

Contents

## 目录

01

背景概述 (06P)

02

关键挑战 (04P)

03

详细介绍 (23P)

04

整体总结 (01P)

# 数据处理技术演进与行业趋势





# 数据资产定义：一切皆数据，数据是一切

## 哪些是数据资产？

各业务域和数据域在数据生产消费过程中形成的数据沉淀及其载体都是数据资产。

**广义概念** 企业生产经营中产生的信息及其载体都属于数据资产。

**狭义概念** 可确定归属权，且能被交易、使用的数据。

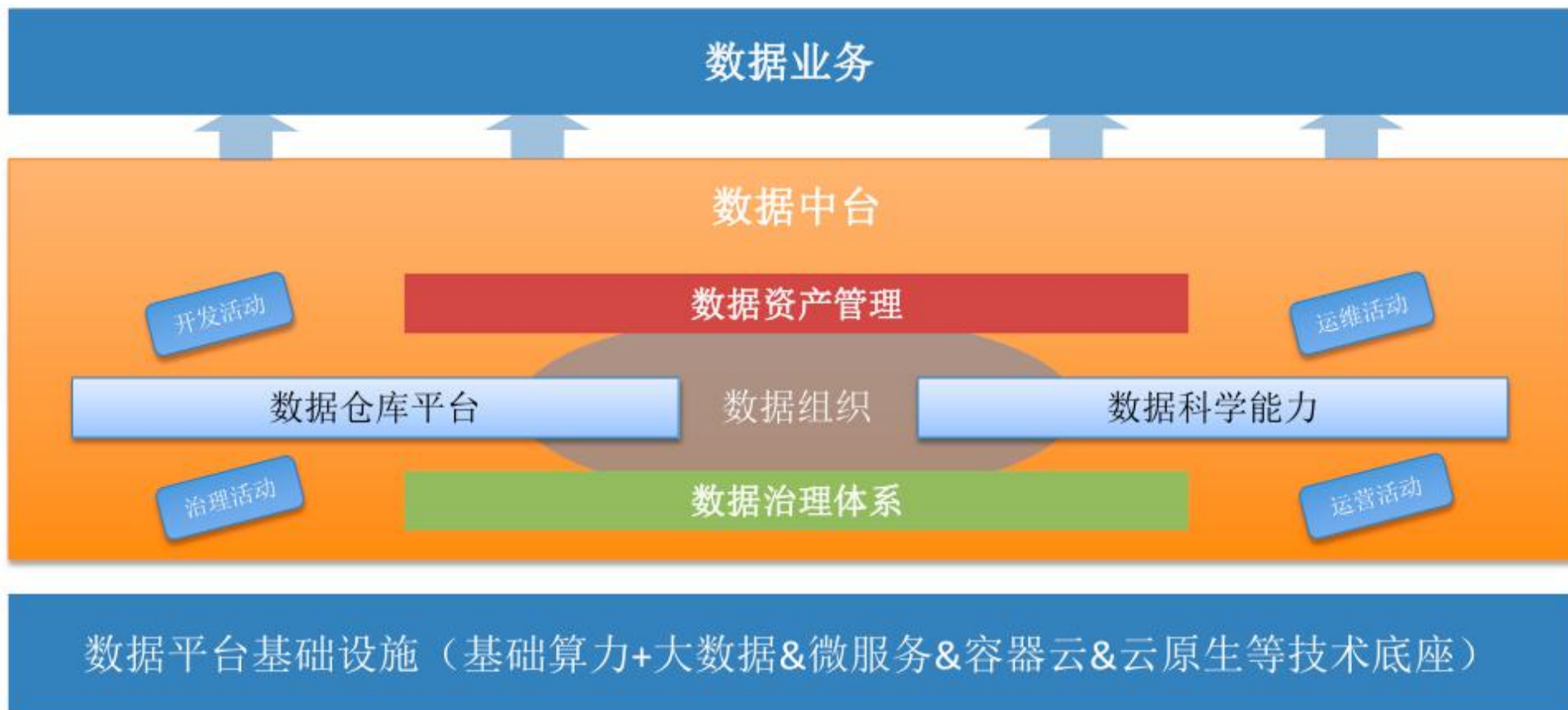


# 大数据资产管理平台边界：

A?

B?

C?



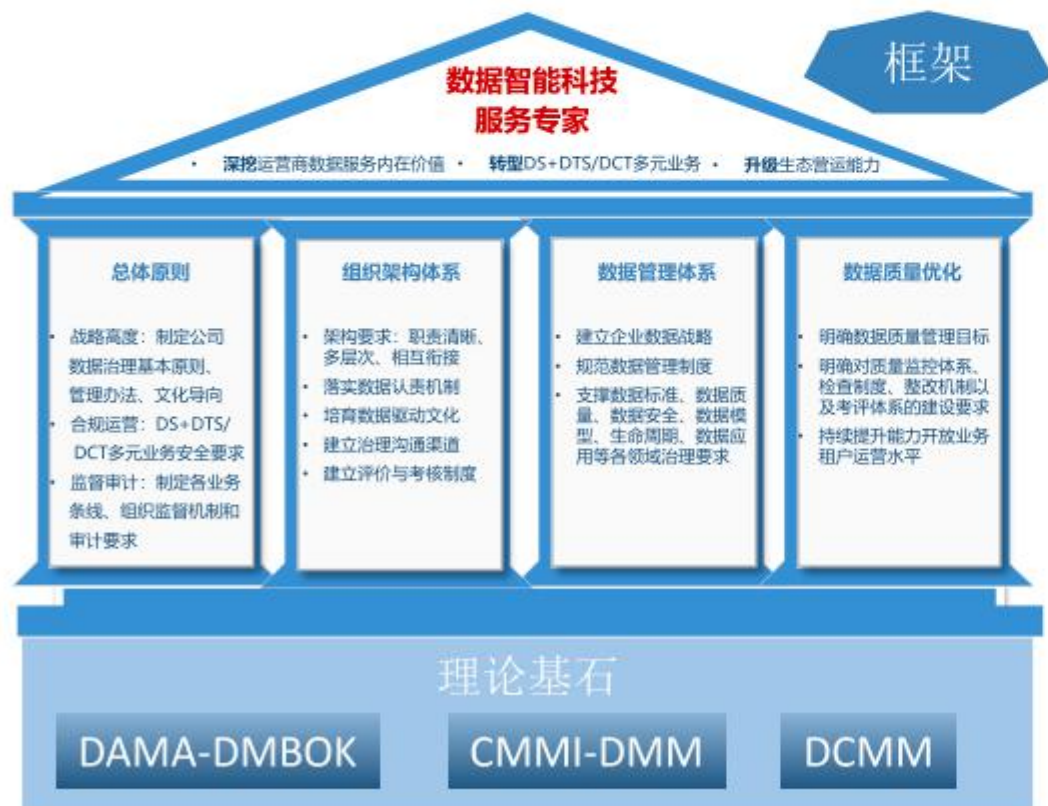
数据资产管理 VS 数据治理 VS DataOps ?

