

MEIZU

IT大咖说
不止于技术

数据探索

魅族技术开放日第七期

● 活动时间:3月18日(周六) 13:30-17:20

活动地点:深圳市南山区科发路1号富利臻大厦3F 思微2.0

联合主办: MEIZU  Flyme msup[®]

揭秘魅族推荐平台架构设计

倪江利 高级软件工程师

联合主办: MEIZU  Flyme msup[®]

- ▶ **推荐介绍**
- ▶ 魅族推荐平台架构演进
- ▶ 魅族推荐平台现状
- ▶ 魅族推荐平台挑战和愿景

推荐能做什么

MEIZU

IT大咖说
不止于智人



在合适的场景, 合适的时机, 通过合适的渠道, 把合适的内容, 推荐给合适的用户

推荐的作用

MEIZU

IT大咖说
不止于技术

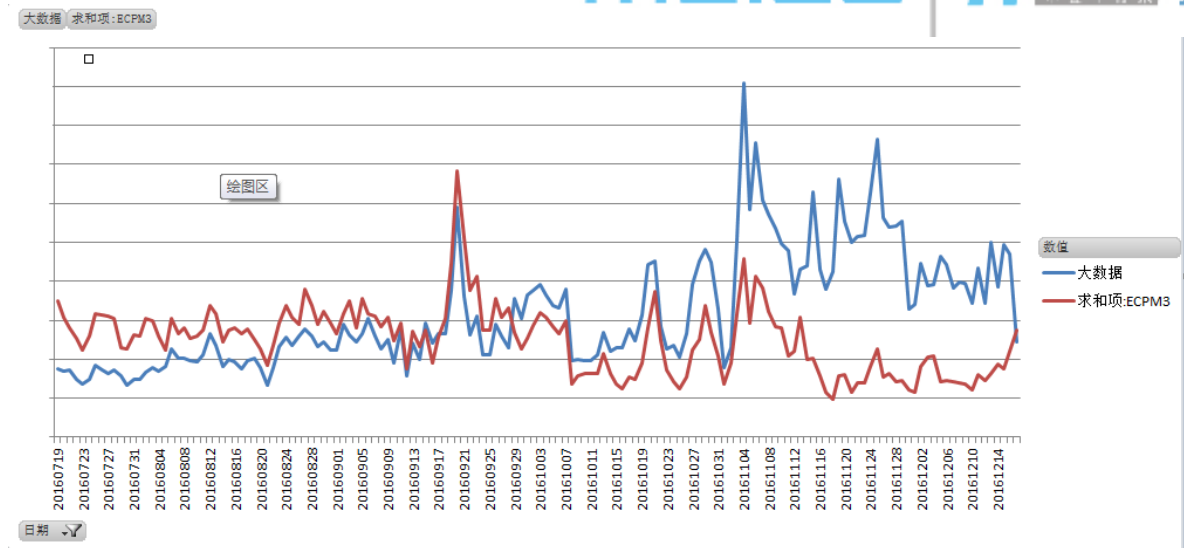
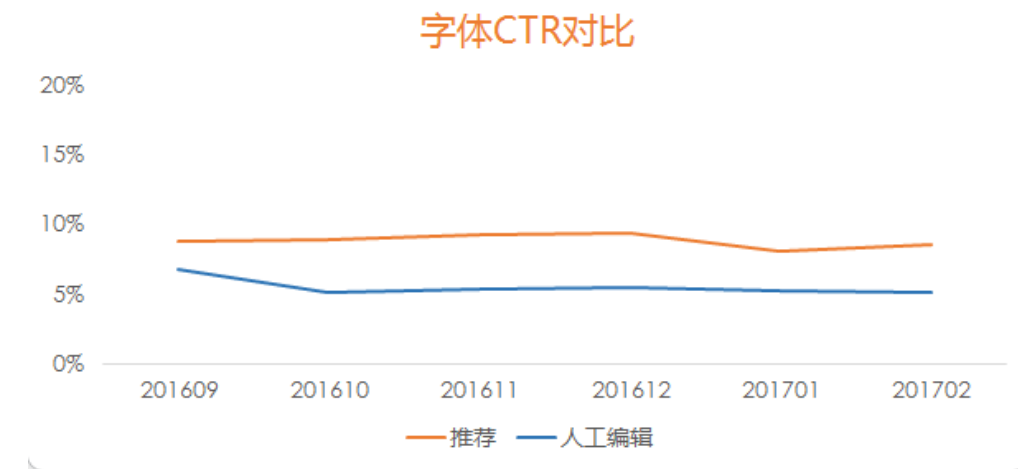
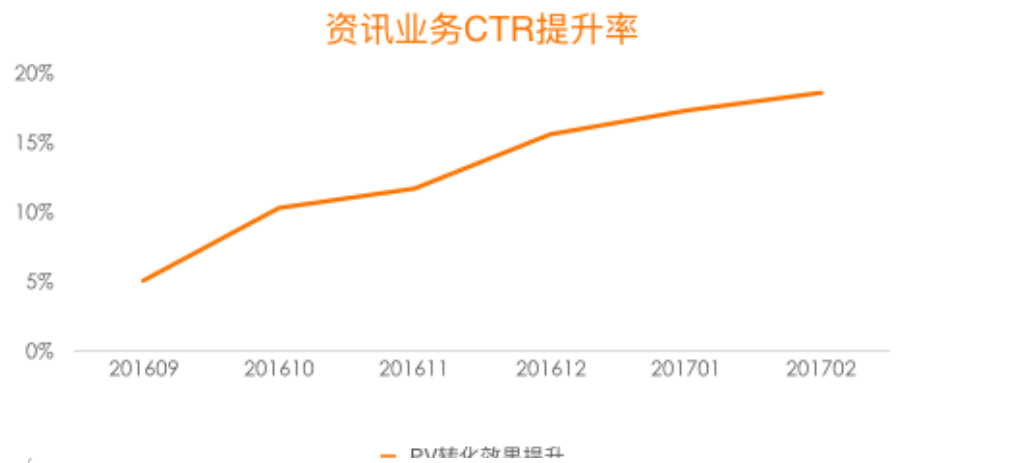


