



DTCC

2016中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2016

数据定义未来

SequeMedia
盛拓传媒

IT168.com

ChinaUnix

ITPUB

【友盟+】数据服务平台介绍

张 金



友盟+



DTCC

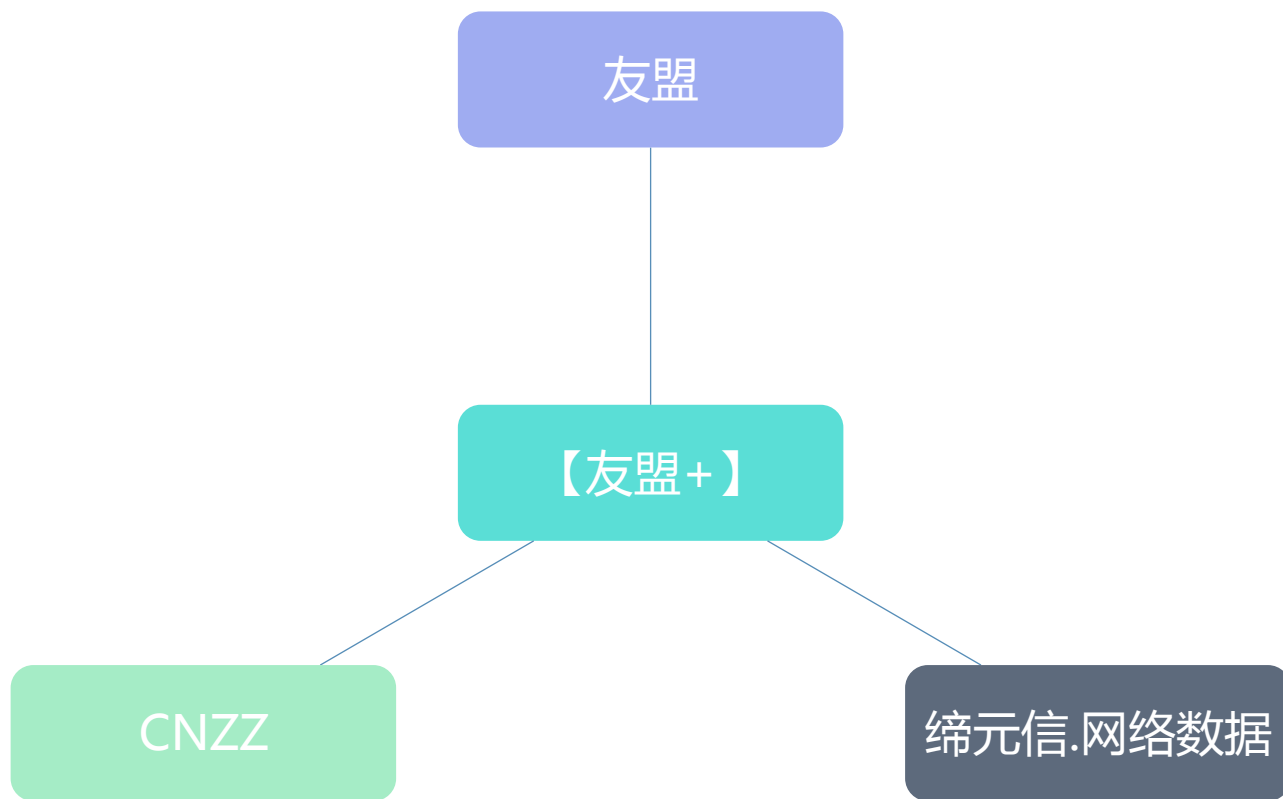
2016年中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