数据中台落地实践过程中的加速器，  
数据平台开发运维和数据治理活动实施的润滑剂，  
数据科学建模分析和数据资产运营管理价值化的催化剂。

# 关于数据资产管理整体框架与方法论

如何制定数据战略？ 收益、安全、质量、架构、流程、效率？

总体方针：以服务公司愿景为导向，以问题为切入点，切实解决业务问题。



# 联通大数据资产管理平台能力架构-全景图



# 联通大数据资产管理平台价值地图



Contents

## 目录

01

背景概述 (06P)

02

关键挑战 (04P)

03

详细介绍 (23P)

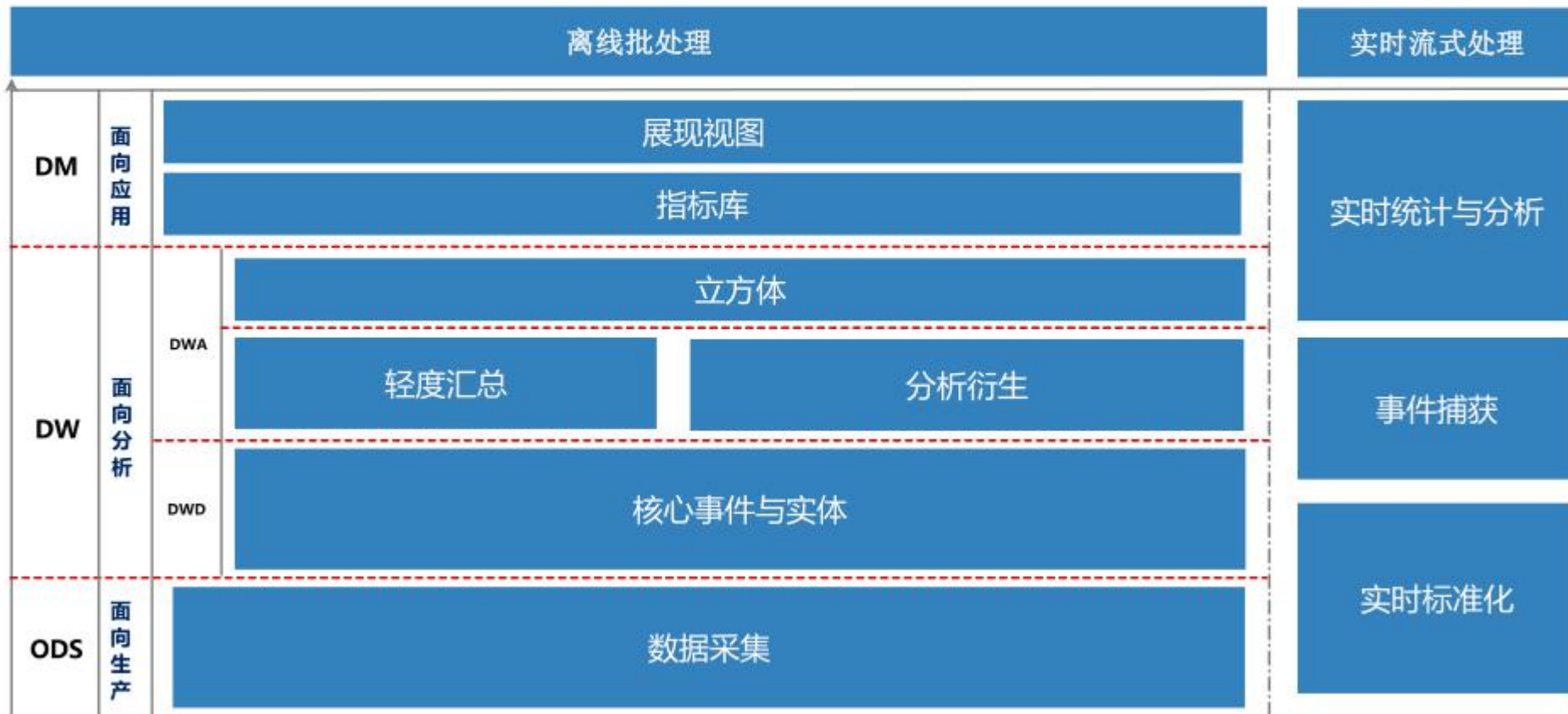
04

整体总结 (01P)



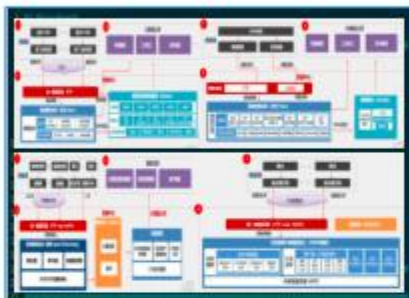
# 数据资产管理平台关键挑战#1 搭建数据仓库统一平台

■ 围绕批量与流式场景分离、业务规范统一、空间效率平衡等原则构建分层分类大数据模型体系



# 数据资产管理平台关键挑战#2 实践数据平台治理管控

## 数据采集链路监控



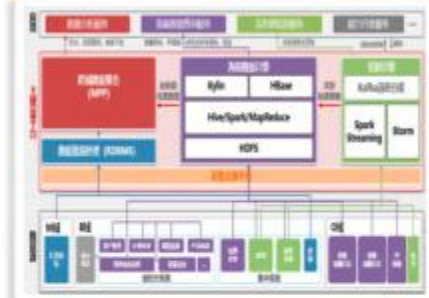
## 数据跨集群同步



## 数据生命周期管理

| 数据源 | 数据类型     | 生命周期  |
|-----|----------|-------|
| 日域  | 平账数据 (日) | X+1个月 |
| 日域  | 平账数据 (月) | X+1个月 |
| 日域  | 加工数据 (日) | X+1个月 |
| 日域  | 加工数据 (月) | XX+1月 |
| 日域  | 平账数据     | X-3天  |
| 日域  | 汇总/分析数据  | X月    |
| 日域  | 平账数据     | X天    |
| 日域  | 标准化解析数据  | X天    |
| 日域  | 汇总/分析数据  | XX天   |
| 日域  | 用户标签数据   | X+1个月 |
| 日域  | 营销推广数据   | X-X个月 |
| 日域  | 平账数据     | X天    |
| 日域  | 汇总/分析数据  | X个月   |
| 日域  | 用户标签数据   | X+月   |