推荐在魅族中的应用

在资讯、视频、应用中心、个性化中心、广告等业务产品线中应用



推荐效果



与人工编辑相比，推荐的优势明显，与行业内的一线互联网公司比我们有超越也有还需要努力的地方

- ▶ 推荐介绍
- ▶ **魅族推荐平台架构演进**
- ▶ 魅族推荐平台现状
- ▶ 魅族推荐平台挑战和愿景

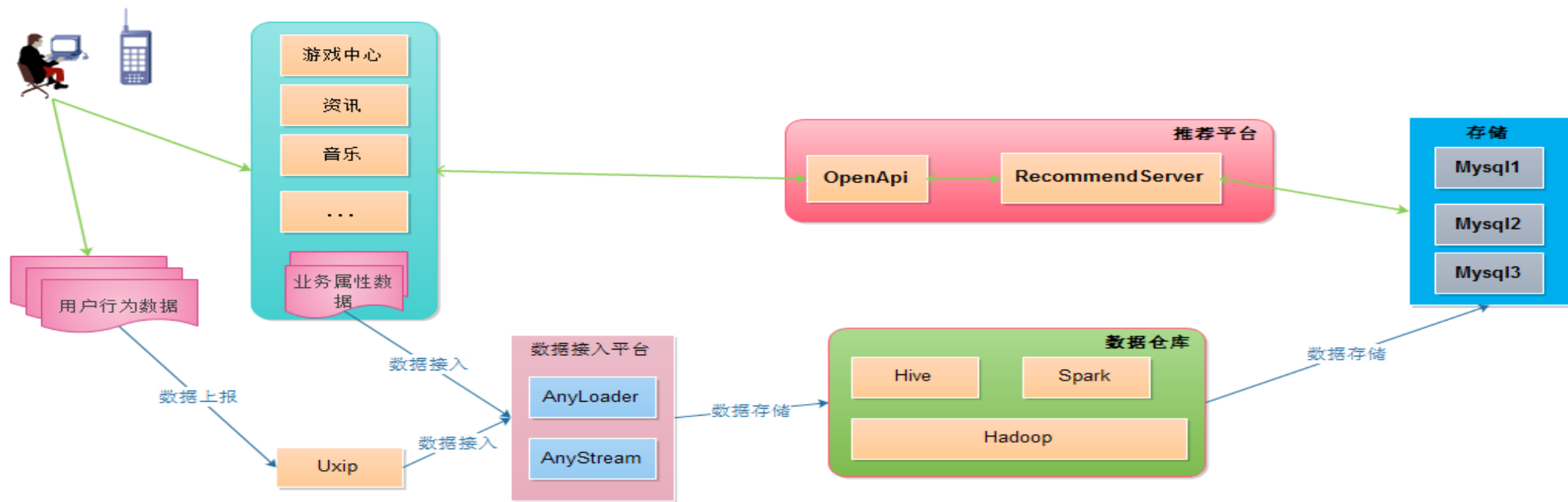
□ 平台的核心需求

支撑5个以上的大产品线的不同场景的推荐业务需求，保证业务稳定运行，可用性达到99.9%，推荐场景当次请求响应在200毫秒以内，广告预测场景当次请求响应在100毫秒以内，一天需要支撑级亿级别的PV量

□ 技术难点

- 1、针对于每一个用户的一次推荐需要从万级甚至是十万级别以上的物品中进行挑选用户可能感兴趣的物品
- 2、每一次推荐需要同时计算十个甚至是数十个算法数据，一个算法需要计算成百上千个维度
- 3、一天需要实时处理上亿条行为日志，进行百亿到千亿次计算
- 4、每天需要访问数据存储上十亿次以上
- 5、每天需要支撑上百个数据模型在线更新及实验

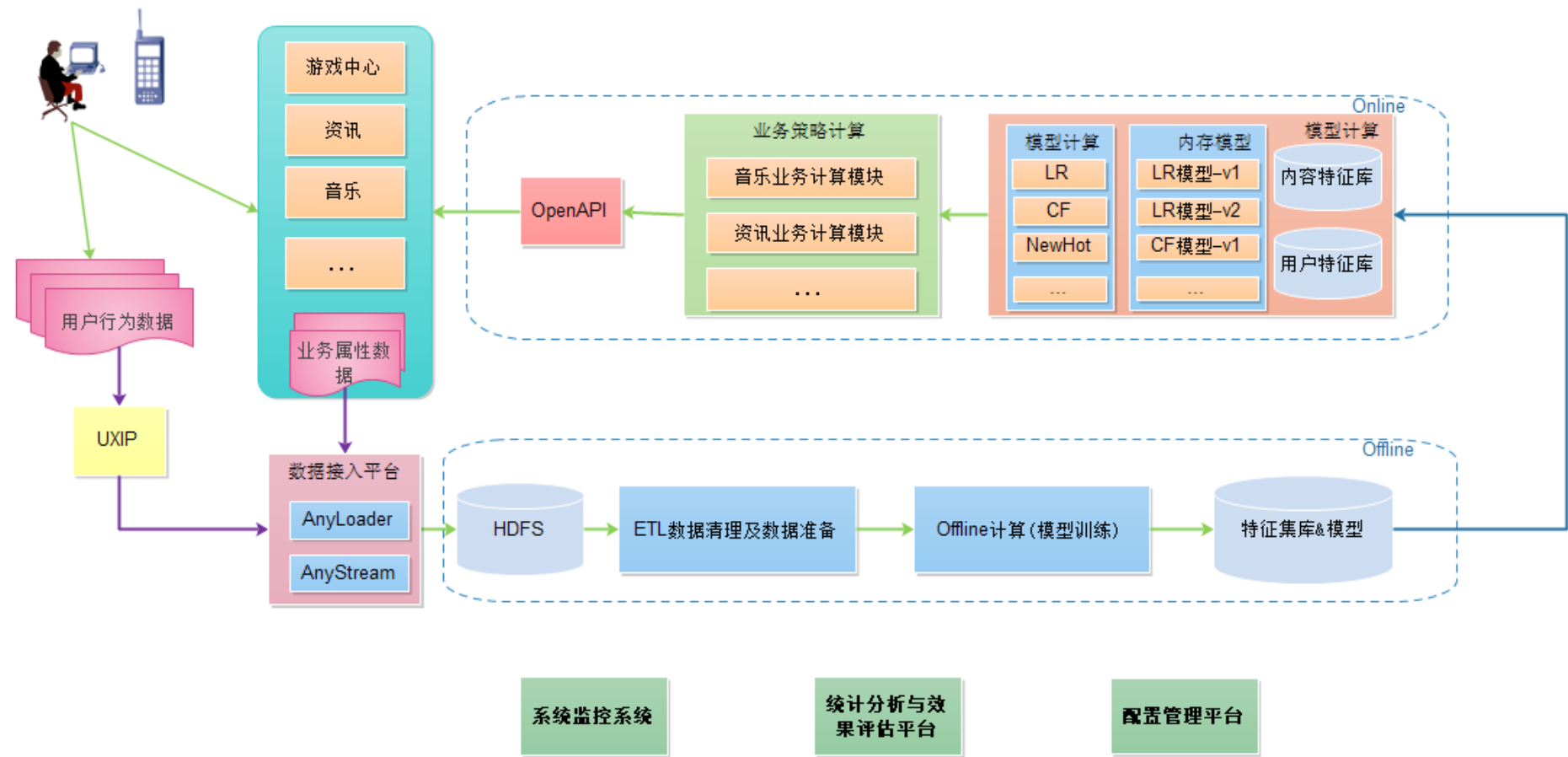
推荐平台第一代架构



■缺点:

