



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

腾讯个性化精准推荐

TEG/数据平台部/实时计算中心
regzhang(张磊)
5/17/16



目录

- 推荐业务介绍
- 精准推荐架构的演进
- 推荐引擎设计与实现
- 总结与未来展望

推荐业务介绍



DTCC

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

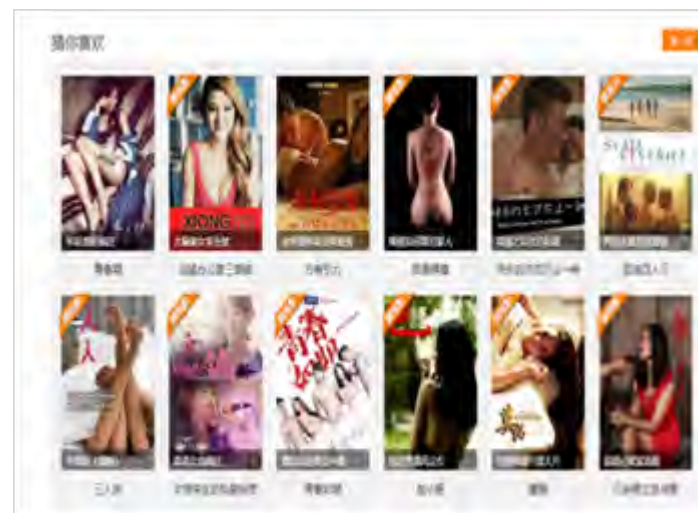
SequeMedia

IT168

ChinaUnix

mpub

推荐业务介绍



DTCC

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

SequeMedia

IT168

ChinaUnix

ITPUB

我们所做的

- 预测未来
- 预测每个人的未来
- 实时预测每个人的未来

核心技术挑战：

如何在100ms的时间内，面对8亿活跃用户中的任意一人，预测其在未来下一秒内，最可能点击10,000个广告主提供的250,000个广告中的哪一个？

难点：

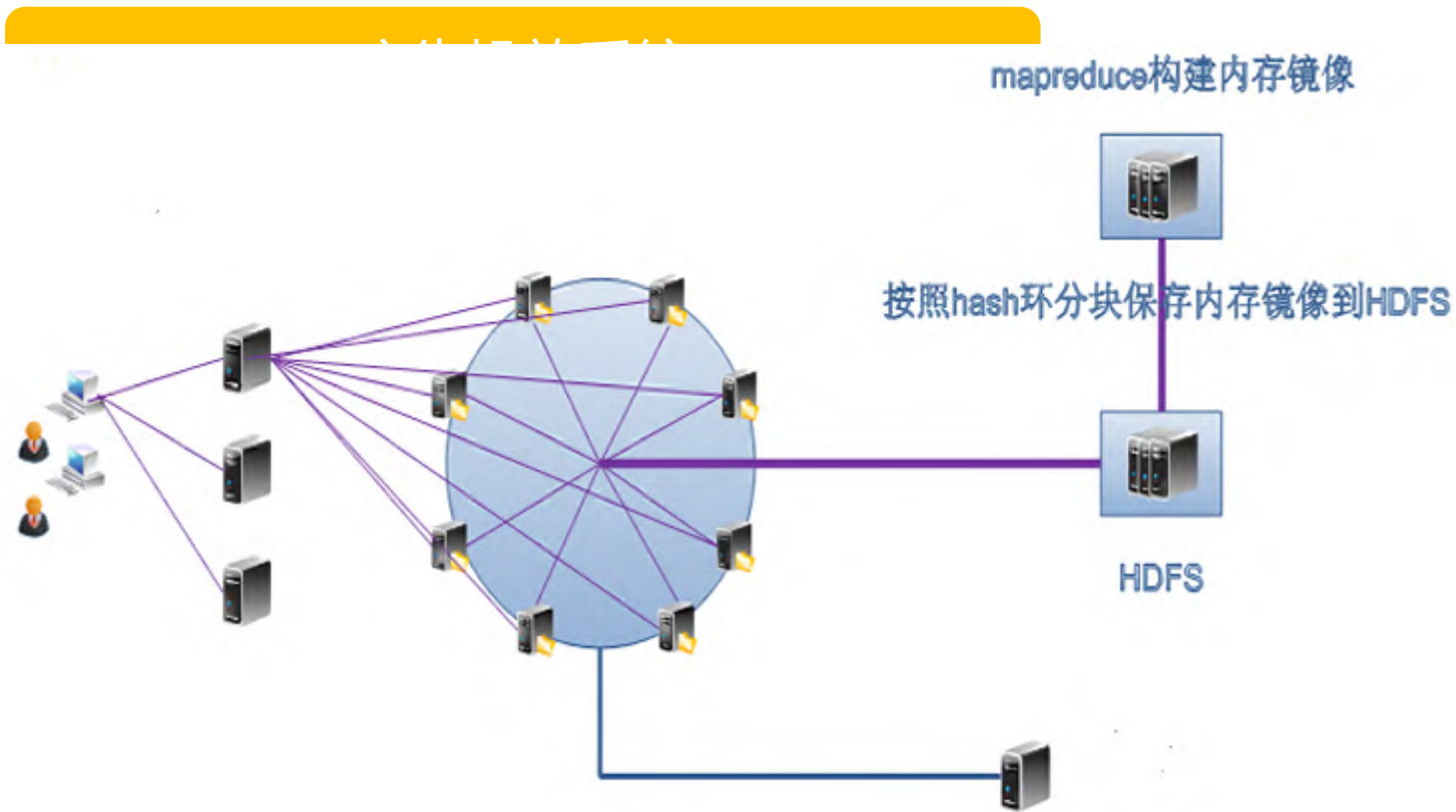
- 1、从1万个广告主、25万个广告中挑出最合适的一个。最合适？
- 2、8亿活跃用户，每个人的结果都是个性化的。为什么？
- 3、100ms内完成，移动端50ms
- 4、效果广告-广点通每天即执行120亿次实时预测请求