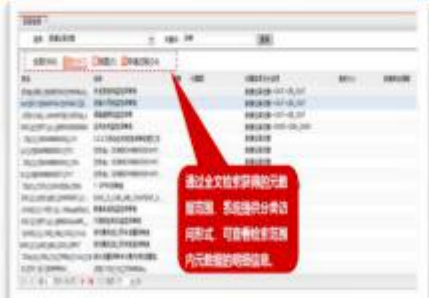
## 数据开发&运维



## 数据ETL调度系统



## 元数据管理系统



## 数据血缘分析



## 数据质量保障



元数据&数据模型&数据元&数据标准概念对齐

通用的商业化元数据管理、数据质量工具的问题

# 数据资产管理平台关键挑战#2 实践数据平台治理管控

## 数据脱敏系统

- 实现敏感数据的定义和识别,如手机号、身份证号等敏感,对数据进行分类分级管理,实施相应的访问控制。
- 支持流式数据动态脱敏,支持敏感文件的脱敏识别。
- 脱敏算法支持敏感信息化、掩码、格式保持、加密脱敏等多种方案。

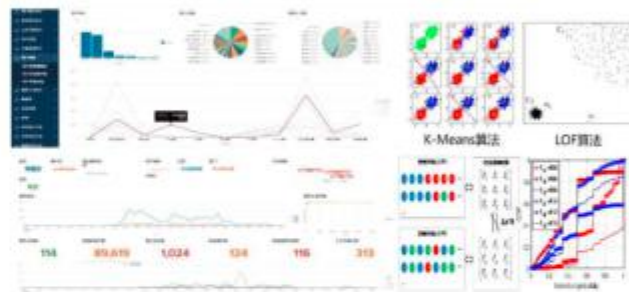


## 数据安全监测与审计系统

- 用户操作行为为核心,通过全网日志提取、关联分析,实现大数据平台的数据资产安全审计。
- 基于零信任安全模型,以动态策略管理为核心的安全审计。



## 行为基线和用户画像系统



## 数据追踪溯源系统

- 自主研发,基于区块链技术和数字水印技术,针对大数据特性,对大数据内容进行标识,一旦发生数据泄露,能够快速追溯并溯源到具体的存储位置。
- 基于内容溯源,支持Word、Excel文本,数据导出文件多种格式,不需要安装客户端软件,不依赖中间平台系统,数据可自由流转,审计信息实时同步,数据导出数据流转过程中,溯源可以成功溯源。



## 数据出口管控系统

- 作为对外数据出口的第一道门,集中管控,集中审计,确保对外数据出口的安全性。
- 实现数据跨境输出审计,数据输出监控与策略自动控制,日志管理,权限管理,安全管理,数据缓存,多租户管理,安全策略配置等功能。
- 针对敏感数据的智能识别策略,如手机号、身份证号、银行卡号等敏感信息的识别,防止敏感数据的外泄。

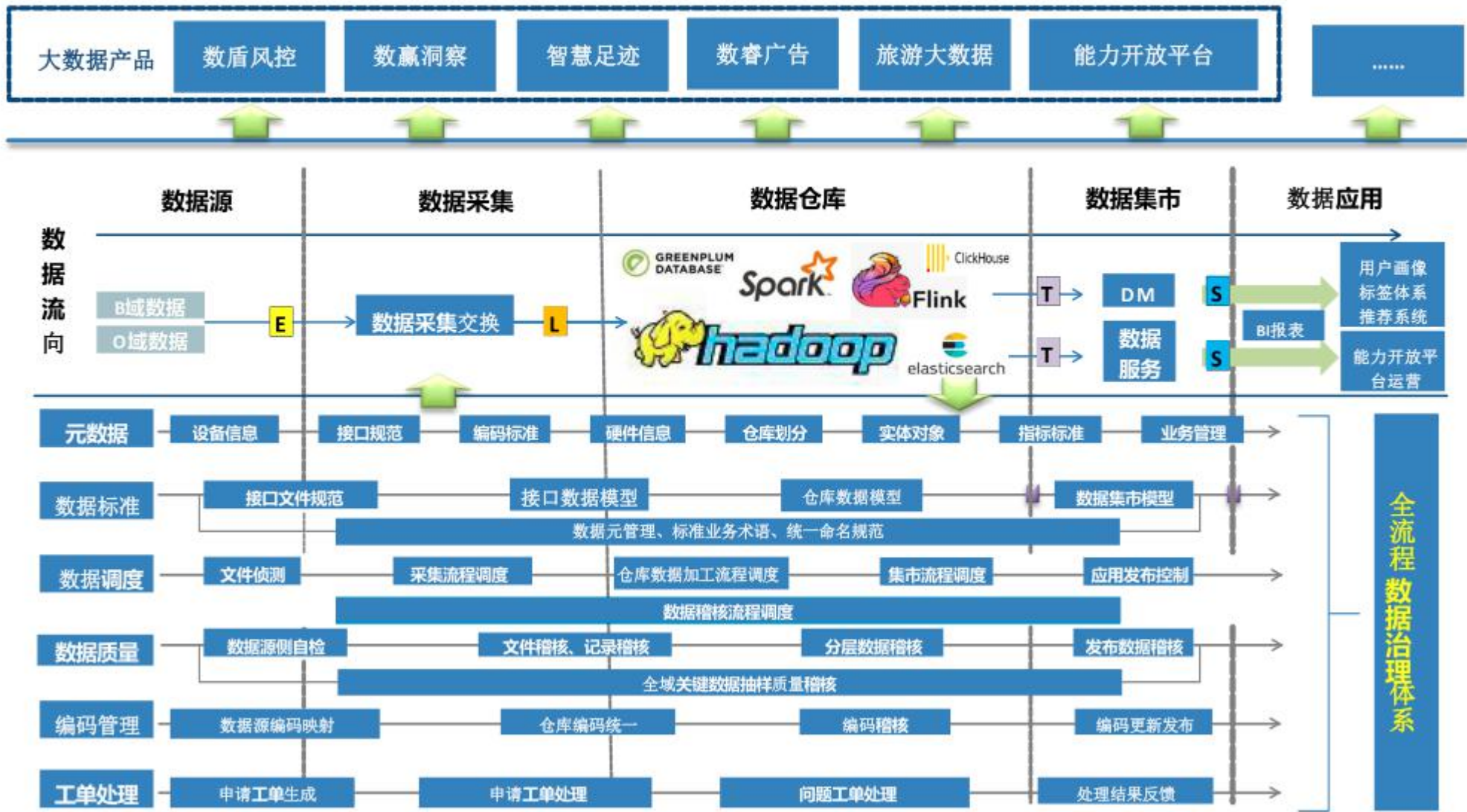


## 平台统一访问控制和审计系统

- 对平台及系统的所有服务器、数据库等资产进行集中管理,通过建立统一访问控制策略系统统一用户身份统一认证、授权。
- 建立“敏感信息风险审计”机制,针对特殊人员的9A4操作风险行为、Hadoop信息操作行为进行监控,并发布风险告警信息。