- 离线计算量大，需要将所有用户的数据进行结果计算,同时浪费机器资源
- 结果数据更新困难，大批量数据更新对数据库的冲击大，可能直接造成用户访问超时,服务不可用
- 数据更新延时大，超大数据量计算基本上只能实现T+1的方式进行数据更新，所以数据推荐都是基于旧行为数据进行预测
- 数据库的瓶颈直接影响算法结果数据输出频率，算法调优困难
- 扩展困难，所有结果数据已经固定输出，很难插入一些业务上特定的需求

推荐平台第二代架构



- 用户推荐数据实时根据用户请求进行计算，减少离线计算量及减少数据存储空间
- 原始模型的输出到线上比起结果数据输出更轻便，对线上性能影响更小，更方便于算法在线调优
- 模块分离，业务个性化处理与模型计算分离，系统更抽象化可复用度越高及可扩展性越好

■问题:

- 模型离线训练，用户实时产生的行为无法反馈到模型当中
- 业务混布，各业务之间相互影响
- 由于把离线的部分计算放到线上进行计算，在请求过程中计算量增大，系统响应时长挑战增大
- 业务接入越多，模型会越来越多，单台机器已经无法装载所有的模型

- ▶ 推荐介绍
- ▶ 魅族推荐平台架构演进
- ▶ 魅族推荐平台现状**
- ▶ 魅族推荐平台挑战和愿景

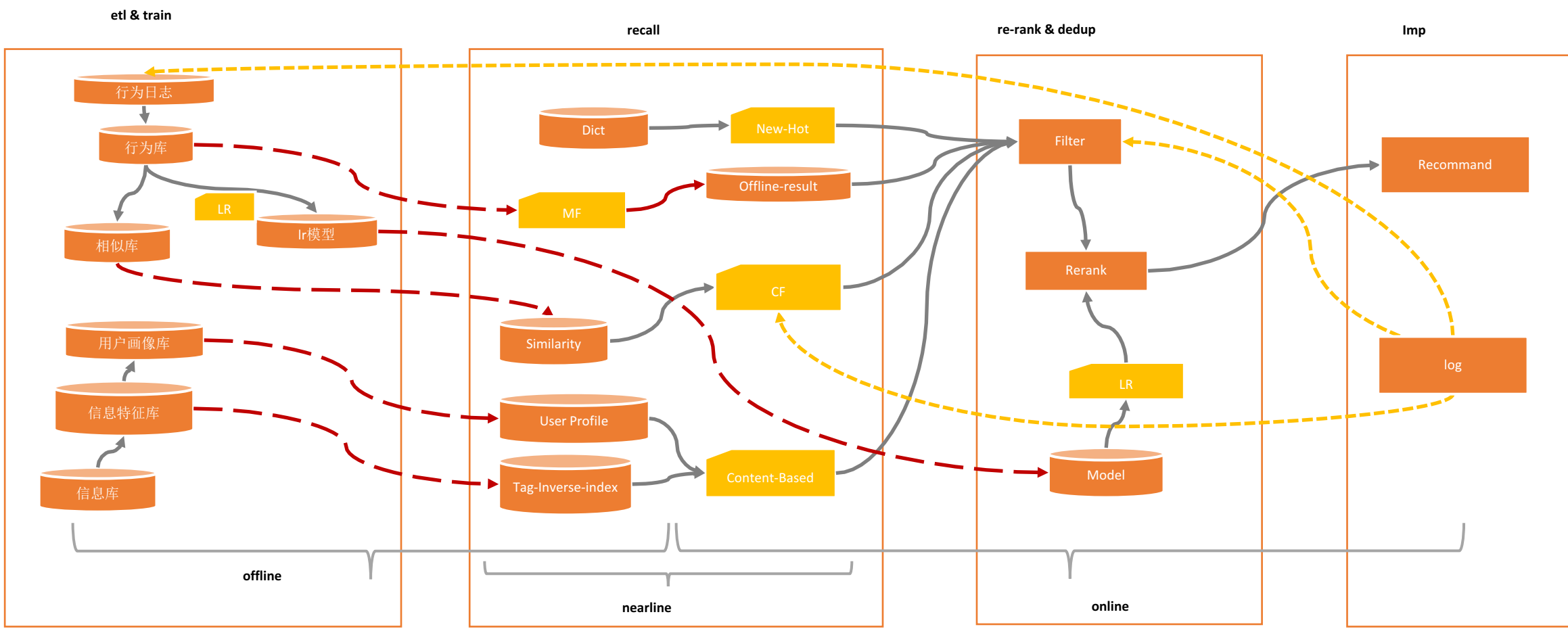
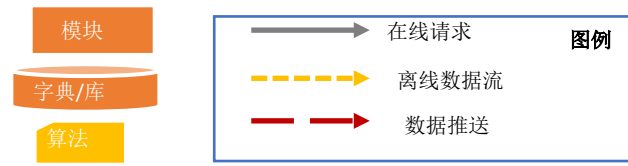
三代架构的核心需求

MEIZU

IT大咖说
不止于技术

- 集群资源动态管理，解决模型存储及计算资源利用率的问题
- 用户行为数据能够实时的进行计算，并最终反馈到模型，提高推荐结果的准确性
- 优法算法模型训练过程，将大部分工作能通过可视化的方式完成，提高工作效率
- 解决业务之间的相互影响
- 优化高效的性能及稳定性

