

国家环境信息与统计能力建设项目

环境信息数据仓库系统技术规定

Technical regulations of environmental data warehouse system

（征求意见稿）

《环境信息数据仓库系统技术规定》编制组

2010年10月

目 次

目 次.....	11
环境信息数据仓库系统技术规定.....	1
1 适用范围.....	1
2 术语和定义.....	1
3 缩略语.....	4
4 环境信息数据仓库系统架构.....	5
5 环境信息数据仓库系统数据源规定.....	5
6 环境信息数据仓库系统数据采集层（ETL）技术要求.....	6
7 环境信息数据仓库系统数据层技术要求.....	8
8 BI 层技术要求.....	11
9 环境信息数据仓库系统元数据技术要求.....	11
10 环境信息数据仓库系统标准规范体系要求.....	12
11 环境信息数据仓库系统基础设施要求.....	12
12 环境信息数据仓库系统数据安全要求.....	12
13 环境信息数据仓库系统运行维护要求.....	12
参考文献.....	13

环境信息数据仓库系统技术规定

1 适用范围

本技术规定对环境信息数据仓库系统体系架构及主要建设环节的工作方法和建设内容进行了规范。本技术规定适用于国家环境信息与统计能力建设项目数据仓库系统的设计、建设和运行维护。

2 术语和定义

2.1 数据仓库 data warehouse

一种用以支持经营管理中的决策制定过程，面向主题，涵盖不同时期的集成而稳定的数据集合。把信息加以整理归纳，并及时提供给相应的管理决策人员，是数据仓库的根本任务。数据仓库的工作流程一般包括收集数据、过滤数据、存储数据，以及把数据应用于分析、研究等应用程序。

[HJ/T 416—2007 术语和定义 7.129]

2.2 环境数据仓库 environmental data warehouse

主要用于存储环境主题分析相关的各类事实数据、维度数据，并提供各类环境主题分析功能。

2.3 数据集市 data mart

数据集市是数据仓库的一个子集，主要面向部门级业务，并且只面向某个特定的主题。数据集市存储为特定用户预先计算好的数据，从而满足用户对性能的需求。数据集市可以在一定程度上缓解访问数据仓库的瓶颈。

2.4 决策支持系统 decision making supporting system

采用信息技术、空间地理技术和多媒体技术等手段，以数据库为支撑，利用核心政务系统和其它政务系统收集相关决策信息，进行分析、预测和评估，提供智能化决策支持的一种应用系统。

[HJ/T 416—2007 术语和定义 7.63]

2.5 脏数据 dirty data

进行“清洗 Clean”时需要抛弃的数据，包括：

- a) 关键信息不完整的数据
- b) 错误的数据库
- c) 重复的数据
- d) 不符合业务逻辑关系的数据

2.6 数据采集 ETL

E (Extract)、T (Transform)、L (Load) 代表“抽取”、“转换”、“加载”，完成数据从数据源向目标数据仓库转化的过程，是实施数据仓库的重要步骤。基本的 ETL 过程如下所示：

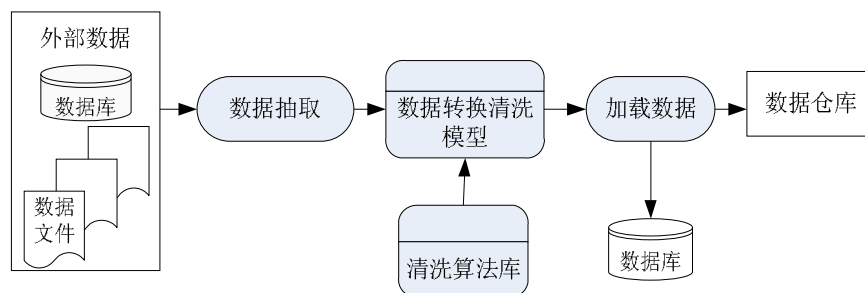


图 1 数据采集过程示意图

- e) 抽取 Extract: 将数据从各种原始的业务系统中读取出来;
- f) 清洗 Clean: 去除脏数据;
- g) 转换 Transform: 按照预先设计好的规则将抽取得数据进行转换, 使本来异构的数据格式能统一起来;
- h) 装载 Load: 将转换完的数据按计划增量或全部导入到数据仓库中。

2.7 在线分析处理 online analytical processing (OLAP)

一种使分析人员、管理人员或执行人员能够从不同的角度对信息进行快速、一致、交互式存取, 从而获得对数据的更深入了解的软件技术。其目标是为决策者提供决策支持, 以及满足在多维环境下的查询和报表需求。

注: 当今的数据处理大致可分成两大类: 在线事务处理OLTP与在线分析处理OLAP。OLAP的概念最早由关系数据库之父E. F. Codd于1933年提出。OLTP是传统的关系型数据库, 主要用来进行基本的、日常的事务处理, 例如银行交易。OLAP则是数据仓库系统, 主要用来支持复杂的分析操作, 侧重于决策支持, 并且能够提供直观易懂的查询结果。

