详细设计说明书

融通物业项目数据及服务子项目组

编写说明

标题：详细设计说明书

版本历史：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **版本** | **编写** | **日期** | **修改说明** | **审核** | **批准** |
| V1.0 | 任泽华 | 20210728 | 创建初始版本  完成设计规范、DWI模型设计、数据采集功能模块设计 |  |  |
| V1.1 | 任泽华 | 20210802 | 完成前言、总体架构、DWR模型设计、主题库数据处理设计、数据服务设计、调度服务设计、集成服务设计、数据管理设计 |  |  |
| V1.2 |  |  |  |  |  |

目录

[编写说明 1](#_Toc40969017)

[目录 i](#_Toc40969018)

[1. 引言 1](#_Toc40969019)

[1.1. 编写目的 1](#_Toc40969020)

[1.2. 背景 1](#_Toc40969021)

[1.3. 定义 2](#_Toc40969022)

[1.4. 参考资料 3](#_Toc40969023)

[2. 总体架构 3](#_Toc40969024)

[2.1. 应用架构 3](#_Toc40969025)

[2.2. 数据架构 4](#_Toc40969026)

[3. 设计规范引用 5](#_Toc40969027)

[3.1. 建模规范 5](#_Toc40969028)

[3.1.1. 建模流程 5](#_Toc40969029)

[3.1.2. 模型命名规范 6](#_Toc40969030)

[3.2. 开发规范 10](#_Toc40969031)

[3.2.1. 数据处理批量脚本开发规范 10](#_Toc40969032)

[3.2.2. 数据批量采集开发规范 35](#_Toc40969033)

[3.2.3. 数据处理实时消息任务开发规范 38](#_Toc40969034)

[3.2.4. 数据服务开发规范 46](#_Toc40969035)

[3.2.5. 附录 48](#_Toc40969036)

[4. 模型设计 51](#_Toc40969037)

[4.1. 通用设计 51](#_Toc40969038)

[4.1.1. 入湖原则 51](#_Toc40969039)

[4.1.2. 字段定义转换规则 52](#_Toc40969040)

[4.1.3. 默认审计字段 52](#_Toc40969041)

[4.2. DWI模型设计 54](#_Toc40969042)

[4.2.1. 环保系统DWI模型设计 54](#_Toc40969043)

[4.2.2. 能耗监测系统DWI模型设计 56](#_Toc40969044)

[4.2.3. 智慧消防系统DWI模型设计 58](#_Toc40969045)

[4.2.4. 国有资产管理系统DWI模型设计 58](#_Toc40969046)

[4.2.5. 智慧安防系统DWI模型设计 59](#_Toc40969047)

[4.2.6. 园区服务平台DWI模型设计 61](#_Toc40969048)

[4.2.7. 园区管理平台（飞企）-DWI模型设计 63](#_Toc40969049)

[4.2.8. 园区管理平台（普天）-DWI模型设计 64](#_Toc40969050)

[4.3. DWR模型设计 67](#_Toc40969051)

[4.3.1. 主题模型设计思路 67](#_Toc40969052)

[4.3.2. 安环监管主题模型设计 68](#_Toc40969053)

[4.3.3. 企业主题模型设计 69](#_Toc40969054)

[4.3.4. 企业服务主题模型设计 69](#_Toc40969055)

[4.3.5. 设备主题模型设计 69](#_Toc40969056)

[4.3.6. 能耗主题模型设计 70](#_Toc40969057)

[4.3.7. 事件主题模型设计 70](#_Toc40969058)

[4.3.8. 产业主题模型设计 71](#_Toc40969059)

[4.3.9. 项目主题模型设计 71](#_Toc40969060)

[4.3.10. 招商主题模型设计 71](#_Toc40969061)

[4.3.11. 办公主题模型设计 72](#_Toc40969062)

[4.3.12. 资源主题模型设计 72](#_Toc40969063)

[4.3.13. 资产主题模型设计 73](#_Toc40969064)

[4.3.14. 应急主题模型设计 73](#_Toc40969065)

[4.3.15. 工单主题模型设计 73](#_Toc40969066)

[4.3.16. 组织主题模型设计 74](#_Toc40969067)

[4.3.17. 人员主题模型设计 74](#_Toc40969068)

[4.3.18. 时空主题模型设计 75](#_Toc40969069)

[5. 数据采集功能设计 75](#_Toc40969070)

[5.1. 环保采集 75](#_Toc40969071)

[5.1.1. 功能 75](#_Toc40969072)

[5.1.2. 数据采集方式及策略 75](#_Toc40969073)

[5.1.3. 数据Mapping 77](#_Toc40969074)

[5.1.4. FDI设计 77](#_Toc40969075)

[5.2. 能耗监测采集 77](#_Toc40969076)

[5.2.1. 功能 77](#_Toc40969077)

[5.2.2. 数据采集方式及策略 77](#_Toc40969078)

[5.2.3. 数据Mapping 78](#_Toc40969079)

[5.2.4. FDI设计 78](#_Toc40969080)

[5.3. 消防采集 78](#_Toc40969081)

[5.3.1. 功能 78](#_Toc40969082)

[5.3.2. 数据采集方式及策略 78](#_Toc40969083)

[5.3.3. 数据Mapping 78](#_Toc40969084)

[5.3.4. FDI设计 78](#_Toc40969085)

[5.4. 国有资产采集 78](#_Toc40969086)

[5.4.1. 功能 78](#_Toc40969087)

[5.4.2. 数据采集方式及策略 79](#_Toc40969088)

[5.4.3. 数据Mapping 79](#_Toc40969089)

[5.4.4. FDI设计 79](#_Toc40969090)

[5.5. 智慧安防采集 79](#_Toc40969091)

[5.5.1. 功能 79](#_Toc40969092)

[5.5.2. 数据采集方式及策略 80](#_Toc40969093)

[5.5.3. 数据Mapping 80](#_Toc40969094)

[5.5.4. FDI设计 81](#_Toc40969095)

[5.6. 园区服务平台采集 81](#_Toc40969096)

[5.6.1. 功能 81](#_Toc40969097)

[5.6.2. 数据采集方式及策略 81](#_Toc40969098)

[5.6.3. 数据Mapping 82](#_Toc40969099)

[5.6.4. FDI设计 83](#_Toc40969100)

[5.7. 园区管理平台（飞企）采集 83](#_Toc40969101)

[5.7.1. 功能 83](#_Toc40969102)

[5.7.2. 数据采集方式及策略 83](#_Toc40969103)

[5.7.3. 数据Mapping 84](#_Toc40969104)

[5.7.4. FDI设计 84](#_Toc40969105)

[5.8. 园区管理平台（普天）采集 84](#_Toc40969106)

[5.8.1. 功能 84](#_Toc40969107)

[5.8.2. 数据采集方式及策略 84](#_Toc40969108)

[5.8.3. 数据Mapping 87](#_Toc40969109)

[5.8.4. FDI设计 87](#_Toc40969110)

[5.9. 测试计划 87](#_Toc40969111)

[6. 主题库数据处理设计 88](#_Toc40969112)

[6.1. 环保主题 88](#_Toc40969113)

[6.1.1. 功能 88](#_Toc40969114)

[6.1.2. 数据Mapping 88](#_Toc40969115)

[6.1.3. DLF设计 88](#_Toc40969116)

[6.2. 能耗主题 88](#_Toc40969117)

[6.2.1. 功能 88](#_Toc40969118)

[6.2.2. 数据Mapping 88](#_Toc40969119)

[6.2.3. DLF设计 89](#_Toc40969120)

[6.3. 企业主题 89](#_Toc40969121)

[6.3.1. 功能 89](#_Toc40969122)

[6.3.2. 数据Mapping 89](#_Toc40969123)

[6.3.3. DLF设计 89](#_Toc40969124)

[6.4. 办公主题 89](#_Toc40969125)

[6.4.1. 功能 89](#_Toc40969126)

[6.4.2. 数据Mapping 89](#_Toc40969127)

[6.4.3. DLF设计 89](#_Toc40969128)

[6.5. 产业主题 90](#_Toc40969129)

[6.5.1. 功能 90](#_Toc40969130)

[6.5.2. 数据Mapping 90](#_Toc40969131)

[6.5.3. DLF设计 90](#_Toc40969132)

[6.6. 工单主题 90](#_Toc40969133)

[6.6.1. 功能 90](#_Toc40969134)

[6.6.2. 数据Mapping 90](#_Toc40969135)

[6.6.3. DLF设计 90](#_Toc40969136)

[6.7. 企业服务主题 90](#_Toc40969137)

[6.7.1. 功能 90](#_Toc40969138)

[6.7.2. 数据Mapping 91](#_Toc40969139)

[6.7.3. DLF设计 91](#_Toc40969140)

[6.8. 人员主题 91](#_Toc40969141)

[6.8.1. 功能 91](#_Toc40969142)

[6.8.2. 数据Mapping 91](#_Toc40969143)

[6.8.3. DLF设计 91](#_Toc40969144)

[6.9. 设备主题 91](#_Toc40969145)

[6.9.1. 功能 91](#_Toc40969146)

[6.9.2. 数据Mapping 91](#_Toc40969147)

[6.9.3. DLF设计 91](#_Toc40969148)

[6.10. 时空主题 92](#_Toc40969149)

[6.10.1. 功能 92](#_Toc40969150)

[6.10.2. 数据Mapping 92](#_Toc40969151)

[6.10.3. DLF设计 92](#_Toc40969152)

[6.11. 事件主题 92](#_Toc40969153)

[6.11.1. 功能 92](#_Toc40969154)

[6.11.2. 数据Mapping 92](#_Toc40969155)

[6.11.3. DLF设计 92](#_Toc40969156)

[6.12. 项目主题 92](#_Toc40969157)

[6.12.1. 功能 92](#_Toc40969158)

[6.12.2. 数据Mapping 93](#_Toc40969159)

[6.12.3. DLF设计 93](#_Toc40969160)

[6.13. 应急主题 93](#_Toc40969161)

[6.13.1. 功能 93](#_Toc40969162)

[6.13.2. 数据Mapping 93](#_Toc40969163)

[6.13.3. DLF设计 93](#_Toc40969164)

[6.14. 招商主题 93](#_Toc40969165)

[6.14.1. 功能 93](#_Toc40969166)

[6.14.2. 数据Mapping 93](#_Toc40969167)

[6.14.3. DLF设计 93](#_Toc40969168)

[6.15. 资产主题 94](#_Toc40969169)

[6.15.1. 功能 94](#_Toc40969170)

[6.15.2. 数据Mapping 94](#_Toc40969171)

[6.15.3. DLF设计 94](#_Toc40969172)

[6.16. 资源主题 94](#_Toc40969173)

[6.16.1. 功能 94](#_Toc40969174)

[6.16.2. 数据Mapping 94](#_Toc40969175)

[6.16.3. DLF设计 94](#_Toc40969176)

[6.17. 组织主题 94](#_Toc40969177)

[6.17.1. 功能 94](#_Toc40969178)

[6.17.2. 数据Mapping 95](#_Toc40969179)

[6.17.3. DLF设计 95](#_Toc40969180)

[6.18. 测试计划 95](#_Toc40969181)

[7. 数据服务设计 96](#_Toc40969182)

[8. 调度服务设计 97](#_Toc40969183)

[8.1. 现有作业分析 97](#_Toc40969184)

[8.2. 作业处理时间设计 97](#_Toc40969185)

[8.3. 调度依赖关系描述 98](#_Toc40969186)

[8.4. 调度组织规划 98](#_Toc40969187)

[8.5. 调度配置 100](#_Toc40969188)

[8.6. 容错处理 101](#_Toc40969189)

[9. 集成服务设计 101](#_Toc40969190)

[9.1. 应用ID设计 101](#_Toc40969191)

[9.2. 业务开发APIC设计 101](#_Toc40969192)

[9.3. MQS设计 102](#_Toc40969193)

[9.4. 测试计划 102](#_Toc40969194)

[10. 数据管理设计 103](#_Toc40969195)

[10.1. 数据存储 103](#_Toc40969196)

[10.2. 数据清理 103](#_Toc40969197)

[10.3. 数据备份与恢复 104](#_Toc40969198)

[10.4. 数据安全性 104](#_Toc40969199)

[11. 待解决问题 105](#_Toc40969200)

[12. 附录：数据资产一览表 105](#_Toc40969201)

# 引言

## 编写目的

本文档为“融通地产物业项目-数据及集成服务项目-详细设计说明书”，旨在对本项目的数据及集成服务详细设计工作进行描述。

本文档预期读者：开发人员、测试人员、技术评审人员、项目技术负责人。

## 背景

建设统一数据服务平台、数据可视化平台和工单调度平台，以本地化部署模式、进行统一数据采集、数据传输、数据服务、数据展现和工单调度。

1、数据交换功能

设计并实现适应性良好、随需应变的数据服务平台，统一数据交换标准，打通系统之间的信息孤岛，实现总线型的系统功能集成和信息共享。

2、数据加工存储功能

将不同类型和来源的数据关联起来，为数据分析应用场景提供数据支撑。

3、数据分析功能

构建数据分析挖掘模型，综合运用报表查询和数据挖掘等技术，实现对企业数据的深度探索与模型建立，通过对模型实践的持续更新优化，实现数据的分析、统计和应用，为运行管理提供数据保障，增强数据服务能力。

4、综合分析应用功能

借助数据分析功能提供的支持，构建面向企业管理层的决策支持功能，为管理展开日常经营管理工作提供全面、直观的数据分析服务，实现“一图观全貌”，满足各级领导的不同决策需要，充分发挥数据核心价值，为信息化支撑企业战略决策提供切实有效的技术支撑。

5、工单调度功能

构建工单调度平台，打破业务系统间工单协同壁垒，实现不同业务系统间的工单协同调度。

## 定义

**DWR**（Data Warehouse Report）：主题库，又叫数据仓库的数据中间层。对上支撑DM层多维分析需求和明细查询需求，对下屏蔽DWI层多源的变化，并将三范式模型转为维度模型。DWR层内又可分为DWD层和DWS层。DWD是数据明细层，DWS是数据汇总层。DWS是基于DWD的轻度汇总，例如按日、月、年进行预统计。

**DWI** （Data Warehouse Integration）：

数据湖，又叫贴源层，是按业务数据模型的统一标准和规范将源系统数据进行整合。对下屏蔽源系统不同的数据库类型和业务模型的变化，对上提供稳定的、完整的源系统数据。

**ROMA：**

ROMA是华为公司发布的多云时代融合集成平台，聚焦应用和数据连接，提供轻量化消息、数据、API等集成能力，简化企业上云，支持云上云下、跨区域集成，打通IT与OT，连接企业与生态伙伴，助力行业数字化转型。

**FDI（Fast Data Integration）：**

是ROMA的数据集成组件，支持文本、消息、API、结构和非结构化数据等多种数据源之间的灵活、快速、无侵入式的数据集成，可以实现跨机房、跨数据中心、跨云的数据集成方案，并能自助实施、运维、监控集成数据。

**MQS（Message Queue Service）：**

是ROMA平台的消息集成组件，支持多种主流消息中间件接入，前后端应用无感知。

**API Connect：**

是ROMA的API管理组件，涵盖API网关、LiveData和API市场三个子产品。聚焦在API轻量化集成，存量系统服务化改造，全球跨云跨数据中心路由等核心功能，实现从API设计、开发、管理到发布的全生命周期管理和端到端集成。

**DLF（Data lake Factory）：**

数据湖工厂（Data lake Factory）提供一站式的大数据开发环境，全托管的大数据调度能力。用户可进行数据建模、数据集成、脚本开发、工作流编排、运维监控等操作，轻松完成整个数据的处理分析流程，极大降低用户使用大数据的门槛。

## 参考资料

《重庆璧山高新区智慧园区项目数据专业服务RFP\_V1.4》

《重庆璧山高新区智慧园区项目数据专业服务和数字平台集成 BOQ v1.3》

《重庆璧山项目需求V1.0》

《重庆璧山高新区智慧园区-数据及集成服务项目-架构设计.docx》

《重庆璧山高新区智慧园区-数据及服务-需求分析说明书-V1.0.X.docx》

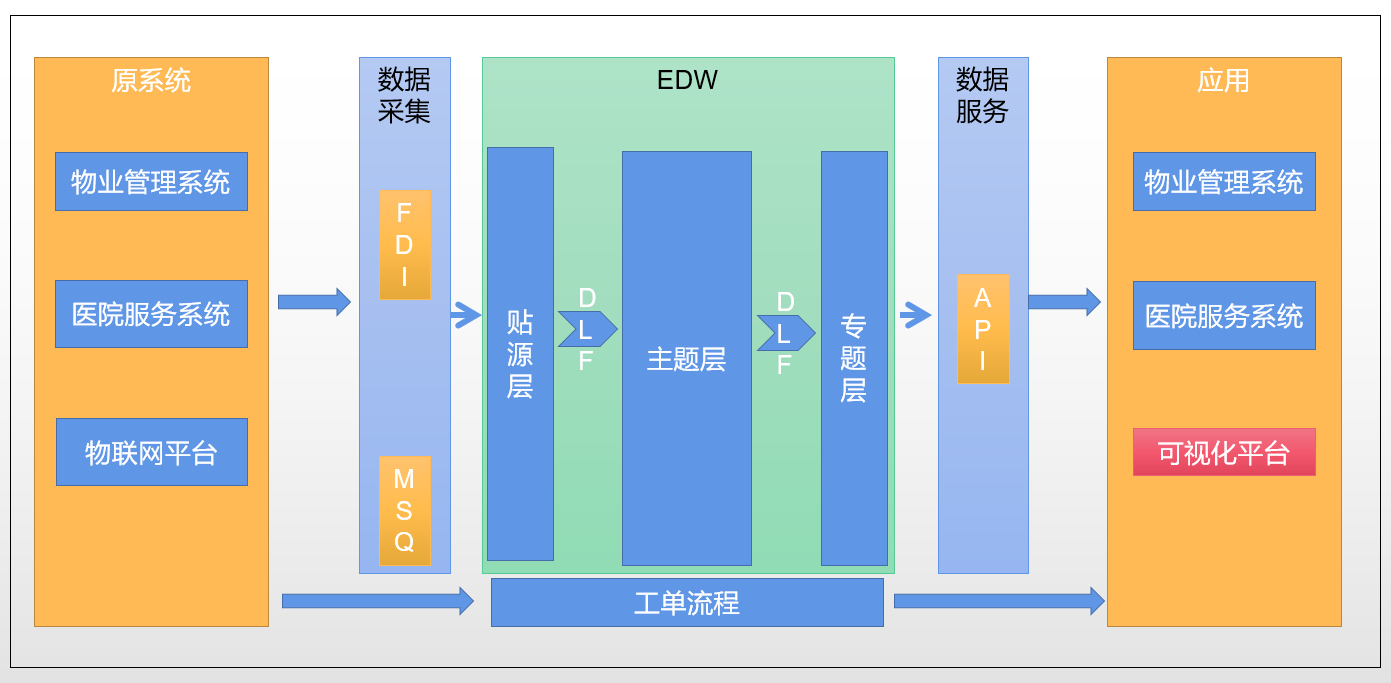
# 总体架构

## 应用架构

应用架构如下图所示：

大数据平台负责抽取物业服务系统，管护系统，医院专项服务系统，资产运营系统，物联网设备的基础数据，并进行清洗、标准化，按主体域进行模型构建，形成园区统一的大数据视图。工单调度平台和IOC平台则从大数据平台获取主题数据，构建为自身应用服务的专题库，从而实现对数据的消费。