# 数据资产管理平台关键挑战#3 探索数据资产运营体系



# 目录

01

背景概述 (06P)

02

关键挑战 (04P)

03

详细介绍 (23P)

04

整体总结 (01P)

# 数据资产管理平台基本概述

## 定位与目标:

基于数据驱动的理念，让数据资产开发、运维、治理、运营过程更安全、敏捷、精益、自动化、服务化和智能化。

1. 借鉴DevOps持续集成与交付方法论，建立数据治理基础框架，实现数据采集、加工、运维、服务过程一站式、体系化、规范化、透明化的流水线管理模式，消除数据生命周期管理、数据质量管理、数据安全等方面的隐患。

2. 基于大数据集群健康评分机制，实现数据平台降本增效，从数据计算、存储、调度、分析挖掘等各个方面分析程序效能、数据资源冗余画像，确保数据平台、模型和应用的快速部署与开发，整合各类数据处理框架，为上层数据分析和智能化应用提供高效的算力和算法支持。

3. 实现数据驱动业务，建立数据资产运营体系，让数据资产开发、运维、治理、能力开放运营过程更加安全、敏捷、精益、服务化和智能化。

平台定位：实现一站式端到端的数据资产开发与标准化管控治理。



# 模块1：数据集成平台功能需求

## 定位与目标：

把企业内外部数据快速整合到一起，提供大数据平台与生产应用系统的双向通信能力，方便构建数据开发、运维、治理、运营闭环系统。

核心模块1：数据采集交换平台

核心模块2：数据工作流程调度

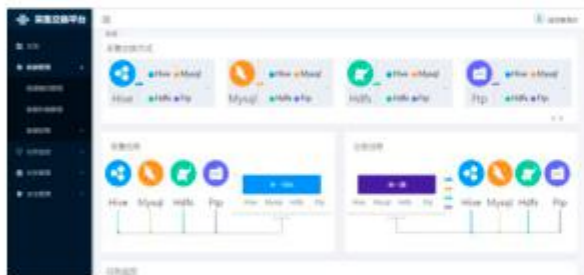
核心模块3：数据应用代理程序



# 模块1：数据集成平台产品设计

基于运营商领域多年数据平台建设经验

跨地域多源异构数据集成



图形化的数据流设计器



可视化数据转换功能



跨系统跨平台统一调度



自动化任务调度机制



节点级全方位实时监控





# 模块1：数据集成平台研发实践

产品设计

技术研发

运营实践

## ✓ 高度分散的异构的数据源

数据大量分散在企业的不同业务系统、数据库、甚至企业外部的第三方系统中；数据源类型、结构、模式不尽相同，必须经过采集、清洗与标准化才能进入数据仓库。

## ✓ ETL调度流程编排问题 (运营商 VS 互联网)

数据处理的流程大量依赖各种脚本程序，难以理解与修改维护；数据处理流程经常无法复用，缺乏统一管理；因为数据**断传、漏传、补传**造成的**数据重跑**问题突出。

## ✓ 开发脚本的复杂性 (如接口机+Shell模式)

数据加工过程一般通过执行复杂冗长晦涩的脚本来完成，要求开发人员必须有较高的专业技能；数据加工过程的逻辑错误、语法错误也不容易捕捉；集群作业提交参数的合理性问题。

## ✓ 飞速增长的数据量和非结构化数据类型

随着5G+物联网场景超大规模数据的输入；数据持续不断的到达，数据集成应当具备PB级实时或准实时数据处理能力；需要支持结构化、半结构化、非结构化等不同数据类型。

易用性 VS 功能完备性

稳定性 与 兼容性

数据平台的基础模块

# 模块2：数据开发平台功能需求

## 定位与目标：

支撑数据模型设计、数据加工脚本标准化，实现开发、测试、上线过程自动化，保障数据模型与生命周期管理标准落地。

核心模块1：数据加工模型设计

核心模块2：数据脚本过程管理

核心模块3：数据生命周期管理



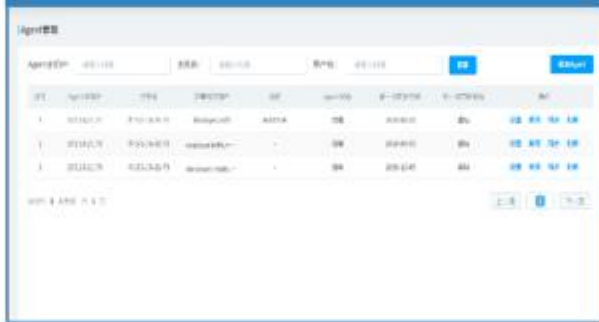
# 模块2：数据开发平台产品设计

基于企业数据治理标准化需求，形成数据加工脚本从生成、测试、审批、到部署的全流程管控

## 标准化脚本模版规则



## 统一的脚本执行机器管控



## 开发脚本标准化创建



## 脚本运行测试



## 审批环节



## 自动化部署上线



## 模块2：数据开发平台运营实践

产品设计

技术研发

运营实践

### ✓ 数据开发平台与数据治理 (从源头治理)

数据开发平台支撑数据治理文化落地，数据治理过程强调组织、文化、工具、流程的全方位协同，数据开发平台仅仅是数据治理工具体系的一部分。

### ✓ IT墙、组织墙问题 (推广成本)

数据开发平台研发背景往往跟生产环境实际痛点相关，涉及组织较多，在向其他部门或项目组推广时，因不同组织绩效目标差异，通常会遇到不同程度的IT墙和部门组织墙问题。

### ✓ 产线环境安全与便捷实用性的博弈（安全）

为了保障产线环境下数据加工脚本执行的安全性，平台需要覆盖脚本模板配置、脚本创建、审核、测试、部署上线的完整流程，整体使用复杂度相对提升，需要配套的运营流程。

### ✓ 从脚本标准化自动化到在线IDE（一体化）

企业各开发团队的技术栈和开发习惯差异较大；平台脚本模板很难兼容所有团队的灵活需求；在线IDE需兼顾数据治理标准落地和个性化需求开发的要求。

# 模块3：元数据管理平台功能需求

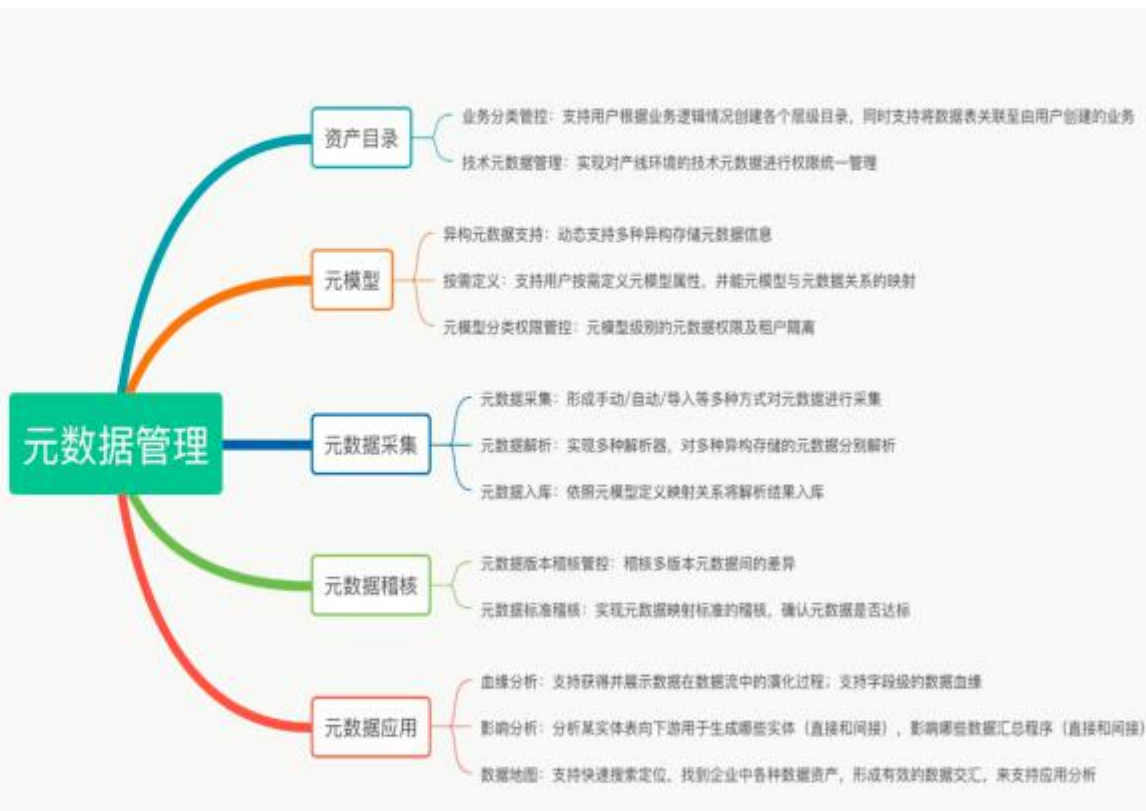