SequeMedia
数据工场

IT168

ChinaUnix

IT-PUB



新挑战

不同的技术栈

数据分散

相同服务，不同实现

数据价值

海量数据



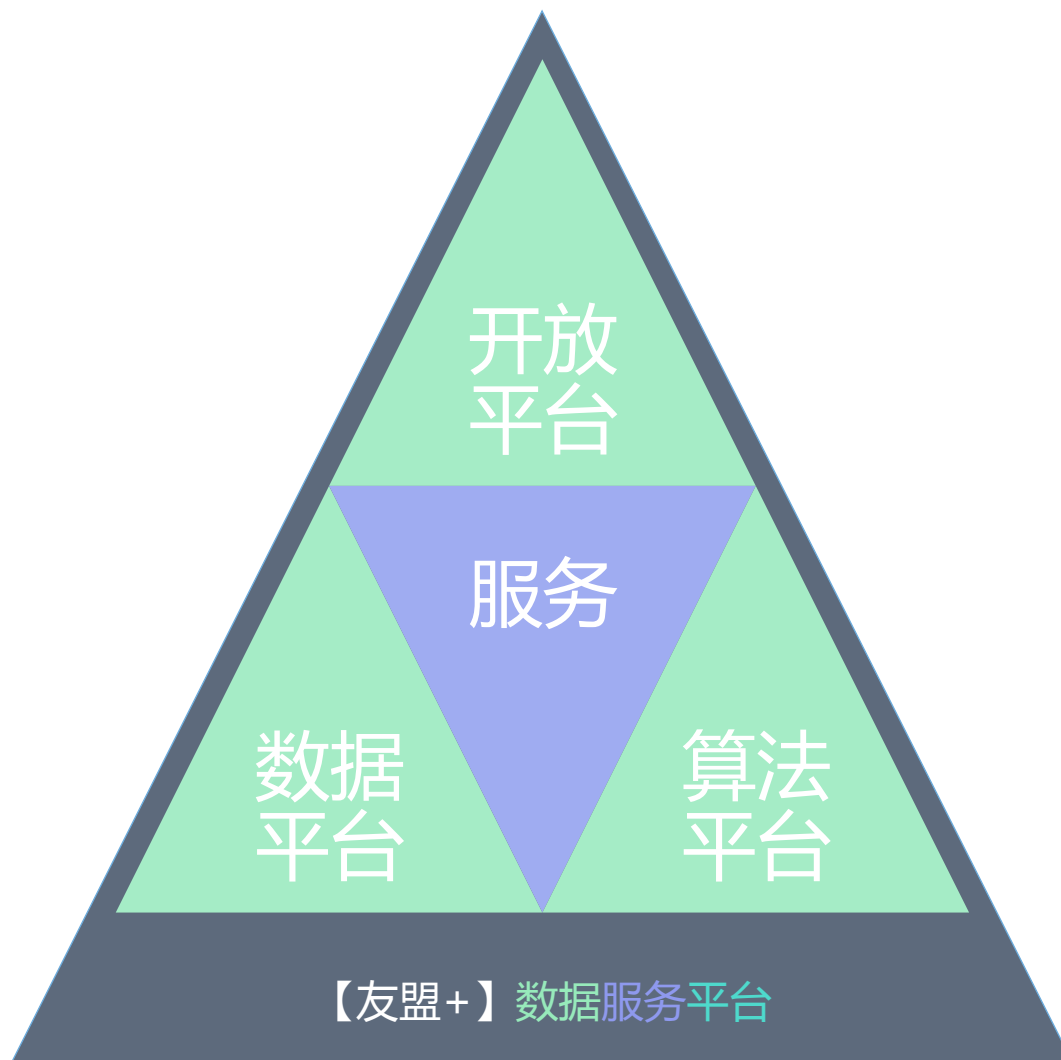
数据 指标 服务

统一

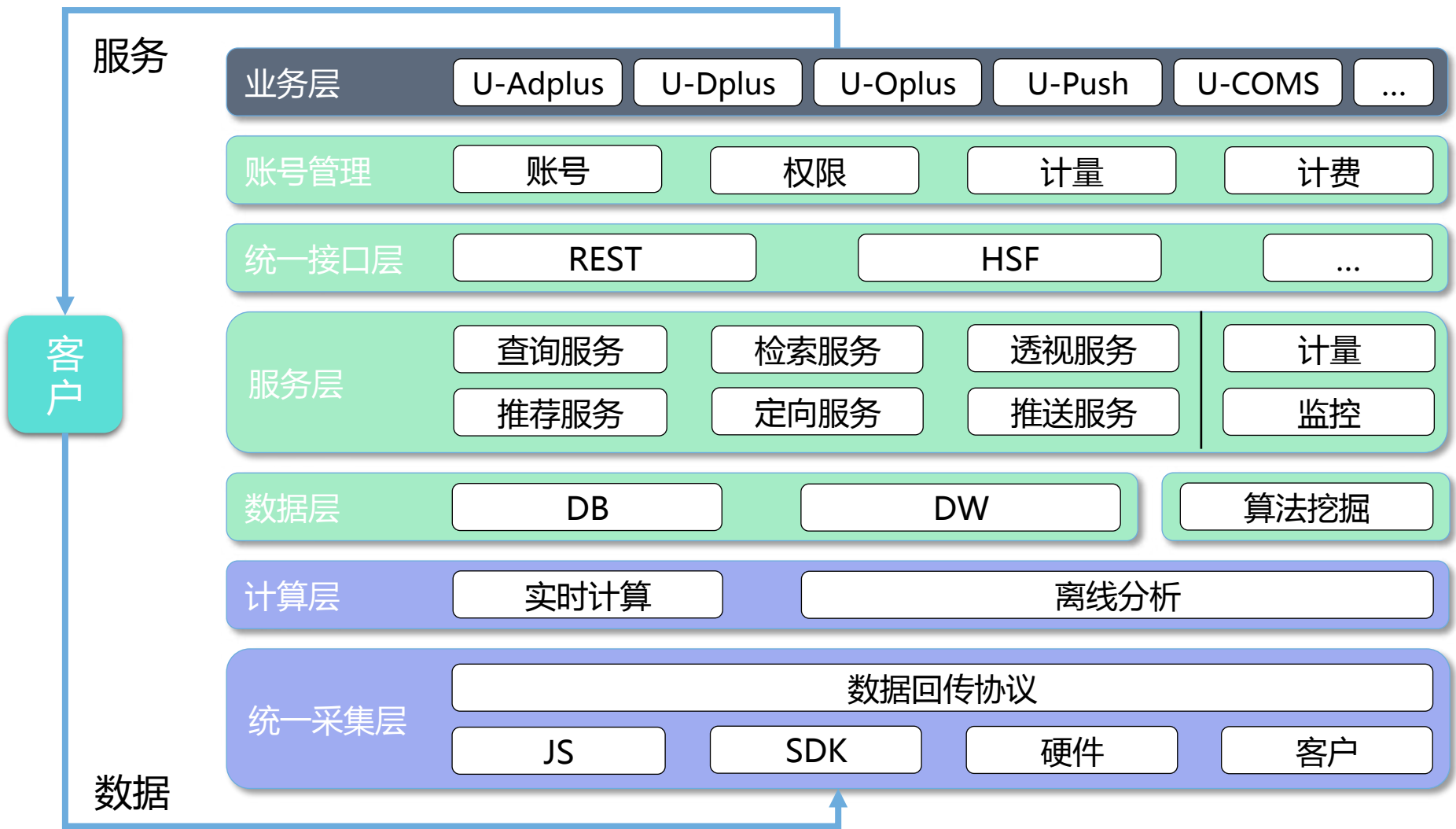
采集 账号 计算

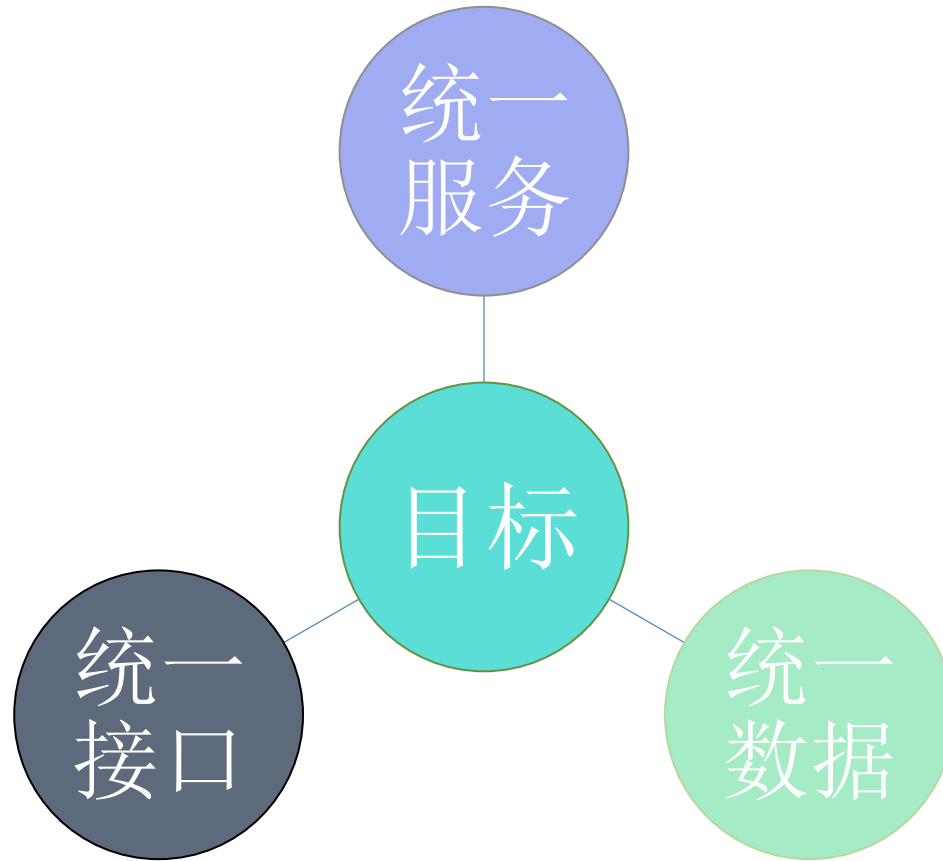
接口 ID

服务才是客户关心的



定位





重点介绍

□ 数据：基础数据体系构建

□ 接口：统一查询服务