## 数据架构

大数据平台整体的数据架构如上图所示。其中：

源系统：主要指提供数据源的业务系统。

数据采集：通过FDI或MQS方式，实现源系统连接，数据访问与抽取工作。

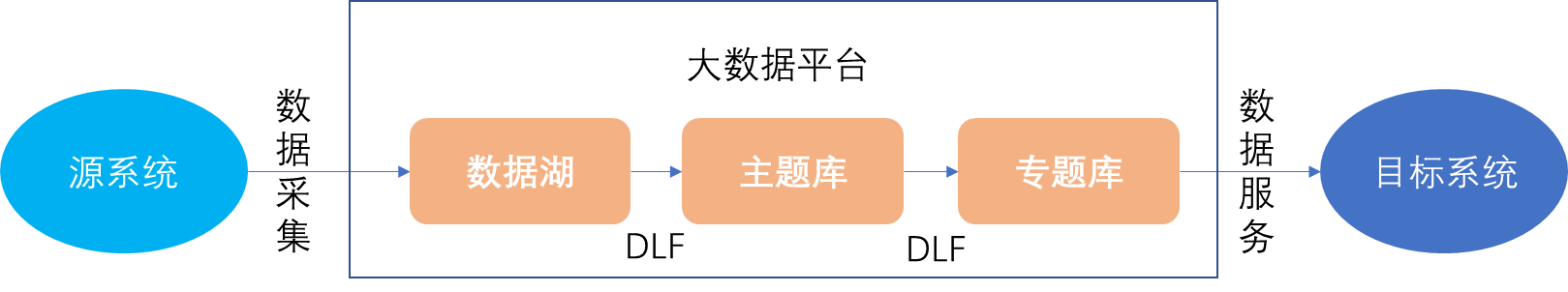
数据湖，又称贴源层：负责存储源系统需要入湖的数据，原则上数据湖不对源系统的数据进行清洗、标准化和加工处理，保持数据原貌，以便于数据追溯，避免数据失真。

主题库：基于业务需求，按主题域模型进行数据的业务分类，汇聚整合，逻辑加工等数据处理，实现主题层数据的构建。主题域中需定义明确的主题名称、数据范围。原则上，主题层不直接对应用系统提供数据服务。

专题库：业务系统根据需要构建基于业务实现的专题库，大数据平台提供数据存储、数据ETL工具。

数据服务：为应用系统提供数据访问服务，包括MQS、APIC或其他方式。

数据加工流程：如下图所示。

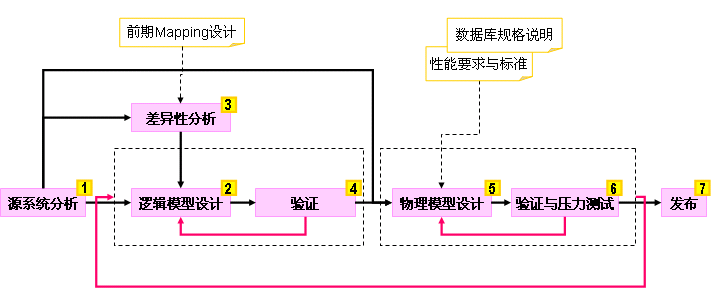


# 设计规范引用

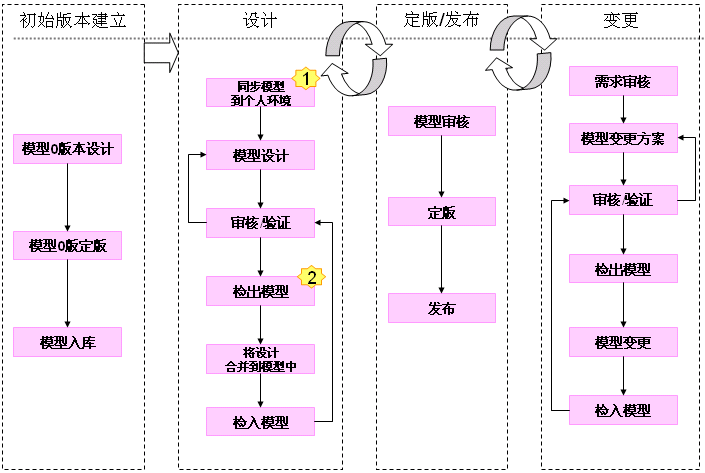
## 建模规范

### 建模流程

#### 建模总体过程

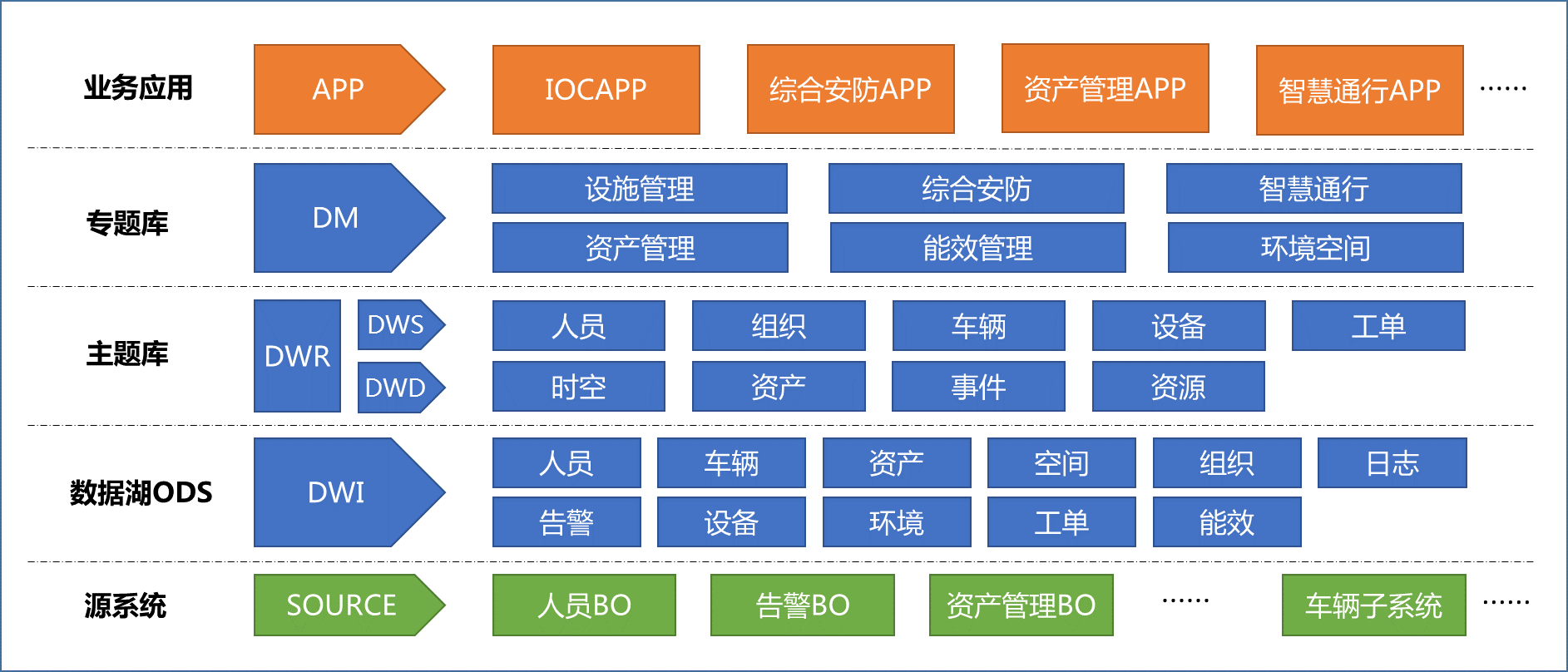


#### 协同工作总体过程



### 模型命名规范

分析模型分为三层，专题库（DM），主题库（DWR）和数据湖（DWI）。



DM（Data Market）：专题库，又叫数据集市。

DWR（Data Warehouse Report）：主题库，又叫数据中间层。

DWD（Data Warehouse Detail）：数据明细层。

DWS（Data Warehouse Summary）：数据汇总层。

DIM（Dimension）：维度模型。

DWI（Data Warehouse Integration）：数据湖，或叫贴源层。

SOURCE：数据源系统。

规范：

* 名称的长度不超过32个字符。
* 名称采用英文单词、英文单词缩写和数字，单词之间用“\_”分隔。不得采用“\_”作为名称的起始字母和终止字母。
* 命名能够望文知意，尽量避免使用自定义缩写（可以使用通用的术语缩写进行命名）。例如，在命名中可以使用具有实际业务含义的英文词汇。
* 名称不得与数据库管理系统保留字冲突。
* 不要在对象名的字符之间留空格。
* schema命名定义为数据分层+业务主题名，例如dwi\_person。
* schema名全部采用小写。

#### 贴源层命名规范

##### 源系统名称定义

融通物业项目涉及源系统名称及简称命名如下表所示：

|  |  |  |
| --- | --- | --- |
| 源系统名称 | 英文名称 | 简称 |
| 物业服务系统 | Property service system | pss |
| 医院专项服务系统 | Hospital special service system | hss |
| 安全管护系统 | Safety management system | sms |
| 物联网设备 | Internet of things | iot |
| 能耗管理系统 | Energy consumption management | ecm |
| 资产运营系统 | Asset operation system | aos |

##### 表名规范

* DWI层表名规范:dwi\_[domain].dwi\_[domain]\_表义名\_[Suffix]

domain = {ems、inf、ecm、nas、sin、vse、psp、dep、pmp}。

Suffix：来源是业务平台时，Suffix不需要；来源为其他子系统时，Suffix为子系统名称。

##### 字段命名规范

* 全部小写。
* 采用有意义的列名，为实际含义的英文单词，多个单词间“\_”作为分割符。
* 不要使用“ID”等与系统保留关键字冲突的单词作为列名。
* 常用字段命名见[常用字段](https://support.huawei.com/bescloud/Smart%20Campus/20.0/topic/TOPIC_0200882375)。
* 基本沿用原表字段名称

##### 索引命名

* 索引的命名为：“表名（或缩写）+\_+列名+\_idx”。
* 其中多单词组成的属性列列名取前几个单词首字符，加末单词组成。

##### 存储过程命名

* 存储过程命名由“sp+\_+存储过程标识(缩写)”组成。
* 存储过程标识要以实际含义的英文单词构成，并用下划线分割各个组成部分。

##### 函数命名

* 函数命名由“func+\_+函数标识”组成。

#### 主题命名规范

##### 主题名称定义

融通物业项目主题名称及简称命名如下表所示：

|  |  |
| --- | --- |
| **业务主题名称** | **缩写** |
| 客户服务主题 | customer |
| 巡检主题 | osi |
| 隐患主题 | danger |
| 资产运营主题 | ams |
| 医疗服务主题 | hss |
| 车辆主题 | parking |
| 能耗主题 | energy |
| 设备主题 | facility |
| 工单主题 | order |
| 人员主题 | person |
| 组织主题 | org |

##### 表名规范

* DWR层表名规范:dwr\_[domain].[Prefix]\_[domain]\_表义名\_[Suffix]

domain = {enterprise、industry、ser、office、item、investment、res、asset、energy、ses、facility、event、order、emergency、person、org、space}。

Suffix：来源是业务平台时，Suffix不需要；来源为其他子系统时，Suffix为子系统名称。

|  |  |  |
| --- | --- | --- |
| **表类型** | **Prefix** | **Suffix** |
| 周期快照事实表（天） | dws | d\_f |
| 周期快照事实表（周） | dws | w\_f |
| 周期快照事实表（月） | dws | m\_f |
| 累积快照事实表 | dws | h\_f |
| 明细事实表 | dwd | f |
| 主题维度表 | dim | d |

##### 字段命名规范

* 全部小写。
* 采用有意义的列名，为实际含义的英文单词，多个单词间“\_”作为分割符。
* 不要使用“ID”等与系统保留关键字冲突的单词作为列名。
* 常用字段命名见[常用字段](https://support.huawei.com/bescloud/Smart%20Campus/20.0/topic/TOPIC_0200882375)。
* 基本沿用原表字段名称

##### 索引命名

* 索引的命名为：“表名（或缩写）+\_+列名+\_idx”。
* 其中多单词组成的属性列列名取前几个单词首字符，加末单词组成。

##### 存储过程命名

* 存储过程命名由“sp+\_+存储过程标识(缩写)”组成。
* 存储过程标识要以实际含义的英文单词构成，并用下划线分割各个组成部分。

##### 函数命名

函数命名由“func+\_+函数标识”组成。

## 开发规范

### 数据处理批量脚本开发规范

#### DLF数据连接命名规范

目前只使用两个数据连接：

* smart\_campus\_dbadmin：使用dbadmin用户，作为开发时必要的管理操作，不作为交付件。
* smart\_campus\_data\_warehouse：使用data\_service用户，所有数据仓库开发工作都使用这个数据连接。
* 名称中仅能存在小写英文字母，并以下划线分割。
* 以“smart\_campus\_”开头。
* 名称有一定的自注释能力。

#### DLF脚本开发规范

##### 命名规范

###### 目录命名规范

* 名称中仅能存在小写英文字母，并以下划线分割。
* 名称有一定的自注释能力。

###### 目录结构规范

* 根目录：/smart\_campus
* 二层目录：/smart\_campus/数据仓库分层，例如：
  + /smart\_campus/dm
  + /smart\_campus/dwr
* 三层目录：所有脚本都放在三层目录下。
  + /smart\_campus/dm/应用名称，例如：

1. /smart\_campus/dm/security
2. /smart\_campus/dm/facility
3. /smart\_campus/dm/env
4. /smart\_campus/dm/energy
5. /smart\_campus/dm/asset
6. /smart\_campus/dm/access
7. ..
   * /smart\_campus/dwr/数据仓库17主题库，例如：
8. /smart\_campus/dwr/person
9. /smart\_campus/dwr/org
10. /smart\_campus/dwr/facility
11. /smart\_campus/dwr/vehicle
12. /smart\_campus/dwr/space
13. /smart\_campus/dwr/asset
14. /smart\_campus/dwr/event
15. /smart\_campus/dwr/res
16. /smart\_campus/dwr/order
17. ..

* 管理目录：/smart\_campus\_dbadmin，目录（不创建子目录）存放管理操作的脚本，该目录下脚本不作为交付件。

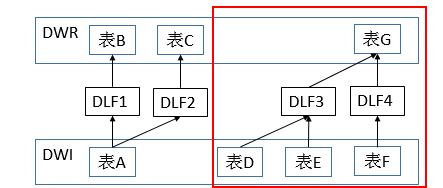
###### 脚本命名规范

* 只包含小写字母，并以下划线分割。
* 以“script\_”开头，中间为目标表名（小写），以来源结尾，如：script\_{to-table-name}\_{source} 。超长的脚本名可以适当使用缩写。
* source具有一定的自注释能力。
* 包含的表名为加载数据的目标表名。
* 脚本名称长度不能超过64位，名称过长的可以适当使用简写。
  + to-table-name解读：

目标表的表名，如资产位置表：script\_dwr\_asset\_position\_f\_{source}。

* + source解读：

在实际情况中，目标表的数据来源可能来自不同的系统，比如人员位置信息表中的数据可以来自进出门的闸机，也可以来自监控系统，那么人员位置信息表就需要从闸机表和监控表中获取，对应下图的红框中的场景：



请务必理解，该场景不同于目标表数据字段来源于不同表的字段，该场景为两张源表都可以为目标表提供一行完整的记录。

所以，我们在往人员位置信息表中加载数据时，需要两个DLF任务，一个从闸机系统读取数据，另一个从监控系统读取，我们使用source来区分，如script\_dwr\_person\_position\_f\_gate、script\_dwr\_person\_position\_f\_monitor。

###### 脚本编写规范

* DLF功能最小、单一。

以数据来源能加载到目标表形成一条完整记录为目标，但是需要区分数据来源的类型，即上面提到的“source”。

* 增量加载一律使用调度时间，不要使用当前时间，即${job\_plan\_time}。
* 数据来源规范。
  + - DWI->DIM(\_D)
    - DWI、DIM、DWD->DWD(\_F)
    - DIM、DWD、DWS->DWS(\_D\_F、\_W\_F、\_M\_F、\_H\_F)
    - DIM、DWD、DWS、DM->DM(\_F)

1. 当数据来源于同层库时，只能来自同一应用的，比如dm层的数据可以来自dm层，但是安防的只能来自安防的，不能来自其他应用的，dim、dwd、dws也是这样。
2. DWS->DWS的场景：日统计数据->周、月、年统计数据等汇总事实统计。
3. DM->DM的场景：日统计数据->周、月、年统计数据等汇总事实统计。
4. 实时数据的场景待定，实时数据场景（只需要实时数据，不需要历史数据）解决方案暂时定为：dm层有一张实时表供update，DLF只执行一次，将历史数据计算后导入实时表作为初始数据，实时表的update不在DLF中，而是flink去做，所以实时场景的DLF只执行一次作为初始数据，不作为更新数据的作用。

##### SQL编码排版规范

* 字体规范：

SQL语句中关键字用大写，列名和表名采用小写。

* 别名规范：

涉及多张表关联时每张表必须有别名，物理表别名采用“abc…”依次往后命名，中间表、临时表别名采用“t1、t2、t3…”依次往后命名

* 格式编排：

CREATE、BEGIN、END、EXCEPTION独立成行，顶格书写不缩进，但是作为程序块嵌套在程序中时仍然需要缩进4格。

【错误示例】：

CREATE OR REPLACE FUNCTION f\_functionname (…) RETURN XXX IS

…

BEGIN

…

EXCEPTION WHEN OTHERS THEN

…

END f\_functionname;

【正确示例】：

CREATE OR REPLACE FUNCTION f\_functionname

(…

) RETURN XXX IS

…

DECLARE

BEGIN

…

EXCEPTION

WHEN OTHERS THEN

…

END f\_functionname;

* 避免将SQL语句写到同一行，再短的语句也要在关键字和谓词处换行。

【错误示例】：

SELECT column1 INTO str\_l\_variable1 FROM t\_tablename

WHERE column2 = i\_l\_variable2

AND column3 = i\_l\_variable3;