仅为了这120亿次预测：

- 1、每天仅实时点击预估算法预测的广告数即超过12000亿个
- 2、每天实时接入2PB、50000亿条数据
- 3、每天流式计算的事件数超过36000亿次
- 4、每天实时数据存储层访问超过5000亿次



精准推荐第一代架构



□核心特征：

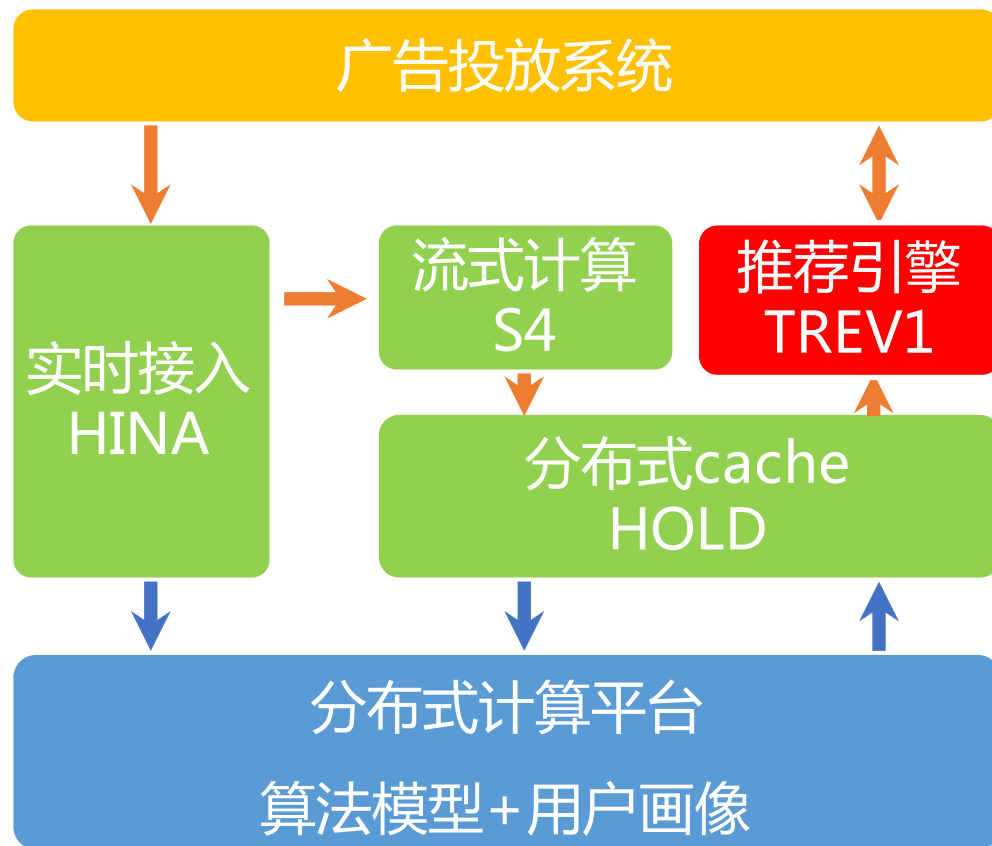
- 离线文件传输
- 批量暴力计算
- 实时匹配查询

□存在的问题：

- 数据时延高
- 人群聚类粗
- 扩展性差



精准推荐第二代架构