□ 服务：DMP

重点介绍

✓ 基础数据体系构建

□ 统一查询服务

□ DMP

问题

存储

数据分散，使用不便

冗余重复，浪费资源

标准

指标计算

标签体系

ID互通

管理

权限管理

资产情况

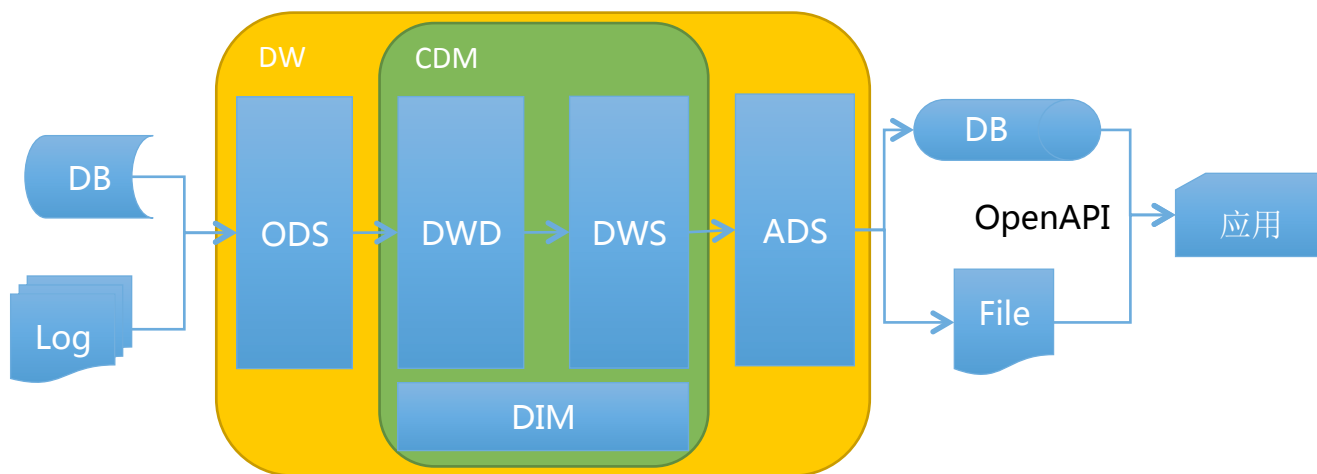
方案

- ✧ 数据迁移整合
 - 建立数据流转通路
- ✧ 制定数据规范
 - 统一研发标准
 - 统一命名规则
- ✧ 用户ID打通
 - ID mapping系统
- ✧ 统一指标和标签体系
 - 确立标准，消除二义性
- ✧ 建立日常数据透视
 - 随时了解数据状况