应编写如下：

SELECT column1

INTO str\_l\_variable1

FROM t\_tablename

WHERE column2 = i\_l\_variable2

AND column3 = i\_l\_variable3;

* SQL语句中关键字右对齐。

【错误示例】：

SELECT column1, column2, column3

INTO i\_l\_variable1, i\_l\_variable2, i\_l\_variable3

FROM t\_tablename

WHERE column4 = value

ORDER BY column2;

【正确示例】：

SELECT column1, column2, column3

INTO i\_l\_variable1, i\_l\_variable2, i\_l\_variable3

FROM t\_tablename

WHERE column4 = value

ORDER BY column2;

IMG_256

ORDER BY、GROUP BY等具有两个关键的子句，只需要第一个关键字（order、group）和其他关键字右对齐即可。

* 缩进格数统一为4个空格，禁止使用Tab键。

缩进4格的前提是“满足SQL语句关键字的右对齐”，如例子中，SELECT相对BEGIN缩进4格，但into不能相对SELECT再缩进4格，into要与SELECT右对齐。SQL语句关键字右对齐规则的优先级较高。

* SQL语句长度应尽量简明精炼，空格只在必要的时候才使用，不得随意添加空格，更不要出现连续多个空格的现象。关键字、保留字、连接符or、in、and等前后有一个空格，左括号之前有一个空格，右括号之后有一个空格，=、>=、<=、，(逗号)前后有有一个空格。

【错误示例】：

SELECT column1

INTO str\_l\_variable1

FROM t\_tablename

WHERE (column2 =i\_l\_variable2

AND column3>=i\_l\_variable3);

【正确示例】：

SELECT column1

INTO str\_l\_variable1

FROM t\_tablename

WHERE ( column2 = i\_l\_variable2

AND column3 >= i\_l\_variable3 );

* IF后的条件要用括号括起来

【错误示例】：

IF i\_l\_variable1 = 2 OR i\_l\_variable2 = 4 OR i\_l\_variable3 = 6 OR i\_l\_variable4 = 8

OR i\_l\_variable5 = 10 OR i\_l\_variable6 = 12 OR i\_l\_variable7 = 14

THEN

…;

END IF;

【正确示例】：

IF ( i\_l\_variable1 = 2 OR i\_l\_variable2 = 4 OR i\_l\_variable3 = 6

OR i\_l\_variable4 = 8 OR i\_l\_variable5 = 10 OR i\_l\_variable6 = 12

OR i\_l\_variable7 = 14

) THEN

…;

END IF;

IMG_257

为让if语句更加整齐，程序的可读性强，不论含有几个条件都需要用括号括起来。另外，建议括号内的条件个数最多不要超过三个条件。

* 对于INSERT into…VALUES，INSERT into…SELECT语句，要求一行只写一个字段。
* INSERT into…VALUES时，表名写在into关键字的后面，换行写字段名字，“（”与into关键字右对齐，“）”与“（”对齐，字段名与表名左对齐，into关键字和VALUES、SELECT关键字要右对齐。
* INSERT into…SELECT时，SELECT语句中的关键字仍然要保证右对齐，VALUES后字段必须与INSERT后字段逐行一一对应,逗号放在字段前。

【错误示例】：

INSERT INTO t\_tablename

( column1 --注释

,column2 --注释

)

VALUES

( fieldvalue1,

fieldvalue2

);

【正确示例】：

INSERT INTO t\_tablename

( column1 --注释

,column2 --注释

)

VALUES

( column1\_value --注释

,column2\_value --注释

);

【错误示例】：

INSERT INTO t\_tablename1

( column1, column2, column3, column4, column5,

column6, column7 …

)

SELECT column1\_value, column2\_value, column3\_value,

column4\_value, column5\_value, column6\_value,

column7\_value …

FROM t\_tablename2;

【正确示例】：

INSERT INTO t\_tablename1

( column1 --注释

,column2 --注释

,column3 --注释

…

)

SELECT column1\_value --注释

,column2\_value --注释

,column3\_value --注释

…

FROM t\_tablename2;

IMG_258

要求一行一个字段，保证每个字段有字段注释。

* UPDATE语句，所有关键字要右对齐，仍然要求一行一个字段。

【错误示例】：

UPDATE t\_tablename set column1 = new\_value1,

column2 = new\_value2,

column3 = new\_value3

WHERE column4 > fix\_value1 AND column < fix\_value2;

【正确示例】：

UPDATE t\_tablename

set column1 = new\_value1, --注释

column2 = new\_value2, --注释

column3 = new\_value3 --注释

WHERE column4 > fix\_value1

AND column5 < fix\_value2;

* 匿名块、存储过程、函数的变量定义部分，类型关键字要左对齐；变量后的注释也要左对齐。

【错误示例】：

CREATE OR REPLACE FUNCTION f\_functionname

(…

) RETURN XXX IS

str\_l\_name VARCHAR(20); --注释

i\_l\_var NUMBER; --注释

dt\_l\_time DATE; --注释

BEGIN

…

END f\_functionname;

【正确示例】：

CREATE OR REPLACE FUNCTION f\_functionname

(…

) RETURN XXX IS

str\_l\_name VARCHAR(20); -- 注释

i\_l\_var NUMBER; -- 注释

dt\_l\_time DATE; -- 注释

BEGIN

…

END f\_functionname;

* 匿名块、存储过程、函数中的字段要左对齐；字段后的注释也要左对齐；逗号写在字段前并且逗号与字段间保留一个空格。

【正确示例】：

INSERT INTO t\_tablename1

( column1 -- 注释

, column2 -- 注释

, column3 -- 注释

…

)

SELECT column1\_value -- 注释

, column2\_value -- 注释

, column3\_value -- 注释

…

FROM t\_tablename2;

* 条件选择语句(CASE)编写规范。

【错误示例】：

SELECT CASE WHEN column1 = column2 THEN … WHEN column3 = column4 THEN …

ELSE … END AS column5

,column6

FROM t\_tablename ;

【正确示例】：

SELECT CASE WHEN column1 = column2 THEN …

WHEN column3 = column4 THEN …

ELSE

END AS column5

,column6

FROM t\_tablename ;

* 存储过程、函数中的IN 、OUT参数应按其类别分开书写，不允许交叉。

【错误示例】：

CREATE OR REPLACE PROCEDUREf\_functionname

(

str\_putinparam1 IN t\_tablename1.column1%type, --注释

str\_outputparam1 OUT t\_tablename1.column3%type, --注释

str\_putinparam2 IN t\_tablename1.column1%type, --注释

str\_outputparam2 OUT t\_tablename2.column2%type, --注释

str\_putinparam3 IN t\_tablename2.column3%type, --注释

str\_putinparam4 IN t\_tablename4.column1%type, --注释

str\_outputparam3 OUT t\_tablename2.column3%type, --注释

) …

【正确示例】：

CREATE OR REPLACE PROCEDUREf\_functionname

(

str\_putinparam1 IN t\_tablename1.column1%type, --注释

str\_putinparam2 IN t\_tablename1.column1%type, --注释

str\_putinparam3 IN t\_tablename2.column3%type, --注释

str\_putinparam4 IN t\_tablename4.column1%type, --注释

str\_outputparam1 OUT t\_tablename1.column3%type, --注释

str\_outputparam2 OUT t\_tablename2.column2%type, --注释

str\_outputparam3 OUT t\_tablename2.column3%type, --注释

) …

* 注释编排：
* 匿名块、存储过程、函数要增加上注释，描述功能。
* 自定义参数需添加注释，注释与内容之间保持两个空格。
* 在逻辑相对独立的程序块之间必须用空行隔开，并加以注释。注释在内容前一行，格式对齐，单独一行。
* 脚本中关键性操作需添加注释，注释与内容之间保持两个空格。
* 脚本中所有使用到的表需添加注释，注释与内容之间保持两个空格。
* 脚本中涉及的字段需添加注释，注释与内容之间保持两个空格。

【示例】：

CREATE OR REPLACE FUNCTION f\_functionname

(…

) RETURN XXX IS

str\_l\_name VARCHAR(20); --自定义参数注释

i\_l\_var NUMBER; --自定义参数注释

dt\_l\_time DATE; --自定义参数注释

/\*

模块功能描述：

目的：

目标表 ： 表名：schema.表名; 表中文名：表中文名

源表 ： 表名：schema.表名; 表中文名：表中文名

作者 ： admin

创建日期 ： 2019-10-29

审核人 ： admin

审核时间 ： 2019-10-25

历史修改记录：

修改时间 修改作者 备注

\*/

BEGIN

…

/\*

WM\_CONCAT 自定义函数作用说明

\*/

-- 独立逻辑功能描述

Create temp table tmp\_t\_tablename1

As

SELECT a.cloumn1 --注释字段描述

,a.cloumn2 --注释字段描述

,a.cloumn3 --注释字段描述

,b.cloumn4 --注释字段描述

,b.cloumn5 --注释字段描述

to\_char(b.cloumn6,'YYYYMMDDHH24') -- 关键性操作描述

FROM t\_tablename1 a --注释表名描述

INNER JOIN t\_tablename2 b --注释表名描述

ON a.cloumn1 = b.cloumn1

;

-- 独立逻辑功能描述

INSERT INTO target\_tablename

( column1 --注释

, column2 --注释

, column3 --注释

…

)

SELECT column1\_value --注释

, column2\_value --注释

, column3\_value --注释

…

FROM tmp\_t\_tablename1 a --注释表名描述

LEFT JOIN t\_tablename3 a --注释表名描述

ON a.cloumn1 = b.cloumn1

;

END;

##### SQL编码语法规范

###### 表设计规范

* 字段设计规范
* 字段高效性原则：选择最高效的类型存储数据，这可以提高数据库的有效容量及查询执行性能。如能整型就不用浮点型，能用整型就不用字符型。
* 字段最小原则：使用满足需求的最新数值类型。如果int或smallint够用，那么选择bigint就会浪费空间。
* 当多个表存在逻辑关系时，表示同一含义的字段应该使用相同的数据类型。
* 字段最短原则：建议使用变长字符串数据类型，并指定最大长度。
* 请务必确保指定的最大长度大于需要存储的最大字符数，避免超出最大长度时出现字符截断现象。除非明确知道数据类型为固定长度字符串，否则，不建议使用CHAR(n)、BPCHAR(n)、NCHAR(n)、CHARACTER(n)。
* 行列存设计
* 建表时，对于有很多字段的大型表，如果单个查询只需访问较少字段，建议使用列存表。
* 如果对压缩比要求比较高，建议使用列存表。
* 数据仓库的分析型业务中，如果SELECT访问少量字段或者在少量字段上执行聚合计算，则建议使用列存表。
* 交易型业务中更新和插入频繁，建议使用行存表。
* 指定某列值的精确查询场景，建议使用行存储。
* 如果需要同时访问宽表的很多字段时，建议使用行存储。
* 如果大多数字段会出现在 SELECT 列表中或者 WHERE 子句中，建议使用行存储。
* 建表时，建议小表和维度表选择定义为复制表。
* 建表时，建议数据量较大的表定义为Hash表。
* **分布键设计**
* 建Hash表时，指定分布建，禁止使用默认分布建。
* 数据均匀分布原则：一般情况下，选择一个能够将数据均匀分布到所有DN实例上的字段做分布键，如果单个字段不能实现数据均匀分布，则考虑使用两个字段做分布键。作为分布键的字段最好不要超过两个。使用更多的字段做分布列，会耗费更多的时间计算哈希值。
* 本地操作原则：在保证均匀分布的情况下，选择join列、group by列做分布列，可以减少计算过程中数据重分布，提升查询性能。
* 不要使用日期或者时间做分布键。
* **分区设计**
* 为大表设置分区，小表不需要设置分区。
* 查询条件可以实现分区剪枝。
* 如果对数据有生命周期管理（如只保留一年数据，往年数据删除或者归档），需要使用分区。
* 对于列存储的表，慎用过多的分区。
* 将分区上边界的分区值定义为MAXVALUE，以防止可能出现的数据溢出。
* **约束设计**
* 给明确不存在NULL值的字段加上NOT NULL约束，优化器会在特定场景下对其进行自动优化。
* 如果能够从业务层面补全字段值，那么，就不建议使用DEFAULT约束，避免数据加载时产生不符合预期的结果。
* 行存表支持唯一约束，而列存表不支持。
* 行存表支持主键约束，而列存表不支持。
* 行存表支持检查约束，而列存表不支持。
* 从命名上明确标识唯一约束，例如，命名为“UNI+构成字段”。

###### 语句规范

* SELECT**语句**要列出具体字段名，严禁使用SELECT \* 方式。

【错误示例】：

SELECT \*

INTO str\_l\_column1, str\_l\_column2, str\_l\_column3

FROM t\_tablename

WHERE column4 = value;

【正确示例】：

SELECT column1, column2, column3

INTO str\_l\_column1, str\_l\_column2, str\_l\_column3

FROM t\_tablename

WHERE column4 = value;

* INSERT语句要列出具体字段名，严禁使用INSERT INTO tablename VALUES ( … )方式，应使用INSERT INTO tablename ( … ) VALUES ( … )方式。

【错误示例】：

INSERT INTO t\_tablename

VALUES ( value1, value2, value3, value4);

【正确示例】：

INSERT INTO t\_tablename

( column1, --注释

column2, --注释

column3, --注释

column4 --注释

)

VALUES

( value1,

value2,

value3,

value4

);

* 禁止使用不下推函数。

对脚本中逐段SQL进行explain，有“\_REMOTE\_TABLE\_QUERY\_”告警的即为不下推，需更换函数或写法，另外单独一段SQL下推不能保证整体SQL下推，所以还要针对所有SQL进行explain检。

* 不同类型的操作符混合使用时，使用括号括起来进行隔离。

【错误示例】：

IF 'abc' || 'bcd' = 'abcbcd ' THEN

...

END IF;

【正确示例】：

IF ( ( 'abc' || 'bcd' ) = 'abcbcd' ) THEN

...

END IF;

* 固定集合的，使用in（not in），动态集合使用 exists （not exists）。

【正确示例】：

SELECT column1

,column2

,column3

,column4

,column5

FROM t\_tablename;

WHERE t1.column6 NOT IN(‘1’, ‘2’, ‘3’);

SELECT first\_name

, last\_name

FROM emp a

WHERE not exists ( SELECT department\_id

FROM dept b

WHERE b.department\_id = a.department\_id

);

* **如果没有去重的需求，优先使用UNION ALL，UNION具有去重和排序过程，union all 没有，所以速度相对比UNION要快。**
* **减少关联表的次数。**
* **排序取值和去重使用row\_number()，尽量少使用distinct、group by很多字段。**
* **脚本中不要频繁使用order by，如有必要在最后一步进行。**
* **脚本中禁止使用硬编码。**
* **不要使用太多的嵌套查询，可创建临时表。**

##### 实现增量加载

样例：

WITH maxtime AS (

SELECT MAX(insert\_date) FROM plan\_time\_test.plan\_time\_data\_to

)

insert into plan\_time\_test.plan\_time\_data\_to (id,name,data\_date,insert\_date) select id,name,data\_date,LOCALTIMESTAMP(0) from plan\_time\_test.plan\_time\_data\_from

where (SELECT \* FROM maxtime) < insert\_date OR (SELECT \* FROM maxtime) IS NULL;

#### DLF作业开发规范

##### 命名规范

###### 目录命名规范

* 名称中仅能存在小写英文字母，并以下划线分割。
* 名称有一定的自注释能力。

###### 目录结构规范

* 所有脚本都放在三层目录下。

| **层级** | **说明** | **示例** |
| --- | --- | --- |
| 根目录 | /smart\_campus | /smart\_campus |
| 二层目录 | /smart\_campus/数据仓库分层 | * /smart\_campus/dm * /smart\_campus/dwr |
| 三层目录 | /smart\_campus/dm/应用名称 | * /smart\_campus/dm/security * /smart\_campus/dm/facility * /smart\_campus/dm/env * /smart\_campus/dm/energy * /smart\_campus/dm/asset * /smart\_campus/dm/access |
| /smart\_campus/dwr/数据仓库17大主题库 | * /smart\_campus/dwr/person * /smart\_campus/dwr/org * /smart\_campus/dwr/facility * /smart\_campus/dwr/vehicle * /smart\_campus/dwr/space * /smart\_campus/dwr/asset * /smart\_campus/dwr/event * /smart\_campus/dwr/res * /smart\_campus/dwr/order * .. |

###### 作业命名规范

* 只包含小写字母，并以下划线分割。
* 以“job\_”开头，其他部分与对应脚本名称一样。

###### 新建job填写规范



##### 脚本节点配置规范

* 数据连接一律选择“smart\_campus\_data\_warehouse”。
* 数据库一律选择“data\_service”。
* 节点名称与脚本名称一致。
* 失败重试：是。
* 最大重试次数：3。
* 重试间隔时间（秒）：60。
* 其他使用默认即可。

**图1 脚本节点配置图**  


##### 调度配置规范

* 目前先按照每小时的0分周期性调度。
* 依赖属性按照来源配置。
* DWI->DIM(\_D)
* DWI、DIM、DWD->DWD(\_F)
* DIM、DWD、DWS->DWS(\_D\_F、\_W\_F、\_M\_F、\_H\_F)
* DIM、DWD、DWS、DM->DM(\_F)

如：DWS中的表依赖了DIM、DWD中的表，所以当配置DWS中表的调度时，就需要配置两个依赖。

1. 依赖的作业失败后，当前作业处理策略选择“挂起”。

* “等待依赖作业的上一周期结束，才能运行”不勾选。
* “跨周期依赖”选择“自依赖（等待上一调度周期结束才能继续运行）”。

##### 调试

* 调试时调度周期可以设置得短一点。
* 调试之前需要有新数据在依赖的底层的表中加载，否则调试视为无效。
* 开发完成后按照规范调整调度策略，新数据加载的调度可适当提前，需保证有新数据加载。

### 数据批量采集开发规范

#### BingoETL开发规范

当前通过BingoETL做数据从业务平台到数据平台的批量采集。BingoETL的任务需要优先于DLF的任务执行。举例来说BingoETL任务在每小时的05分执行，采集需要花费5分钟，那么后续DLF任务可以在每小时的15分启动，保障BingoETL已经采集完成数据。

需要注意的是，当前版本BingoETL不支持设置定时任务的起始时间，目前是以点击启动任务的时间作为每个周期的起始时间的。所以这个版本开发时，需要手动控制任务启动时间，来保证和DLF之间的执行顺序正确。

##### 目录命名规范

* 目录结构名称中仅能存在小写英文字母，并以下划线分割。
* 名称有一定的自注释能力。

##### 规范

| **层级** | **说明** | **示例** |
| --- | --- | --- |
| 根目录 | /smart\_campus | /smart\_campus |
| 二层目录 | /smart\_campus/数据仓库分层 | /smart\_campus/dwi |
| 三层目录 | /smart\_campus/dwi/数据仓库17大主题库 | * /smart\_campus/dwi/person * /smart\_campus/dwi/org * /smart\_campus/dwi/facility * /smart\_campus/dwi/vehicle * /smart\_campus/dwi/space * /smart\_campus/dwi/asset * /smart\_campus/ dwi/alarm * /smart\_campus/ dwi /res * /smart\_campus/ dwi /order * .. |

