,051.000		17.952 %	17.952	-1,683.8%
炼铁厂炼铁	240,000.000	8,460.000	3,759.650	19.645 %	19.645	697.0%
炼钢厂第一作业区	147,500.000	4,401.000	4,248.000	16.518 %	16.518	-4,183.8%
炼钢厂第二作业区	140,700.000	3,850.000	2,829.590	18.005 %	18.005	-1,899.8%
棒材厂第一作业区	85,000.000	2,988.000	2,841.380	21.013 %	21.013	1,409.8%



## 第三节 企业数据可视化

- 图形展示
- 报表展示
- OLAP展示



## 3.3 OLAP展示

### A.钻取 (追溯)

产量兑现率							
物料分类第一层							
物料分类第七层							
物料分类第三层							
物料分类第九层							
物料分类第二层							
物料分类第五层							
物料分类第八层							
物料分类第六层							
物料分类第四层							
列							
关键指标							
自由特性							
MRP控制员							
外部物料组							
库存地							
日历年/月							
日历日							
生产物料							
生产调度员							
半成品	664,920.000	40,917.690	592,162.525	270,427.041	89.058		89.05
废钢料		2.625	40.035	40.035	X		X
钢卷	110,420.000	5,237.435	73,975.920	20,546.888	66.995		67.00
正品	110,420.000	5,176.925	73,825.850	20,396.818	66.859		66.86
冷硬卷	6,000.000	687.065	6,167.325	3,264.099	97.211		102.79
平整卷	24,420.000	1,220.040	13,745.030	1,928.901	56.286		56.29
热卷	80,000.000	2,873.570	49,502.460	10,792.783	61.878		61.88
低合金高强度钢	5,410.000		271.580	-2,346.162	5.020		5.02
低碳镇静钢	36,900.000		26,796.550	8,941.711	72.619		72.62
碳素结构钢	37,690.000	2,873.570	22,434.330	4,197.233	59.523		59.52
Q235A		455.230	2,978.800	2,978.800			X
Q195C			162.180	162.180	X		X
Q235B	37,690.000	2,418.340	19,236.550	999.453	51.039		51.04
3<H≤4	2,040.000	792.830	1,033.400	46.303	50.657		50.66
热轧	2,040.000	792.830	1,033.400	46.303	50.657		50.66
未分配的	2,040.000	792.830	1,033.400	46.303	50.657		50.66
热卷 Q235B 3.5*1275*c	320.000	243.940	322.670	167.831	99.186		100.83
热卷 Q235B 3.75*1275*c	880.000			-416.129	0.000		
热卷 Q235B 4.0*1275*c	860.000	548.890	710.730	294.601	82.643		82.64

## 3.3 OLAP展示

### B.切片

数据分析 图形显示 信息 信息传输

产量兑现率报表 上次数据更新: 2008.10.25 07:11:11

变量屏幕 输出到 Excel 输出到 CSV

工厂	月计划产量	产量(当日)	产量(当月)	月进度比	品种兑现率	月计划完成率(%)
总计结果	159,300.000	4,733.170	87,912.320	10,831.675	55.187	55.19
股份公司工厂	159,300.000	4,733.170	87,912.320	10,831.675	55.187	55.19
150方坯	74,300.000	2,374.300	42,337.640	6,388.027	56.982	56.98
HRB335F			112.350	112.350	X	X
HRB400F		786.190	3,559.140	3,559.140	X	X
HRB400FT	100.000		3,421.180	3,372.793	-3,221.180	3,421.18
HRB335T	9,800.000		2,761.670	-1,980.265	28.180	28.18
HRB335FT	14,200.000		7,037.100	166.132	49.557	49.56
HRB335			408.550	408.550	X	X
HRB500	1,500.000		213.100	-512.706	14.207	14.21
HRB400	2,700.000		3,299.800	1,993.348	77.785	122.21
60	3,600.000		1,842.710	100.775	51.186	51.19
45	4,300.000	773.430	2,887.710	807.065	67.156	67.16

## 3.3 OLAP展示

### C.行列互换 (数据旋转)

数据分析

图形显示

信息

信息传播

产里兑现率报表

上次数据更新: 2008.10.25 07:11:11

变量屏幕

输出到 Excel

输出到 CSV

行

关键指标

列

工厂

物料

物料分类第七层

物料分类第六层

物料分类第四层

自由特性

工厂	总计结果	股份公司工厂	150方坯	HRB335F	HRB400F	HRB400FT	HRB335T
月计划产量	159,300.000	159,300.000	74,300.000			100.000	9,800.000
产量(当日)	4,733.170	4,733.170	2,374.700		786.190		
产量(当月)	87,912.320	87,912.320	42,337.640	112.350	3,559.140	3,421.180	2,761.670
月进度比	10,831.675	10,831.675	6,386.027	112.350	3,559.140	3,372.793	-1,980.265
品种兑现率	55.187	55.187	56.982	X	X	-3,221.180	28.180
月计划完成率(%)	55.19	55.19	56.98	X	X	3,421.18	28.18

## 小结



1. 介绍了一个企业数据仓库系统的设计与实现过程。
2. 主要对企业数据仓库的背景做了概述，讨论了企业数据仓库的需求分析和模型设计，并讨论了具体实现过程，最后给出了数据可视化的探讨。





张冬松  
[dszhang@nudt.edu.cn](mailto:dszhang@nudt.edu.cn)



# 谢谢! Q&A

THANKS FOR YOUR ATTENTION