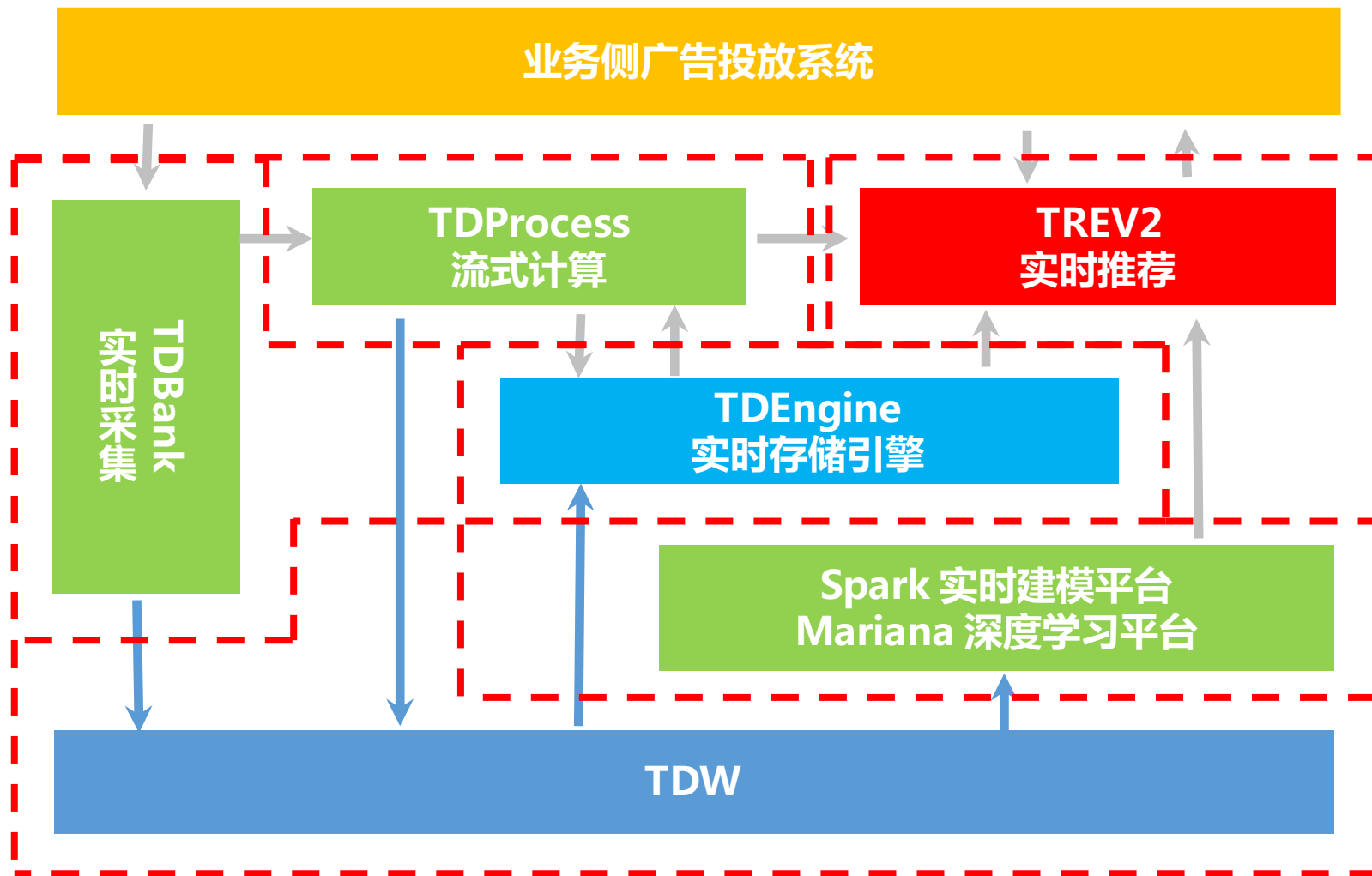


- 特征：
 - 实时数据传输
 - 在线实时计算

- 优点：
 - 秒级延迟
 - 低耦合
 - 易扩展



精准推荐第三代架构



用户画像



□ 用户群体

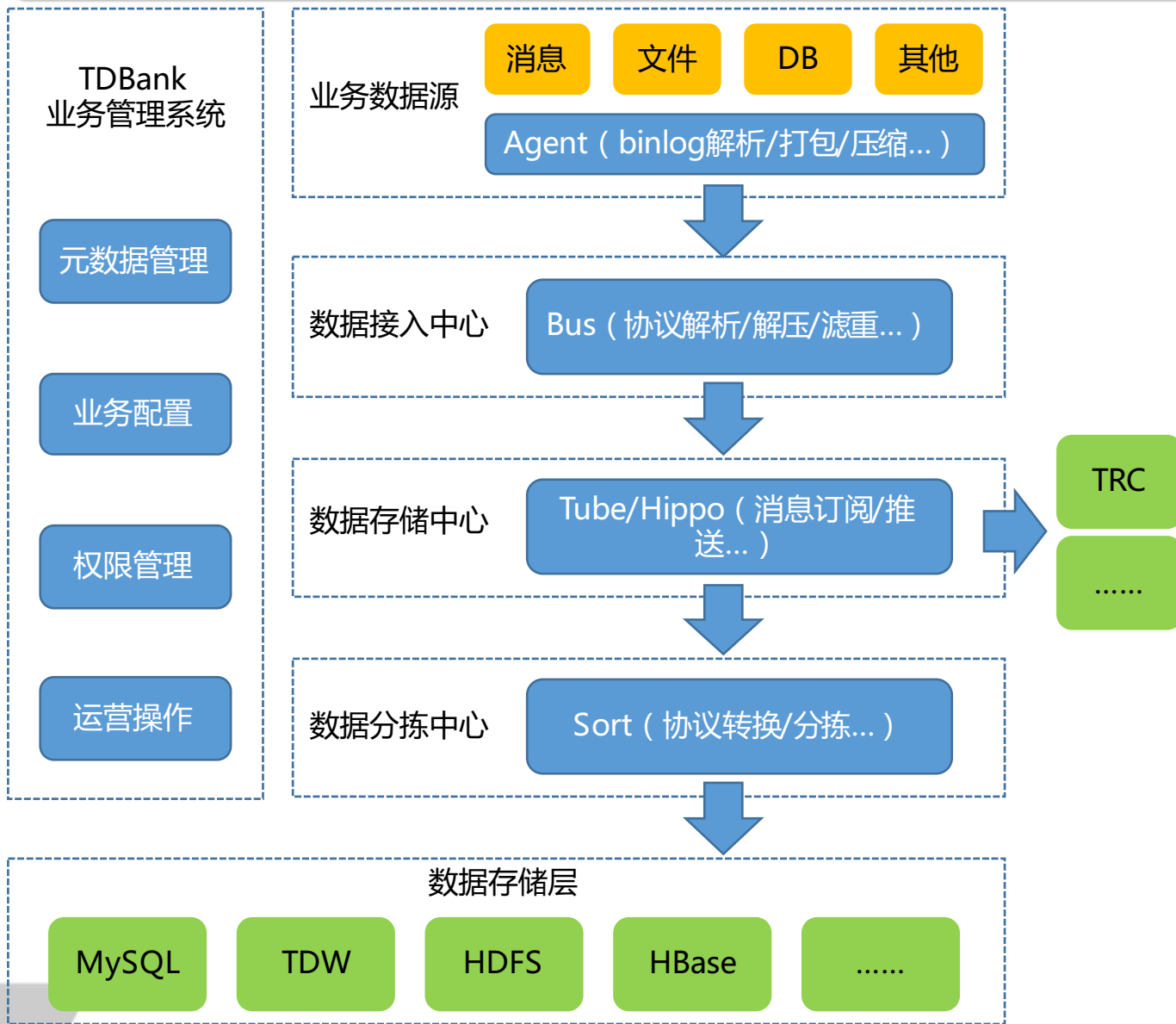
- QQ月活8.5E
- QQ智能终端月活6.42E
- 微信和WeChat月活6.97E
- QQ空间活跃6.4E
- QQ空间智能终端月活5.73E