[HJ/T 416—2007 术语和定义 7.83]

2.8 OLAP 数据库 OLAP database

即多维数据库 (Multidimensional Database), 是一种在某个预定义维度结构中 (包括预先计算好的聚集) 进行数据管理的数据库, 和关系型数据库中使用的表相对应。

2.9 业务事项 business items

业务事项是部门履行职能的核心和基础, 是指成某一业务功能的业务处理事项。每一个业务事项, 代表完成某一个功能的具体事项, 具有可执行性, 能够产生明确的业务结果, 如: 污染物排放行政许可事项、监督管理业务事项等重点业务事项。业务事项具有工作流程相对稳定、信息密集、实时性强等特点。

2.10 业务线 business lines

业务线是以业务功能为主线, 通过优化流程, 形成逻辑相关的一组业务事项, 是业务事项的功能聚类。

2.11 星型模式 star schema

星型模式可以描述为一个简单的星型: 中央表包含事实数据, 多个表以中央表为中心呈放射状分布, 它们通过数据库的主键和外键相互连接。在星型模式实施中, 将所有维度级别的维度数据存储存储在单个表或视图中。典型的星型模型示意图如下:

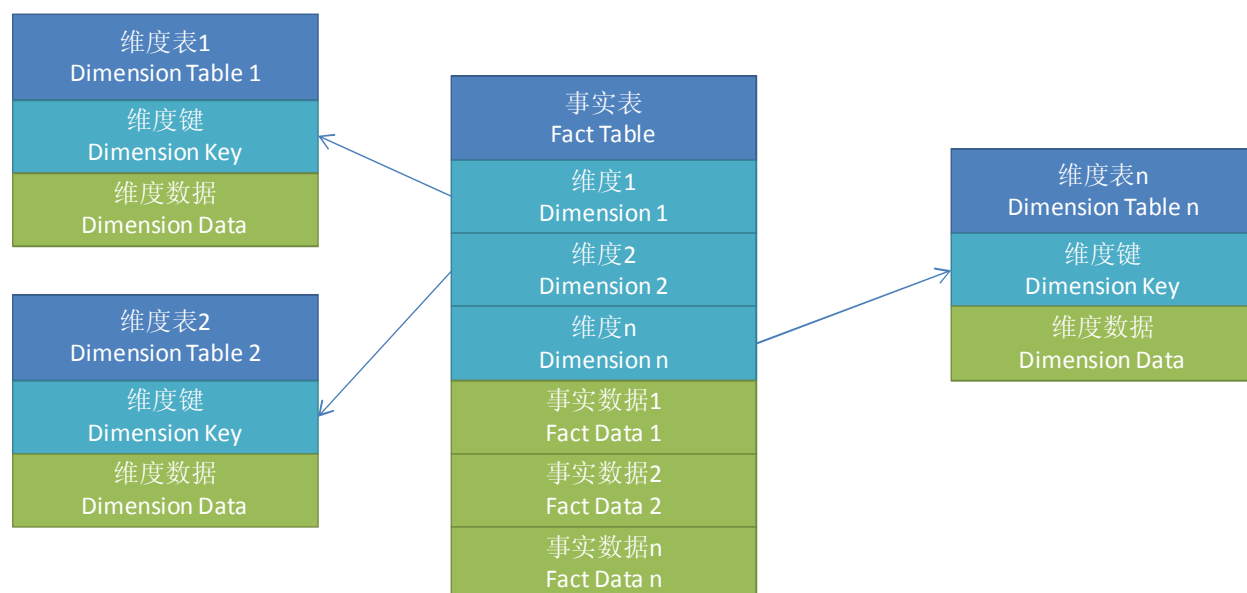


图 2 星型模型示意图

2.12 雪花模式 snow schema

雪花型模式表示一种维度模型，该模型也是由一个中央事实表和一组成员维度表组成，这些维度表可进一步规范化为子维度表。在雪花型模式实施中，将使用多个表或视图来存储维度数据。单独的数据表或视图存储与维中每个级别相关的数据。典型的雪花模型如下图所示：