## 定位与目标：

基于技术元数据、业务元数据和管理元数据的采集与分析，实现数据血缘、影响分析和全链分析，解决企业内部数据资产统一盘点和运营问题。

核心模块1：元数据采集

核心模块2：元数据分析

核心模块3：元数据应用



# 模块3：元数据管理平台产品设计

元数据管理功能规划和原则、元数据管理的统一方法

## 资产发布



## 多种数据源管控



## 元数据变更稽核



## 数据血缘



## 数据地图



## 元数据获取



# 模块3：元数据管理平台运营实践

产品设计

技术研发

运营实践

## ✓ 多种元数据采集方案的抉择 (Metadata API & 公有云平台的坑)

针对不同数据源/集群的采集有多种方案，站在交付目标、无侵入性的角度考虑，综合权衡好安全性、性能和扩展性要求。

## ✓ 数据血缘分析准确性、完整性、实用性

由于企业数仓存储介质、加工方式、调度手段多样，在采集多种元数据后，整合血缘分析的困难度较高，建议定向对指定场景进行血缘分析。(Hive→HBase: 跨库跨集群跨源全链路端到端分析)

## ✓ 元数据管理应用作为内部推广抓手 (价值驱动)

从数据资产目录、数据视图、元数据检索、元数据稽核、数据地图、数据血缘、影响分析、全链分析、活性分析、数据价值图谱等应用方向寻找内部推广应用的突破口。

## ✓ 元数据稽核与数据标准、数据质量协同

元模型管理、元数据属性填充率、贯标落地统计、生产环境最新版本与资产管理平台、测试环境双向稽核。  
(业务&IT协同: 自治文化)

## 模块4：数据质量平台功能需求

### 定位与目标：

针对数据进行稽核来确保数据的质量，覆盖及时性、完整性、准确性、一致性、唯一性及合理性等，及各系统之间数据的统一性。建立标准化度量系统，方便系统性改进质量问题。

核心模块1：数据源、数据对象、元数据分类管理

核心模块2：数据质量检测模型、方案、规则管理

核心模块3：数据质检任务调度、报告、流程管理





# 模块4：数据质量平台产品设计

## 数据质量模型管理

The screenshot displays a table with columns for '名称' (Name), '模型类型' (Model Type), '模型描述' (Model Description), '业务系统名称' (Business System Name), '责任人' (Responsible Person), and '创建时间' (Creation Time). The table lists several data quality models.

## 数据质量方案管理

The screenshot shows a form titled '数据质量方案管理 - 新增质检方案' (Data Quality Scheme Management - Add Quality Check Scheme). It includes a dropdown menu for '数据源名称' (Data Source Name) and a text input field for '方案名称' (Scheme Name).

## 数据质量规则管理

The screenshot displays a table with columns for '名称' (Name), '规则类型' (Rule Type), '规则描述' (Rule Description), '业务系统名称' (Business System Name), '责任人' (Responsible Person), and '创建时间' (Creation Time). The table lists several data quality rules.

## 数据源质检对象管理

The screenshot displays a table with columns for '名称' (Name), '数据源名称' (Data Source Name), '模型名称' (Model Name), and '模型类型' (Model Type). A modal window is open, showing details for 't\_text\_table\_1' data source model.

## 数据质检SQL看板

The screenshot shows a dashboard for '数据质检SQL看板' (Data Quality SQL Dashboard). It includes a search bar, a dropdown for '数据源名称' (Data Source Name), and a text input field for 'SQL语句' (SQL Statement).