□ 用户画像

- 用户基础属性
- 行为兴趣属性



TDBank实时接入



□ 核心需求

- 海量数据接入
- 秒级接入延时
- 低成本、高效率、强安全
- 方便数据管理和使用

□ 主要难点

- 数据总量巨大
- 数据源种类繁多
- 数据格式各异
- 数据分布IDC众多

□ 特色功能

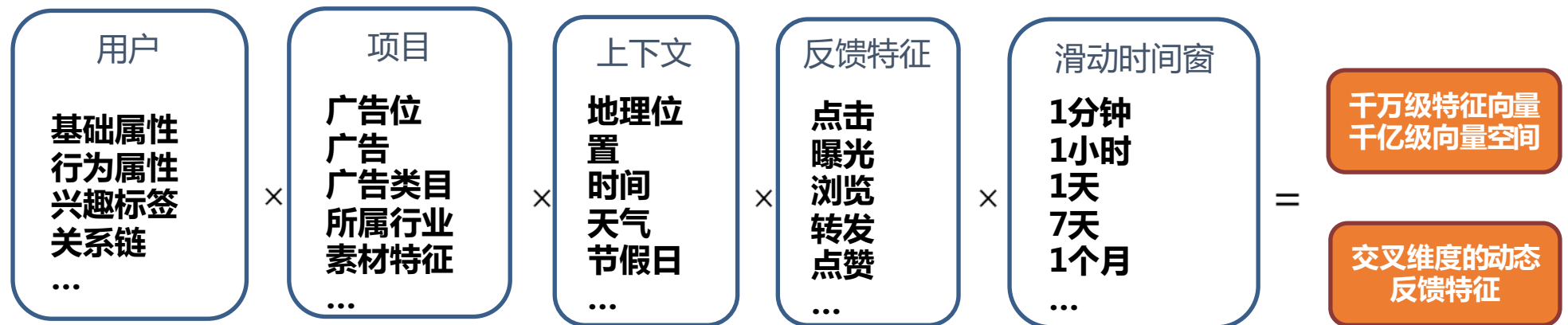
- 配置驱动
- 多种数据源适配
- 平衡读写
- 支持多种序列化协议
- Hippo支持事务