##### 转换作业命名规范

* 只包含小写字母，并以下划线分割。
* 以“trans\_”开头，以目标表结尾，如：trans \_{to-table-name}。超长的脚本名可以适当使用缩写。
* to-table-name具有一定的自注释能力。

##### 作业开发规范

* 作业命名规范

1. 只包含小写字母，并以下划线分割。
2. 以“job\_”开头，以目标表结尾，如：job \_{to-table-name}。超长的脚本名可以适当使用缩写。
3. to-table-name具有一定的自注释能力。

* 作业开发规范

作业中转换组件的命名需跟使用的转换作业名字一致。

#### FDI开发规范

##### ROMA应用的命名规范

dataservice.etl.xxx（其中xxx为贴源层的名称，人员：person，车辆：vehicle，告警：alarm，设备：facility，工单：order，资产：asset，资源：res，时空：space，环境：env，组织：org）。

##### FDI的任务名称命名规范

* 实时任务：
  + 只包含小写字母，并以下划线分割。
  + 以“job\_real”开头，以cs作业任务名为依据，如：
    - cs作业为：cs\_res\_datalog\_all
    - FDI实时任务为：job\_real\_res\_datalog\_all
* 批量任务：
  + 只包含小写字母，并以下划线分割。
  + 以“job\_batch”开头，以bingoETL作业任务名为依据，如：
    - bingoETL作业为：job\_dwi\_person\_type
    - FDI实时任务为：job\_batch\_dwi\_person\_type

##### FDI的MQS节点名称命名规范

* 统一写当前任务对应的TOPIC名称。

##### FDI的API节点名称命名规范

* 统一写：read

##### FDI的transform节点名称命名规范

* 只包含小写字母，并以下划线分割。
* 以“transform\_”开头，以目标表结尾，如：作业往目标表dwi\_person\_type写入数据，transform节点命名为：transform\_dwi\_person\_type。

##### FDI的GaussDB 200节点名称命名规范

* 只包含小写字母，并以下划线分割。
* 以“write\_”开头，以目标表结尾，如：作业往目标表dwi\_person\_type写入数据，GaussDB 200节点命名为：write\_dwi\_person\_type。

### 数据处理实时消息任务开发规范

#### MQS开发规范

DataService的交付件只涉及北向MQS的配置，不包含MQS消息的接入。

##### 命名规范

* MQS名称以“T\_XXX\_MSGOUT”为原型输出，大写，以“\_MSGOUT”结尾，XXX为具体场景。

#### CS开发规范

##### 命名规范

* CS作业命名规范：job\_cs\_{模块名}\_作用，描述里要写清楚作业的具体作用，并且脚本里要有关键注释。

##### 存放路径

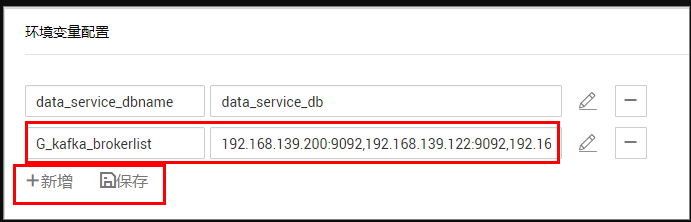
* /smart\_campus/data\_service/{模块名称}

##### 全局变量配置及使用

* 全局变量命名。

命名规范：G\_xxx，例如：G\_kafka\_brokerlist。

所有公共变量、app秘钥等都要配置为全局变量。



* 全局变量引用。

引用方式：#{Env.get("G\_kafka\_brokerlist")}

CREATE SOURCE STREAM sensor\_monitoring\_infos (eventTime TIMESTAMP, deviceId STRING, temperature DOUBLE, humidity DOUBLE, pm25 DOUBLE, noise DOUBLE)

WITH (

type = "kafka",

kafka\_bootstrap\_servers = '#{Env.get("G\_kafka\_brokerlist")}',

kafka\_group\_id = "whgroup1",

kafka\_topic = "T\_ENV\_Source",

encode = "json",

json\_config = "eventTime=eventTime;deviceId=deviceId;temperature=data.InfoService\_Temperature;humidity=data.InfoService\_Humidity;pm25=data.InfoService\_PM25;noise=data.InfoService\_Noise"

) TIMESTAMP BY proctime.proctime;

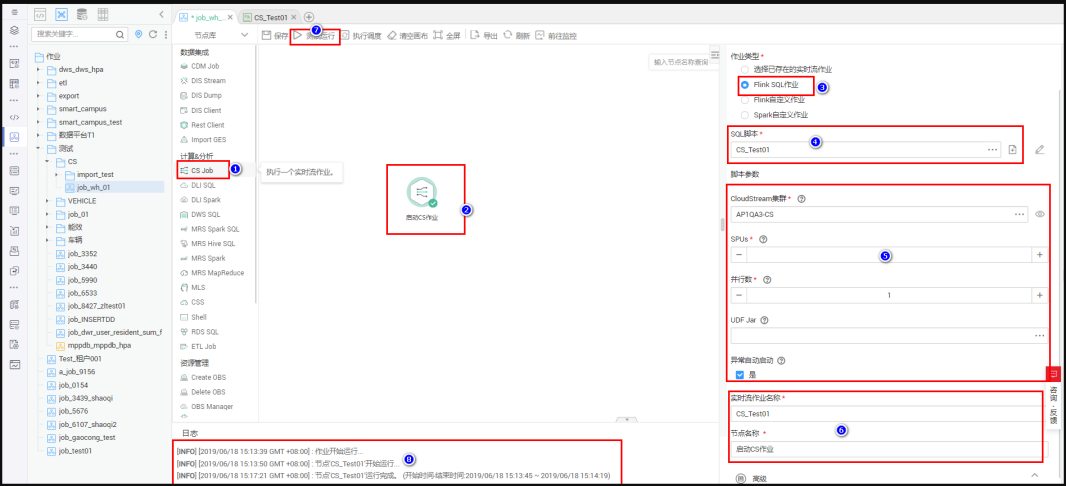
* 注意事项。

变量表达式使用单引号括起来，表达式中变量名使用双引号括起来。

##### 任务配置&调度配置

* DLF任务创建。

作业名称和节点名称保持一致。可以与CS作业名称一致。UDF jar路径配置为：#{Env.get("G\_obs\_flink\_jar\_path")}。

**图1 作业配置**  


* 调度配置。

**图2 调度配置**  


##### 归档

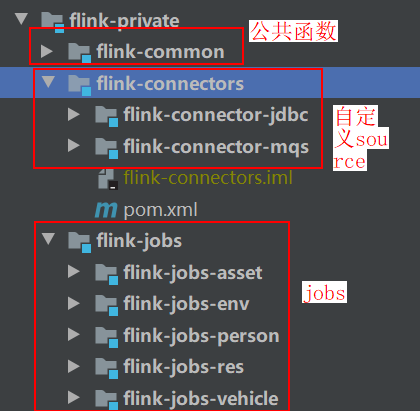
* 脚本归档位置：ParkCore\DataService\dlf\scripts
* 作业归档位置：ParkCore\DataService\dlf\jobs

#### Flink(java)开发规范

##### 实时代码开发规范

###### 模块命名规范

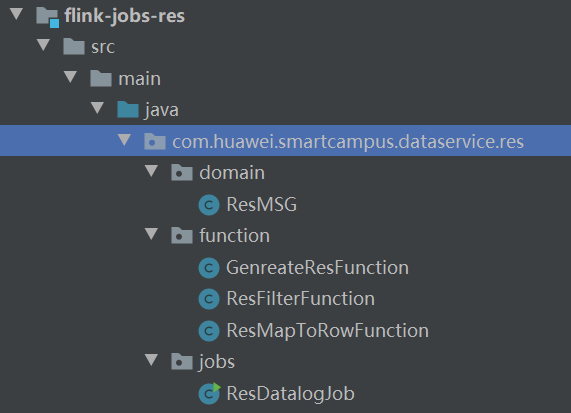
* 每个模块是一个maven子工程，父工程为flink-private。
* 公共函数工具类等放在flink-private/flink-common子工程下。
* 自定义source等连接器放在flink-private/flink-connectors下，目录命名参考flink-connector-jdbc。
* Job工程放在flink-private/flink-jobs/flink-jobs-{模块名}下。



###### 包命名规范

任务包名可根据实际情况自行定义。最底层的三个目录建议分层：

* domain:存放任务用到的消息对应的Java POJO对象。
* function：存放flink任务处理过程中的自定义转换函数。
* jobs：存放本模块下所有的flink任务。



###### 类命名规范

* 任务的job入口类，命名以{业务语义}Job为佳，例如：ResDatalogJob，PeopleHeatJob
* 任务用到的自定义Function，建议名称为{功能}Function，例如：ResFilterFunction
* Java POJO类命名可根据消息类型自行定义。

###### 其他规范

* 代码中需要方法要给出注释。
* 关键步骤要添加注释。
* 关键处需要打印日志，利于定位问题。
* 异常要打印日志，及时处理。
* 变量命名要规范，不要随意命名。

##### Java代码通用规范

###### 代码风格

* 命名
  + 为包、类、方法、变量取一个合适的名字，使代码易于理解。
  + 禁止使用魔鬼数字
  + 直接使用数字，造成代码难以理解，也难以维护。应采用有意义的静态变量或枚举来代替。
  + 例外情况，有些特殊情况下，如循环或比较时采用数字0，-1，1，这些情况可采用数字。
  + 常量命名，由全大写单词组成，单词间用下划线分隔，且使用 static final修饰。
  + 变量、属性命名，使用名词，并采用首字母小写的驼峰命名法。
  + 方法的命名，用动词和动宾结构，并采用首字母小写的驼峰命名法，格式如下：
    - get + 非布尔属性名()
    - is + 布尔属性名()
    - set + 属性名()
    - has + 名词/形容词()
    - 动词()
    - 动词 + 宾语()
  + 类和接口的命名，采用首字母大写的驼峰命名法。
  + 包的命名，由一个或若干个单词组成，所有的字母均为小写。
* 注释
  + 注释应解释代码的意图，而不是描述代码怎么做的。
  + 保证注释与代码一致，避免产生误导。
  + 注释应与其描述代码位置相邻，放在所注释代码上方或右方，并与代码采用同样缩进。
  + 一般单行注释用//，块注释用/\* \*/，JavaDoc注释用/\*\* \*/
  + 排版
  + 团队应遵守一致的排版风格。
  + 将排版风格固化到IDE的代码格式化配置文件中，并让整个团队使用。
  + 在不同的概念之间，增加空行。
  + 将逻辑紧密相关的代码放在一起。
  + 控制一行的宽度，不要超过120个字符。
  + 在不同的概念间（关键字、变量、操作符等）增加空格，以便清楚区分概念。
  + 采用缩进来区分不同层次的概念。
  + 将局部变量的作用域最小化。
  + 给if、for、do、while、switch等语句的执行体加大括号{}。
  + 控制文件的长度，最好不要超过500行。

###### 变量和类型

* 避免随意进行类型强制转换，应改善设计，或在转换前用instanceof进行判断。
* 需要精确计算时不要使用float和double。
* 不能用浮点数作为循环变量。
* 浮点型数据判断相等不能直接使用==。
* 避免同一个局部变量在前后表达不同的含义。
* 不要在单个的表达式中对相同的变量赋值超过一次。
* 基本类型优于包装类型，注意合理使用包装类型。

###### 方法

* 方法设计的第一原则是要短小。
* 方法设计应遵循单一职责原则（SRP），一个方法仅完成一个功能。
* 方法设计应遵循单一抽象层次原则（SLAP）。

IMG_256

SLAP原则，是指让一个方法中所有的操作处于相同的抽象层。否则跳跃的代码的抽象层次破坏了代码的流畅性。

* 方法设计应遵循命令与查询职责分离原则（CQRS）。
* 不要把方法的入参当做工作变量/临时变量，除非特别需要。
* 应明确规定对接口方法参数的合法性检查由调用者负责还是由接口方法本身负责。
* 方法的参数个数不宜过多，不要超过七个。
* 谨慎使用可变数量参数的方法。

###### 异常和日志

* 异常
  + 只针对真正异常的情况才使用exception机制；不要用exception机制来做流程控制。exception机制只应该用于处理罕见的、意料之外的、导致正常流程无法继续执行的行为，而不是取代正常的业务逻辑判断。
  + 在抛出异常的细节信息中，应包含能捕获失败的信息。
  + 对可恢复的情况使用受检异常（checked exception），对编程错误使用运行时异常（runtime exception）。
  + 方法注释和文档中要包含所抛出异常的说明。
  + 对第三方API抛出大量各类异常进行封装。
  + 在finally块中不要使用return、break或continue使finally块非正常结束。
  + 不要直接捕获受检异常的基类Exception。
  + 一个方法不应抛出太多类型的异常。
* 日志
  + 日志信息准确、繁简得当，满足快速定位的需要。
  + 日志的记录，不要使用 System.out 与 System.err 进行控制台打印，应该使用专用的日志工具(比如：slf4j+logback)进行处理。
  + 日志工具对象logger应声明为private static final。
  + 日志应分等级；日志可分为以下级别：debug、info、warn、error、fatal。
  + 日志中不要记录敏感信息。

### 数据服务开发规范

#### ROMA应用开发规范

##### ROMA应用的命名规范

* dataservice.api.xxx

其中xxx为专题库的名称 (security、facility、env、energy、asset、access、office)。

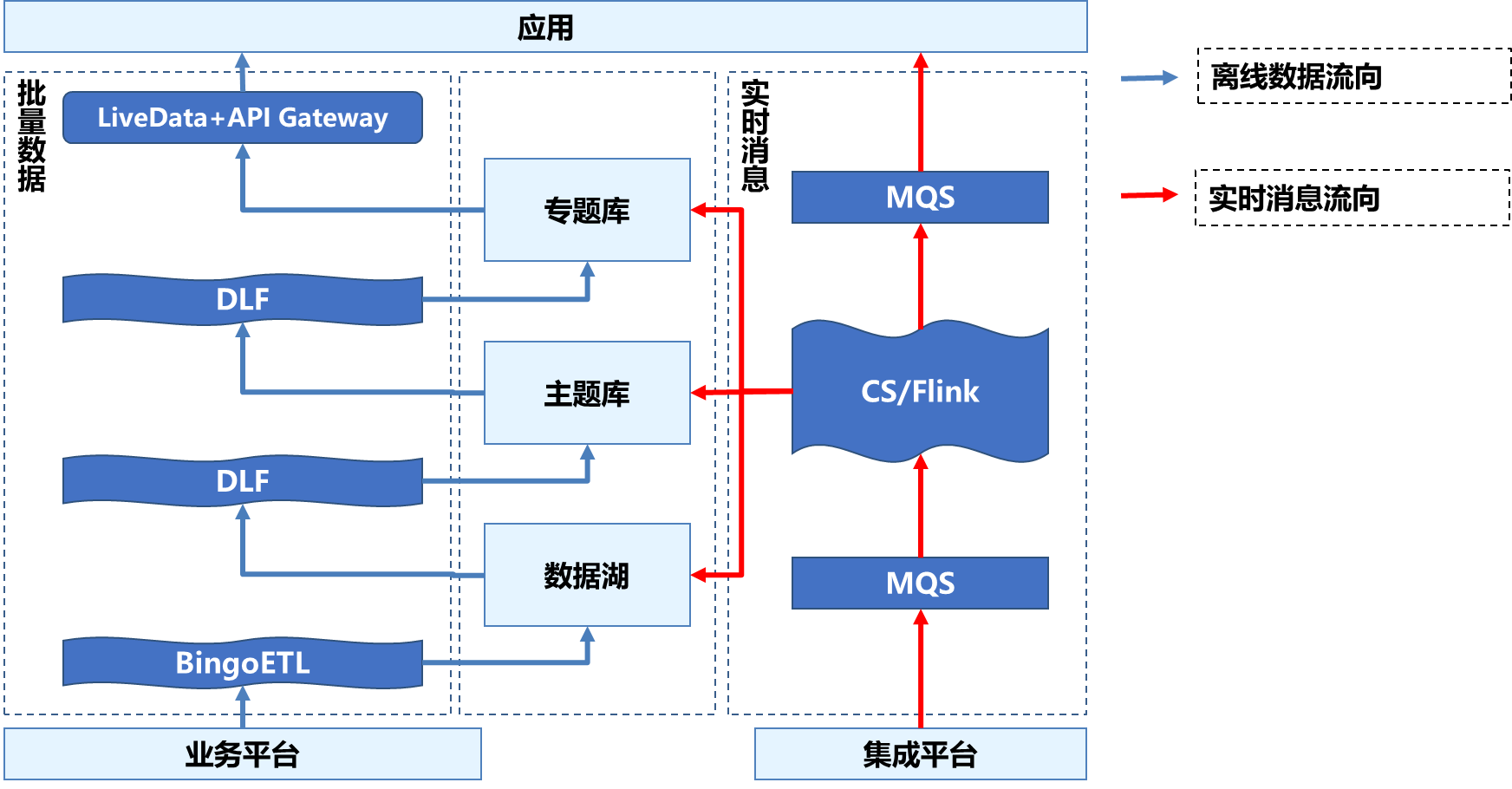
* dataservice.msg.yyy

其中yyy为11个贴源层的名称（facility、alarm、org、space、person、order、vehicle、asset、env、log、res），如果不需要实时消息，则不创建。

##### ROMA中不同应用的开发内容

* dataservice.api.xxx仅仅包含LiveData，为上层应用提供服务。
* dataservice.msg.yyy包含MQS配置。

数据流向图如[图1](https://support.huawei.com/bescloud/Smart%20Campus/20.0/topic/view.do?projectid=328572906&lang=zh&pidid=pid_bookmap_0175250265&topicid=TOPIC_0200882365#ZH-CN_TOPIC_0200882365__fig2196111572211)，数据平台负责虚线框中的部分：

**图1 数据流向图**  


#### LiveData开发规范

##### URL命名规范

* 统一采用“/dataservice/api/模块名/关键词”的形式。
* 模块名与关键词规范。
  + 名词组合。
  + 各种名词以“-”分割。
  + 单词一律小写。
  + 模块名为应用名称（security、facility、env、energy、asset、access、office）。
  + 关键词有一定的自注释能力。

##### LiveData数据来源优先级

* DM>DWS>DWD>DWI

##### LiveData命名规范

* 名称有一定的自注释能力。
* 名称使用动宾结构。
* 统计类接口以“（时间）”结尾，如统计在园人数（实时）、统计车辆通过数量（天）、统计用电量（月）。
* 描述信息是对接口的解释，让读者知道接口的作用。