数据规范

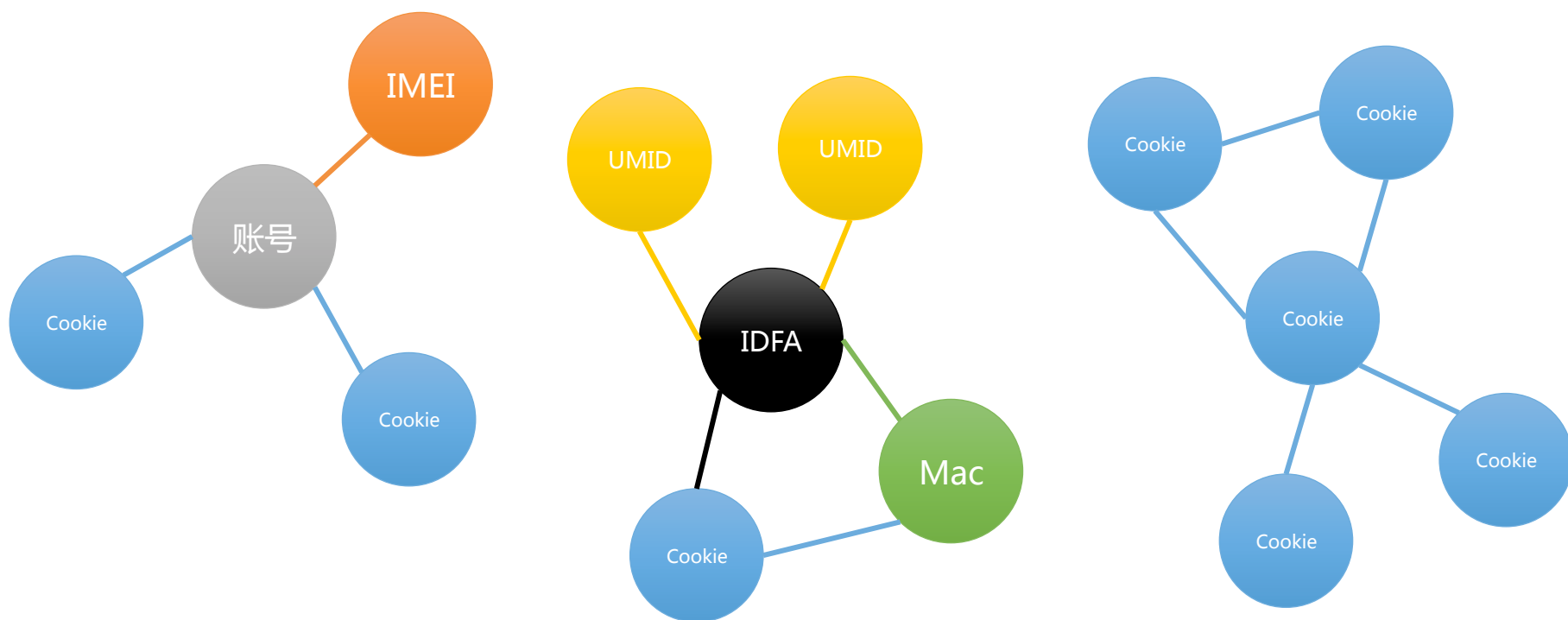
• 数据仓库分层管理

- ODS层：业务系统的原始底层数据
- CDM层：公共中间层数据
 - 明细粒度事实层（DWD）
 - 汇总粒度事实层（DWS）
 - 公共维度层（DIM）
- ADS层：应用数据服务层