TDProcess流式处理

预测用户A最可能点击广告，如何准备好预测相关数据？

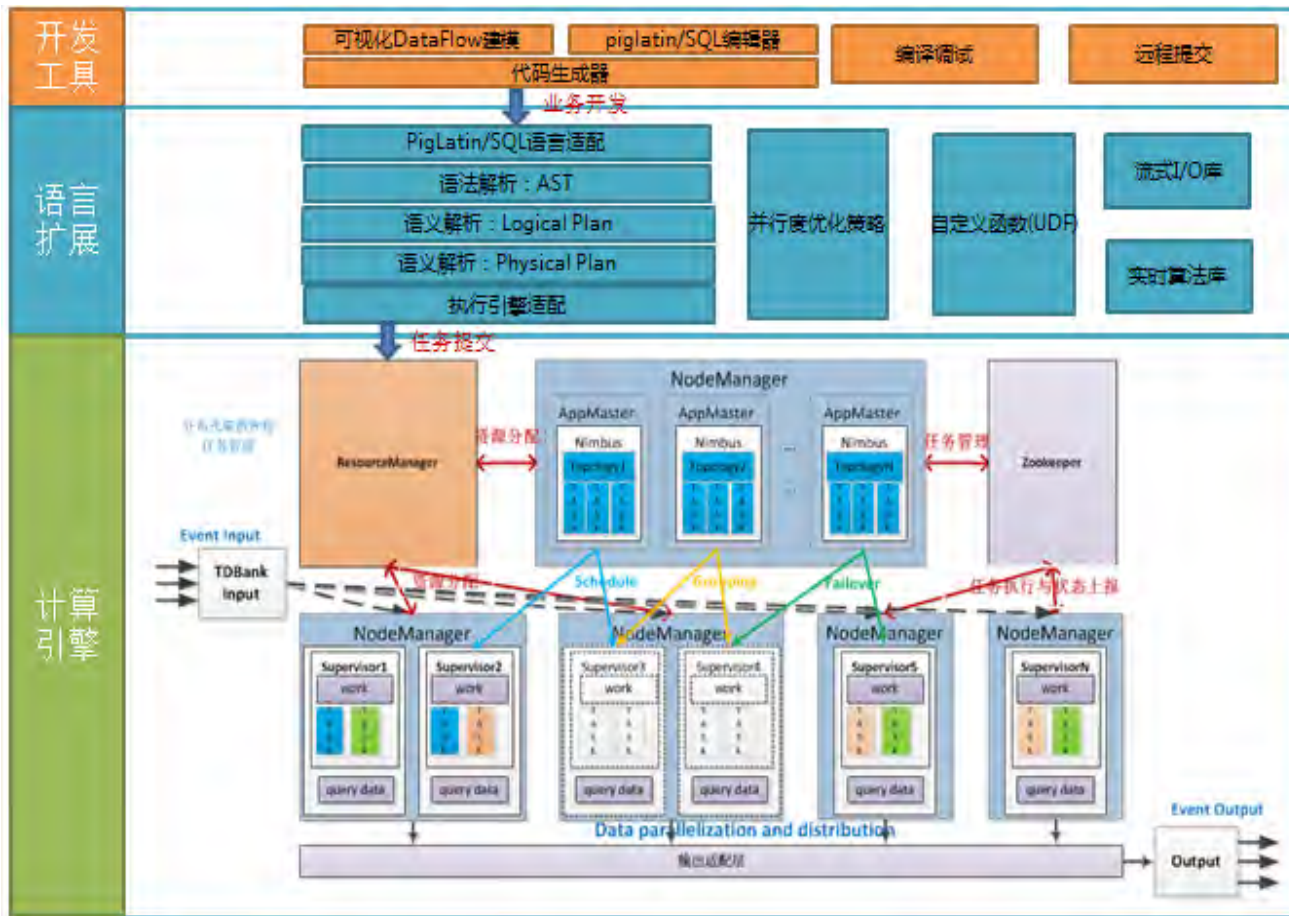
- 对每个广告，实时计算“用户*广告*时间”多个不同维度组合的指标



- 广点通120亿请求对应的每1条曝光日志，平均计算50多种交叉特征
- 仅广点通业务每天实时计算量超过6000亿次
- 整个集群的计算量超过36000亿次
- 秒级滑动时间窗，每一秒的数据都在变化



TDProcess流式处理



□ 核心需求

- 基于消息的流式处理
- 分布式（方便线性扩展，应对数据的爆炸式增长计算）
- 实时(数据消费计算延时毫秒级)

