

友乾营

专注数据技术的社群

报表和BI的区别

让工具选型不再迷茫



11 期

www.raqsoft.com.cn/yqy



BI (Business Intelligence) 也就是商务智能，它是一套完整的解决方案，用来将企业中现有的数据进行有效的整合，快速准确的展现报表并提出决策依据，帮助企业做出明智的业务经营决策。



数据仓库

存储数据

ETL

加工数据

OLAP

分析数据

报表

统计数据

BI基本概念—数据仓库

数据仓库是一个面向主题的、集成的、相对稳定的、反映历史变化的数据集合，用于支持管理决策。

与数据库的区别

- (1) 数据库是面向事务的设计，数据仓库是面向主题设计的。
- (2) 数据库一般存储实时数据，数据仓库存储的一般是历史数据。
- (3) 数据库设计是尽量避免冗余，数据仓库在设计是有意引入冗余。
- (4) 数据库是为捕获数据而设计，数据仓库是为分析数据而设计。

BI基本概念—ETL

ETL是将业务系统的数据经过抽取、清洗转换之后加载到数据仓库的过程，目的是将企业中的分散、零乱、标准不统一的数据整合到一起，为企业的决策提供分析依据。

数据抽取

来源：

数据库
文件

清洗转换

去错
转换格式
聚合汇总

加载数据

处理后的数
据加载到数
据(仓)库中

▶ BI基本概念—OLAP

联机分析，专门设计用于支持复杂的分析操作，侧重对决策人员和高层管理人员的决策支持，可以根据分析人员的要求快速、灵活地进行大数据量的复杂查询处理，并且以一种直观而易懂的形式将查询结果提供给决策人员，以便他们准确掌握企业（公司）的经营状况。

基本功能：

- WEB端在线操作
- 切片、切块
- 钻取、上卷
- 旋转
- 自定义指标
-

BI基本概念—报表

报表就是用表格、图表等格式来动态显示数据

当月销售业绩表 (1997-5)

工号	姓名	本月销售额	同比上月	与第一名的差距	与前一名的差距	排名
3	郑建杰	¥6580.00	-34.38%	0	0	1
8	张雷雷	¥4790.00	1964.66%	-1790	-1790	2
4	赵军	¥3955.00	-65.85%	-2625	-835	3
1	张颖静	¥3670.00	0%	-2910	-285	4
2	王伟	¥2953.00	11.22%	-3627	-717	5
7	刘高斌	¥2532.00	-38.93%	-4048	-421	6
6	金士鹏	¥765.00	-73.77%	-5815	-1767	7
9	张三	¥140.00	-76.74%	-6440	-625	8
5	孙林	¥0.00	0%	-6580	-140	9