ID Mapping

利用匹配对数据，通过算法对所有ID进行关联。

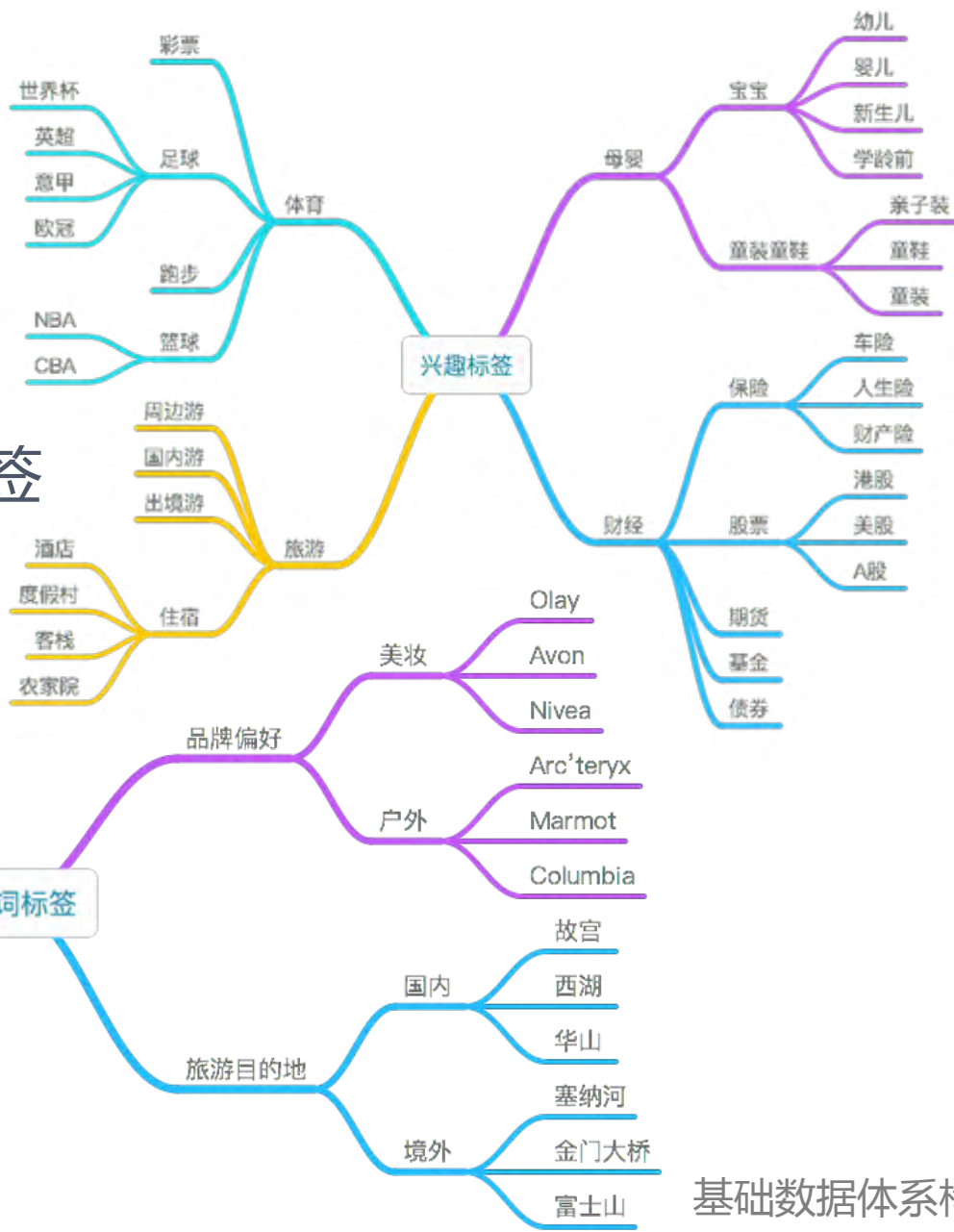


标签体系

• 根据分类梳理标签

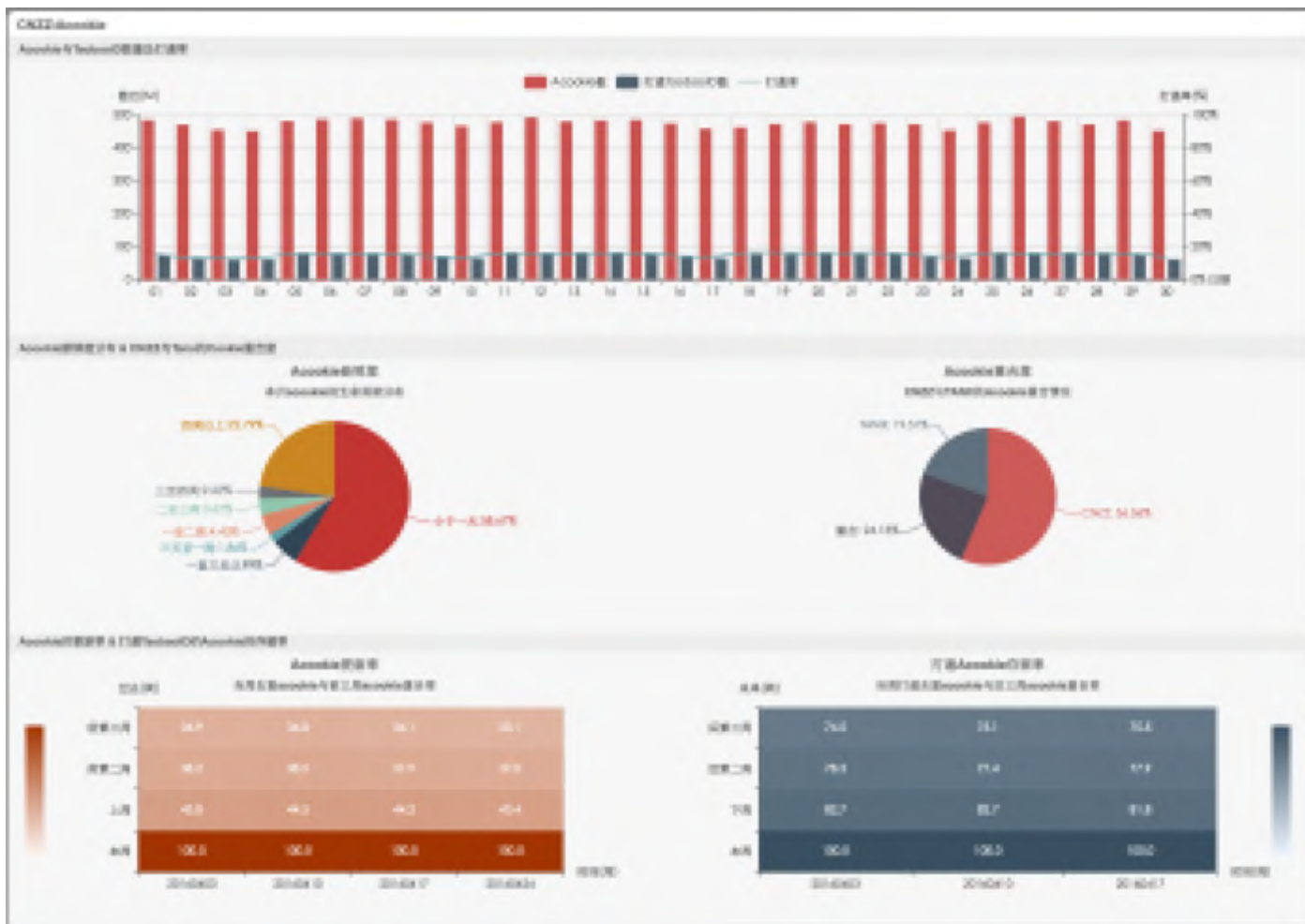
- 人口属性
- 长期爱好
- 短期兴趣

• 类目化标签 vs. 关键词标签



数据洞察

随时掌握数据情况



数据Wiki

快速检索数据表



The screenshot displays the Data Wiki interface for a database named 'test'. It features a search bar at the top right, a navigation menu with tabs like '数据表', '数据表', '数据表', '数据表', and '数据表', and a main table listing various data tables. The table has columns for '表名', '表类型', '表大小', '表空间', '表分区', '表索引', '表主键', '表外键', '表注释', and '表操作'. The table lists 14 data tables with their respective details.