##### HTTP方法规范

* 使用GET方法。

##### 请求参数规范

* 具有一定的自注释能力。
* 单词一律小写，单词之间以“\_”连接。
* 例如“园区编码：park\_code”。

##### 返回类型规范

* 使用JSON。

##### LiveData脚本

* SQL嵌套不能超过2层。
* 不能使用“select \*”，使用具体字段。
* 关键逻辑添加注释。
* 尽量减少函数的使用。

### 附录

常用字段

|  |  |  |  |
| --- | --- | --- | --- |
| **字段** | **含义** | **类型** | **长度** |
| device\_id | 设备id | character varying | 100 |
| device\_code | 设备code | character varying | 255 |
| park\_code | 园区code | character varying | 100 |
| park\_name\_cn | 园区中文名 | character varying | 500 |
| count\_entry | 进园数 | integer | - |
| count\_exits | 出园数 | integer | - |
| report\_date | 记录日期（天） | timestamp without time zone | - |
| record\_time | 工单记录时间 | timestamp without time zone | - |
| alarm\_level\_info | 提示 | character varying | 100 |
| alarm\_level\_normal | 一般 | character varying | 100 |
| alarm\_level\_major | 重要 | character varying | 100 |
| alarm\_level\_critical | 严重 | character varying | 100 |
| entry\_time | 进场时间 | timestamp without time zone | - |
| exits\_time | 出场时间 | timestamp without time zone | - |
| gender | 性别 | character varying | 50 |
| person\_name | 人员姓名 | character varying | 100 |
| asset\_code | 资产编码 | character varying | 255 |
| nationality | 国籍 | character varying | 255 |
| effective\_date | 生效时间 | timestamp without time zone | - |
| failure\_time | 失效时间 | timestamp without time zone | - |
| expiration\_date | 截止时间 | timestamp without time zone | - |
| complete\_time | 完成时间 | timestamp without time zone | - |
| definition\_code | 规格编码 | character varying | - |
| category\_name/type\_name | 分类名称 | character varying | 100 |
| action\_type | 操作类型 | character varying | 100 |
| brand | 产品品牌 | character varying | 255 |
| manufacturer | 产品厂商 | character varying | 255 |
| status | 产品状态 | character varying | 100 |
| description | 描述 | character varying | 500 |
| type\_description\_lang | 全类型描述 | character varying | 500 |
| path | 路径 | character varying | 500 |
| date\_of\_birth | 出生日期 | timestamp without time zone | - |
| direction | 人员进出： | character varying | 50 |
| In：进 |
| Out：出 |
| passing\_result | 通行结果 | character varying | 100 |
| space\_code | 空间代码 | character varying | 100 |
| definition\_code | 规格编码 | character varying | 255 |
| altitude | 海拔 | character varying | 100 |
| longitude | 经度 | character varying | 100 |
| latitude | 纬度 | character varying | 100 |
| line\_code | 进线编码 | - | 100 |
| dw\_creation\_by | 数据创建者 | character varying | 100 |
| dw\_creation\_date | 数据创建时间 | timestamp without time zone | - |
| dw\_last\_update\_by | 数据最后更新者 | character varying | 100 |
| dw\_last\_update\_date | 数据最后更新时间 | timestamp without time zone | - |
| dw\_batch\_number | 数据批次号 | Bigint | - |

# 模型设计

## 通用设计

大数据量处理（分布键、行列存储模式、索引构建、数据清理\备份和归档策略）

增加数据入湖、入主题时间戳

### 入湖原则

* **系统级入湖原则：**
  + 原则1：本身不产生数据，而是由其他系统提供数据的不进，如IOC、物联网管理平台
  + 原则2：没有明确下游系统需求使用的数据源系统不入
  + 原则3：属于集市应用，将来有可能变成大数据下游应用的不进
* **表级入湖原则**
  + 原则1：技术控制类表不进
  + 原则2：中间数据表不进
  + 原则3：临时表不进
  + 原则4：备份数据表不进
  + 原则5：作废（失效）表不进
  + 原则6：没有明确含义的表不进
  + 原则7：管理流程类表不进
* **字段级入湖原则**
  + 原则1：冗余字段不进
  + 原则2：长文本(超过5000汉字)信息字段不进
  + 原则3：非结构化信息字段不进
  + 原则4：中间计算结果字段不进
  + 原则5：未启用字段不进
  + 原则6：没有明确含义的字段不进

### 字段定义转换规则

|  |  |  |  |
| --- | --- | --- | --- |
| **Oracle源字段类型** | **MySql源字段类型** | **高斯字段类型** | **备注** |
| varchar2 | varchar | character varying | 默认使用类型 |
| date | date | timestamp | 一般默认长度为6 |
| datetime | datetime | timestamp | 一般默认长度为6 |
| int | int | integer | 不要加长度 |
| char | char | char |  |
| double | double | number |  |
| float | float | number |  |
| timestamp | timestamp | timestamp | 一般默认长度为6 |
| bigint | bigint | bigint | 不要加长度 |
| tinyint | tinyint | tinyint | 不要加长度 |
| decimal | decimal | number |  |
| text | text | text |  |

### 默认审计字段

在数据平台入湖和主题物理模型设计时，默认添加一些默认审计和控制字段。

* **数据入湖**

在物理表末尾添加入湖时间字段，具体信息如下：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **字段名称** | **字段名** | **字段描述** | **类型** | **长度** |
| 数据创建者 | dw\_creation\_by | 数据创建者 | character varying | 100 |
| 数据创建时间 | dw\_creation\_date | 数据创建时间  具体格式为：YYYY-MM-DD HH:mm:ss.SSS（年-月-日 小时:分钟:秒.毫秒） | timestamp | 6 |
| 数据最后更新者 | dw\_last\_update\_by | 数据最后更新者 | character varying | 100 |
| 数据最后更新时间 | dw\_last\_update\_date | 数据最后更新时间  具体格式为：YYYY-MM-DD HH:mm:ss.SSS（年-月-日 小时:分钟:秒.毫秒） | timestamp | 6 |
| 数据批次号 | dw\_batch\_number | 数据批次号 | bigint |  |
| 数据来源子系统 | dw\_data\_source\_id | 描述此数据来源业务系统，字段值参考下表 | character varying | 5 |
| 数据来源子系统名称 | dw\_data\_source | 描述此数据来源业务系统名称 | character varying | 50 |

* 子系统对应值

|  |  |
| --- | --- |
| **子系统ID** | **子系统** |
| 1 | 物业管理系统 |
| 2 | 医院专项服务系统 |
| 3 | 管护系统 |
| 4 | 物联网平台 |
| 5 | 能耗监测 |
| 6 | 工单调度平台 |
| 7 | IOC |
| 8 | 资产运营系统 |

* **主题层**

在物理表添加系统ID字段，具体信息如下：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **字段名称** | **字段名** | **字段描述** | **类型** | **长度** |
| 系统ID | id | 系统唯一标识 | serial |  |

在物理表末尾添加审计和控制字段，具体信息如下：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **字段名称** | **字段名** | **字段描述** | **类型** | **长度** |
| 数据创建者 | dw\_creation\_by | 数据创建者 | character varying | 100 |
| 数据创建时间 | dw\_creation\_date | 数据创建时间  具体格式为：YYYY-MM-DD HH:mm:ss.SSS（年-月-日 小时:分钟:秒.毫秒） | timestamp | 6 |
| 数据最后更新者 | dw\_last\_update\_by | 数据最后更新者 | character varying | 100 |
| 数据最后更新时间 | dw\_last\_update\_date | 数据最后更新时间  具体格式为：YYYY-MM-DD HH:mm:ss.SSS（年-月-日 小时:分钟:秒.毫秒） | timestamp | 6 |
| 数据批次号 | dw\_batch\_number | 数据批次号 | bigint |  |

## DWI模型设计

### 物业服务系统DWI模型设计

#### 数据入湖分析

| **源系统名称** | **数据资源名称** | **描述** | **是否入湖** | **不入湖原因** |
| --- | --- | --- | --- | --- |
|  |
| 污染源 | PS\_POLL\_POLLUTION\_INFO | 污染源基本信息 | 是 |  |  |
| 污染源 | PS\_POLL\_ENTEROUTLET\_INFO | 属于排口 | 是 |  |  |
| 污染源 | PS\_POLL\_WATEROUTLET\_INFO | 属于排口 | 是 |  |  |
| 污染源 | PS\_POLL\_GASOUTLET\_INFO | 属于排口 | 是 |  |  |
| 污染源 | PS\_POLL\_NOISE\_EMISSIONS | 属于排口 | 是 |  |  |
| 污染源 | PS\_POLL\_SOLID\_STORE | 固危废 | 是 |  |  |
| 污染源 | PS\_POLL\_RISK\_UNIT | 风险单元 | 是 |  |  |
| 污染源 | PS\_POLL\_RISK\_GAS\_PROTECT | 风险单元 | 否 | 作废（失效）表 |  |
| 污染源 | PS\_POLL\_EMERGENCY\_MATERIAL | 应急物资 | 是 |  |  |
| 污染源 | PS\_POLL\_EMERGENCY\_PLAN | 应急预案 | 是 |  |  |
| 污染源 | PS\_POLL\_SELF\_MONITOR | 自行监测 | 是 |  |  |
| 污染源 | PS\_POLL\_WATER\_EMISSIONS\_RECORD | 废水排放 | 是 |  |  |
| 污染源 | PS\_POLL\_GAS\_EMISSIONS\_RECORD | 废气排放 | 是 |  |  |
| 污染源 | PS\_POLL\_SOLID\_EMISSIONS\_RECORD | 固废排放 | 是 |  |  |
| 污染源 | PS\_POLL\_NOISE\_EMISSIONS\_RECORD | 噪声排放 | 是 |  |  |
| 污染源 | PS\_POLL\_SOLID\_WASTE | 一般固体废物 | 是 |  |  |
| 污染源 | PS\_POLL\_DANGER\_WASTE | 危废 | 是 |  |  |
| 污染源 | BC\_T\_PETITION\_LETTER | 投诉 | 是 |  |  |
| 污染源 | PS\_POLL\_PUNISH | 行政处罚 | 是 |  |  |
| 污染源 | PS\_POLL\_HIDDENIINVEST | 隐患排查 | 是 |  |  |
| 物联网信息 | iotconnect\_ext\_enterprise\_info | 企业信息 | 是 |  |  |
| 物联网信息 | iotconnect\_point\_info | 点位基本信息 | 是 |  |  |
| 物联网信息 | iotconnect\_dev\_info | 设备 | 是 |  |  |
| 物联网信息 | EDA\_REALTIME\_DATA | 监控数据 | 是 |  |  |
| 物联网信息 | IOT\_MINUTE\_DATA | 监控数据 | 是 |  |  |
| 物联网信息 | IOT\_HOUR\_DATA | 监控数据 | 是 |  |  |
| 物联网信息 | IOT\_DAY\_DATA | 监控数据 | 是 |  |  |
| 任务和问题 | sls\_event\_current | 环保告警信息 | 是 |  |  |
| 任务和问题 | sls\_event\_original | 问题源信息 | 是 |  |  |
| 任务和问题 | sls\_task | 由环保问题生成的工单 | 是 |  |  |
| 任务和问题 | sls\_event\_process | 问题处理过程明细 | 是 |  |  |
| 任务和问题 | sls\_task\_dynamic\_operation | 任务处理过程明细 | 是 |  |  |
| 任务和问题 | sls\_task\_subsequent\_process | 后续处置任务关系 | 是 |  |  |

#### 物理模型设计

详见：…\详细设计\贴源层\环保\01-模型设计\《DWI-环保系统-物理模型设计V1.0.xlsx》。

### 能耗监测系统DWI模型设计

#### 数据入湖分析

| **源系统名称** | **数据资源名称** | **描述** | **是否入湖** | **不入湖原因** |
| --- | --- | --- | --- | --- |
|
|  |
| 能耗采集 | dx\_device\_activity\_status | 设备激活状态 | 是 |  |
| 能耗采集 | dx\_device\_info | 设备信息 | 是 |  |
| 能耗采集 | dx\_notify\_bind\_device | 绑定设备信息 | 是 |  |
| 能耗管理 | electric\_data\_collect\_hour | 用电数据采集（小时） | 是 |  |
| 能耗管理 | electric\_data\_day | 用电数据汇总（天） | 是 |  |
| 能耗管理 | electric\_data\_month | 用电数据汇总（月） | 是 |  |
| 能耗管理 | electric\_data\_quarter | 用电数据汇总（季） | 是 |  |
| 能耗管理 | electric\_data\_year | 用电数据汇总（年） | 是 |  |
| 能耗管理 | water\_data\_collect\_day | 用水数据采集（天） | 是 |  |
| 能耗管理 | water\_data\_month | 用水数据汇总（月） | 是 |  |
| 能耗管理 | water\_data\_quarter | 用水数据汇总（季） | 是 |  |
| 能耗管理 | water\_data\_year | 用水数据汇总（年） | 是 |  |
| 能耗管理 | device\_info | 仪表基础信息 | 是 |  |
| 能耗管理 | energy\_warning\_record | 能耗预警记录 | 是 |  |
| 能耗采集 | dx\_auth | 鉴权 | 否 | 技术控制类表 |
| 能耗采集 | dx\_device\_service | 设备数据变化体 | 否 | 中间数据表 |
| 能耗采集 | dx\_error\_def | 错误定义 | 否 | 中间数据表 |
| 能耗采集 | dx\_error\_log | 错误日志 | 否 | 中间数据表 |
| 能耗采集 | dx\_notify\_device\_added | 注册设备通知 | 否 | 中间数据表 |
| 能耗采集 | dx\_notify\_device\_data\_changed | 设备数据变化通知 | 否 | 中间数据表 |
| 能耗采集 | dx\_notify\_device\_info\_changed | 设备信息变化通知 | 否 | 中间数据表 |
| 能耗采集 | dx\_reg\_direct\_device | 注册 | 否 | 中间数据表 |
| 能耗采集 | ammeter\_info | 电表信息 | 否 | 中间数据表 |
| 能耗采集 | pos\_positive\_power | 电表用电量 | 否 | 中间数据表 |
| 能耗管理 | business\_info | 企业基础信息 | 否 | 管理流程类表 |
| 能耗管理 | warning\_threshold\_configuration | 预警阀值配置 | 否 | 技术控制类表 |
| 能耗管理 | word\_order\_info | 工单记录 | 否 | 技术控制类表 |
| 能耗管理 | operation\_daily | 操作日志 | 否 | 技术控制类表 |
| 能耗管理 | system\_daily | 系统日志 | 否 | 技术控制类表 |
| 能耗管理 | sl\_xt\_user | 用户管理（与华为IAM表对应） | 否 | 管理流程类表 |
| 能耗管理 | sl\_xt\_role | 角色管理 | 否 | 管理流程类表 |
| 能耗管理 | sl\_xt\_power | 权限管理 | 否 | 管理流程类表 |
| 能耗管理 | note\_notice | 短信通知 | 否 | 作废（失效）表 |
| 能耗管理 | alarm\_message | 告警基本信息 | 否 | 作废（失效）表 |

#### 物理模型设计

详见：…\详细设计\贴源层\能耗监测\01-模型设计\《DWI-能耗监测-物理模型设计V1.0.xlsx》。

### 医院专项服务系统DWI模型设计

#### 数据入湖分析

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **源系统名称** | **数据资源名称** | **描述** | **是否入湖** | **不入湖原因** |
|
| 消防 | xf\_device | 消防设备数据信息 | 是 | 消防 |

#### 物理模型设计

详见：…\详细设计\贴源层\消防\01-模型设计\《DWI-智慧消防-物理模型设计V1.0.xlsx》。

### 资产运营系统DWI模型设计

#### 数据入湖分析

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **源系统名称** | **数据资源名称** | **描述** | **是否入湖** | **不入湖原因** |
|
| 国有资产 | udip\_dict\_item | 数据字典基本信息 | 是 |  |
| 国有资产 | udip\_operation\_log | 操作日志基本信息 | 否 | 技术控制类表 |
| 国有资产 | land | 土地资源基本信息 | 是 |  |
| 国有资产 | room | 厂房办公房基本信息 | 是 |  |
| 国有资产 | room\_repair | 用房维修基本信息 | 是 |  |
| 国有资产 | owner\_ship | 产权产籍基本信息 | 是 |  |
| 国有资产 | office\_equipment | 办公设备基本信息 | 是 |  |
| 国有资产 | car | 车辆信息基本信息 | 是 |  |
| 国有资产 | car\_maintenance | 维保信息基本信息 | 是 |  |
| 国有资产 | car\_insurance | 保险信息基本信息 | 是 |  |
| 国有资产 | car\_repair | 维修记录基本信息 | 是 |  |
| 国有资产 | car\_fault | 故障记录基本信息 | 是 |  |