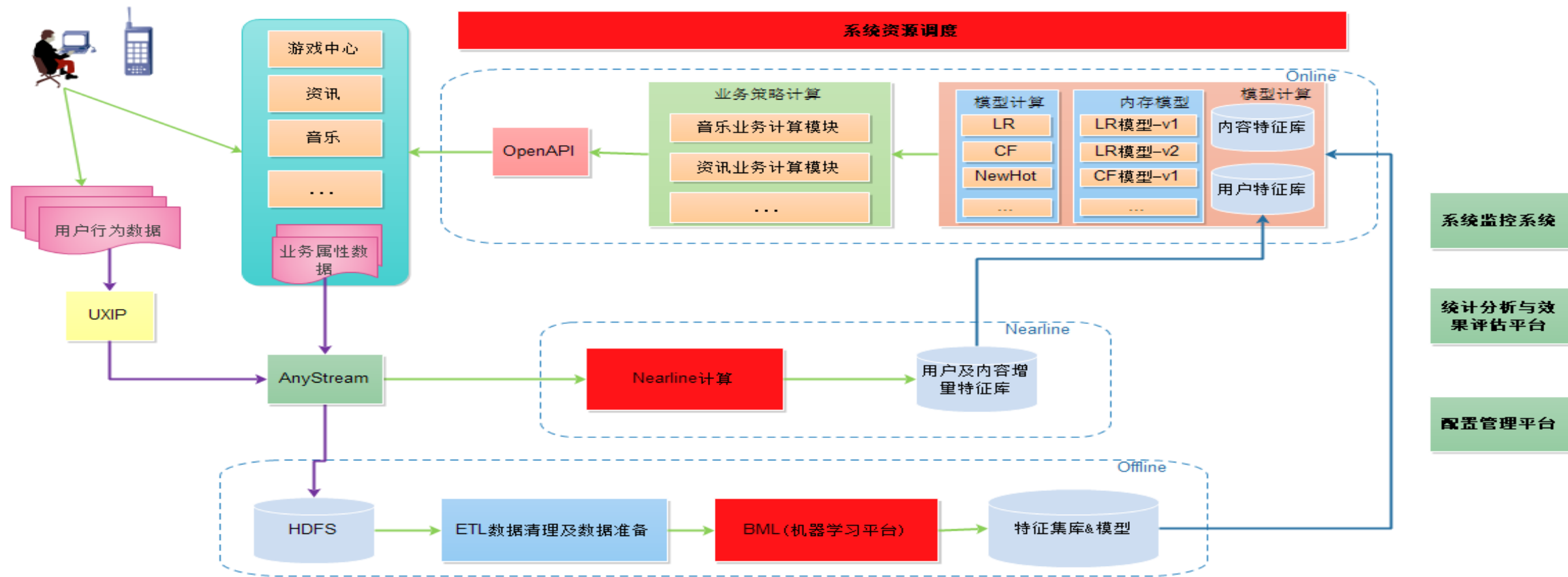
推荐平台数据处理模型



推荐平台现有架构

MEIZU

IT大咖说

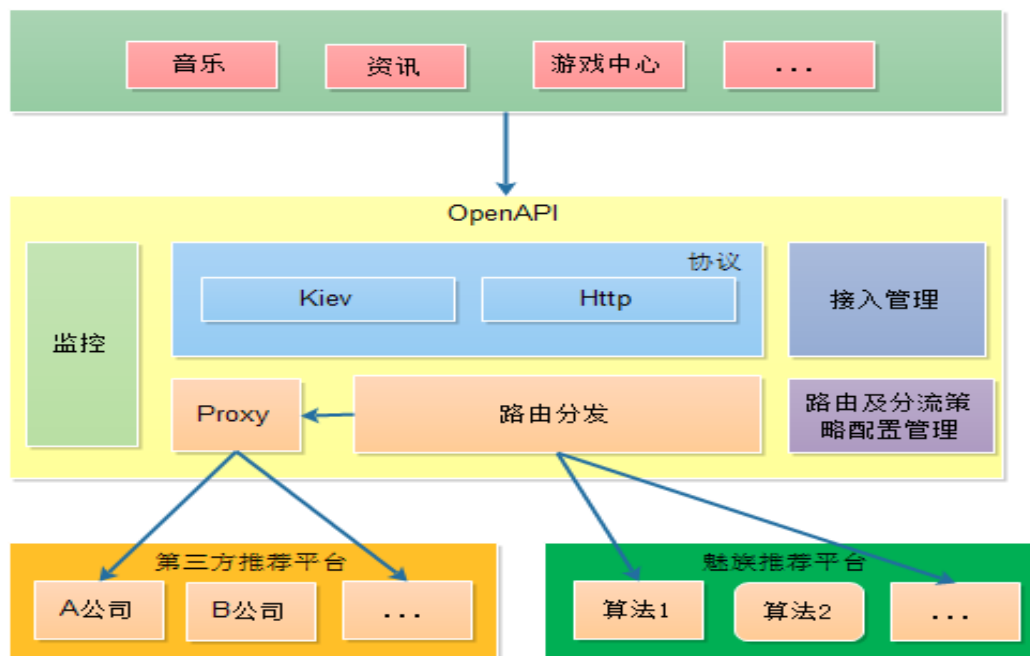


推荐平台架构分层

推荐系统被分为三层：

- **offline**运算层：该层主要是离线对海量的数据进行建模加工，生产有价值的数据，如Item相似库、user相关库、CF离线推荐结果等。
- **nearline**运算层：该层主要是利用流式处理的技术对用户实时产生的行为日志进行加工，利用一些高效、高性能的算法生产有价值的数据
- **online**运算层：该层主要处理一些相对简单的运算逻辑，在线进行计算。