表名	表类型	表大小	表空间	表分区	表索引	表主键	表外键	表注释	表操作
1	varchar	varchar							
11	varchar	varchar							
12	varchar	varchar							
13	varchar	varchar							
14	varchar	varchar							
15	varchar	varchar							
16	varchar	varchar							
17	varchar	varchar							
18	varchar	varchar							
19	varchar	varchar							
20	varchar	varchar							
21	varchar	varchar							
22	varchar	varchar							
23	varchar	varchar							
24	varchar	varchar							

重点介绍

□ 基础数据体系构建

✓ 统一查询服务

□ DMP

问题

U-API和业务开发需求

- 指标查询

历史遗留

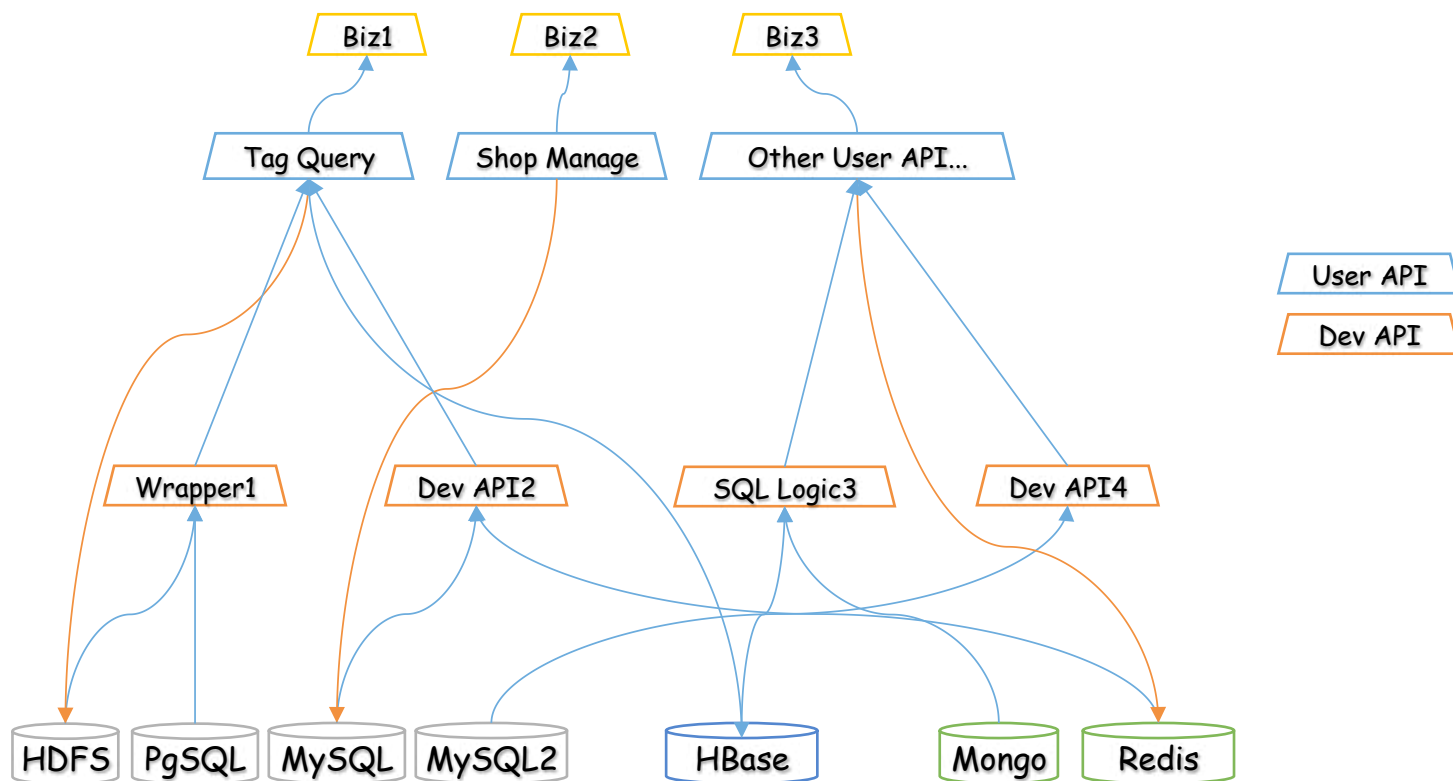
- 指标数据存在各种不同的DB中
- 数据库异构：SQL，NoSQL
- 数据库的权限管理复杂