#### 物理模型设计

详见：…\详细设计\贴源层\国有资产\01-模型设计\《DWI-国有资产-物理模型设计V1.0.xlsx》。

### 物联网DWI模型设计

#### 数据入湖分析

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **源系统名称** | **数据资源名称** | **描述** | **是否入湖** | **不入湖原因** |
|
| 告警BO | AL\_Alarm | 告警表 | 是 |  |
| 告警BO | AL\_AlarmDef | 告警事件规格 | 是 |  |
| 告警BO | AL\_AlarmDefCategory | 告警实例附件历史表 | 是 |  |
| 告警BO | AL\_AlarmDefAttr | 告警实例历史操作历史表 | 是 |  |
| 告警BO | AL\_AlarmAttr | 告警实例操作历史表 | 是 |  |
| 告警BO | AL\_AttrDef | 告警实例附件表 | 是 |  |
| 告警BO | AL\_TypeUnit | 告警历史表 | 是 |  |
| 告警BO | AL\_AlarmAttrHistory | 告警规格分类 | 是 |  |
| 告警BO | AL\_AlarmAttachmentHistory | 告警规格属性 | 否 | 技术控制类表 |
| 告警BO | AL\_AlarmHandleLogHistory | 告警事件实例属性值 | 否 | 技术控制类表 |
| 告警BO | AL\_AlarmOperationLog | 告警事件实例属性值定义 | 否 | 技术控制类表 |
| 告警BO | AL\_AlarmAttachment | 告警属性值类型单位表 | 否 | 技术控制类表 |
| 告警BO | AL\_AlarmDefDisplay | 告警实例属性值历史表 | 否 | 技术控制类表 |
| 告警BO | AL\_AlarmDefSource | 告警事件规格界面显示 | 否 | 技术控制类表 |
| 告警BO | AL\_AlarmHistory | 告警事件规格来源设备消息类型 | 否 | 技术控制类表 |
| 设备BO | DE\_DeviceDefCategory | 设备规格 | 是 |  |
| 设备BO | DE\_DeviceDef | 设备产品表 | 是 |  |
| 设备BO | DE\_DeviceProduct | 产品表 | 是 |  |
| 设备BO | DE\_DeviceDefService | 设备规格分类 | 否 | 技术控制类表 |
| 设备BO | DE\_DeviceProductAttr | 设备规格具有的服务表 | 否 | 技术控制类表 |
| 设备BO | DE\_Devices | 设备 | 是 |  |
| 设备BO | DE\_ExternalChannel | 设备渠道表 | 否 | 技术控制类表 |
| 设备BO | DE\_DeviceDefMessage | 设备规格具有的消息表 | 否 | 技术控制类表 |
| 设备BO | DE\_DeviceLocation | 设备位置表 | 否 | 技术控制类表 |
| 设备BO | DE\_DeviceDefAttri | 设备规格属性表 | 否 | 技术控制类表 |
| 设备BO | DE\_DeviceDefDisplay | 设备规格展示表 | 否 | 技术控制类表 |
| 设备BO | DE\_DeviceProductAttachment | 设备产品附件表 | 否 | 技术控制类表 |
| 设备BO | DE\_AttrDefCategoryMember | 设备属性类别成员表 | 否 | 技术控制类表 |
| 设备BO | DE\_FuncDefCategory | 设备功能类别表 | 否 | 技术控制类表 |
| 设备BO | DE\_DeviceProductService | 设备产品支持的服务表 | 否 | 技术控制类表 |
| 设备BO | DE\_DeviceProductMessage | 设备产品支持上报的消息表 | 否 | 技术控制类表 |
| 设备BO | DE\_TriggerInstanceParam | 设备触发器实例参数表 | 否 | 技术控制类表 |
| 设备BO | DE\_TriggerDefParam | 设备触发器规格参数表 | 否 | 技术控制类表 |
| 设备BO | DE\_TriggerDef | 设备触发器规格表 | 否 | 技术控制类表 |
| 设备BO | DE\_TriggerInstance | 设备触发器实例表 | 否 | 技术控制类表 |
| 设备BO | DE\_DeviceAttr | 设备属性表 | 否 | 技术控制类表 |
| 设备BO | DE\_AttrDef | 设备属性定义表 | 否 | 技术控制类表 |
| 设备BO | DE\_AttrType | 设备属性类型表 | 否 | 技术控制类表 |
| 设备BO | DE\_AttrTypeUnit | 设备属性值类型单位表 | 否 | 技术控制类表 |
| 设备BO | DE\_DeviceRelationType | 设备关联关系类型 | 否 | 技术控制类表 |
| 设备BO | DE\_DeviceRelation | 设备关联关系 | 否 | 技术控制类表 |
| 设备BO | DE\_DeviceRegisterRequest | 设备自注册请求 | 否 | 技术控制类表 |
| 工单BO | FWO\_FlowStatusMeta | 现场服务单摘要信息 | 是 |  |
| 工单BO | FWO\_TemplateMeta | 工单方案元数据 | 是 |  |
| 工单BO | FWO\_FlowTransitionMeta | 流程状态表 | 否 | 技术控制类表 |
| 工单BO | FWO\_DefaultSchemaConfig | 工单模板表 | 否 | 技术控制类表 |
| 工单BO | FWO\_FlowConditionMeta | 工单转换表 | 否 | 技术控制类表 |
| 工单BO | FWO\_TemplateElementRelation | 设置默认工单方案 | 否 | 技术控制类表 |
| 工单BO | EQM\_FileMeta | 工单条件表 | 否 | 技术控制类表 |
| 工单BO | EQM\_FieldMeta | 工单模板元素关系表 | 否 | 技术控制类表 |
| 工单BO | FWO\_ElementMeta | 设备文件元数据 | 否 | 技术控制类表 |
| 工单BO | FWO\_FlowMeta | 设备字段元数据 | 否 | 技术控制类表 |
| 工单BO | FWO\_SchemeMeta | 元素元数据 | 否 | 技术控制类表 |
| 工单BO | FWO\_CheckListRelation | 流程元数据 | 否 | 技术控制类表 |
| 工单BO | FWO\_Summary | 服务单与检查清单关系表 | 否 | 技术控制类表 |
| 工单BO | FWO\_WorkorderRelation | 服务单的母子关系 | 否 | 技术控制类表 |
| 工单BO | CLK\_CheckListResult | 检查清单项实例值 | 否 | 技术控制类表 |
| 工单BO | CLK\_CheckListInfo | 检查清单 | 否 | 技术控制类表 |
| 工单BO | CLK\_CheckListOldItem | 检查清单项历史 | 否 | 备份数据表 |
| 工单BO | CLK\_CheckListItem | 检查清单项 | 否 | 技术控制类表 |

#### 物理模型设计

详见：…\详细设计\贴源层\智慧安防\01-模型设计\《DWI-智慧安防-物理模型设计V1.0.xlsx》。

### 安全管护系统DWI模型设计

#### 数据入湖分析

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **源系统名称** | **数据资源名称** | **描述** | **是否入湖** | **不入湖原因** |
|
| 告警BO | AL\_Alarm | 告警表 | 是 |  |
| 告警BO | AL\_AlarmDef | 告警事件规格 | 是 |  |
| 告警BO | AL\_AlarmDefCategory | 告警实例附件历史表 | 是 |  |
| 告警BO | AL\_AlarmDefAttr | 告警实例历史操作历史表 | 是 |  |
| 告警BO | AL\_AlarmAttr | 告警实例操作历史表 | 是 |  |
| 告警BO | AL\_AttrDef | 告警实例附件表 | 是 |  |
| 告警BO | AL\_TypeUnit | 告警历史表 | 是 |  |
| 告警BO | AL\_AlarmAttrHistory | 告警规格分类 | 是 |  |
| 告警BO | AL\_AlarmAttachmentHistory | 告警规格属性 | 否 | 技术控制类表 |
| 告警BO | AL\_AlarmHandleLogHistory | 告警事件实例属性值 | 否 | 技术控制类表 |
| 告警BO | AL\_AlarmOperationLog | 告警事件实例属性值定义 | 否 | 技术控制类表 |
| 告警BO | AL\_AlarmAttachment | 告警属性值类型单位表 | 否 | 技术控制类表 |
| 告警BO | AL\_AlarmDefDisplay | 告警实例属性值历史表 | 否 | 技术控制类表 |
| 告警BO | AL\_AlarmDefSource | 告警事件规格界面显示 | 否 | 技术控制类表 |
| 告警BO | AL\_AlarmHistory | 告警事件规格来源设备消息类型 | 否 | 技术控制类表 |

#### 物理模型设计

详见：…\详细设计\贴源层\智慧安防\01-模型设计\《DWI-智慧安防-物理模型设计V1.0.xlsx》。

## DWR模型设计

### 主题模型设计思路

主题模型将基于项目中各应用需求的数据诉求进行设计，对所需数据入湖后，大数据平台主题划分，进行数据抽象、分类建模，从而构成主题数据库。

主题模型设计时采用星型模型构建，遵循3NF约束，从应用出发，适度考虑数据访问的方便性，提供冗余设计。

主题库构建所需包括的主题域和主题内容，如下表所示：

|  |  |
| --- | --- |
| **主题名称** | **主题描述** |
| **巡检主题** | 人员巡检任务分类、状态、数量、时间等、发现问题后后，进行派发及处理的工单信息 |
| **客户服务主题** | 针对客户反映的报修、投诉、咨询、报事、隐患进行信息统计，并且跟进客户处理反馈。 |
| **隐患主题** | 来源于管护系统与隐患报告（客户报告的）发现的隐患分类信息，进行统计分类，包括综合管理、消防检查、电气安全、特种设备、危险作业、结构安全等 |
| **能耗主题** | 园区企业使用的能耗资源信息，包含水、电、热力。 |
| **设备主题** | 管辖区域的物联设备信息，包含各类型物联设备的基础信息（名称、类型、位置，所属企业等）、设备的状态信息（在线状态）。 |
| **工单主题** | 日常工作中进行派发及处理的工单信息，包含工单基本信息、状态，类型等。 |
| **人员主题** | 业务过程涉及的人员信息，包含办公人员、巡检人员、值班人员等。 |
| **组织主题** | 组织结构信息，含公司组织结构、部门组织结构。 |
| **资产运营主题** | 用于统计资产租赁信息与空间资源信息，并收集日常业务中营收台账，企业服务信息以及生活服务信息（订单，成交额，注册量） |
| **医院服务主题** | 涉及医疗废物的管理与治理，医疗运送的专项服务 |
| **车辆主题** | 出入车辆信息，停车位总数，停车位剩余数量，出入车流量统计 |

### 客户服务主题模型设计

#### 逻辑模型

详见：

…\03、系统设计\详细设计\主题层\客户服务\01逻辑模型设计\客户服务.erwin

#### 物理表设计

详见：

…\03、系统设计\详细设计\主题层\客户服务\02物理模型设计\DWR-客户服务-物理模型设计-V1.0.xlsx

…\03、系统设计\详细设计\主题层\客户服务\02物理模型设计\DWR-客户服务-标准代码表设计-V1.0.xlsx

### 设备主题模型设计

#### 逻辑模型

详见：

…\03、系统设计\详细设计\主题层\设备\01逻辑模型设计\设备.erwin

#### 物理表设计

详见：

…\03、系统设计\详细设计\主题层\设备\02物理模型设计\DWR-设备-物理模型设计-V1.0.xlsx

…\03、系统设计\详细设计\主题层\设备\02物理模型设计\DWR-设备-标准代码表设计-V1.0.xlsx

### 能耗主题模型设计

#### 逻辑模型

详见：

…\03、系统设计\详细设计\主题层\能耗\01逻辑模型设计\用电.erwin

…\03、系统设计\详细设计\主题层\能耗\01逻辑模型设计\用水.erwin

#### 物理表设计

详见：

…\03、系统设计\详细设计\主题层\能耗\02物理模型设计\DWR-能耗-物理模型设计-V1.0.xlsx

…\03、系统设计\详细设计\主题层\能耗\02物理模型设计\DWR-能耗-标准代码表设计-V1.0.xlsx

### 巡检主题模型设计

#### 逻辑模型

详见：

…\03、系统设计\详细设计\主题层\巡检\01逻辑模型设计\巡检.erwin

#### 物理表设计

详见：

…\03、系统设计\详细设计\主题层\巡检\02物理模型设计\DWR-巡检-物理模型设计-V1.0.xlsx

…\03、系统设计\详细设计\主题层\巡检\02物理模型设计\DWR-巡检-标准代码表设计-V1.0.xlsx

### 隐患主题模型设计

#### 逻辑模型

详见：

…\03、系统设计\详细设计\主题层\隐患\01逻辑模型设计\隐患.erwin

#### 物理表设计

详见：

…\03、系统设计\详细设计\主题层\隐患\02物理模型设计\DWR-隐患-物理模型设计-V1.0.xlsx

…\03、系统设计\详细设计\主题层\隐患\02物理模型设计\DWR-隐患-标准代码表设计-V1.0.xlsx

### 车辆主题模型设计

### 资产主题模型设计

#### 逻辑模型

详见：

…\03、系统设计\详细设计\主题层\资产\01逻辑模型设计\资产.erwin

#### 物理表设计

详见：

…\03、系统设计\详细设计\主题层\资产\02物理模型设计\DWR-资产-物理模型设计-V1.0.xlsx

…\03、系统设计\详细设计\主题层\资产\02物理模型设计\DWR-资产-标准代码表设计-V1.0.xlsx

### 医疗主题模型设计

#### 逻辑模型

详见：

…\03、系统设计\详细设计\主题层\医疗\01逻辑模型设计\中央运送.erwin

…\03、系统设计\详细设计\主题层\医疗\01逻辑模型设计\医废处理.erwin

#### 物理表设计

详见：

…\03、系统设计\详细设计\主题层\医疗\02物理模型设计\DWR-医疗-物理模型设计-V1.0.xlsx

…\03、系统设计\详细设计\主题层\医疗\02物理模型设计\DWR-医疗-标准代码表设计-V1.0.xlsx

### 工单主题模型设计

#### 逻辑模型

详见：

…\03、系统设计\详细设计\主题层\应急\01逻辑模型设计\工单模型.erwin

#### 物理表设计

详见：

…\03、系统设计\详细设计\主题层\应急\02物理模型设计\DWR-工单-物理模型设计-V1.0.xlsx

…\03、系统设计\详细设计\主题层\应急\02物理模型设计\DWR-工单-标准代码表设计-V1.0.xlsx

### 组织主题模型设计

#### 逻辑模型

详见：

…\03、系统设计\详细设计\主题层\组织\01逻辑模型设计\组织.erwin

#### 物理表设计

详见：

…\03、系统设计\详细设计\主题层\组织\02物理模型设计\DWR-组织-物理模型设计-V1.0.xlsx

…\03、系统设计\详细设计\主题层\组织\02物理模型设计\DWR-组织-标准代码表设计-V1.0.xlsx

### 人员主题模型设计

#### 逻辑模型

详见：

…\03、系统设计\详细设计\主题层\人员\01逻辑模型设计\人员.erwin

#### 物理表设计

详见：

…\03、系统设计\详细设计\主题层\人员\02物理模型设计\DWR-人员-物理模型设计-V1.0.xlsx

…\03、系统设计\详细设计\主题层\人员\02物理模型设计\DWR-人员-标准代码表设计-V1.0.xlsx

# 数据采集功能设计

## 物业服务采集

### 功能

主要完成环保系统的数据采集功能中数据采集方式及策略定义、数据Mapping、FDI设计等工作。

### 数据采集方式及策略

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **源系统名称** | **数据资源名称** | **采集频率** | **采集策略** | **数据规模** | |
| **存量（记录数）** | **增量（记录数）** |
| 污染源 | 污染源基本信息 | 7-日 | 增量 | 1000 | <100 |
| 污染源 | 进水口（排口） | 7-日 | 增量 | 100 | <100 |
| 污染源 | 废水排口（排口） | 7-日 | 增量 | 100 | <100 |
| 污染源 | 废气排口（排口） | 7-日 | 增量 | 100 | <100 |
| 污染源 | 噪声排口（排口） | 7-日 | 增量 | 0 | 0 |
| 污染源 | 固危废 | 7-日 | 增量 | 1000 | <100 |
| 污染源 | 风险单元主表 | 7-日 | 增量 | 1000 | <100 |
| 污染源 | 副表-气风险单元 | 7-日 | 增量 | 1000 | <100 |
| 污染源 | 应急物资 | 7-日 | 增量 | 1000 | <100 |
| 污染源 | 应急预案 | 7-日 | 增量 | 1000 | <100 |
| 污染源 | 自行监测 | 7-日 | 增量 | 1000 | <100 |
| 污染源 | 废水排放 | 7-日 | 增量 | 1000 | <100 |
| 污染源 | 废气排放 | 7-日 | 增量 | 1000 | <100 |
| 污染源 | 固废排放 | 7-日 | 增量 | 1000 | <100 |
| 污染源 | 噪声排放 | 7-日 | 增量 | 1000 | <100 |
| 污染源 | 一般固体废物 | 7-日 | 增量 | 1000 | <100 |
| 污染源 | 危废 | 7-日 | 增量 | 1000 | <100 |
| 污染源 | 投诉 | 7-日 | 增量 | 1000 | <100 |
| 污染源 | 行政处罚 | 7-日 | 增量 | 1000 | <100 |
| 污染源 | 隐患排查 | 7-日 | 增量 | 1000 | <100 |
| 物联网信息 | 企业信息 | 7-日 | 增量 | 1000 | <100 |
| 物联网信息 | 点位 | 7-日 | 增量 | 1000 | <100 |
| 物联网信息 | 设备 | 7-日 | 增量 | 100 | <100 |
| 任务和问题 | 问题 | 4-小时 | 增量 | 0 | 100 |
| 任务和问题 | 问题源信息 | 4-小时 | 增量 | 0 | <100 |
| 任务和问题 | 任务 | 4-小时 | 增量 | 0 | 30 |
| 任务和问题 | 问题流水 | 4-小时 | 增量 | 0 | 1000 |
| 任务和问题 | 任务流 | 4-小时 | 增量 | 0 | 1000 |
| 任务和问题 | 后续处置任务关系表 | 4-小时 | 增量 | 0 | 1000 |
| 物联网信息 | 环保设备监控数据\_实时 | 分钟 | 增量 | 0 | >5000 |
| 物联网信息 | 环保设备监控数据\_分钟 | 分钟 | 增量 | 0 | >5000 |
| 物联网信息 | 环保设备监控数据\_小时 | 4-小时 | 增量 | 0 | <2000 |
| 物联网信息 | 环保设备监控数据\_天 | 7-日 | 增量 | 0 | <1000 |

### 数据Mapping

详见：…\详细设计\贴源层\环保\ 02-Mapping\《DWI-环保系统-Mapping分析表-V1.0.xlsx》。

### FDI设计

详见：…\详细设计\贴源层\环保\ 03-FDI设计\《DWI-环保系统-FDI设计V1.0.xlsx》。

## 能耗管理采集

### 功能

主要完成能耗监测系统的数据采集功能中数据采集方式及策略定义、数据Mapping、FDI设计等工作。

### 数据采集方式及策略

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **源系统名称** | **数据资源名称** | **采集频率** | **采集策略** | **数据规模** | |
| **存量（记录数）** | **增量（记录数）** |
| 能耗采集 | [设备激活状态](#设备激活状态!A1) | 7-日 | 增量 | 500 | 10 |
| 能耗采集 | [设备信息](#设备信息!A1) | 7-日 | 增量 | 500 | 10 |
| 能耗采集 | [绑定设备信息](#设备绑定信息!A1) | 7-日 | 增量 | 500 | 10 |
| 能耗管理 | [用电数据采集（小时）](#'用电数据采集（小时）'!A1) | 4-1小时 | 增量 | 0 | 1200 |
| 能耗管理 | [用电数据汇总（天）](#'用电数据汇总（天）'!A1) | 7-日 | 增量 | 0 | 300 |
| 能耗管理 | [用电数据汇总（月）](#'用电数据汇总（月）'!A1) | 9-月 | 增量 | 0 | 10 |
| 能耗管理 | [用电数据汇总（季）](#'用电数据汇总（季）'!A1) | 9-月 | 增量 | 0 | 10 |
| 能耗管理 | [用电数据汇总（年）](#'用电数据汇总（年）'!A1) | 9-月 | 增量 | 0 | 10 |
| 能耗管理 | [用水数据采集（天）](#'用水数据采集（天）'!A1) | 7-日 | 增量 | 0 | 300 |
| 能耗管理 | [用水数据汇总（月）](#'用水数据采集（月）'!A1) | 9-月 | 增量 | 0 | 10 |
| 能耗管理 | [用水数据汇总（季）](#'用水数据采集（季）'!A1) | 9-月 | 增量 | 0 | 10 |
| 能耗管理 | [用水数据汇总（年）](#'用水数据采集（年）'!A1) | 9-月 | 增量 | 0 | 10 |
| 能耗管理 | [仪表基础信息](#'仪表基础信息 '!A1) | 7-日 | 增量 | 500 | 10 |
| 能耗管理 | [能耗预警记录](#能耗预警记录!A1) | 7-日 | 增量 | 0 | 100 |