在线模块-OpenAPI



■ 统一接入规范

所有应用接入按照统一规范进行接入，所有提供出去的接口模式统一，这样大大降低接入方的难度。

■ 路由

根据用户标识、版本、服务器IP以及权重规则路由到不同的Online计算插件服务。这样一来可以实现实现流量分流、A/B Test、灰度发布的目的、接口代理

■ 接入权限管理

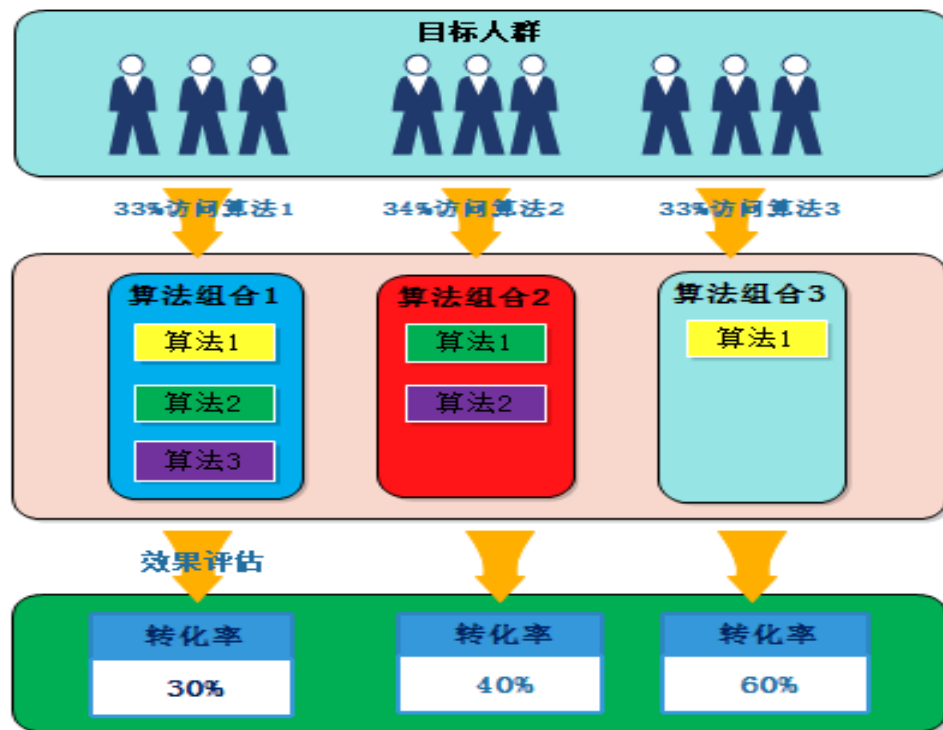
统一管理接口调用权限

■ 统一监控

统一进行业务设用监控，如业务调用量、QPS、响应时长、业务设用失败告警等

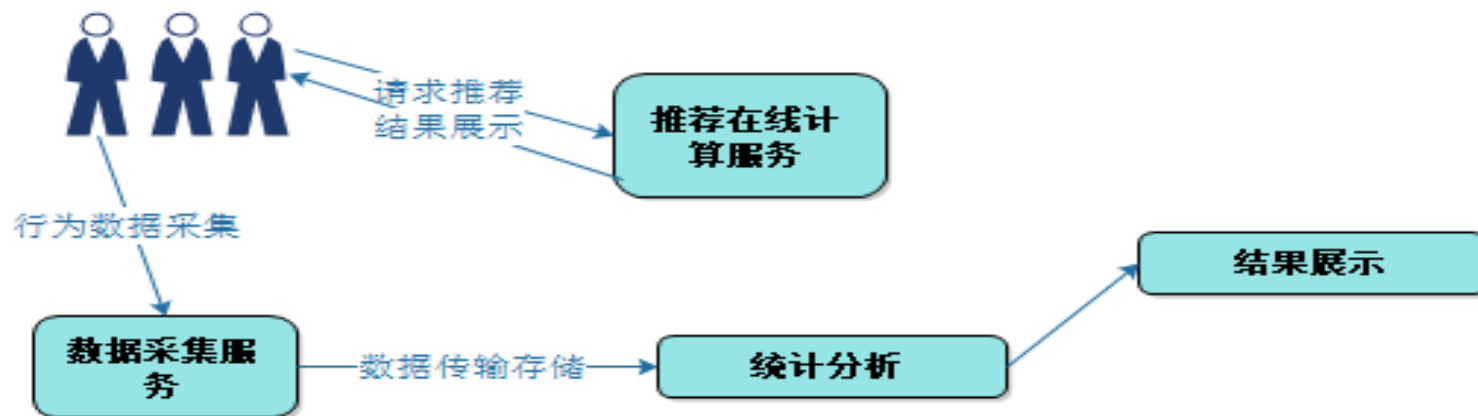
A/B测试模块

在推荐平台中最重要的一个功能就是A/B测试，A/B测试主要是对用户进行抽样分流到不同的算法组合当中，最后通过评估数据来驱动算法工程师对算法效果不断的进行调优



A/B测试效果评估过程

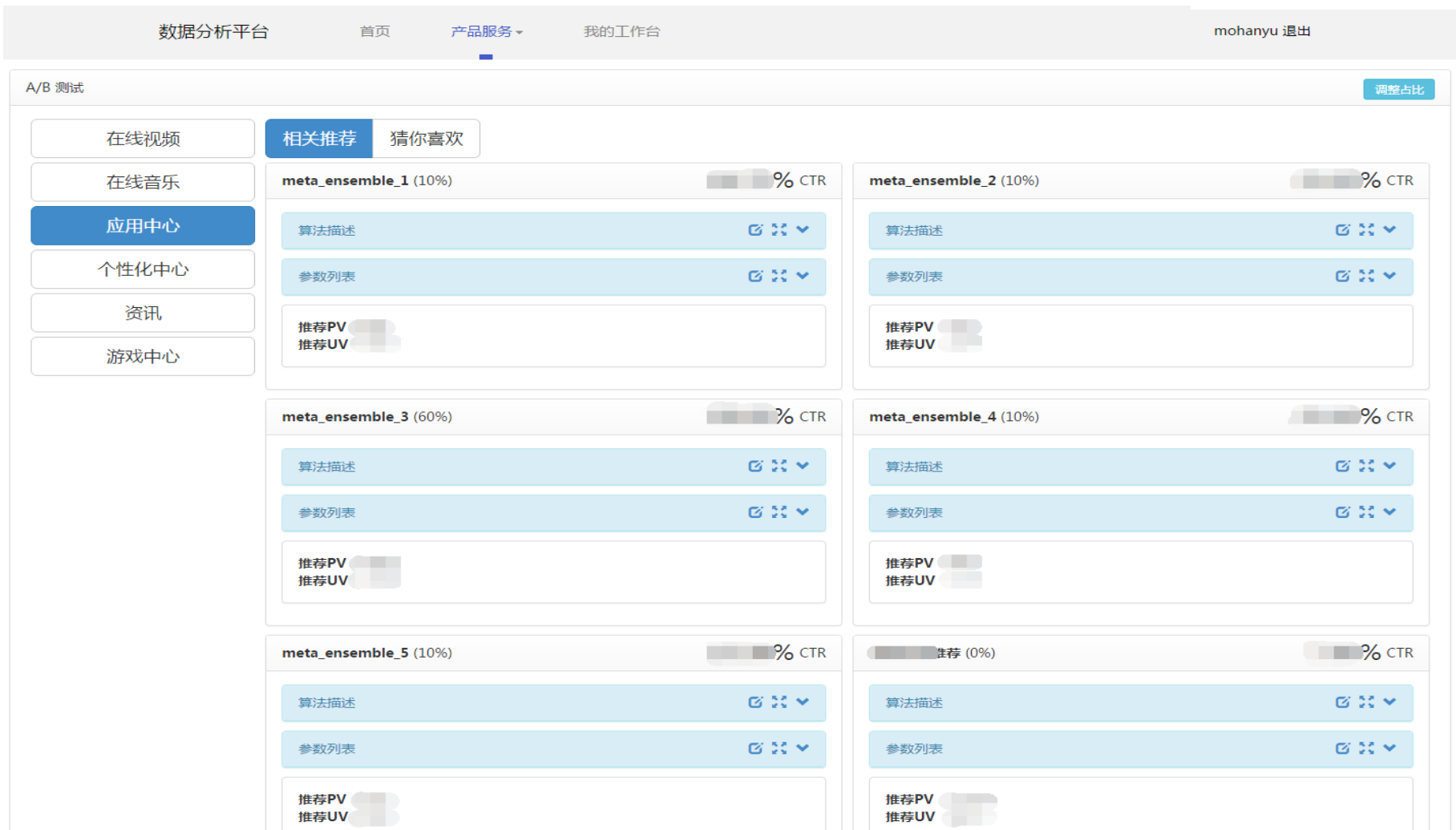
用户请求数据后App端及Web对用户看到的推荐数据所产生的一系列行为进行上报，数据采集服务端对日志数据进行收集并通过流平台将数据进行归并，同时对部分的实时数据进行在线统计分析最终产生效果评估数据