## 数据质量任务调度管理

The screenshot displays a table with columns for '名称' (Name), '数据源名称' (Data Source Name), '模型名称' (Model Name), '模型类型' (Model Type), '业务系统名称' (Business System Name), '责任人' (Responsible Person), and '创建时间' (Creation Time). The table lists several data quality tasks.

# 模块4：数据质量平台运营实践

产品设计

技术研发

运营实践

## ✓ 数据质量稽核投入产出比问题（性能和成本）

当关键业务域数据体量太大(如每日新增超过百TB)、集群规模较大(无法建立对等测试环境), 总体质量稽核成本过高。  
(通用SQL稽核方式带来的资源浪费和性能问题)

## ✓ 数据稽核对象和策略的选择

针对省分不同账期不同主题域数据, 如何根据业务要求和实时流代码埋点处理流程进行抽样, 选择性做质量稽核? 优先解决采集链路质量监控、数据断传补传漏传、波动性监测等基础层稽核问题。然后解决业务层稽核问题。  
(数仓开发、数据应用业务冲突)

## ✓ 数据质量报告的问题（项目交付）

常态化的质量稽核统计报表, 无法给非技术口领导层直观呈现, 需要结合业务领域知识和组织结构做进一步封装。

## ✓ 与元数据、数据标准、调度系统协同

数据源目录分类和质检对象来源于元数据系统, 数据表模型质检要求来自于数据标准系统, 数据质量任务执行通常要跟 workflow 调度系统对接。

# 模块5：数据标准平台功能需求

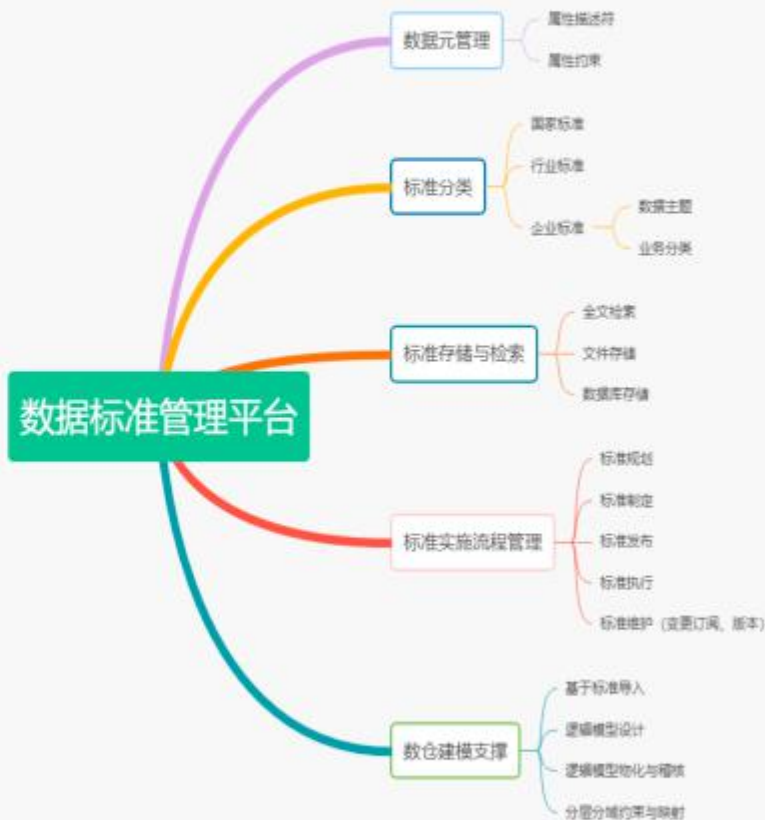
## 定位与目标：

数据标准是大数据治理生态中重要的一环，与数据过程管理、元数据管理、质量管理等模块进行协作，组成完整工具集，促进公司、组织内数据处理、交换相关流程、功能的标准化，有效提高数仓平台建设和数据管理的质量和效率，加速数据流转，从而促进业务创新。

核心模块1：数据元、代码集、标准术语管理

核心模块2：数据标准分类检索、实施流程管理

核心模块3：数仓建模管理 (逻辑模型设计与物化)



# 模块5：数据标准平台产品设计

## 数据模型标准



## 数据标准化构成

| 区分    | 说明   |
|-------|--|
| 标准单元  | <ul style="list-style-type: none"> <li>实体(名词属性/形容词的中文名称英文名称的构成要素, 实体名/属性名/形容词/一定源自单词的组合)</li> <li>业务上使用的词语和名称只能使用标准单元中定义的单词组合而成</li> </ul> |
| 标准术语  | <ul style="list-style-type: none"> <li>由标准单元组合而成的业务用语和技术用语</li> </ul>  |
| 标准数据域 | <ul style="list-style-type: none"> <li>按属性及其属性特征的数据分类及数据可应用的范围</li> <li>构成: 域名 + 数据类型 (数据类型+字长)</li> </ul>                                 |
| 信息类型  | <ul style="list-style-type: none"> <li>命名规则: (修饰语) + 分类语</li> <li>部分域中包含的数据设计时划分, 在构建数据模型时, 要转换成各自域中使用的数据类型及长度</li> </ul>                  |

## 数据命名标准

| 区分  | 内容   | 实例                  |
|-----|--|---------------------|
| 实体  | <ul style="list-style-type: none"> <li>(修饰语 + 主题词 + 修饰语) + 实体分类语 -&gt; 可省略修饰语</li> <li>主要含中文和数字组合</li> <li>使用单个名词或名词短语</li> <li>实体分类语 Ex: 订单号, 姓名, 性别, 职位, 地址, 姓名, 电话, 事实</li> </ul> | 客户账户                |
| 属性  | <ul style="list-style-type: none"> <li>(修饰语 + 主题词 + 修饰语) + 域(域+修饰语+分类语)</li> <li>中文单词的构成或词组</li> <li>有对属性的和域的名称, 但按一般情况予以省略, 按形式进行定义</li> </ul>                                      | 客户名称                |
| 表   | <ul style="list-style-type: none"> <li>表名由标准单元和英文字母按顺序自动生成</li> <li>一般: 领域+规则+主题词+实体</li> </ul>  | DWD_D_CUS_CUST_INFO |
| 字段  | <ul style="list-style-type: none"> <li>字段名由标准单元和英文字母按顺序自动生成</li> <li>在英文单词后加下划线 + 字段名称</li> </ul>  | CUST_ID             |
| 键值码 | <ul style="list-style-type: none"> <li>键值码按照上述数字表示, 如需要前缀以英文+数字</li> </ul>   | A11_101             |