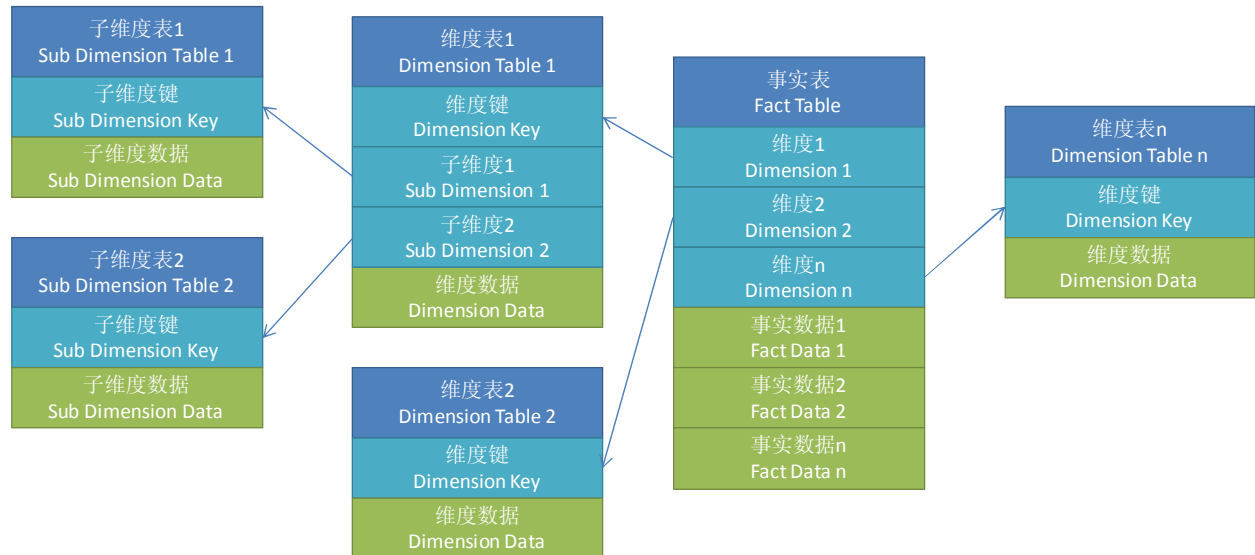


图 3 雪花模型示意图

2.13 星系模型 galaxy schema

当多个主题之间具有公共的维时，我们可以把围绕这些主题组织的星型模式通过共享维表，把事实表相互连接。这种多个事实表共享维表的星型模式集称为星系模式。这些共享的维度构成了数据仓库的“总线”，通过这些“总线”可以将各个独立的数据集市关联起来，从而为用户提供一个统一的视图。星系模型的示意图如下：

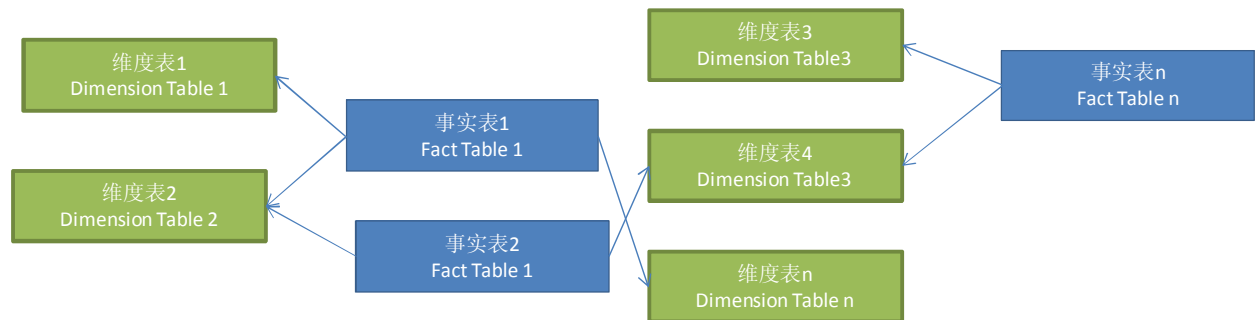


图 4 星系模型示意图

2.14 维度建模 dimensional modeling

维度建模是数据仓库建设中的一种数据建模方法。Kimball 最先提出这一概念。这种方法的最被人广泛知晓的名字就是星型模式。维度建模在物理上将用于量化业务的量度（也称为事实）与用于描述业务和对业务进行分类的描述性元素（也称为维）分隔开。

2.15 事实表 fact table

事实表存储用于量度业务（如指标）的事实。事实表还包含指向维表的外键。这些外键使事实表中的每个数据行与其对应的维和级别相关。

2.16 度量值 measurement

在多维数据集中，度量值是一组值，这些值基于多维数据集的事实数据表中的一列，而且通常为数字。此外，度量值是所分析的多维数据集的中心值。即，度量值是最终用户浏览多维数据集时重点查看的数字数据。所选择的度量值取决于最终用户所请求的信息类型。

2.17 维度 dimension

是 OLAP 数据库的结构特性。它们是事实数据表中用来描述数据的分类的有组织层次结构（级别）。这些分类和级别描述了一些相似的成员集合，用户将基于这些成员集合进行分析。

2.18 维度属性 dimension attribute

维度属性包括用于描述业务和对业务进行分类，如日期，行政区等。

2.19 维表 dimension table

OLAP 数据库中的表，其条目描述事实数据表中的数据。维表包含创建维度所基于的数据。

2.20 级别 level

级别是维度层次结构的一个元素。级别描述了数据的层次结构。

2.21 粒度 granularity

事实表所捕捉的内容的细节层次。

2.22 表分区 partition

表分区技术把数据分散存放到不同的物理磁盘中，提高这些磁盘的并行处理性能以优化查询性能。

2.23 商业智能 Business Intelligence

指用现代数据仓库技术、OLAP技术、数据挖掘和数据展现技术进行数据分析以实现商业价值。

2.24 数据挖掘 data mining

从大量数据中提取出可信、新颖、有效、能被人理解的信息的高级处理过程。

[HJ/T 416—2007 术语和定义 7.130]

