

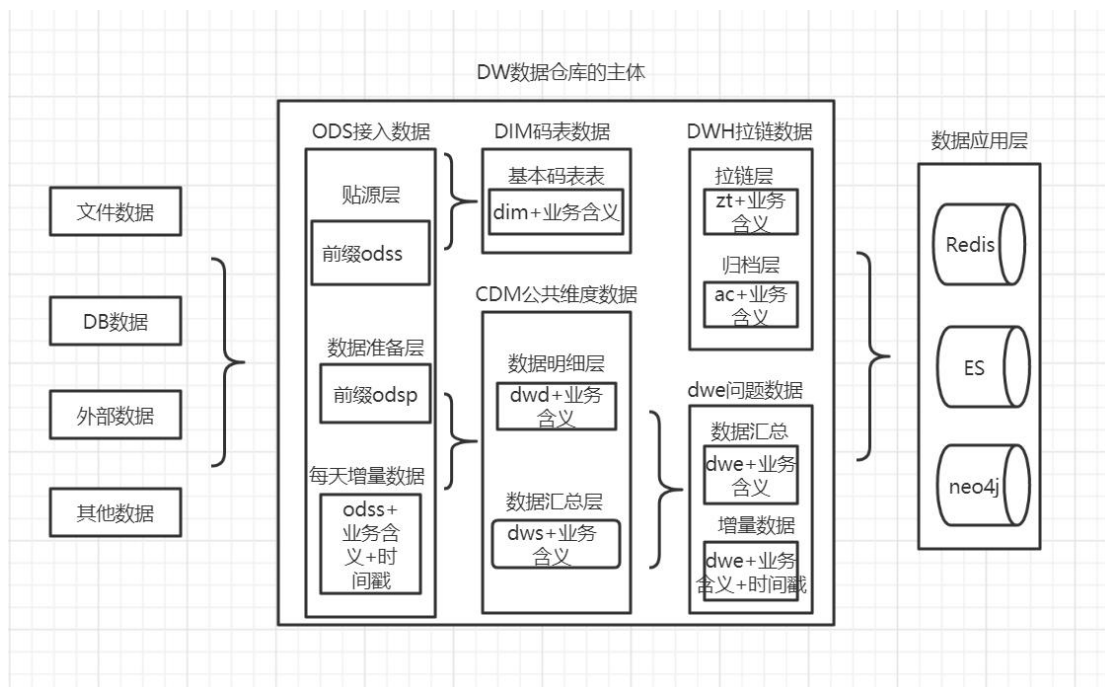
Greenplum 数据仓库

目录

- Greenplum 数据仓库.....1
- 目录..... 1
- 1 数据仓库流程图.....1
- 2 数仓的命名规范..... 2
- 3 命名规范实例.....3
 - 3.1 基本命名原则.....3
 - 3.2 表命名实例.....3
 - 3.3 字段命名实例.....3
 - 3.4 索引命名.....3
 - 3.5 约束命名.....4
 - 3.6 触发器命名.....4
 - 3.7 函数过程命名.....4
 - 3.8 序列命名.....4

1 数据仓库流程图

以下为数据的建设的整个流程图，包括数据的入库，数仓的设计以及数据与应用层的交互



2 数仓的命名规范

以下为创建 schema 与表的规范

数据模型规范要求的层次	GP Schemam	物理模型层次名称	物理表前缀
ODS（数据接入层）	ODS	贴源层	ODSS_
		数据准备层	ODSP_
CDM（公共数据层）	CDM	DWD 数据明细层	DWD_
		DWS 数据汇总层	DWS_
DWH（历史数据层）	DWH	ZT 拉链层	ZT_
		AC 归档层	AC_
DIM（公共纬度层）	DIM	CT 码表	CT_
		SYS 系统层	SYS_
DWE（问题数据处理层）	DWE	DWE 异常数据	DWE_

3 命名规范实例

说明：指数据库对象如表（TABLE）、序列（SEQUENCE）、过程（PROCEDURE）、触发器（TRIGGER）等的命名约定。

3.1 基本命名原则

规则 1:命名使用具有意义的英文词汇，词汇中间以下划线分隔。

规则 2:命名只能使用英文字母，数字，并以英文字母开头,英文字母必须是小写

规则 3 表名，学列，过程，触发器不可以以下划线开头

规则 4:避免用 GP 的保留字如 desc，关键字如 index。

3.2 表命名实例

实例：

1、odss_e_ent_baseinfo_20180506（贴源层-E 数据源-企业基本信息表-2018 年 5 月 6 日的增量数据）

2、zt_ent_invinfo(企业投资信息拉链表)

3、odsp_e_ent_person_20180506（数据准备层-E 数据源-企业高管表-20180506 处理的数据）

4、ct_e_map(E 数据源使用的码表)

5、dwd_ent_baseinfo（明细层-企业信息表）

注意：

1、同一个模块的表尽可能使用相同的前缀，表名称尽可能表达含义。

2、长度不超过 40 个字符。

3、表名应该以数据层名称开头，如果需要区分数据源要以下划线分割，如果是按日期区分的表应该在结尾加 yyyymmdd 日期数字

3.3 字段命名实例

1、表达其实际含义的英文单词或简写。

2、各表之间相同意义的字段应同名。

3.4 索引命名

1、表名称_字段名_IND

示例：employee 表 first_name 上的 index：employee_fid_ind

3.5 约束命名

- 1、主键约束：表名称_PK。
- 2、唯一约束：表名称_字段名_UK。

3.6 触发器命名

- 1、TRG_表名_操作。

3.7 函数过程命名

- 1、数据处理的过程使用 proc_开头
- 2、公共方法等函数使用 fn_开头
- 3、采用动词+名词的形式表达其含义。

3.8 序列命名

- 1、SEQ_表名