A/B测试效果评估

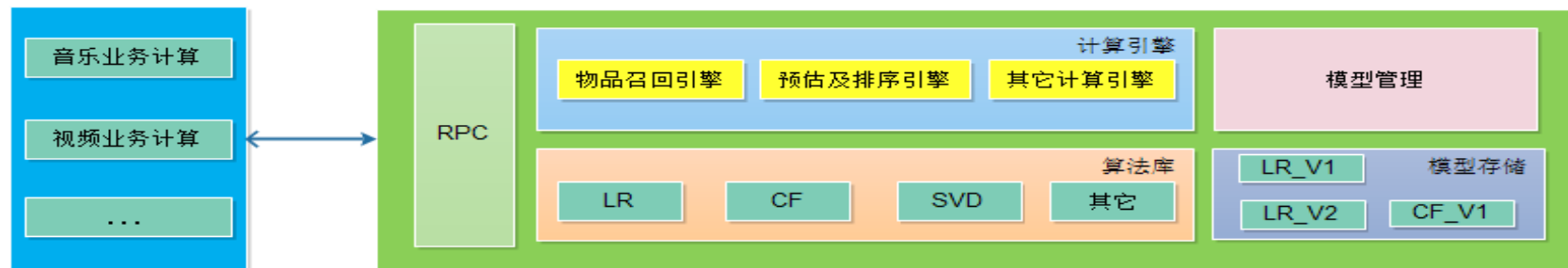
MEIZU

IT大咖说
不止于技术



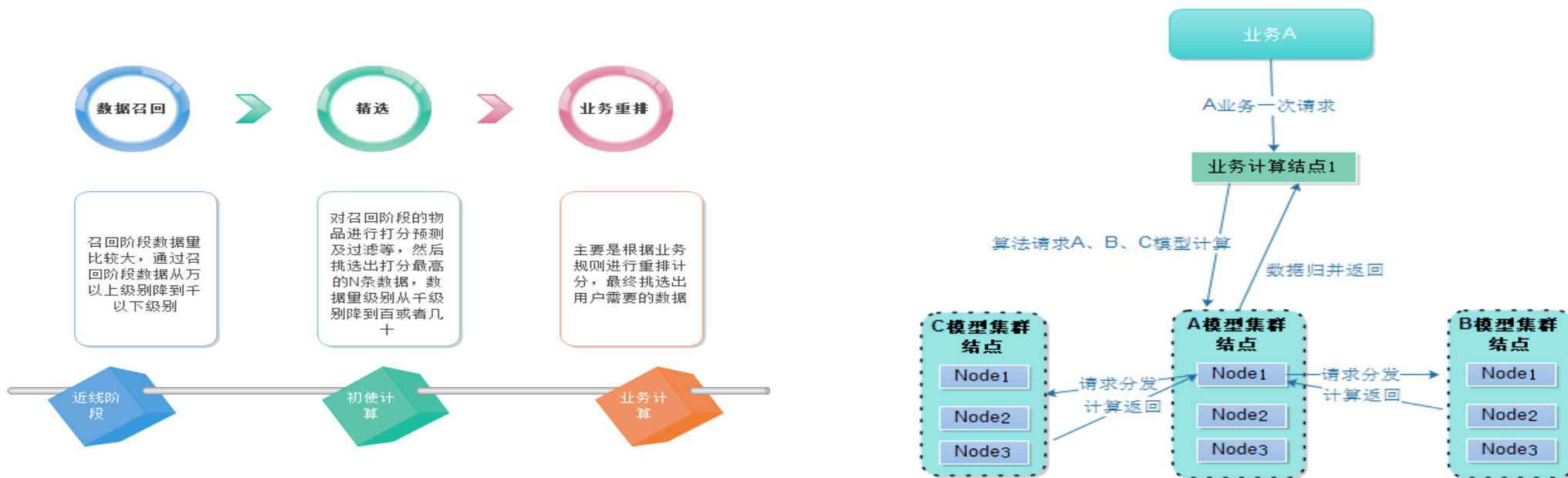
在线模块-计算模块

- ◆ 业务策略计算主要是处理业务相关的一些排序、过滤、人工干预竞价排名等与具体业务相关的逻辑，不同的业务个性化需求采用插件化的方式进行接入
- ◆ 初始化模块主要是对物品进行精选相关的计算，同时管理对新的算法的支持及模型的存储