2.25 聚集 clustering

数据挖掘的一种模式，聚集是对记录分组，把相似的记录在一个聚集里。

2.26 分类 classification

数据挖掘的一种模式，分类是从数据中选出已经分好类的训练集，在该训练集上运用数据挖掘分类的技术，建立分类模型，对于没有分类的数据进行分类。聚集和分类的区别是聚集不依赖于预先定义好的类，不需要训练集。

2.27 估算 estimation

数据挖掘的一种模式，估算与分类类似，不同之处在于，分类描述的是离散型变量的输出，而估算处理连续值的输出；分类的类别是确定数目的，估算的量是不确定的。

2.28 预测 prediction

数据挖掘的一种模式，预测是通过分类或估算起作用的，也就是说，通过分类或估算得出模型，该模型用于对未知变量的预测。

3 缩略语

本文中所涉及的缩写规定如下：

BI	商业智能	Business Intelligence
DW	数据仓库	Data Warehouse
DM	数据挖掘	Data Mining
DSS	决策支持系统	Decision Support System
EDW	环境数据仓库	Environmental Data Warehouse
ETL	抽取、清洗、转换、加载	Extract、Clean、Transform、Load
HOLAP	混合型联机分析处理	Hybrid On-Line Analytical Processing
MOLAP	多维型联机分析处理	Multidimensional On-Line Analytical Processing
ODS	操作型数据存储	Operational Data Store
OLAP	在线分析处理	On-Line Analytical Processing
ROLAP	关系型联机分析处理	Relational On-Line Analytical Processing