□ 难点挑战

- 高度容错(分布式系统，宕机是常态)
- 任务间资源隔离（多任务并发不相互影响）
- 基于资源消耗自动扩容/缩容
- 业界无成熟方案

□ 特色功能

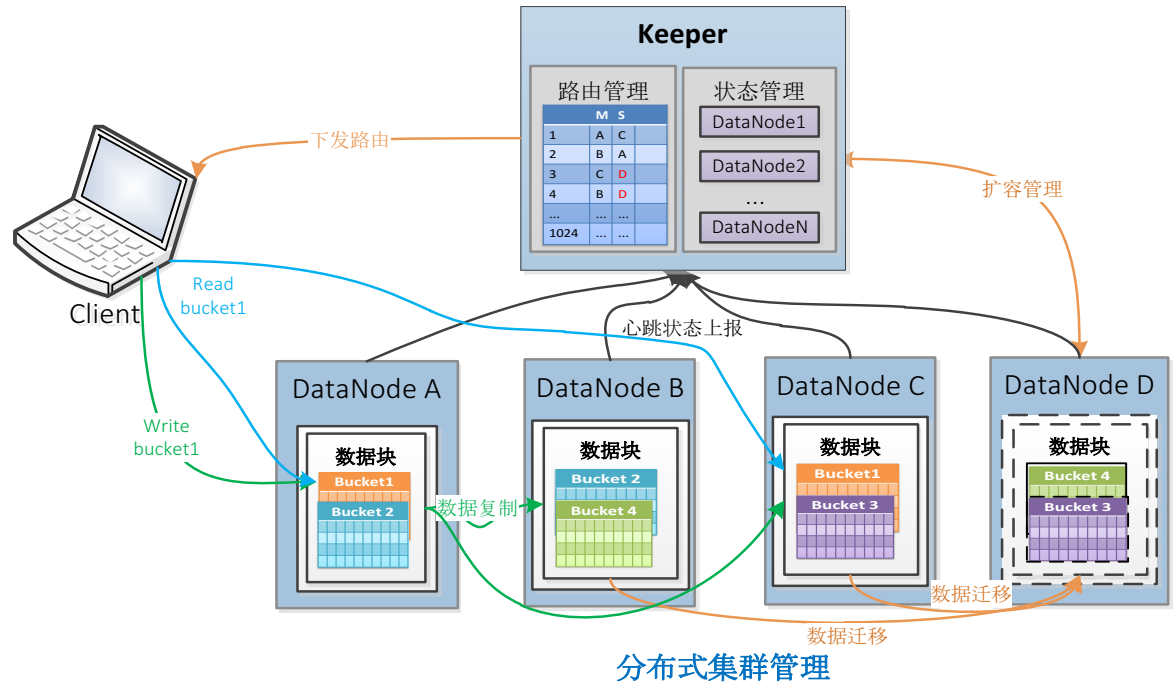
- 服务级功能抽象
- 统一的资源管理
- 多种编程接口



TDEngine存储引擎

□ 核心需求

- 高并发，低延迟
- 高可用性，数据安全
- 关注成本，关注资源利用率
- 线性扩展



□ 特色功能

- 支持多副本数据备份，确保数据安全
- 主备机同时提供服务，提升集群资源利用率
- 集群高可用，容灾切换过程中仍然提供读写服务