开发工程师之痛

- 不同业务需要面对不同数据源

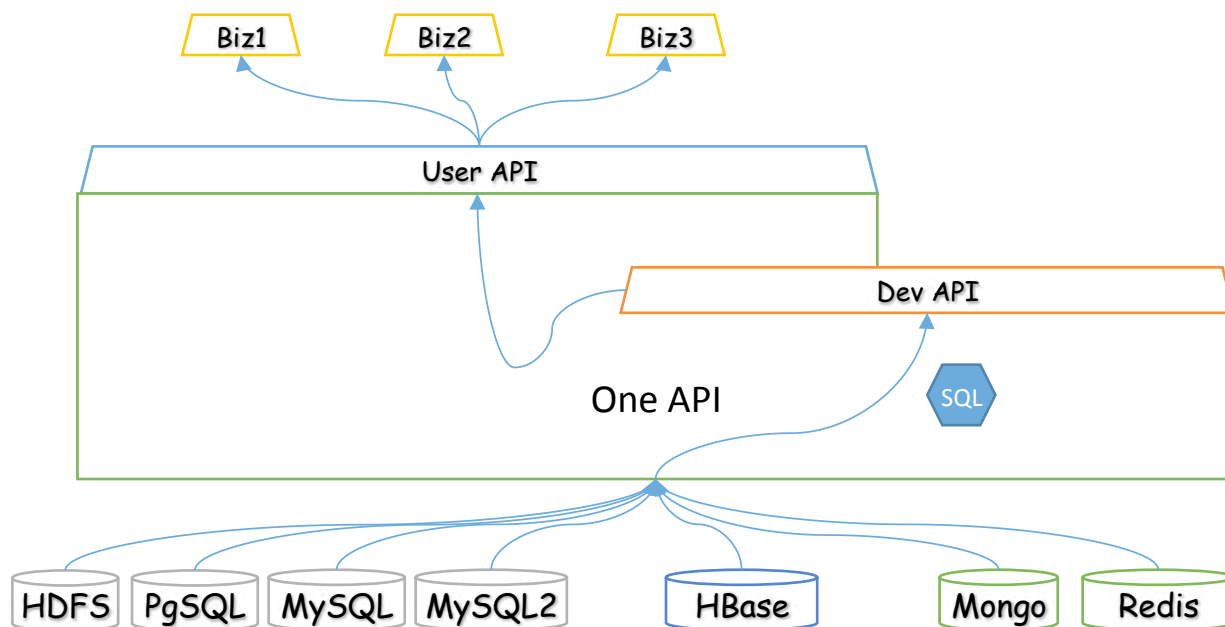
老方法

- 为每一个功能包装一个接口
 - 混乱的接口设计
 - 错综复杂的调用关系
 - 苦、不灵活、重复劳动



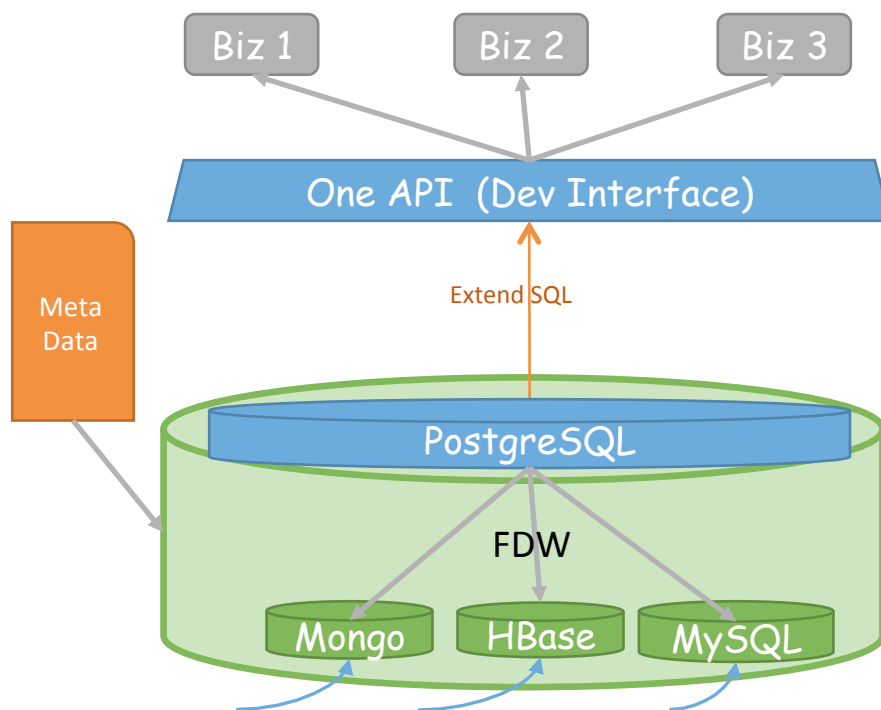
新办法

- 业务开发工程师只需要关心
 - 统一的调用方式
 - 清晰的逻辑数据表



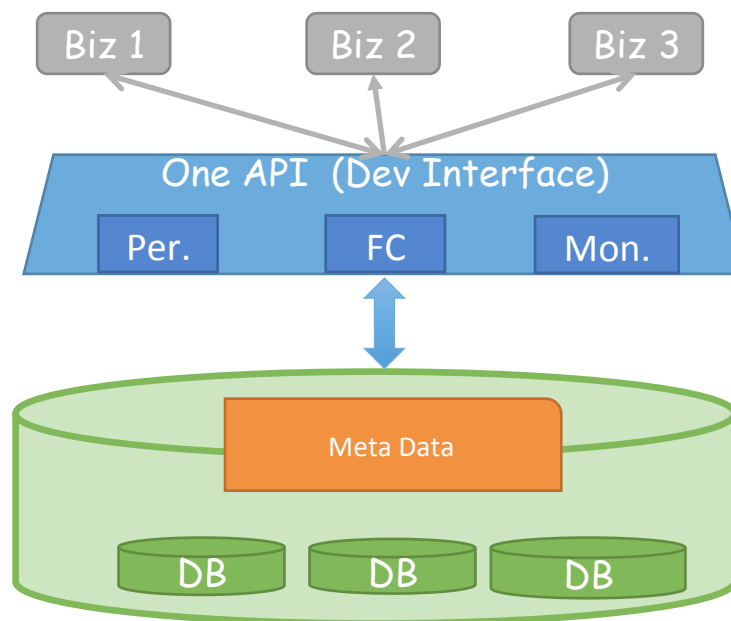
方案

- 接口方式
 - PostgreSQL
- 多种数据库链接
 - FDW(Foreign Data Wrapper)
- 逻辑表建设
 - 数据公共层梳理



后续

- 完整的中间件支持
 - 读写
 - 更多数据库
 - 权限管理
 - 流量控制



重点介绍

□ 基础数据体系构建

□ 统一查询服务

✓ DMP