雇员ID: 1

姓名: 张颖静

职务: 副总经理

出生日期: 1968-12-31

地区: 华北

家庭电话: (010) 65559857

地址: 复兴门245号

城市: 北京

邮政编码: 100085



雇员ID: 3

姓名: 郑建杰

职务: 销售代表

出生日期: 1968-09-19

地区: 华北

家庭电话: (010) 65558122

地址: 前门大街789号

城市: 北京

邮政编码: 198052



雇员ID: 2

姓名: 王伟

职务: 销售代表

出生日期: 1962-02-19

地区: 华北

家庭电话: (010) 65559482

地址: 罗马花园890号

城市: 北京

邮政编码: 109801



雇员ID: 4

姓名: 赵军

职务: 销售经理

出生日期: 1965-03-04

地区: 华北

家庭电话: (010) 65554848

地址: 学院路78号

城市: 北京

邮政编码: 100090





BI基本概念—报表

特点：格式复杂、样式丰富、贴近用户需求



**报表只是BI中的一个组成模块，
并不完全等同于BI。**

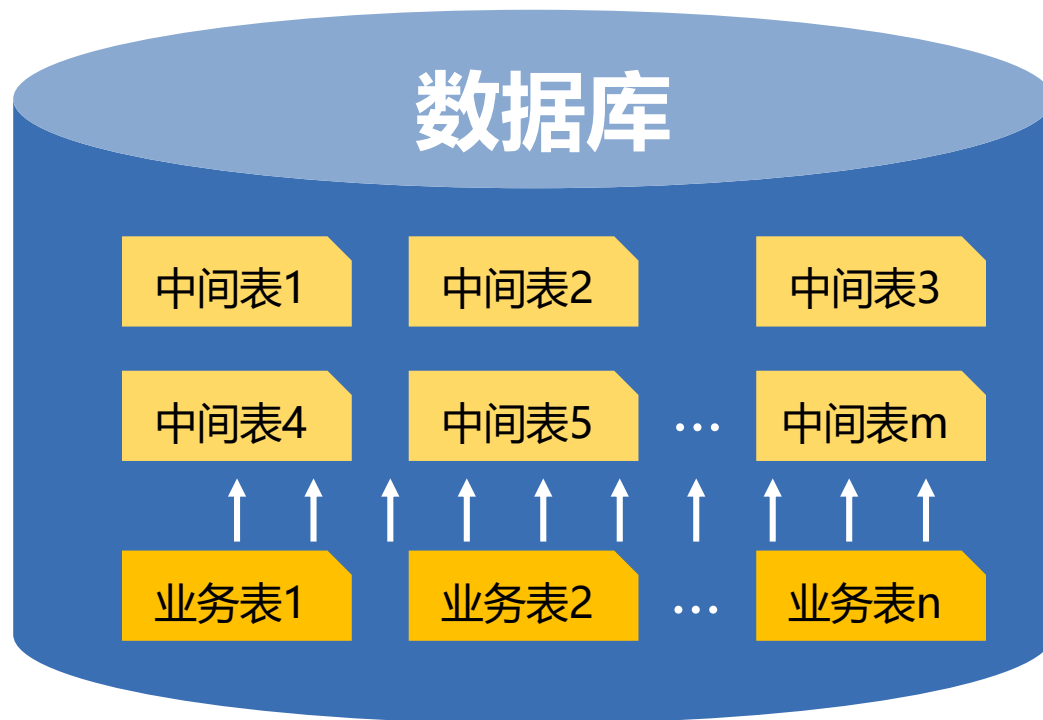
那么报表是否和BI完全没有关系？ ？ ？

数据量大、计算复杂带来的性能问题。

业务人员的自助报表需求。

中间表的使用

基于业务表，经过相应的数据预处理，将处理结果以数据表的形式存储在数据库中，形成数据库中间表



订单年汇总表

区域汇总表

订单季汇总表

人员汇总表

订单月汇总表

.....