4 环境信息数据仓库系统架构

环境信息数据仓库系统体系架构由数据采集层、数据层、BI层、元数据管理、标准规范、安全管理和系统运维组成，具体架构如下图所示：

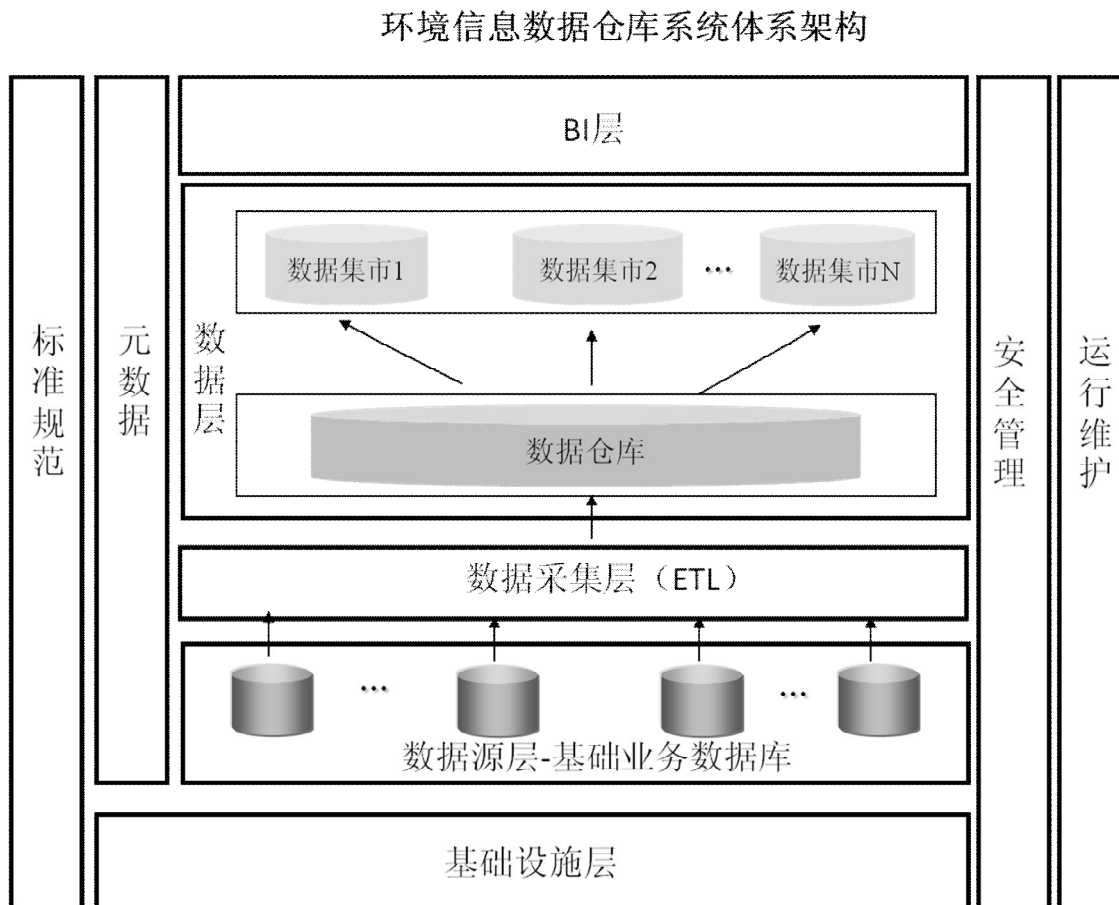


图 5 环境信息数据仓库系统架构示意图

5 环境信息数据仓库系统数据源规定

环境信息数据仓库系统的数据源为基础数据库，基础数据库分类如下：

- 1) 污染源监督性监测基础数据库：主要用于存储环境保护部和各省环保厅污染源监督性监测的业务数据，包括污染源废水监督性监测数据、污染源废气监督性监测数据、污染源噪声监督性监测数据等。
- 2) 污染源在线监测基础数据库：主要用于存储环境保护部和各省环保厅污染源在线监测的业务数据，包括污水排放口小时数据、废气排放口小时数据、污染治理设施数据、污染治理设施运行日数据等。
- 3) 污染源常规监测基础数据库：主要用于存储环境保护部和各省环保厅污染源常规监测的业务数据，包括污染源废气监测采样数据、污染源废气监测项目数据、污染源废水监测采样数据、污染源废水监测项目数据等。
- 4) 排污申报基础数据库：主要用于存储环境保护部和各省环保厅排污申报的业务数据，包括污染源废气申报数据、污染源废水申报数据、污染源噪声申报数据、建筑施工噪声申报数据等。
- 5) 排污收费基础数据库：主要用于存储环境保护部和各省环保厅排污收费的业务数据，包括污染源废气缴费数据、污染源废水缴费数据、污染源噪声缴费数据、建筑施工噪声缴费数

据、排污费减免处理数据等。

- 6) 环境统计基础数据库：主要用于存储环境保护部和各省环保厅环境统计的业务数据，包括工业企业污染排放处理利用数据、火电企业污染排放及处理租用数据、工业企业排放废水/废气中污染物监测数据等。
- 7) 建设项目管理基础数据库：主要用于存储环境保护部和各省环保厅建设项目管理的业务数据，包括建设项目基本数据、工业类项目申请数据、土地类项目申请数据、饮食娱乐类项目申请数据、项目跟踪数据。
- 8) 三同时验收基础数据库：主要用于存储环境保护部和各省环保厅三同时验收的业务数据，包括三同时项目审批数据、三同时项目试运行数据、三同时项目验收数据、三同时项目环评报告数据。
- 9) 环境质量基础数据库：主要用于存储环境保护部和各省环保厅环境质量的业务数据，包括大气监测数据、地表水监测数据、地下水监测数据、降尘监测数据、降水监测数据、区域噪声监测数据、道路交通噪声监测数据等。
- 10) 文档数据库：存储减排相关法律法规、应用标准等文档信息。
- 11) 信息分类及标准代码库：主要包括减排业务中的各类标准化代码，如污染源类别等。

