

sunny62520的专栏

目录视图 摘要视图 RSS 订阅

个人资料



sunny62520

访问：13484次
积分：235分
排名：千里之外

原创：6篇 转载：29篇
译文：0篇 评论：3条

文章搜索

文章分类

- javaSE (3)
- oracle (4)
- SQL Server (1)
- android (3)
- web (5)
- SSH (6)
- 杂 (1)
- Linux (5)
- 数据仓库 (1)
- sql (3)
- hadoop (1)

文章存档

- 2013年09月 (2)
- 2013年07月 (1)
- 2013年04月 (3)
- 2013年03月 (4)
- 2013年02月 (4)

展开

阅读排行

- oracle 视图的定义修改查 (1981)
- android完全退出 (关闭) (1742)
- box2d刚体移动的方法 (1130)
- 给JList中的Item添加右键 (1076)

【大声说出你的爱】CSDN社区情人节特别活动 专访李云：从通讯行业的架构师到互联网“新兵” 电子版《程序员》杂志免费领

星型模型和雪花模型（数据仓库模型）

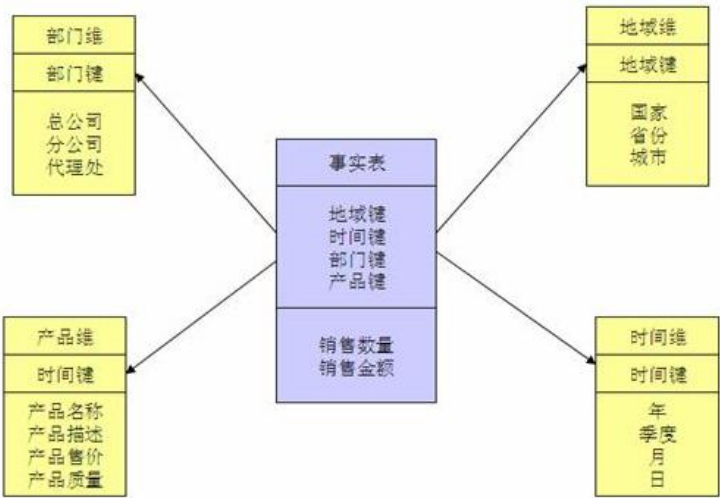
分类：数据仓库 2013-02-28 10:50 1034人阅读 评论(0) 收藏 举报

一、概述

在多维分析的商业智能解决方案中，根据事实表和维度表的关系，又可将常见的模型分为星型模型和雪花型模型。在设计逻辑型数据的模型的时候，就应考虑数据是按照星型模型还是雪花型模型进行组织。当所有维表都直接连接到“事实表”上时，整个图解就像星星一样，故将该模型称为星型模型，如图 1。

星型架构是一种非正规化的结构，多维数据集的每一个维度都直接与事实表相连接，不存在渐变维度，所以数据有一定的冗余，如在地域维度表中，存在国家 A 省 B 的城市 C 以及国家 A 省 B 的城市 D 两条记录，那么国家 A 和省 B 的信息分别存储了两次，即存在冗余。

图1. 销售数据仓库中的星型模型



当有一个或多个维表没有直接连接到事实表上，而是通过其他维表连接到事实表上时，其图解就像多个雪花连接在一起，故称雪花模型。雪花模型是对星型模型的扩展。它对星型模型的维表进一步层次化，原有的各维表可能被扩展为小的事实表，形成一些局部的“层次”区域，这些被分解的表都连接到主维度表而不是事实表。如图 2，将地域维表又分解为国家，省份，城市等维表。它的优点是：**通过最大限度地减少数据存储量以及联合较小的维表来改善查询性能**。雪花型结构去除了数据冗余。