## 数仓逻辑模型设计与物化



## 模型在线编辑

The screenshot shows a software interface for online editing a data warehouse model. It displays a table with columns for '编号' (ID), '名称' (Name), '编码' (Code), '类型' (Type), '长度' (Length), '精度' (Precision), '主键' (Primary Key), '外键' (Foreign Key), '分区' (Partition), and '备注' (Remarks). The table contains the following data:

| 编号 | 名称   | 编码         | 类型  | 长度 | 精度 | 主键 | 外键 | 分区 | 备注 |
|----|------|------------|-----|----|----|----|----|----|----|
| 1  | 客户编号 | USER_ID    | 字符串 | 48 | 7  | 是  |    |    |    |
| 2  | 客户名称 | AREA_ID    | 字符串 | 20 | 8  |    |    |    |    |
| 3  | 订单编号 | ORDER_ID   | 数字  | 1  | 8  |    |    |    |    |
| 4  | 商品编号 | PRODUCT_ID | 数字  | 7  | 8  |    |    |    |    |
| 5  | 数量   | QUANTITY   | 字符串 | 8  | 8  |    |    |    |    |

## 数据标准导入



# 模块5：数据标准平台运营实践

产品设计

技术研发

运营实践

## ✓ 数据标准分类管理问题

国家标准、行业标准、企业标准同时管理成本较大，其实践层面的指导意义待深入探索研究和试错。

## ✓ 数据标准制定相关的业务梳理工作

数据标准的制定往往依赖于领域业务能手、IT架构专家等组织团队的通力协作，相关的业务梳理工作工程量很大。

## ✓ 数据标准的内部推广应用问题

数据标准管理工具在内部推广应用的实际困难往往会超出预期，需要持续迭代，离不开一把手的支持和长期的努力。

## ✓ 基于数据标准做数仓全局规划和落地（一体化设计）

主要是从逻辑模型设计和物化入手，逐渐完善数据仓库分层分域、数据质量、数据指标标准等数据架构规范的落地。  
(大规模数据平台匹配度较低，适合推到重来)

# 模块6：集群治理平台功能需求

## 定位与目标：

基于Hadoop集群底层组件运行机制和大数据开发运维等组织活动进行多维交叉洞察，以降本增效为中心，向下保障大规模Hadoop集群算力，向上指导数据治理动作实施和业务连续性。

核心模块1：集群治理数据采集

核心模块2：集群治理分析引擎

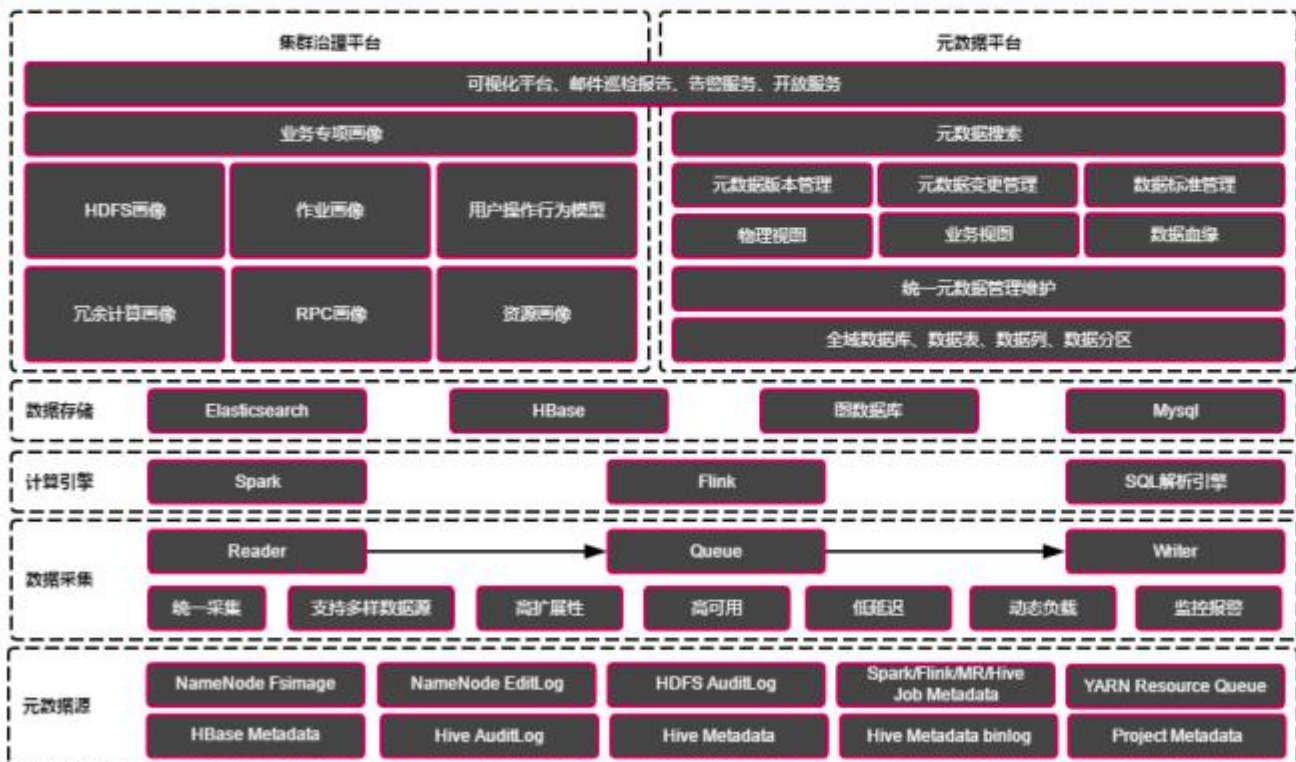
核心模块3：集群治理平台应用



# 模块6：集群治理平台技术架构

## 设计原则

自底向上，用大数据的技术手段解决大规模数据集治理层面的各种问题



**统一实时、离线元数据自动化采集**  
多种元数据数据源，研发可配置式支持不同引擎作业元数据采集、审计日志采集、变更日志采集、分布式存储元数据采集、Hive元数据采集等统一采集工具

无侵入性

**统一离线、流式分布式计算引擎**  
采用业界成熟的分布式实时、离线数据处理引擎，保证高吞吐、低延迟、高性能数据处理

实时审计

**融合多种存储服务满足各种场景**  
采用高可用分布式多种存储系统，分布式列数据库存储、分布式搜索引擎、高可用图数据库、高并发时序数据库

# 模块6：集群治理平台产品设计

大数据集群治理平台针对大数据集群和数据加工流程进行精细化监控分析，从资源画像、存储画像、作业画像、RPC画像、冗余计算挖掘、数据血缘分析、用户行为告警八大维度几十个小维度交叉洞察可优化作业和业务处理流程，不断驱动集群资源优化。