订单日汇总表

产品汇总表

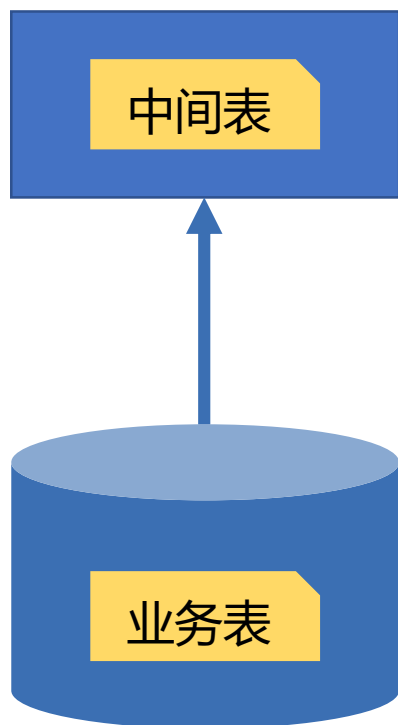
为什么需要中间表

功能

计算复杂，需要预存中间结果，以方便后续运算外部
数据无法计算运算，需要存成中间表

性能

基于基础表查询太慢，事先预处理形成中间表

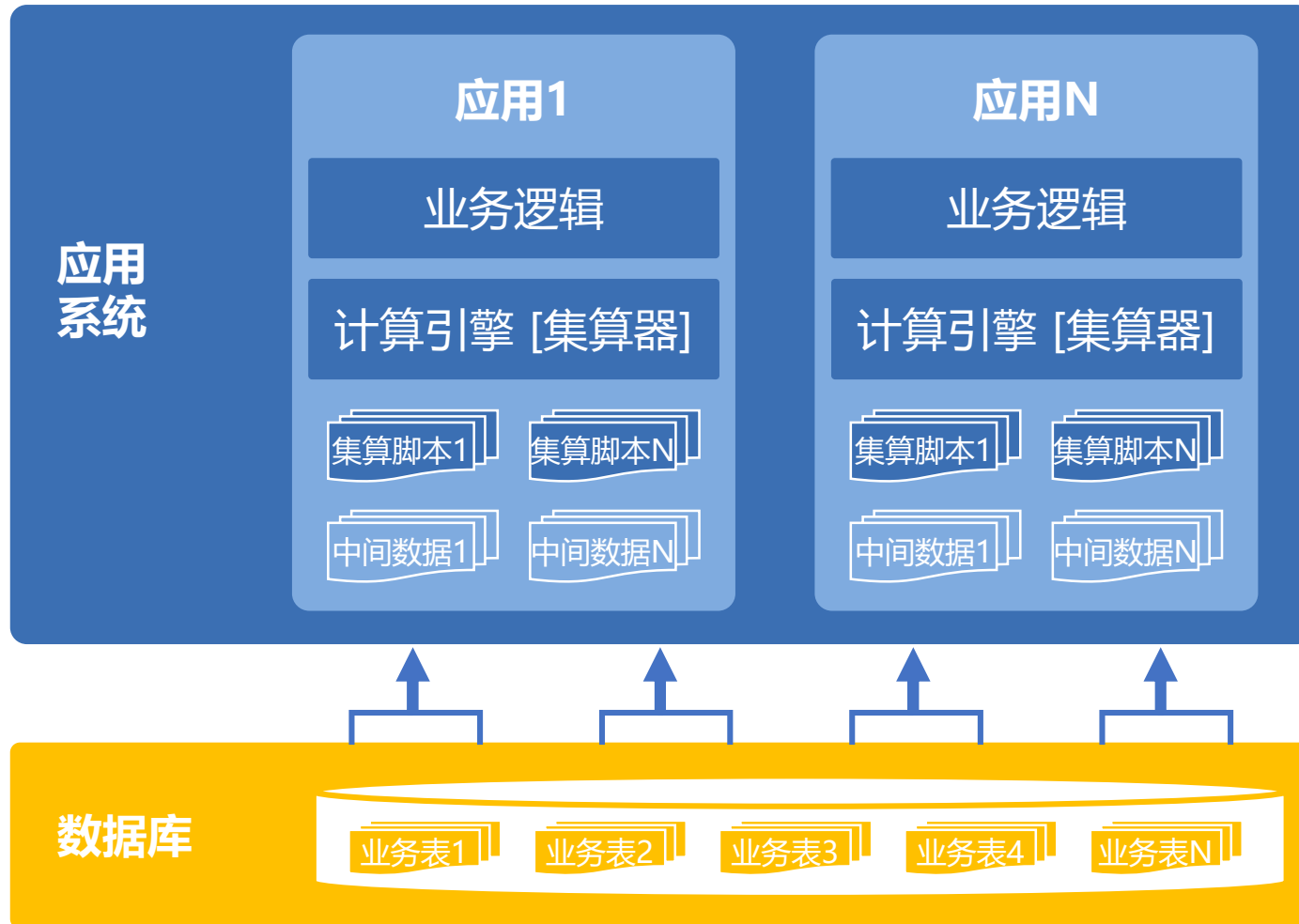


将中间表跑到库外!