6 环境信息数据仓库系统数据采集层（ETL）技术要求

6.1 环境信息数据仓库 ETL 过程规定

环境信息数据仓库 ETL 过程如下图

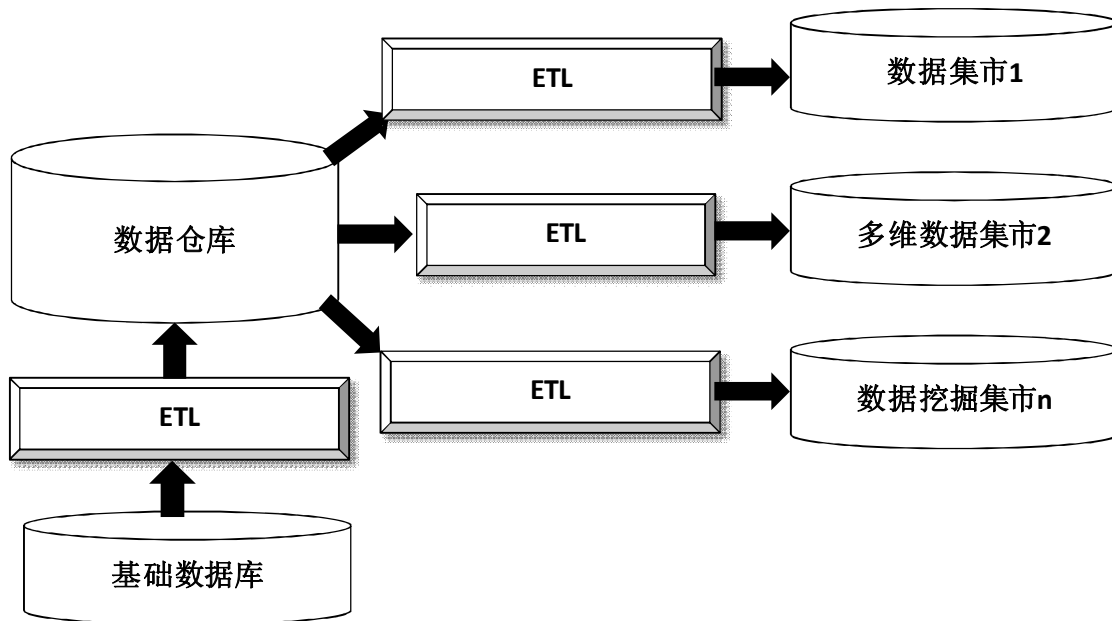


图 6 ETL 示意图

各 ETL 过程主要任务如下表

序号	源系统	目标数据库	ETL主要任务
1	基础数据库	数据仓库	抽取，转模，加载，去除重复数据
2	数据仓库	数据集市	汇聚加载，满足BI分析要求

6.2 ETL 主要工作步骤规定

6.2.1 源数据现状分析和映射

每做一步ETL工作之前都应对源数据进行全面分析，包括如下内容：

- 数据的业务范围及业务含义

-
- 数据所在平台，包括系统平台和数据库平台
 - 数据结构
 - 数据更新周期
 - 数据更新方式
 - 数据量

源数据分析完成之后，应将源数据与目标数据模型进行匹配，即源数据映射。

国家环境信息与统计能力建设项目的数据仓库的源数据应为基础数据库。新增源数据应先进入基础数据库再ETL至数据仓库。

6.2.2 数据抽取

数据抽取从数据源提取数据，供数据清洗和转换环节进行处理，数据抽取方式包括批量抽取和单次抽取两种方式，批量抽取可将所有历史数据一次性抽取完成；单次抽取根据规则要求进行抽取。

控制单次抽取可采用手工和自动两种方式，手工方式通过人为方式启动任务，自动方式基于调度程序，定期定时执行抽取任务。

抽取的检验：在抽取完成后，进行数据源与目标数据库的对比，发现并解决抽取过程产生的错误。

抽取的日志记录：在抽取过程中要全程记录抽取日志，至少包括：数据记录条数、开始时间、完成时间，错误信息等。

6.2.3 数据清洗

数据清洗的任务主要进行脏数据处理，应遵守的主要规则如下：

a) 清洗规则

- 列筛选。根据值域设置，对每个列中的数据进行测试。
- 结构筛选。主要检查主键/外键关系。
- 业务规则筛选。按照业务规则进行数据检查，如空气质量等级与污染源监测结果是否一致。

b) 对数据清洗转换过程的日志记录，包括：

- 原数据记录
- 脏数据记录
- 运用的清洗规则
- 错误信息

6.2.4 数据转换

数据转换的任务主要进行不一致的数据转换、数据粒度的转换，以及一些商务规则的计算。应遵守的主要规则如下：

a) 转换规则

- 拆分数据：将一个数据分解为多个值
- 处理空值，捕获空值，根据规则替换为对应数据
- 格式化数据：统一数据源中同类数据的格式，具体包括时间、数值、字符、计量单位等数据
- 替换数据：根据规则用标准数据替换原来的数据，例如用标准编码替换业务系统自定的编码
- 数据粒度的转换：将业务系统数据按照数据仓库粒度进行聚合。
- 保证转换后的误差在规定的范围内。

c) 对数据转换过程的日志记录，包括：

- 原数据记录
- 转换后数据记录
- 运用的转换规则
- 转换的时间

- 错误信息

6.2.5 数据加载

数据加载工作实现转换后数据的入库。在数据加载后，还要更新元数据，以反映刚完成的数据加载活动。数据的加载活动应使用标准方法和通用工具，例如，在关系数据库管理系统中可以使用SQL 或专门用于管理数据仓库的数据库管理系统加载工具。在数据装入数据仓库数据库后，应验证维度和事实数据表之间的引用完整性，确保所有的记录都与其他表中的适当记录相关。

6.2.6 数据缓存区规定

数据缓存区用来支持数据抽取、清理、转换并加载到数据仓库，应为数据缓存区创建单独的数据库。数据缓存区应包括传入数据的表、帮助实现代理键的表以及容纳转换数据的表，以及其他表以协调来自不同数据源的数据。

数据缓存区可做归档，以备后查。

6.2.7 ETL 周期设置

根据粒度设置要求对 ETL 周期进行设置。

6.2.8 ETL 管理要求

对 ETL 管理主要包括 ETL 的调度管理和 ETL 的监控管理。

ETL 调度管理要求包括三个方面的统一：统一界面、统一配置、统一调度。在界面方面，应实现统一的图形界面方式，且清晰易懂；在配置方面，应具备以下要素的配置要求：自动调度时间的设置、各调度要素之间逻辑关系的设置、指定调度要素的设置(包含指定数据源系统、功能模块、接口表等)、最大任务数的设置；在统一调度方面，应通过统一界面进行统一配置后，能够统一的进行调度程序的启动和停止。在不同阶段调用相应的资源进行处理，以支撑 ETL 的整个过程。ETL 的调度支持流水线并行数据处理。

ETL 的监控管理包括对 ETL 过程处理状态的监控，出错信息的监控，提供处理开始时间，完成时间，成功失败标识等，同时进行完整的处理日志记录，并提供通过对日志记录的分析达到对整个处理过程的统计和分析能力。一旦发现错误需要提供邮件等多种方式进行报警。