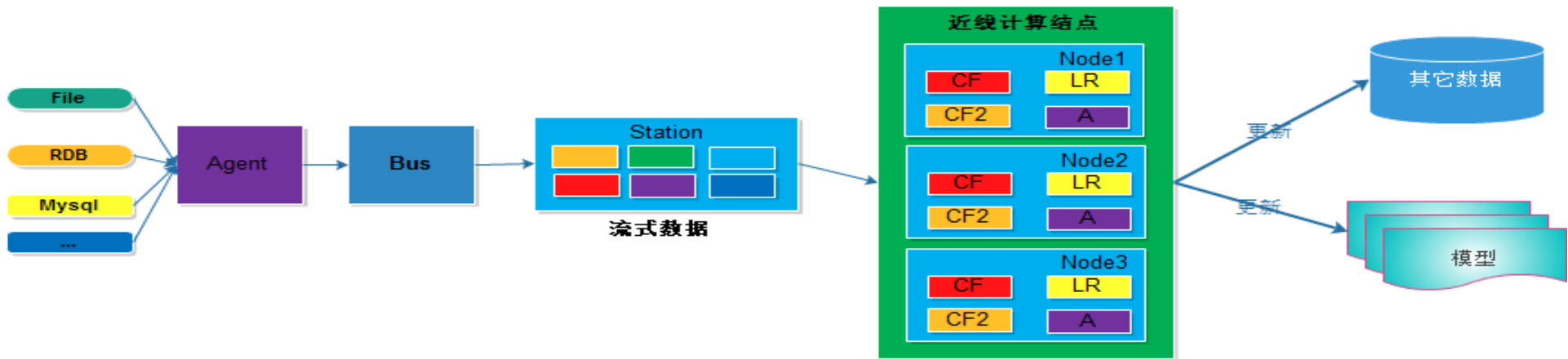
在线模块-计算模块

推荐一般性的数据处理过程从召回阶段到预测再到业务重排阶段数据量依次减少



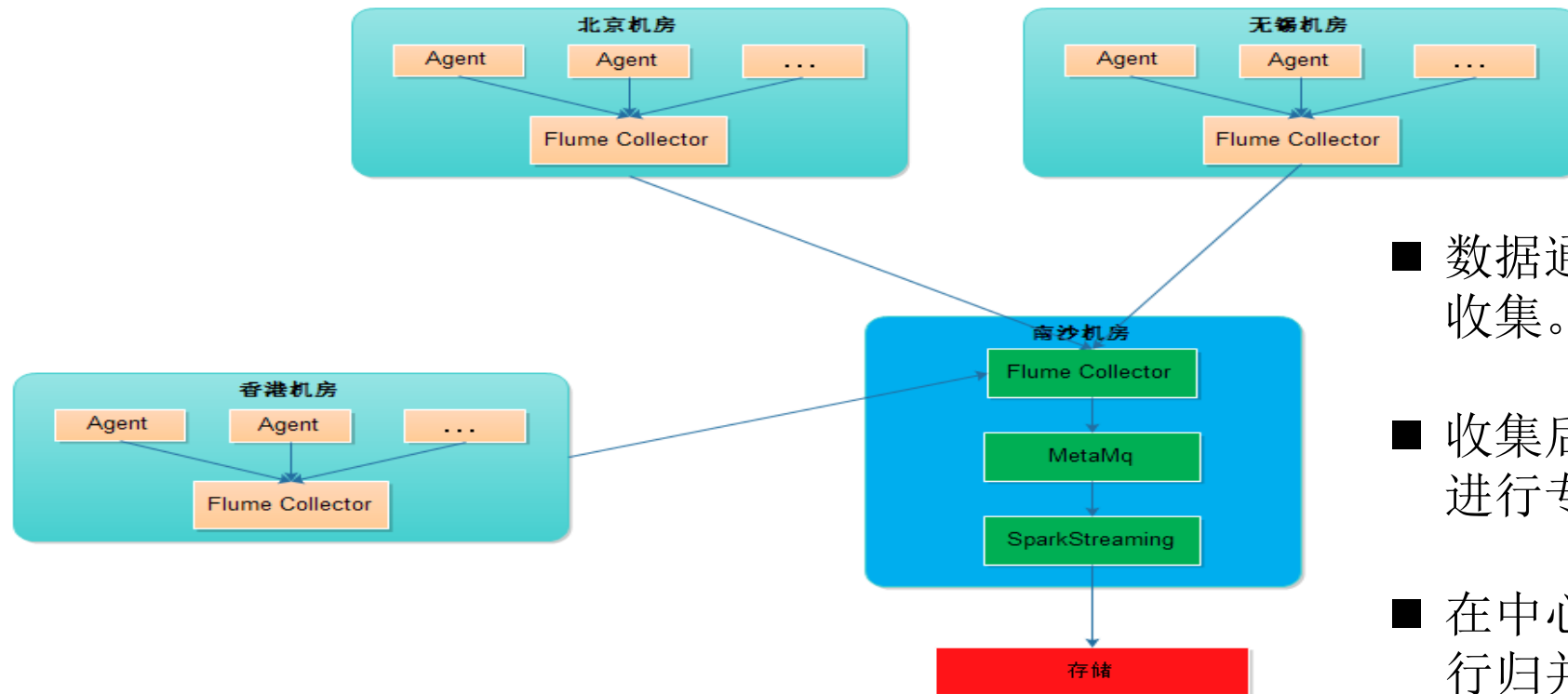
- 精选阶段的数据是来源于召回的数据，有可能同时存在几个或者是十几个召回算法，对不同召回的数据及相关的资源可能存储在不同的机器上或者数据库中，所以请求接收点结在接收请求后需要根据配置将不同的处理请求分发到不同的机器上进行计算然后再归并返回