## HDFS文件画像



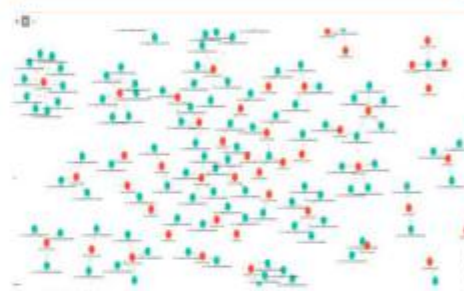
## 数据作业画像



## 集群资源画像



## 数据血缘分析



## 待优化作业排行



## 待优化作业定位



## 冗余计算挖掘



## 用户行为告警





# 模块6：集群治理平台运营实践

## 场景描述

- 公司业务高速发展过程中数据业务需求越来越复杂，所需要的算力也越来越大，进一步导致集群的规模越来越大，承担的产品也越来越多，集群面临资源负载过高、资源抢占严重、RPC请求负载过高等问题，存储系统也面临空文件过多、垃圾文件过多、小文件过多、平均文件大小过小、文件数持续增长等一系列问题，存储系统稳定性面临很大隐患，作业又面临执行耗时过长、耗资源大、数据倾斜严重等问题，直接导致数据加工异常率过高、数据具备时间有延迟风险、产品交付面临很多风险。

## 面临的挑战



## 大数据分析方法

HDFS文件  
存储洞察

Job数据  
作业洞察

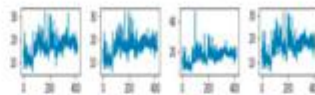
冗余计算  
挖掘分析

开发NameNode元数持久化文件Fsimage和元数操作文件记录文件EditLog的反序列化解析项目，无侵入性洞察全集群，冷温热存储状态，千万级目录精细画像。

实现资源监控与异常作业多维度洞察、高效协同优化。综合几十个小维度进行集群交叉治理并协同各相关组织进行全城治理，使集群逐步向良性健康方向发展。

通过对HDFS JOB BINARY FILE分析，定位疑似冗余计算作业，与组织架构复杂度无关，不依赖上层业务的大量输入，其核心提取出具有相同输入路径的作业，以目录维度视角挖掘作业。

## 实际应用效果



## 价值创新

**精准洞察**：能够精准定位问题根因，提升优化效率  
**数据准确**：从底层采集强关系元数据以及日志，数据准确度高  
**无侵入性**：优化洞察不需要修改产线环境配置，不影响产线环境生产  
**多维画像**：从多个大维度几十小维度交叉洞察潜在问题与可优化方向  
**深入内核**：通过阅读内核源码深入了解底层组件，找到解决方案  
**跨组协同**：根据洞察报告多组协同自发优化

实现了在算力不增加，数据处理量**翻倍**的情况下，算力整体负载下降**20%**以上，每年直接节省固定资产投资**上千万元**成本。  
先后解决集群小文件治理、作业模型高级参数自动化、数据模型治理专题优化等难题，并形成了集群深度治理平台化能力。

## 集群健康分



## 核心技术框架



# 模块7：数据服务平台功能需求

## 定位与目标：

以生产环境的运营支撑和应用开发为主要IT诉求，构建IaaS、PaaS、SaaS三层私有云体系，提供可复用、可隔离的存储计算资源、数据资源、开发组件资源，同时保证多租户安全隔离，方便数据资源开放共享和数据资产运营。

核心模块1：云计算资源池

核心模块2：数据能力商店

核心模块3：多租户控制台



# 模块7：数据服务平台产品设计

统一能力开放云数据服务平台各项业务功能，为租户提供基础设施、云数据库、计算框架、数据指令平台、应用与服务、安全控制等产品的统一浏览、申请、划配，并对租户使用的产品和运营进行统一管理。

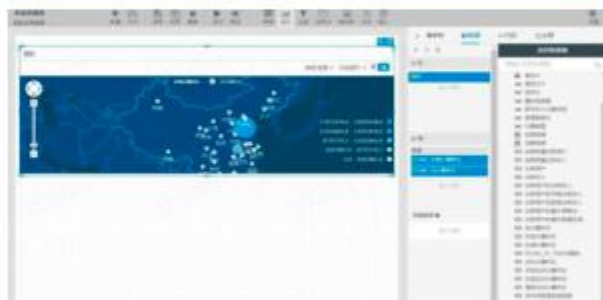
## 数据服务门户系统



## 数据服务API管理



## BI自助分析工具



## 图形化的数据管控套件



## 统一的计费报表管理



## 自助式可视化建模平台



# 模块7：数据服务平台运营实践

数据资产运营=数据能力开放+数据生态合作

- 能力开放平台向租户提供资源、数据以及服务，为租户方便有效的使用平台环境保驾护航。



- 租户基于能力开放平台获取全国样例数据，可结合本地数据进行模型训练，训练后的脚本提交平台反馈至大数据生产服务平台进行加工计算，计算后的结果推送至租户空间，满足租户基于全国数据的分析挖掘。

Contents

## 目录

01

背景概述 (06P)

02

关键挑战 (04P)

03

详细介绍 (23P)

04

整体总结 (01P)

## 总结

在数字经济规模发展的大背景下，疫情加速了各行各业数字化转型的节奏和步伐。**数据治理/数据资产管理**已成为企业IT和数字化转型的核心要素和重要驱动力，企业要结合自身实际经营情况，以新兴数字技术来创造价值，快速响应市场需求，持续构建创新与服务能力，为企业高质量发展注入新的活力要素。



### 数据治理/数据资产管理**失败**的5个主要原因：

1. 在认知战略层面，没有真正上升到数据战略层面，没有一把手牵头去规划和执行，在中途放弃了很多目标。
2. 在组织协同层面，没有建立起高效的数据组织协同机制，没有形成合力，互相推诿扯皮，组织墙问题突出。
3. 在文化价值层面，没有形成精益实用的数据驱动文化，大量历史习惯阻碍了文化的落地，无法突破舒适圈。
4. 在人才体系层面，没有配套的数据人才支撑体系，团队的成长较缓慢，跟不上行业和企业快速发展的节奏。
5. 在技术工具层面，无法驾驭以元数据平台、质量平台、标准平台、数据服务开放平台为代表的先进生产力。