7 环境信息数据仓库系统数据层技术要求

7.1 环境信息数据仓库数据仓库维度建模

采用维度建模方法进行环境信息数据仓库逻辑模型设计。维度建模方法由以下阶段组成：

- 确定业务线
- 粒度设计
- 维度设计
- 度量值设计

7.1.1 确定业务线

参考基础业务库对业务线的分类，业务线由以下几部分组成：

- 污染源监督性监测
- 污染源在线监测
- 排污申报
- 排污收费
- 环境统计
- 建设项目管理
- 三同时验收
- 环境质量监测

注：文档数据库和信息分类及标准代码库不放在业务线分类中。

7.1.2 粒度设计

粒度设计时应参考如下原则：

-
- a) 满足业务需求。
 - b) 数据源应能够提供的数据。
 - c) 粒度设计应尽可能在最细节层次上进行考虑。
 - d) 可进行混合粒度声明。

示例：

环境空气质量监测粒度声明：

- a) 原子粒度声明
每个城市每个自动检测站各种污染物浓度小时均值。
- b) 粗粒度声明
每个城市每个自动检测站各种污染物浓度日均值。

7.1.3 维度设计

维度属性设计应参考如下原则：

- a) 维度属性应由完整的词汇组成
- b) 应由描述性的名词组成
- c) 维度属性设计应是完整的，没有遗漏的
- d) 定义好的维度属性应进行质量检验，如没有错误拼写

示例：

环境空气质量主题一般维度包括：

- a) 时间维度
- b) 行政区维度
- c) 自动检测站维度
- d) 空气质量指数维度
- e) 空气质量级别维度
- f) 首要污染物维度

7.1.4 度量值设计

度量值选择原则：

- a) 度量值的选取根据业务需要选取数值型指标。

示例：环境空气质量监测度量值包括二氧化硫、总悬浮颗粒物、可吸入颗粒物、氮氧化物、二氧化氮、一氧化碳、臭氧、铅、苯并芘、氟化物等污染物浓度。

7.2 数据仓库物理设计与性能设计原则

物理设计按照如下步骤进行设计。

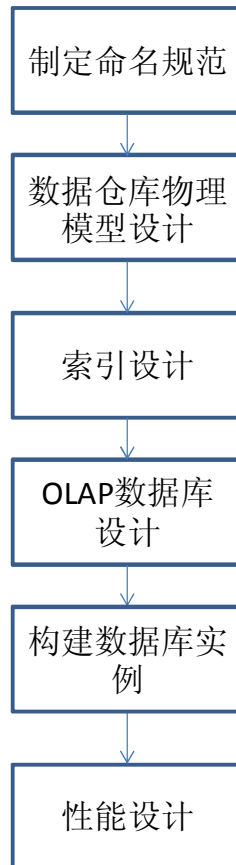


图 7 数据仓库设计流程图

7.2.1 制定命名规范

数据库命名规范参考HJ/T 419-2007。

7.2.2 数据仓库物理数据模型设计

- a) 物理模型应尽量与逻辑模型一致。
- b) 在设计物理数据模型时，可根据所选的数据库管理系统的特点做相应调整，来改善查询性能或提高系统的可维护性。
- c) 物理模型应对物理数据库的特征进行详细的描述。
- d) 物理数据模型除设计包括事实表、维度表以及为 ETL 过程提供的临时库的设计。
- e) 应采用数据建模工具由逻辑数据模型直接生成物理模型。

7.2.3 索引设计

环境信息数据仓库系统的数据更新频率低，应设计多种索引结构或对各个数据存储建立专用的复杂的索引，以获得最高的存储效率。

应为以下表或数据建立索引：

- a) 维度表
- b) 事实表
- c) OLAP 数据库

7.2.4 OLAP 数据库设计

OLAP数据库设计包括以下内容：

- a) 定义 OLAP 数据粒度和钻取深度
- b) 完善 OLAP 维度
- c) 定义 OLAP 计算，即将计算从报表定义移到 OLAP 数据库引擎中完成。

7.2.5 构建数据库实例

- a) 内存，内存比磁盘的访问性能高很多，因此尽可能扩大内存容量。

b) 块大小，数据库块从 2KB 到 32KB，应采用小块。

7.2.6 性能设计

环境信息数据仓库系统主要是对数据的查询和检索，没有修改操作，主要考虑数据查询性能，主要采用以下的策略：

a) 数据存放位置

应根据数据的重要程度、使用频率以及响应时间的要求将不同类的数据分别存储在不同的存储设备中。重要程度高、经常存取并对响应时间要求高的数据存放在高速存储设备上，如硬盘；存取频率低或对存取响应时间要求低的数据则可以放在低速存储设备上，如光盘或磁带。

b) 存储分配

存储分配主要包括块的大小、缓冲区大小和个数等。

c) 采用表分区（Partition）技术对事实表进行分区存储。

d) 对查询分析效率要求比较高的应用采用 OLAP。

e) 采用磁盘阵列（RAID）技术提高数据访问的效率。

f) 将索引数据与表数据存储在不同的表空间。

g) 并行处理

可根据具体情况采用相关的方法对数据存储结构进行优化，以满足数据访问的要求。

7.3 数据集市设计原则

根据数据集市数据库结构复杂程度、数据量大小、查询性能要求等因素，选用相应的OLAP存储方式，ROLAP适合结构复杂、数据量大同时对性能要求不高的数据集市设计，MOLAP适合结构简单、数据量小但对性能要求高的数据集市。