### 数据Mapping

详见：…\详细设计\贴源层\能耗\ 02-Mapping\《DWI-能耗监测-Mapping分析表-V1.0.xlsx》。

### FDI设计

详见：…\详细设计\贴源层\能耗\ 03-FDI设计\《DWI-能耗监测-FDI设计V1.0.xlsx》。

## 医院专项服务采集

### 功能

主要完成智慧消防系统的数据采集功能中数据采集方式及策略定义、数据Mapping、FDI设计等工作。

### 数据采集方式及策略

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **源系统名称** | **数据资源名称** | **采集频率** | **采集策略** | **数据规模** | |
| **存量（记录数）** | **增量（记录数）** |
| 消防 | [设备表](#设备表!A1) | 4-1小时 | 增量 | 5000 | 100 |

### 数据Mapping

详见：…\详细设计\贴源层\消防\ 02-Mapping\《DWI-智慧消防-Mapping分析表-V1.0.xlsx》。

### FDI设计

详见：…\详细设计\贴源层\消防\ 03-FDI设计\《DWI-智慧消防-FDI设计V1.0.xlsx》。

## 资产运营采集

### 功能

主要完成国有资产系统的数据采集功能中数据采集方式及策略定义、数据Mapping、FDI设计等工作。

### 数据采集方式及策略

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **源系统名称** | **数据资源名称** | **采集频率** | **采集策略** | **数据规模** | |
| **存量（记录数）** | **增量（记录数）** |
| 国有资源 | [数据字典](#污染源基本信息!A1) | 7-日 | 增量 | 0 | 1000 |
| 国有资源 | [土地资源](#'废水排口（排口）'!A1) | 7-日 | 增量 | 0 | 1000 |
| 国有资源 | [厂房办公房](#'废气排口（排口）'!A1) | 7-日 | 增量 | 0 | 1000 |
| 国有资源 | [用房维修](#'噪声排口（排口）'!A1) | 7-日 | 增量 | 0 | 1000 |
| 国有资源 | [产权产籍](#固危废!A1) | 7-日 | 增量 | 0 | 1000 |
| 国有资源 | [办公设备](#风险单元主表!A1) | 7-日 | 增量 | 0 | 1000 |
| 国有资源 | [车辆信息](#'副表-气风险单元'!A1) | 7-日 | 增量 | 0 | 1000 |
| 国有资源 | [维保信息](#应急物资!A1) | 7-日 | 增量 | 0 | 1000 |
| 国有资源 | [保险信息](#应急预案!A1) | 7-日 | 增量 | 0 | 1000 |
| 国有资源 | [维修记录](#自行检测!A1) | 7-日 | 增量 | 0 | 1000 |
| 国有资源 | [故障记录](#废水排放!A1) | 7-日 | 增量 | 0 | 1000 |

### 数据Mapping

详见：…\详细设计\贴源层\国有资产\ 02-Mapping\《DWI-国有资产-Mapping分析表-V1.0.xlsx》。

### FDI设计

详见：…\详细设计\贴源层\国有资产\ 03-FDI设计\《DWI-国有资产-FDI设计V1.0.xlsx》。

## 物联网平台采集

### 功能

主要完成智慧安防系统的数据采集功能中数据采集方式及策略定义、数据Mapping、FDI设计等工作。

### 数据采集方式及策略

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **源系统名称** | **数据资源名称** | **采集频率** | **采集策略** | **数据规模** | |
| **存量（记录数）** | **增量（记录数）** |
| 智慧安防 | [告警表](#告警表!A1) | 2-15分钟 | 增量 |  |  |
| 智慧安防 | [告警事件规格](#告警事件规格!A1) | 7-日 | 增量 |  |  |
| 智慧安防 | [告警规格分类](#告警规格分类!A1) | 7-日 | 增量 |  |  |
| 智慧安防 | [告警实例附件历史表](#告警实例附件历史表!A1) | 4-1小时 | 增量 |  |  |
| 智慧安防 | [告警实例历史操作历史表](#告警实例历史操作历史表!A1) | 4-1小时 | 增量 |  |  |
| 智慧安防 | [告警实例操作历史表](#告警实例操作历史表!A1) | 4-1小时 | 增量 |  |  |
| 智慧安防 | [告警实例附件表](#告警实例附件表!A1) | 4-1小时 | 增量 |  |  |
| 智慧安防 | [告警历史表](#告警历史表!A1) | 4-1小时 | 增量 |  |  |
| 智慧安防 | [设备规格分类](#设备规格分类!A1) | 7-日 | 增量 | 以业务涉及设备数量为准 | 以业务产生数量为准 |
| 智慧安防 | [设备规格](#设备规格!A1) | 7-日 | 增量 | 以业务涉及设备数量为准 | 以业务产生数量为准 |
| 智慧安防 | [设备产品表](#设备产品表!A1) | 7-日 | 增量 | 以业务涉及设备数量为准 | 以业务产生数量为准 |
| 智慧安防 | [设备基本信息表](#设备基本信息表!A1) | 4-1小时 | 增量 | 以业务涉及设备数量为准 | 以业务产生数量为准 |
| 智慧安防 | [设备渠道表](#设备渠道表!A1) | 7-日 | 增量 | 以业务涉及设备数量为准 | 以业务产生数量为准 |
| 智慧安防 | [设备位置表](#设备位置表!A1) | 4-1小时 | 增量 | 以业务涉及设备数量为准 | 以业务产生数量为准 |
| 智慧安防 | [现场服务单摘要信息](#现场服务单摘要信息!A1) | 7-日 | 增量 | 0 | 以业务产生数量为准 |

### 数据Mapping

详见：…\详细设计\贴源层\智慧安防\ 02-Mapping\《DWI-智慧安防-Mapping分析表-V1.0.xlsx》。

### FDI设计

详见：…\详细设计\贴源层\智慧安防\ 03-FDI设计\《DWI-智慧安防-FDI设计V1.0.xlsx》。

## 安全管护采集

### 功能

主要完成园区服务平台的数据采集功能中数据采集方式及策略定义、数据Mapping、FDI设计等工作。

### 数据采集方式及策略

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **源系统名称** | **数据资源名称** | **采集频率** | **采集策略** | **数据规模** | |
| **存量（记录数）** | **增量（记录数）** |
| 园区服务平台 | [金融产品](#金融产品!A1) | 7-日 | 增量 | 40000 | 0 |
| 园区服务平台 | [融资申请](#金融申请!A1) | 7-日 | 增量 | 135000 | 0 |
| 园区服务平台 | [融资需求](#融资需求!A1) | 7-日 | 增量 | 135000 | 0 |
| 园区服务平台 | [公司](#公司!A1) | 7-日 | 增量 | 2771752 | 0 |
| 园区服务平台 | [公司内容](#公司内容!A1) | 7-日 | 增量 | 2771752 | 0 |
| 园区服务平台 | [栏目分类](#栏目分类!A1) | 7-日 | 增量 | 16000 | 0 |
| 园区服务平台 | [资讯](#资讯!A1) | 7-日 | 增量 | 116327 | 100 |
| 园区服务平台 | [资讯内容](#资讯内容!A1) | 7-日 | 增量 | 116327 | 100 |
| 园区服务平台 | [专家](#专家!A1) | 7-日 | 增量 | 4600 | 100 |
| 园区服务平台 | [专利库](#专利库!A1) | 7-日 | 增量 | 18600 | 100 |
| 园区服务平台 | [求购](#求购!A1) | 7-日 | 增量 | 22129 | 0 |
| 园区服务平台 | [供应](#供应!A1) | 7-日 | 增量 | 18000 | 0 |
| 园区服务平台 | [供应搜索](#供应搜索!A1) | 7-日 | 增量 | 53572 | 0 |
| 园区服务平台 | [模型库](#模型库!A1) | 7-日 | 增量 | 278456 | 0 |
| 园区服务平台 | [活动](#活动!A1) | 7-日 | 增量 | 15000 | 0 |
| 园区服务平台 | [活动报名](#活动报名!A1) | 7-日 | 增量 | 530000 | 0 |
| 园区服务平台 | [招商](#招商!A1) | 7-日 | 增量 | 240 | 0 |
| 园区服务平台 | [招聘](#招聘!A1) | 7-日 | 增量 | 0 | 0 |
| 园区服务平台 | [应聘工作](#应聘工作!A1) | 7-日 | 增量 | 0 | 0 |
| 园区服务平台 | [简历](#简历!A1) | 7-日 | 增量 | 0 | 0 |
| 园区服务平台 | [人才库](#人才库!A1) | 7-日 | 增量 | 0 | 0 |
| 园区服务平台 | [痛点库](#痛点库!A1) | 7-日 | 增量 | 25000 | 100 |
| 园区服务平台 | [痛点解答](#痛点解答!A1) | 7-日 | 增量 | 175600 | 100 |
| 园区服务平台 | [商城](#商城!A1) | 7-日 | 增量 | 4534705 | 1000 |
| 园区服务平台 | [商城订单](#商城订单!A1) | 7-日 | 增量 | 0 | 0 |
| 园区服务平台 | [订单评论](#订单评论!A1) | 7-日 | 增量 | 0 | 0 |
| 园区服务平台 | [会员](#会员!A1) | 7-日 | 增量 | 2771752 | 100 |
| 园区服务平台 | [科研机构](#科研机构!A1) | 7-日 | 增量 | 400 | 0 |
| 园区服务平台 | [机构人才](#机构人才!A1) | 7-日 | 增量 | 3200 | 0 |
| 园区服务平台 | [科研机构转化](#科研转化!A1) | 7-日 | 增量 | 20000 | 0 |
| 园区服务平台 | [服务](#服务!A1) | 7-日 | 增量 | 18520 | 100 |
| 园区服务平台 | [服务详情](#服务详情!A1) | 7-日 | 增量 | 18520 | 100 |
| 园区服务平台 | [服务需求](#服务需求!A1) | 7-日 | 增量 | 0 | 0 |
| 园区服务平台 | [服务发票信息](#服务发票信息!A1) | 7-日 | 增量 | 0 | 0 |
| 园区服务平台 | [服务订单](#服务订单!A1) | 7-日 | 增量 | 0 | 0 |
| 园区服务平台 | [服务订单评价](#服务订单评价!A1) | 7-日 | 增量 | 0 | 0 |
| 园区服务平台 | [视频](#视频!A1) | 7-日 | 增量 | 4500 | 10 |
| 园区服务平台 | [软件库](#软件库!A1) | 7-日 | 增量 | 8000 | 0 |
| 园区服务平台 | [地区](#地区!A1) | 7-日 | 增量 | 1000 | 0 |
| 园区服务平台 | [模型](#模型!A1) | 7-日 | 增量 | 100 | 0 |
| 园区服务平台 | [物流网点](#物流网点!A1) | 7-日 | 增量 | 12395815 | 10000 |
| 园区服务平台 | [物流线路](#物流线路!A1) | 7-日 | 增量 | 12169 | 1000 |
| 园区服务平台 | [物流需求](#物流需求!A1) | 7-日 | 增量 | 0 | 0 |
| 园区服务平台 | [车源信息](#车源信息!A1) | 7-日 | 增量 | 37693 | 100 |
| 园区服务平台 | [会员组](#会员组!A1) | 7-日 | 增量 | 100 | 0 |
| 园区服务平台 | [软件下载历史](#软件下载历史!A1) | 7-日 | 增量 | 100000 | 100 |

### 数据Mapping

详见：…\详细设计\贴源层\服务平台\ 02-Mapping\《DWI-服务平台-Mapping分析表-V1.0.xlsx》。

### FDI设计

详见：…\详细设计\贴源层\服务平台\ 03-FDI设计\《DWI-服务平台-FDI设计V1.0.xlsx》。

## 测试计划

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **时间点** | **测试类型** | **测试轮次** | **输入数据** | **预期结果** | **测试人员** |
| 开发完成 | 单元测试-自测 | 1 | 源表结构模拟数据 | 1、数据入湖成功； 2、数据格式、类型、值正确 3、FDI任务执行成功 | 开发人员 |
| 单元测试-自测 | 2 | 源表结构模拟数据 | 同上 | 开发人员 |
| 单元测试-自测 | 3 | 源表结构模拟数据 | 同上 | 开发人员 |
| 单元测试 | 1 | 源表结构模拟数据 | 同上 | 测试人员 |
| 单元测试 | 2 | 源表结构模拟数据 | 同上 | 测试人员 |
| 单元测试 | 3 | 源表结构模拟数据 | 同上 | 测试人员 |
| 集成测试阶段 | 集成测试 | 1 | 源表结构真实业务数据 | 1、正确连接源系统； 2、数据入湖成功； 3、数据格式、类型、值正确 4、FDI任务执行成功 5、按采集策略进行正确的数据采集 | 开发人员 测试人员 |
| 集成测试 | 2 | 源表结构真实业务数据 | 同上 | 开发人员 测试人员 |
| 集成测试 | 3 | 源表结构真实业务数据 | 同上 | 开发人员 测试人员 |
| 现场联调阶段 | 现场联调 | 1 | 源表结构真实业务数据 | 1、正确连接源系统； 2、数据入湖成功； 3、数据格式、类型、值正确 4、FDI任务执行成功 5、按采集策略进行正确的数据采集 | 开发人员 测试人员 |
| 现场联调 | 2 | 源表结构真实业务数据 | 同上 | 开发人员 测试人员 |
| 现场联调 | 3 | 源表结构真实业务数据 | 同上 | 开发人员 测试人员 |

# 主题库数据处理设计

## 能耗主题

### 功能

主要完能耗主题的相关数据汇总归档存储，及数据汇总归档存储时的作业调度、数据MAPPING等工作。

### 数据Mapping

详见：…\详细设计\主题层\能耗\ 03Mapping\《DWR-能耗-MAPPING映射-V1.0.xlsx》。

### DLF设计

详见：…\详细设计\主题层\企业\DLF设计\《DWR-DLF设计-企业-V1.0.xlsx》。

## 巡检主题

### 功能

主要完成巡检主题的相关数据汇总归档存储，及数据汇总归档存储时的作业调度、数据MAPPING等工作。

### 数据Mapping

详见：…\详细设计\主题层\巡检\ 03Mapping\《DWR-巡检-MAPPING映射-V1.0.xlsx》。

### DLF设计

详见：…\详细设计\主题层\巡检\DLF设计\《DWR-DLF设计-巡检-V1.0.xlsx》。

## 医疗主题

### 功能

主要完成办公主题的相关数据汇总归档存储，及数据汇总归档存储时的作业调度、数据MAPPING等工作。

### 数据Mapping

详见：…\详细设计\主题层\医疗\ 03Mapping\《DWR-医疗-MAPPING映射-V1.0.xlsx》。

### DLF设计

详见：…\详细设计\主题层\医疗\DLF设计\《DWR-DLF设计-医疗-V1.0.xlsx》。

## 隐患主题

### 功能

主要完成产业主题的相关数据汇总归档存储，及数据汇总归档存储时的作业调度、数据MAPPING等工作。

### 数据Mapping

详见：…\详细设计\主题层\隐患\ 03Mapping\《DWR-隐患-MAPPING映射-V1.0.xlsx》。

### DLF设计

详见：…\详细设计\主题层\产业\DLF设计\《DWR-DLF设计-隐患-V1.0.xlsx》。

## 工单主题

### 功能

主要完成工单主题的相关数据汇总归档存储，及数据汇总归档存储时的作业调度、数据MAPPING等工作。

### 数据Mapping

详见：…\详细设计\主题层\工单\ 03Mapping\《DWR-工单-MAPPING映射-V1.0.xlsx》。

### DLF设计

详见：…\详细设计\主题层\工单\DLF设计\《DWR-DLF设计-工单-V1.0.xlsx》。

## 客户服务主题

### 功能

主要完成企业服务主题的相关数据汇总归档存储，及数据汇总归档存储时的作业调度、数据MAPPING等工作。

### 数据Mapping

详见：…\详细设计\主题层\客户服务\ 03Mapping\《DWR-客户服务-MAPPING映射-V1.0.xlsx》。

### DLF设计

详见：…\详细设计\主题层\客户服务\DLF设计\《DWR-DLF设计-客户服务-V1.0.xlsx》。

## 人员主题

### 功能

主要完成人员主题的相关数据汇总归档存储，及数据汇总归档存储时的作业调度、数据MAPPING等工作。

### 数据Mapping

详见：…\详细设计\主题层\人员\ 03Mapping\《DWR-人员-MAPPING映射-V1.0.xlsx》。

### DLF设计

详见：…\详细设计\主题层\人员\DLF设计\《DWR-DLF设计-人员-V1.0.xlsx》。

## 设备主题

### 功能

主要完成设备主题的相关数据汇总归档存储，及数据汇总归档存储时的作业调度、数据MAPPING等工作。

### 数据Mapping

详见：…\详细设计\主题层\设备\ 03Mapping\《DWR-设备-MAPPING映射-V1.0.xlsx》。

### DLF设计

详见：…\详细设计\主题层\设备\DLF设计\《DWR-DLF设计-设备-V1.0.xlsx》。

## 车辆主题

### 功能

主要完成时空主题的相关数据汇总归档存储，及数据汇总归档存储时的作业调度、数据MAPPING等工作。

### 数据Mapping

详见：…\详细设计\主题层\车辆\ 03Mapping\《DWR-车辆-MAPPING映射-V1.0.xlsx》。

### DLF设计

详见：…\详细设计\主题层\车辆\DLF设计\《DWR-DLF设计-车辆-V1.0.xlsx》。

## 资产主题

### 功能

主要完成资产主题的相关数据汇总归档存储，及数据汇总归档存储时的作业调度、数据MAPPING等工作。

### 数据Mapping

详见：…\详细设计\主题层\资产\ 03Mapping\《DWR-资产-MAPPING映射-V1.0.xlsx》。

### DLF设计

详见：…\详细设计\主题层\资产\DLF设计\《DWR-DLF设计-资产-V1.0.xlsx》。

## 组织主题

### 功能

主要完成组织主题的相关数据汇总归档存储，及数据汇总归档存储时的作业调度、数据MAPPING等工作。

### 数据Mapping

详见：…\详细设计\主题层\组织\03Mapping\《DWR-组织-MAPPING映射-V1.0.xlsx》。

### DLF设计

详见：…\详细设计\主题层\组织\ DLF设计\《DWR-DLF设计-组织-V1.0.xlsx》。

## 测试计划

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **时间点** | **测试类型** | **测试轮次** | **输入数据** | **预期结果** | **测试人员** |
| 开发完成 | 单元测试-自测 | 1 | 主题层表结构模拟数据 | 1、数据归集成功； 2、数据格式、类型、值正确 3、DLF任务执行成功 | 开发人员 |
| 单元测试-自测 | 2 | 主题层表结构模拟数据 | 同上 | 开发人员 |
| 单元测试-自测 | 3 | 主题层表结构模拟数据 | 同上 | 开发人员 |
| 单元测试 | 1 | 主题层表结构模拟数据 | 同上 | 测试人员 |
| 单元测试 | 2 | 主题层表结构模拟数据 | 同上 | 测试人员 |
| 单元测试 | 3 | 主题层表结构模拟数据 | 同上 | 测试人员 |
| 集成测试阶段 | 集成测试 | 1 | 主题层表结构真实业务数据 | 1. 正确连接源系统； 2. 数据归集成功； 3. 数据格式、类型、值正确 4. 前置贴源层数据处理成功； 5. DLF任务执行成功； 6. 按作业策略进行正确的数据处理。 | 开发人员 测试人员 |
| 集成测试 | 2 | 主题层表结构真实业务数据 | 同上 | 开发人员 测试人员 |
| 集成测试 | 3 | 主题层表结构真实业务数据 | 同上 | 开发人员 测试人员 |
| 现场联调阶段 | 现场联调 | 1 | 主题层表结构真实业务数据 | 1. 正确连接源系统； 2. 数据归集成功； 3. 数据格式、类型、值正确 4. 前置贴源层数据处理成功； 5. DLF任务执行成功； 6. 按作业策略进行正确的数据处理。 | 开发人员 测试人员 |
| 现场联调 | 2 | 主题层表结构真实业务数据 | 同上 | 开发人员 测试人员 |
| 现场联调 | 3 | 主题层表结构真实业务数据 | 同上 | 开发人员 测试人员 |