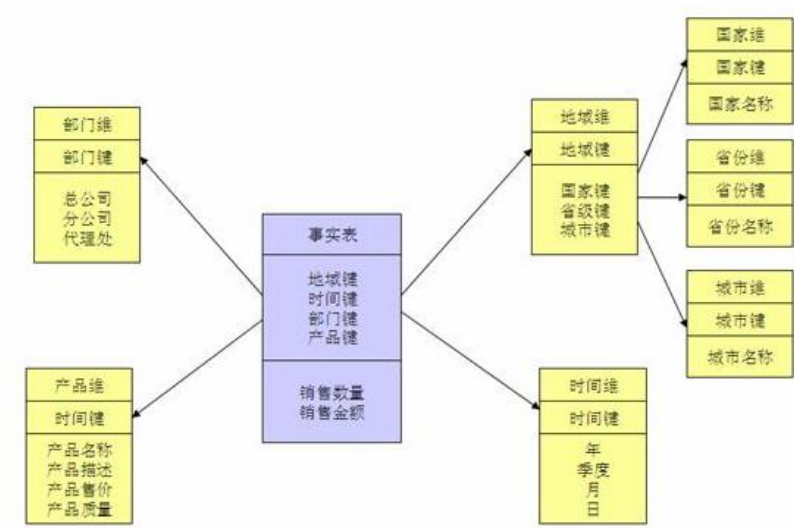
图 2. 销售数据仓库中的雪花型模型

- 星型模型和雪花模型（数 (1034)
- Java双缓冲机制--如何消 (507)
- oracle 索引的定义及使用 (431)
- 转-Java调用本机默认程 (390)
- JSP页面直接显示Blob类 (357)
- 在servlet中输出JS并且解 (269)

- 评论排行
- oracle 索引的定义及使用 (1)
- JSP页面直接显示Blob类 (1)
- android完全退出（关闭 (1)
- 星型模型和雪花模型（数 (0)
- 正确理解ThreadLocal (0)
- 逻辑与和短路与 (0)
- ThreadLocal---线程本地 (0)
- Struts2中ActionError与A (0)
- FreeMarker教程 (0)
- Oracle to_char(date,form (0)

- 推荐文章
- * 【GamingAnywhere源码分析之知识补充四】C操作Windows AD实现用户身份认证
- * Android学习小Demo（10）ToDoList的加强版
- * Axis2+SOAP Header验证方法实践（基于SOAP的Web安全调用机制）
- * Android跨应用间的数据传输
- * 【OpenCV入门教程之二】一览众山小：OpenCV 2.4.8组件结构全解析
- * Java7之集合类型 - 二叉排序树、平衡树、红黑树

- 最新评论
- JSP页面直接显示Blob类型图片波雅汉库克：不错！
- oracle 索引的定义及使用 华清545：简明扼要，~(≥▽≤)/~
- android完全退出（关闭所有Activyxhloen01：privateBroadcastReceiverbroadcastReceiver = new ...



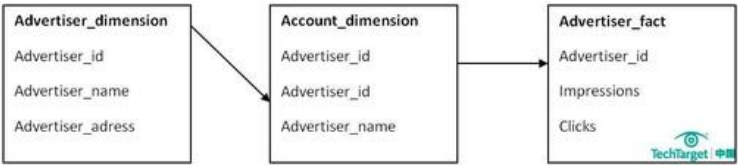
星型模型因为数据的冗余所以很多统计查询不需要做外部的连接，因此一般情况下效率比雪花型模型要高。星型结构不用考虑很多正规化的因素，设计与实现都比较简单。雪花型模型由于去除了冗余，有些统计就需要通过表的联接才能产生，所以效率不一定有星型模型高。正规化也是一种比较复杂的过程，相应的数据库结构设计、数据的 ETL、以及后期的维护都要复杂一些。因此在冗余可以接受的前提下，实际运用中星型模型使用更多，也更有效率。

二、使用选择

星形模型(Star Schema)和雪花模型(Snowflake Schema)是数据仓库中常用到的两种方式，而它们之间的对比要从四个角度来进行讨论。

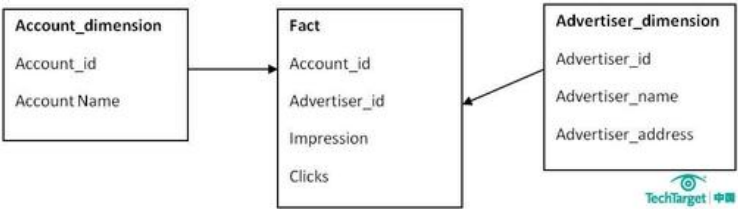
1.数据优化

雪花模型使用的是规范化数据，也就是说数据在数据库内部是组织好的，以便消除冗余，因此它能够有效地减少数据量。通过引用完整性，其业务层级和维度都将存储在数据模型之中。