近线模块



该层主要是利用流式处理的技术对用户实时产生的行为日志进行加工，利用一些高效、高性能的算法生产有价值的数据,如处理算法数据召回、实时数据统计等等

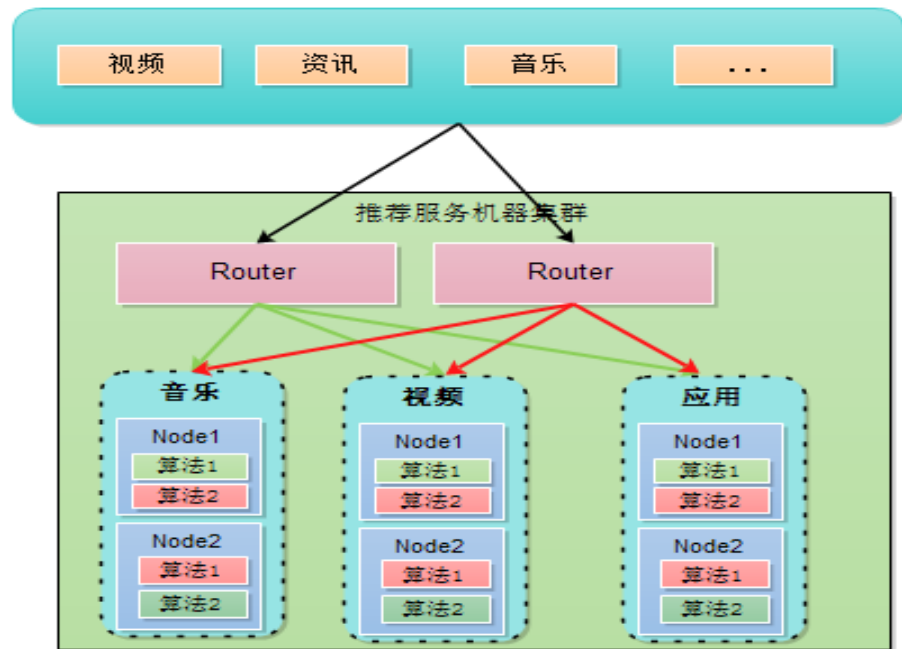
近线模块-流式日志数据传输



- 数据通过Uxip从移动端、Web端进行收集。
- 收集后的数据通过Agent转发到Flume进行专线或公网等网络传输
- 在中心机房的Flume将日志数据等进行归并传输到MetaMq
- 基于MQ消息可以对数据进行流式处理如实时计算、数据落地等等

资源调度管理

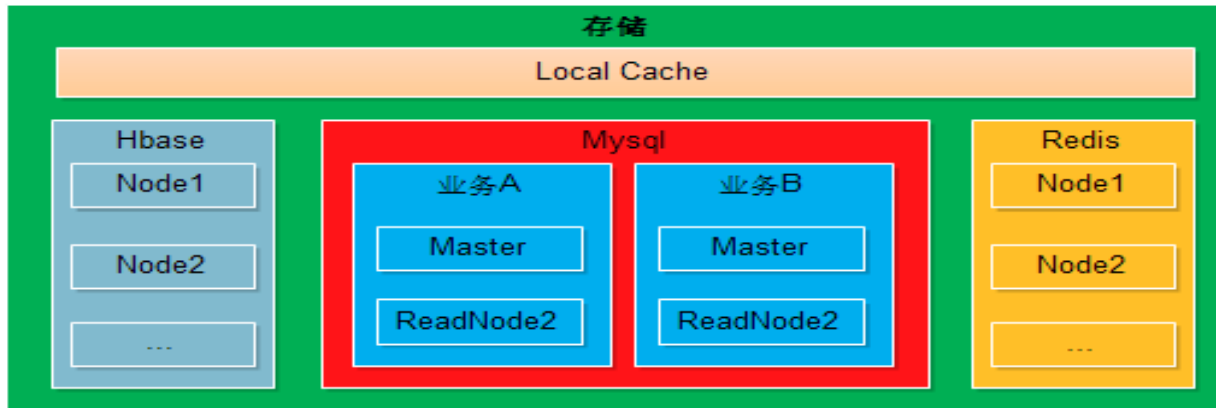
机器动态划分分组，可以按业务进行划分，也可以按照模型资源情况进行划分



- 解决业务之间相互影响,按照业务对性能的要求及复杂对分配不同的硬件机器。同时能够整合资源，不同大小的配置都可以在集群中得到应用
- 解决内存模型存储限制问题，将模型分散到不同的集群中进行横向扩展
- 在请求过程中请求根据Master进行动态调度，大型资源加载过程中机器请求自动调度到其它机器，解决大型资源加载过程中对业务的影响

在线模块-存储

在存储上多样性，不同类型的组合使用，根据不同的场景与性能指标采用不同的存储组合



■ LocalCache

一般用来处理一次请求中访问数据频次超高但数据容量不需要太大的数据，如LR模型数据

■ Mysql、Hbase、Redis

这三种存储的选择一般从性能、和各自的特性出发点来选择最合适的，各自都是集群的方式，MySQL可以按业务数据进行拆分成不同的集群进行访问

离线-机器学习平台

MEIZU

IT大咖说
不止于智人

提供特征工程、统计、训练、评估、预测和模型发布等功能，覆盖机器学习全流程，算法同学可以通过拖拽的方式就能完成模型训练和评估

The screenshot displays the Meizu offline machine learning platform interface. On the left, there is a sidebar with navigation options: 问题类型 (Problem Type), 分类 (Classification), 回归 (Regression), 聚类 (Clustering), and 推荐 (Recommendation). The main area shows a configuration form for training a model, including fields for 数据集ID (Dataset ID), 模型输出路径 (Model Output Path), 模型名字 (Model Name), 训练数据集比例 (Training Dataset Ratio), Driver 内存大小 (Driver Memory Size), 每个 Executor 内存大小 (Memory Size per Executor), 每个 Executor 核数 (Number of Cores per Executor), and Executor 总数 (Total Number of Executors). Below the form, there is a note about algorithm parameters and a button for 逻辑回归LR (Logical Regression LR).

In the center, there is a table titled 模型信息 (Model Information) showing the status of models. It is divided into 正在运行的模型 (Running Model) and 已完成模型 (Completed Model). The table columns include Model Id, Problem Type, and other details.

On the right, there is a window titled 模型ROC曲线 (Model ROC Curve) showing a plot of True Positive Rate (TPR) versus False Positive Rate (FPR). The curve is a red line with circular markers, showing a sharp increase in TPR at low FPR values, indicating good model performance.

At the bottom right, there is a table showing job execution details, including Job Name, Stages, and Start Time. The table lists jobs such as Training ROC, Test ROC, Training PR, and Test PR, along with their respective start times.

- 模型训练及评估界面化，与调度平台无缝集成，使得算法离线模型处理及模型发布上线等更加高效简单
- 系统集成多种算法可进行逻辑回归LR、聚类Kmeans、模型LDA、协同过滤CF等多种模型训练
- 分布式数据处理与计算



- 细粒度性能监控，可以细粒度到具体的业务请求接口，从业务QPS、PV量、响应时长（P50,P70,P99,P999...）等等
- 应用服务器及操作系统各指标监控
- 业务指标监控，如算法效果及其它业务指标监控
- 监控指标可根据具体的需求扩展

- ▶ 推荐介绍
- ▶ 魅族推荐平台架构演进
- ▶ 魅族推荐平台现状
- ▶ **魅族推荐平台挑战和愿景**

挑战 and 愿景

- 挑战10亿/每天以上在线实时计算请求PV数
- 支撑起百亿条/每天的日志进行实时计算，毫秒级别的进行用户模型更新
- 支撑更多的特征集计算，同时在线计算响应时长更加的短
- 支撑更多的魅族产品线业务
- 推荐平台对外开放，能为行业其它的企业提供专业的推荐服务
- 深度学习集成

谢谢！

Q & A