# 数据服务设计

按照初期项目架构原则，大数据平台中主要对外开放的是主题层数据库，其面对下游的业务系统直接访问或下游业务系统利用大数据平台提供的数据服务能力实现专题库的构建。

* **支持的下游业务系统**

1. 园区管理平台
2. 园区服务平台
3. 环保监测系统
4. 智慧消防系统
5. 智慧安防系统
6. 国有资产管理系统
7. 物联网平台
8. IOC平台

* **数据服务方式**

1. 专题库数据库直接访问

适用业务场景：业务系统在大数据平台构建专题库，需要通过直接访问库表的方式获取数据，主要适用于IOC平台和园区管理平台的数据访问。

大数据平台主题库，开放可访问数据视图列表、数据属性说明、数据访问链接供业务系统专题库直接登录访问现有主题库，获取数据。主题库负责进行用户管理和权限控制。

1. 数据集成服务

适用业务场景：业务系统无法/不愿直接访问数据库表方式获取数据时适用，适合数据交互量较小的场景。

大数据平台统一构建公共专题库，主要包括企业基本信息，基于华为ROMA平台实现，通过RestAPI方式实现对外提供数据服务。

# 调度服务设计

## 现有作业分析

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 处理周期 | | | | | |
| 实时 | 15分钟 | 1小时 | 日 | 月 | 合计 |
| FDI任务 | 4 | 0 | 7 | 216 | 6 | **233** |
| DLF任务 | 0 | 1 | 7 | 115 | 0 | **123** |

现有FDI任务233个，DLF任务123个。

其中，月任务类同按日调度处理，FDI任务日调度占比为：95.27%；DLF任务日调度占比为：93.49%。

## 作业处理时间设计

作业平均处理时间预估

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 记录数 | FDI测试数据(S) | DLF测试数据(S) | FDI设计数据（S） | DLF设计数据（S） |
| 1000 | 7.2 | 11 | 30 | 60 |
| 10000 | 17.7 | 31 | 60 | 120 |
| 100000 | 72 | 25 | 180 | 180 |

测试环境描述：北京MySQL数据库--华为公有云GuassDB（数据仅供参考）

测试作业描述：入湖FDI任务，直接一对一Mapping入湖；DLF作业直接从湖表抽取加载到主题表。

源、目标表属性：58个字段，空表，表宽：4179Byte。

设计数据：根据园区现有业务特点、充分考虑任务调度时间影响、冗余处理时间、未来业务扩展。

综合现有源系统、入湖层设计、主题层设计，日增量数据在1000以内，因此FDI任务处理时间设计为30秒，DLF任务设计处理时间为60秒。

## 调度依赖关系描述

从整个数据流向进行切分，大数据平台涉及到入湖、主题和专题3个逻辑层，其中入湖对应ROMA平台的FDI任务，主题和专题对应DAYU平台的DLF任务。

限制：ROMA平台和DAYU平台没有对接，且无法利用ROMA平台提供的系统变量、API接口或其他方式实现任务标识设置（或事件消息的发送），因此，主题无法实现与入湖层进行强依赖关系建立。

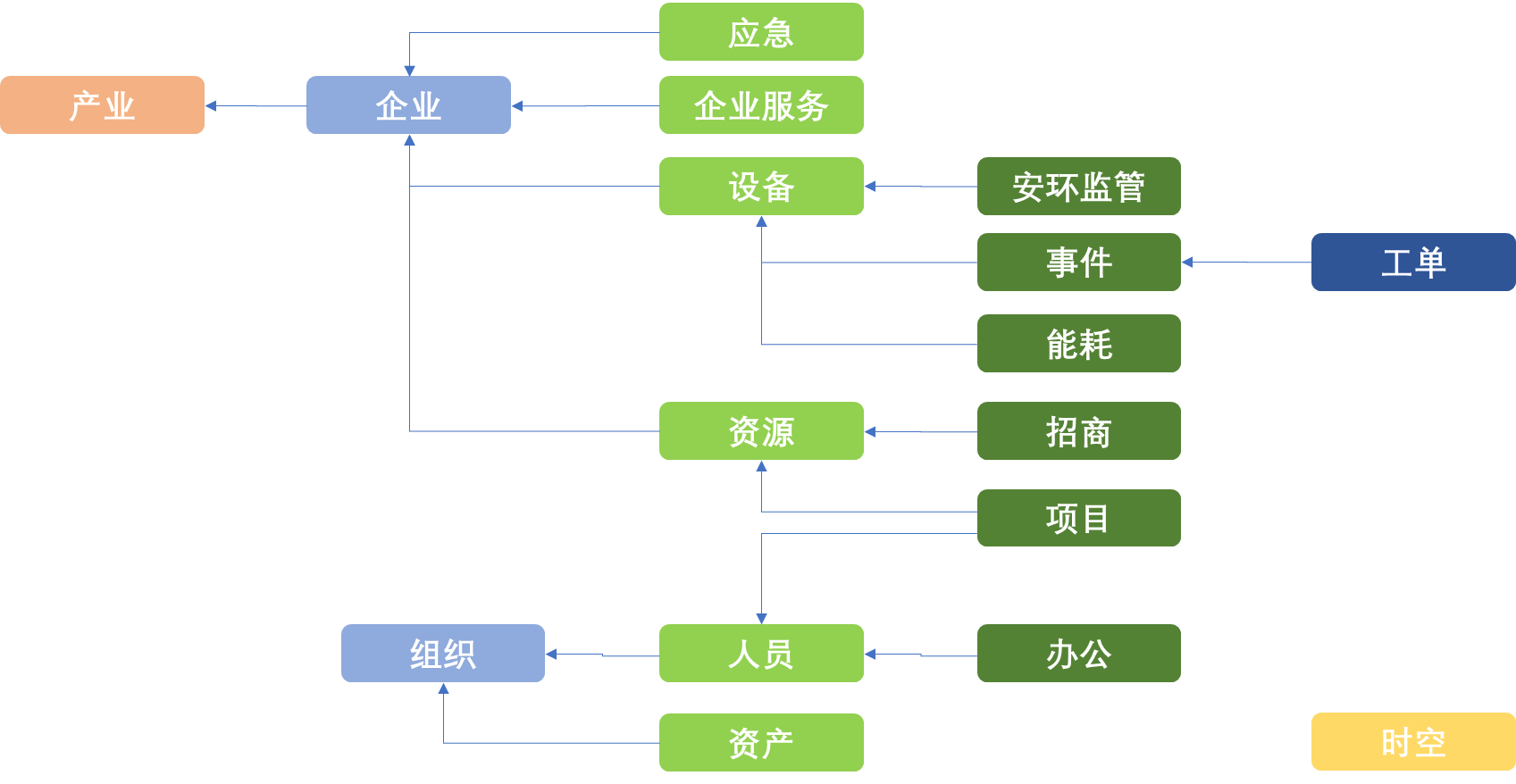
因此，在本设计中，尽量从作业任务调度时间设置上，保障当日FDI任务处理完成后，再启动DLF任务的调度工作。

局限性：无法通过程序逻辑控制数据处理顺序。

风险点：数据时效性无法保障。

专题库调度注意事项：需要在主题库任务调度执行完成后启动。

主题域中主题依赖关系如下图所示：



## 调度组织规划

**调度周期：实时任务**，事件驱动触发任务调度执行。

**调度周期：小时**，FDI任务设置后整点触发，后续关联DLF任务延后30分钟执行。

**调度周期：日**（含月）。

主题任务依赖调度设计如下表所示：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **主题名称** | **依赖主题** | **任务数量** | **计划执行批次** | **合计任务数** |
| 产业 |  | 1 | 1 | 17 |
| 组织 |  | 1 | 1 |
| 企业 | 产业 | 15 | 1 |
| 人员 | 组织 | 1 | 2 | 19 |
| 设备 | 企业 | 10 | 2 |
| 资源 | 企业 | 8 | 2 |
| 企业服务 | 企业 | 9 | 3 | 21 |
| 应急 | 企业 | 7 | 3 |
| 资产 | 组织 | 5 | 3 |
| 事件 | 设备 | 4 | 4 | 21 |
| 项目 | 企业、资源、人员 | 6 | 4 |
| 办公 | 人员、组织 | 11 | 4 |
| 安环 | 企业、设备 | 19 | 5 | 23 |
| 工单 | 设备、事件 | 4 | 5 |
| 能耗 | 设备 | 9 | 6 | 22 |
| 招商 | 企业、资源 | 12 | 6 |
| 时空 |  | 1 | 6 |

FDI任务（入湖）和DLF任务（主题）调度执行详细规划如下表所示：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **时间点** | **任务类型** | **入湖源系统/主题名称** | **任务数** | **任务汇总** |
| 00:00 | 入湖FDI任务 | 国有资源 | 11 | 33 |
| 能耗 | 7 |
| 安防 | 15 |
| 00:10 | 环保 | 23 | 30 |
| 消防 | 1 |
| 能耗（月度） | 6 |
| 00:20 | 园区服务 | 46 | 46 |
| 00:30 | 园区管理（飞企-办公） | 20 | 47 |
| 园区管理（飞企-招商、经济） | 27 |
| 00:40 | 园区管理（普天） | 33 | 33 |
| 00:50 | 园区管理（普天） | 33 | 33 |
| 1:00-1:59 | 缓冲（运维监控、错误处理） | | | |
| 02:00 | 主题DLF任务 | 产业 | 1 | 17 |
| 组织 | 1 |
| 企业（应在产业主题后启动） | 15 |
| 02:10 | 人员 | 1 | 19 |
| 设备 | 10 |
| 资源 | 8 |
| 02:20 | 企业服务 | 9 | 21 |
| 应急 | 7 |
| 资产 | 5 |
| 02:30 | 事件 | 4 | 21 |
| 项目 | 6 |
| 办公 | 11 |
| 02:40 | 安环 | 19 | 23 |
| 工单 | 4 |
| 02:50 | 能耗 | 9 | 22 |
| 招商 | 12 |
| 时空 | 1 |
| 3:00-3:59 | 缓冲（运维监控、错误处理） | | | |
| 4:00-5:00 | 专题DLF任务 | | | |
| 5:00-6:00 | 数据备份窗口 | | | |
| 8:00-23:59 | 业务系统交易时间窗口 | | | |

**并发需求**：50。

**设计容量**：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **处理情况**  **任务类型** | **10分钟节点处理分析** | | | **日节点处理分析** | | | **是否满足调度需求** |
| **需求** | **最小值** | **最大值** | **需求** | **最小值** | **最大值** |
| FDI任务 | 50 | 100 | 2000 | 300 | 600 | 12000 | 是 |
| DLF任务 | 50 | 100 | 1000 | 200 | 600 | 6000 | 是 |
| 备注：FD作业任务处理时间为30秒，DLF作业任务处理时间为60秒，失败重置3次，间隔60秒 | | | | | | | |
|  |

## 调度配置

调度配置时，需明确作业依赖关系、依赖作业失败处理策略、跨周期依赖策略、调度方式、调度周期、启动时间、是否失败重试、最大执行次数、重试间隔时间、失败策略、节点执行的最长时间（分钟、小时、天）、节点状态轮询时间（秒）等要素。

在“调度组织规划”章节中描述了贴源层到主题层，以及主题层内部主题与主题的依赖关系和调度时间节点划分，但主题内，如涉及数据处理的先后顺序，也请定义清楚作业间的依赖关系。

调度配置详情请参见贴源层的“FDI设计”和主题层的“DLF文档”中的详细描述。

## 容错处理

为应对网络闪断、数据库锁表等待等可能导致作业处理失败风险，加强平台自动化运维能力建设，需对作业及作业的调度进行容错处理设计。

1. 作业本身需支持手工和自动重做；
2. 作业调度时，需定义并配置“依赖作业失败处理策略、是否失败重试、最大执行次数、重试间隔时间、失败策略”等内容。
3. ROMA平台和DAYU平台提供作业运维监控功能。

# 集成服务设计

## 应用ID设计

|  |  |  |
| --- | --- | --- |
| **应用ID** | **应用名称** | **备注** |
| iot.device.ecm | ROMA1\_能耗监测 | 水表、电表设备上报数据 |
| iot.device.inf | ROMA1\_智慧消防 | 烟感设备上报数据 |
| data.dep | ROMA2共享交换平台 | 从共享交换平台查询API获取数据 |
| iot.device.ecm | ROMA2\_能耗监测 | 水表、电表设备上报数据 |
| iot.device.inf | ROMA2\_智慧消防 | 烟感设备上报数据 |

## 业务开发APIC设计

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **序号** | **任务号** | **源系统** | **API名称** | **API相对路径** | **HTTP方法** | **返回数据类型（json、xml）** | **源接口相对对路径** | **适配说明** | **备注** |
| 1 | ROMA2-01 | 共享交换平台 | 共享交换平台登录 | /v1/dep/login | POST | json |  | 从共享交换平台登录 |  |
| 2 | ROMA2-02 | 共享交换平台 | 从共享交换平台查询API获取数据 | /v1/dep/data | POST | json |  | 从共享交换平台查询API获取数据 |  |

## MQS设计

详见：…\详细设计\《ROMA集成深化设计-V1.0.xlsx》。

## 测试计划

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **时间点** | **测试类型** | **测试轮次** | **输入数据** | **预期结果** | **测试人员** |
| 开发完成 | 单元测试-自测 | 1 | 结构模拟数据 | 1、API调用成功/MQS产生消费成功； 2、数据返回格式、类型、值正确 | 开发人员 |
| 单元测试-自测 | 2 | 结构模拟数据 | 同上 | 开发人员 |
| 单元测试-自测 | 3 | 结构模拟数据 | 同上 | 开发人员 |
| 单元测试 | 1 | 结构模拟数据 | 同上 | 测试人员 |
| 单元测试 | 2 | 结构模拟数据 | 同上 | 测试人员 |
| 单元测试 | 3 | 结构模拟数据 | 同上 | 测试人员 |
| 集成测试阶段 | 集成测试 | 1 | 结构模拟数据 | 1、API调用成功/MQS产生消费成功； | 开发人员 测试人员 |
| 集成测试 | 2 | 结构模拟数据 | 同上 | 开发人员 测试人员 |
| 集成测试 | 3 | 结构模拟数据 | 同上 | 开发人员 测试人员 |
| 现场联调阶段 | 现场联调 | 1 | 真实业务数据 | 1、API调用成功/MQS产生消费成功； | 开发人员 测试人员 |
| 现场联调 | 2 | 真实业务数据 | 同上 | 开发人员 测试人员 |
| 现场联调 | 3 | 真实业务数据 | 同上 | 开发人员 测试人员 |

# 数据管理设计

## 数据存储

GuassDB提供数据分布式存储、数据分区、列压缩存储等功能。本项目将利用GuassDB的这些数据存储特性，在数据库安装部署规划以及物理表设计时分别进行考虑。

对所有表进行切片分区设计，设计规则为按照表内存储数据类型：流水数据为增量切片分区，非流水数据为全量切片分区。详情见《主题表切片分区存储实现方案\_V1.0》

在大数据平台贴源层和主题层物理表设计中，明确定义每张表的存储方式：

* 是否建切片
* 是否构建历史表
* 存储类型：当前、拉链、增量切片、全量切片
* 存储周期：主档类永久保存、明细档类(保存一定周期进入历史表)、历史档类永久保存
* 分区键
* 分布键
* 行列存储模式

详细设计内容可参见物理表设计中的定义。

## 数据清理

对于平台使用的临时表、数据量较大而需要转存历史表的明细档、超过存储周期的历史表，大数据平台应提供数据自动清理能力。

临时表：由创建者负责数据清理。

需要转存的明细档：由处理该数据表的作业自行完成。

超过存储周期的历史表：平台提供独立的数据清理作业完成。

## 数据备份与恢复

为应对数据丢失或损坏对用户业务造成不利影响，在异常情况下快速恢复系统，GuassDB根据用户业务的需要提供全量备份、增量备份和恢复功能。

* 支持根据备份策略实现周期备份。
* 支持自动备份和手动备份能力。
* 支持集群、库级、表级等不同组件备份。
* 支持全量备份和增量备份。
* 提供数据快速恢复能力。

建议备份策略：日增量、周全量数据库级自动备份，备份时间窗口为2:00-6:00。

## 数据安全性

大数据平台数据的安全性仅限于应用安全，其它安全内容依赖于园区整体安全体系的规划和建设。从应用安全方面，依赖于GuassDB本身提供的安全管控能力实现，包括：

* 身份鉴别和认证
* Web应用安全
* 访问控制
* 审计安全
* 密码安全

应用专题库的建设包括IOC平台、园区管理平台等业务系统，需要在GuassDB上分别为其构建独立的访问用户，对主题层数据仅开放查看（即Select）权限，以便于数据访问与审计。下游应用系统已获取数据的安全性，由下游应用系统进行控制。

大数据平台获取各源系统的数据进行集成和验证性测试，源系统在提供数据时需进行脱敏处理，大数据平台获取的数据严禁进行本地保存、下载、转发导致数据泄密。

# 待解决问题

1. 各个业务系统相关数据信息。

# 附录：数据资产一览表

详见:…\详细设计\《数据资产清单V1.0》。