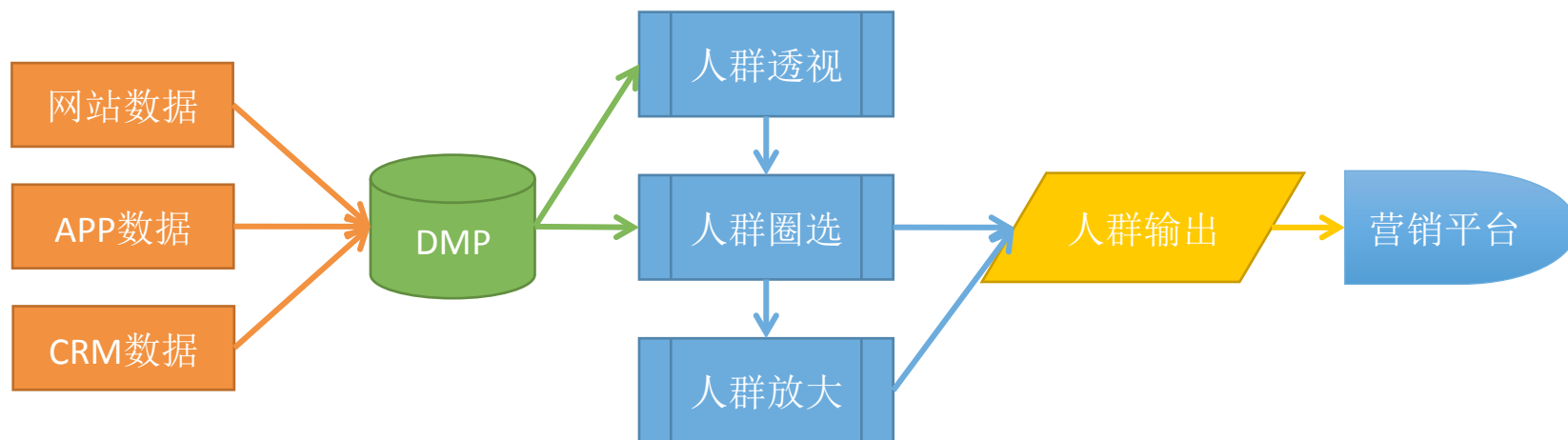
目标

- 让营销者能够轻松洞察并定向目标人群

能力

- 目标人群的透视分析
- 支持运营及广告的投放活动
- 标签数据的购买管理
- 让客户的自有数据产生价值
- 支持客户间的数据购买及交换

典型流程



主要服务

- 人群透视
 - 海量数据实时高并发在线分析
 - Analytic DB
- 人群圈选
 - 白盒投放系统
 - 通过生成新标签提升查询效率
- 人群输出
 - 10亿级用户，数千标签
 - 毫秒级查询 (HBase、Memory DB)
 - 同源机房高速缓存

数据服务平台愿景

以丰富的基础数据服务支持
数据应用的创新孵化

【友盟+】

<http://www.umeng.com/>



友盟全域数据



个人公众号



THANKS
THANKS

SequeMedia
威拓传媒

IT168.com

ChinaUnix

ITPUB



友盟+