▲图1 雪花模型

相比较而言，星形模型实用的是反规范化数据。在星形模型中，维度直接指的是事实表，业务层级不会通过维度之间的参照完整性来部署。



▲图2 星形模型

2.业务模型

主键是一个单独的唯一键(数据属性)，为特殊数据所选择。在上面的例子中，Advertiser_ID就将是一个主键。外键(参考属性)仅仅是一个表中的字段，用来匹配其他维度表中的主键。在我们所引用的例子中，Advertiser_ID将是Account_dimension的一个外键。

在雪花模型中，数据模型的业务层级是由一个不同维度表主键-外键的关系来代表的。而在星形模型中，所有必要的维度表在事实表中都只拥有外键。

3.性能

第三个区别在于性能的不同。雪花模型在维度表、事实表之间的连接很多，因此性能方面会比较低。举个例子，如果你想要知道Advertiser 的详细信息，雪花模型就会请求许多信息，比如Advertiser Name、ID以及那些广告主和客户表的地址需要连接起来，然后再与事实表连接。

而星形模型的连接就少的多，在这个模型中，如果你需要上述信息，你只要将Advertiser的维度表和事实表连接即可。

4.ETL

雪花模型加载数据集市，因此ETL操作在设计上更加复杂，而且由于附属模型的限制，不能并行化。

星形模型加载维度表，不需要再维度之间添加附属模型，因此ETL就相对简单，而且可以实现高度的并行化。

总结

雪花模型使得维度分析更加容易，比如“针对特定的广告主，有哪些客户或者公司是在线的？”星形模型用来做指标分析更适合，比如“给定的一个客户他们的收入是多少？”

更多 1

上一篇：[正确理解ThreadLocal](#)
下一篇：[Oracle to_char\(date,format\)](#)

相关主题推荐 [数据仓库](#) [商业智能](#) [数据存储](#) [解决方案](#) [数据库](#)

相关博文推荐

- fedora9只能进入黑客帝国（字符命令...
通常情况下：好天气多半会让人开心，而不好...
专家分享：IBM Lotus系列专场课件
Spring框架，使用ModelMap传...
- Mysql忘记root密码
11g R1 & R2新特性介绍(针对D...
iOS--加载、滑动翻阅大量图片解决方案...
Oracle RAC Cache Fus...



查看评论

暂无评论

您还没有登录,请[\[登录\]](#)或[\[注册\]](#)

* 以上用户言论只代表其个人观点，不代表CSDN网站的观点或立场

专区推荐内容

- 详解 HTML5 新特性
- 程序员永远的痛之字符编码的奥秘
- 企业安全技术（1）：检查平台AE...
- 使用非 html5 实现 js ...
- 跟燕青一起学Android应用开...
- 苹果的顶级应用程序：设计，语言，...

<< >>

更多招聘职位

我公司职位也要出现在这里

- 【北京佰易迅移动科技有限公司】Android研发工程师
- 【北京佰易迅移动科技有限公司】Android研发工程师
- 【百度（中国）有限公司】Android高级开发工程师
- 【北京极游互动时代科技有限公司】ios高级开发工程师
- 【广州广像信息科技有限公司】.net高级工程师
- 【上海洛熊网络科技有限公司】高级软件工程师

核心技术类目

- 全部主题 [Java](#) [VPN](#) [Android](#) [iOS](#) [ERP](#) [IE10](#) [Eclipse](#) [CRM](#) [JavaScript](#) [Ubuntu](#) [NFC](#)
[WAP](#) [jQuery](#) [数据库](#) [BI](#) [HTML5](#) [Spring](#) [Apache](#) [Hadoop](#) [.NET](#) [API](#) [HTML](#) [SDK](#)
[IIS](#) [Fedora](#) [XML](#) [LBS](#) [Unity](#) [Splashtop](#) [UML](#) [components](#) [Windows Mobile](#) [Rails](#) [QEMU](#)
[KDE](#) [Cassandra](#) [CloudStack](#) [FTC](#) [coremail](#) [OPhone](#) [CouchBase](#) [云计算](#) [iOS6](#) [Rackspace](#)
[Web App](#) [SpringSide](#) [Maemo](#) [Compuware](#) [大数据](#) [apttech](#) [Perl](#) [Tornado](#) [Ruby](#) [Hibernate](#)
[ThinkPHP](#) [Spark](#) [HBase](#) [Pure](#) [Solr](#) [Angular](#) [Cloud Foundry](#) [Redis](#) [Scala](#) [Django](#)
[Bootstrap](#)