使用外部文件系统存储中间表

- 树状结构的文件目录利于管理
- I/O性能更高
- 为数据库减负

应用结构图



中间数据外置存放于应用文件系统中



让文件拥有不依赖于数据库的计算能力

如何生成中间表（数据文件）

多数据源支持

很方便从多数据库或者数据文件中取数

强计算能力

可实施复杂计算，计算体系相对完备；实现简单

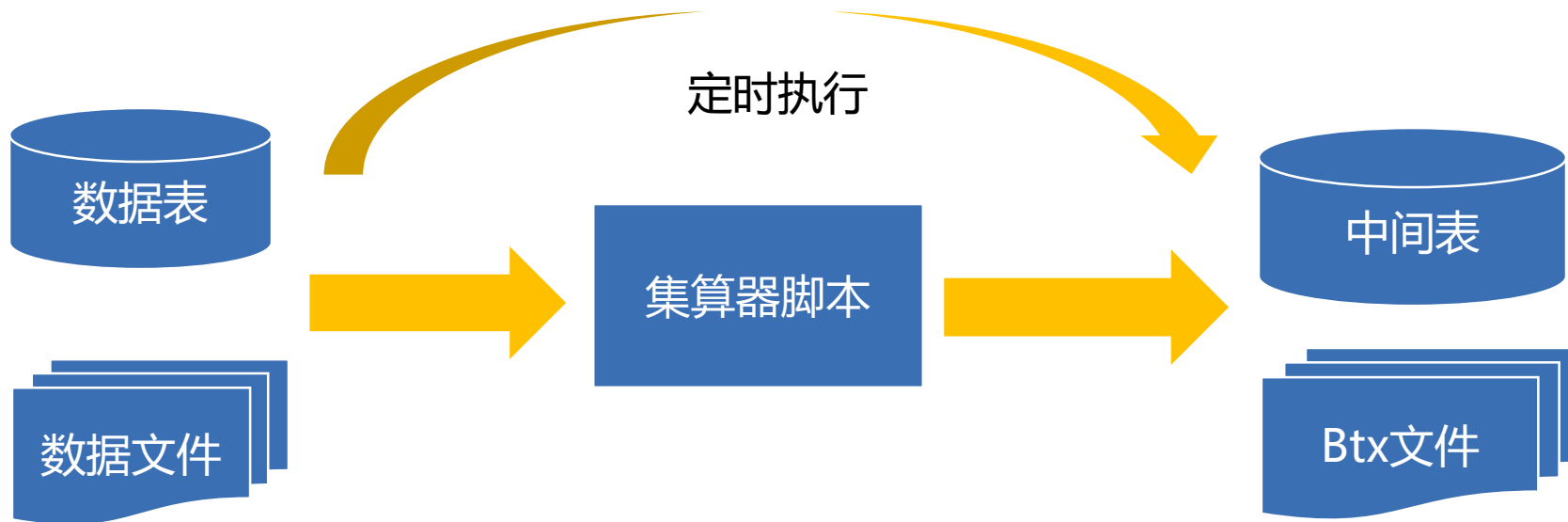
支持多源混合计算

支持异构数据源间的混合计算

自动生成

需要周期性的自动执行

如何生成中间表（数据文件）



普通报表

格式复杂并且相对固定

由技术人员实现

自助报表

格式简单、操作灵活

业务人员自定义



自助报表功能

建模

字典映射

维度拖拽

旋转

切片
切块

下钻

自定义指标

统计图

结果保存

特殊数据源

NOSQL数据库

跨库关联
取数

文本型
数据源

BI

是一套完整解决方案，包括不同的专业工具

报表

侧重数据展现，但是针对一些特定场景，也提供完整的解决方案，比如ETL，并不是专业的ETL工具，但能够实现对应功能



传统BI

展示数据，被动接收数据

智能BI

不仅展示数据，还提供相应结论



	25~30	31~40	41~50	51~60
一厂	20	30	45	60
二厂	30	30	40	50
三厂	60	70	20	10
四厂	70	50	15	10

针对这张表，传统BI或报表一般情况下只是负责数据的呈现，不提供结论。智能BI除了展示数据，还提供决策，例如某工厂直供年龄偏大，需要尽快招人。

好多乾

润乾线上直销系统



好多乾 - 润乾互联网营销

<http://www.raqsoft.com.cn/wx/hdq-strategy.html>