数据集市具体设计原则与数据仓库相同，可参考数据仓库设计方法。

8 BI 层技术要求

BI应用模式如下：

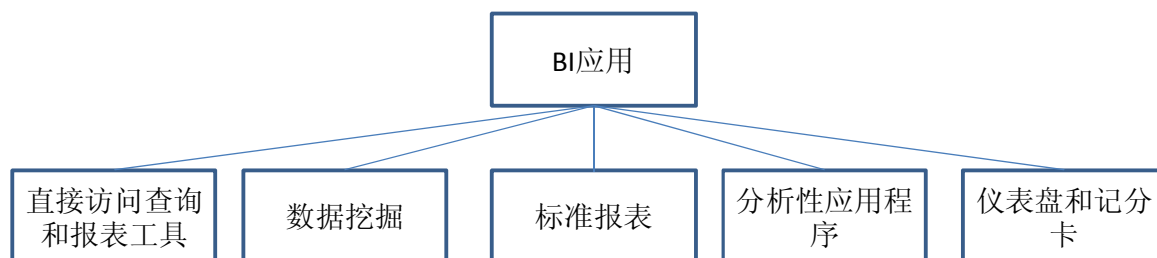


图 8 BI 应用示意图

直接访问查询和报表工具：可直接查询维度模型并定义结果集，可用来进行标准报表开发。

数据挖掘：应提供聚集，分类，估算和预测，相关性分组，以及异常探测等常用挖掘方式。

标准报表：应提供简单的可定义的用户交互，如输入参数，可向下钻取较低层次的细节数据，以及联接到相关报表。

分析性应用程序：应能够嵌入环境领域专业知识从而对某具体业务过程进行分析。

仪表板和记分卡：应由报表和图表组合而成，提供突出异常和向下钻取功能对多个业务过程的数据进行分析。

9 环境信息数据仓库系统元数据技术要求

元数据管理应遵循以下原则：

- 环境信息数据仓库系统元数据编制应符合《环境信息元数据标准》（在编）的具体规定。
- 应对所有元数据进行编目。

- 应使用标准软件库的形式管理元数据，包括版本控制、版本迁移和可靠的备份恢复。
- 为元数据编制文档，形成关于元数据的元数据。

10 环境信息数据仓库系统标准规范体系要求

环境信息数据仓库系统建设必须在标准体系规范指导下进行，标准规范体系包括运维、应用、数据、基础设施等各方面建设内容的标准与规范的综合，用来指导、协调和优化数据仓库的建设与运维。

11 环境信息数据仓库系统基础设施要求

参考《环境信息能力建设评价技术要求》，基础设施包括：

- a) 网络与安全
- b) 服务器及存储设备
- c) 机房及外设
- d) 系统软件与应用支撑

由于环境信息数据仓库系统为环境决策提供支持，在基础设施配置上应满足系统的性能、可靠性、可用性、安全性等方面的要求。

12 环境信息数据仓库系统数据安全要求

数据安全设计需要从用户范围、数据操作权限、用户权限、以及数据自身的安全进行设计。

- a) 用户范围，需要描述使用数据仓库的各类用户，以及他们的需求。
- b) 数据操作权限，包括：
 - 创建
 - 修改
 - 删除
 - 检索
- c) 用户权限，包括：
 - 针对用户范围中所描述的用户，确定每一类用户可以使用的数据
 - 针对用户范围中所描述的用户，确定每一类用户可以进行的操作
- d) 数据自身安全，包括：
 - 数据的备份策略
 - 数据的还原策略
 - 数据加密策略，涉及到必需进行加密的数据包括：
 - 管理员密码
 - 用户密码

环境信息数据仓库系统数据安全参考《环境信息安全技术规范》（在编）执行。

13 环境信息数据仓库系统运行维护要求

环境信息数据仓库系统运行维护体系建设参考《环境信息应用系统运行管理维护技术规定》（在编）。

参考文献

- | | |
|-------------------|-----------------------|
| GB3095-1996 | 环境空气质量标准 |
| GB 8978-1996 | 污水综合排放标准 |
| HJ/T 416 | 环境信息术语 |
| HJT193-2005 | 环境空气质量自动监测技术规范 |
| HJ/T 417-2007 | 环境信息分类与代码 |
| HJ/T 419-2007 | 环境数据库设计与运行管理规范 |
| HJ565-2010 | 环境保护标准编制出版技术指南 |
| SDS/T 2111 – 2004 | 元数据标准化基本原则和方法 |
| GB/T 19488.2-2008 | 电子政务数据元第 2 部分：公共数据元目录 |