- 全内存设计，多引擎支持

每天支撑5000亿数据访问请求



DTCC

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

SequeMedia

IT168

ChinaUnix

ITPUB

TRE推荐引擎挑战

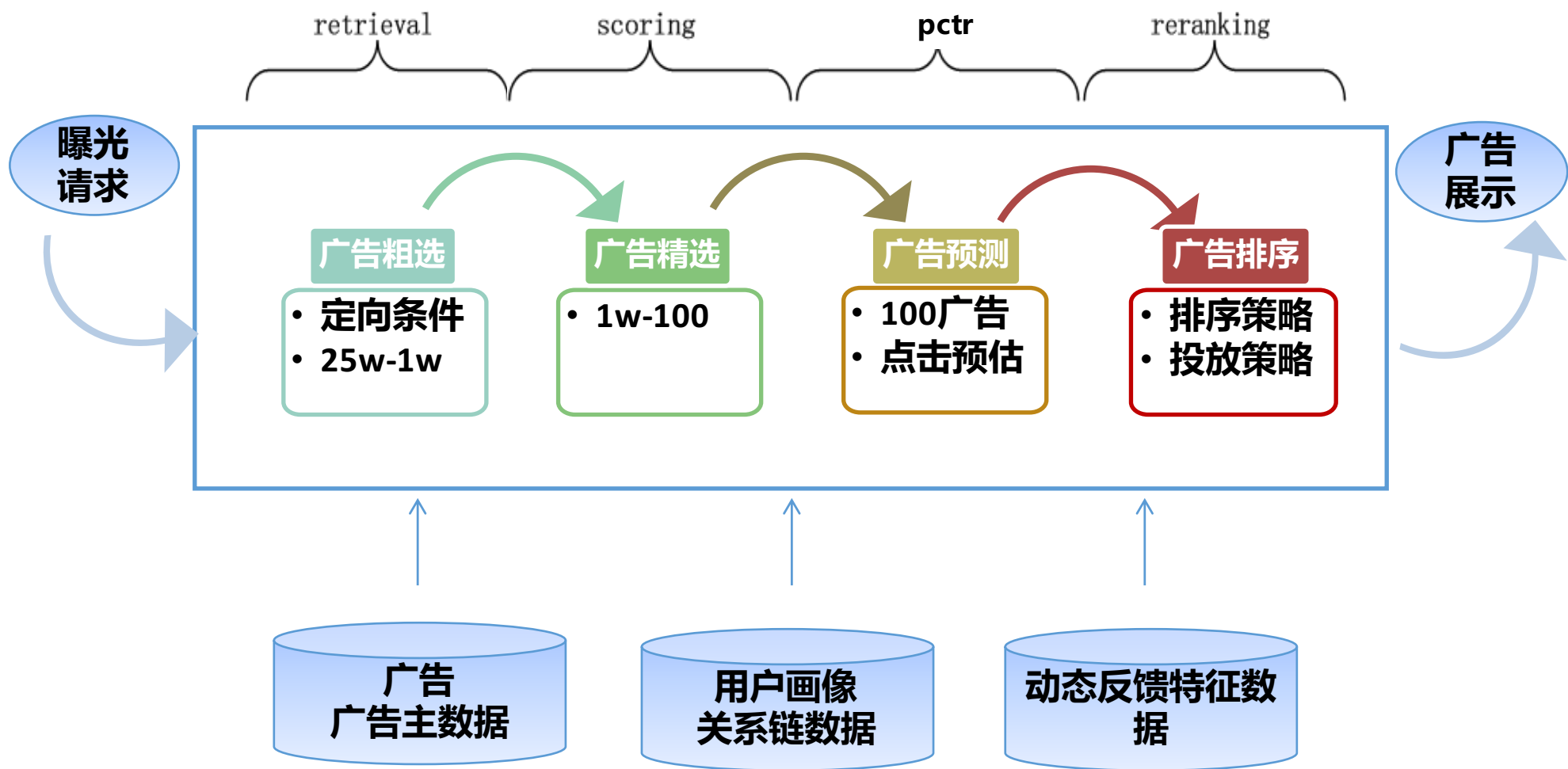
效果广告推荐的核心技术挑战：

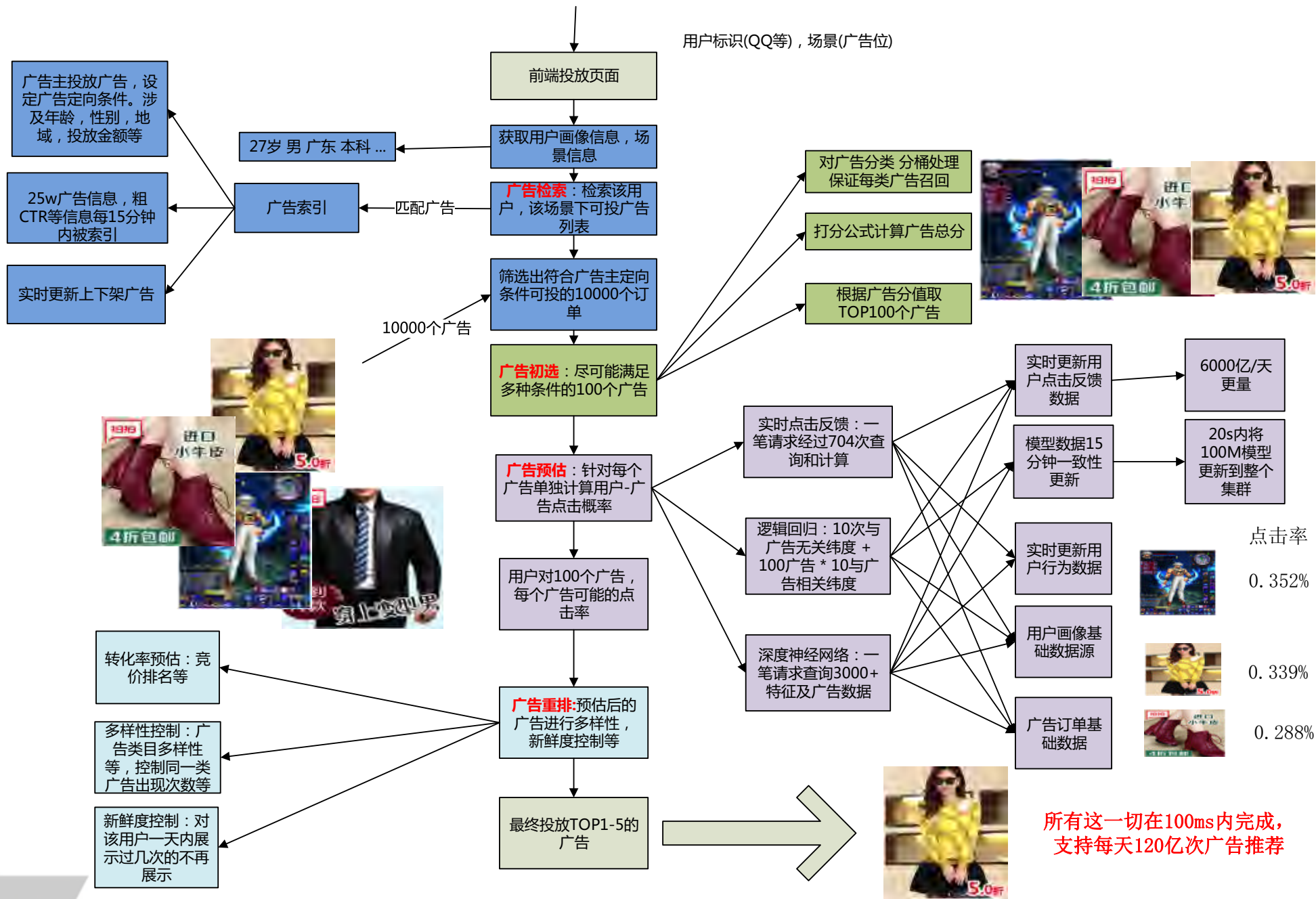
如何在100ms的时间内，面对8亿活跃用户中的任意一人，预测其在未来下一秒内，最可能点击10,000个广告主提供的250,000个广告中的哪一个？

- “8亿用户↔25万实时变化广告”的相关度计算非常复杂
- 100ms内完成，移动端50ms
- 广点通每天执行120亿次这样的预测



效果广告的在线业务流程





TRE推荐引擎难点

业务逻辑复杂

每次请求，经过实时检索，初选，精选，重排等多个业务流程
分布式任务管理，调度

每次请求，对100个广告进行精确预测、排序
预测每个广告点击率，需要访问30多种数据
每秒9亿次数据访问量

每个请求100ms返回
移动端要求50ms

业务流量大

每天为8亿用户执行预测120亿次预测，峰值每秒30万次
每天执行12000亿“用户↔广告”相关度预测(精排序)，峰值每秒3000万次

算法管理

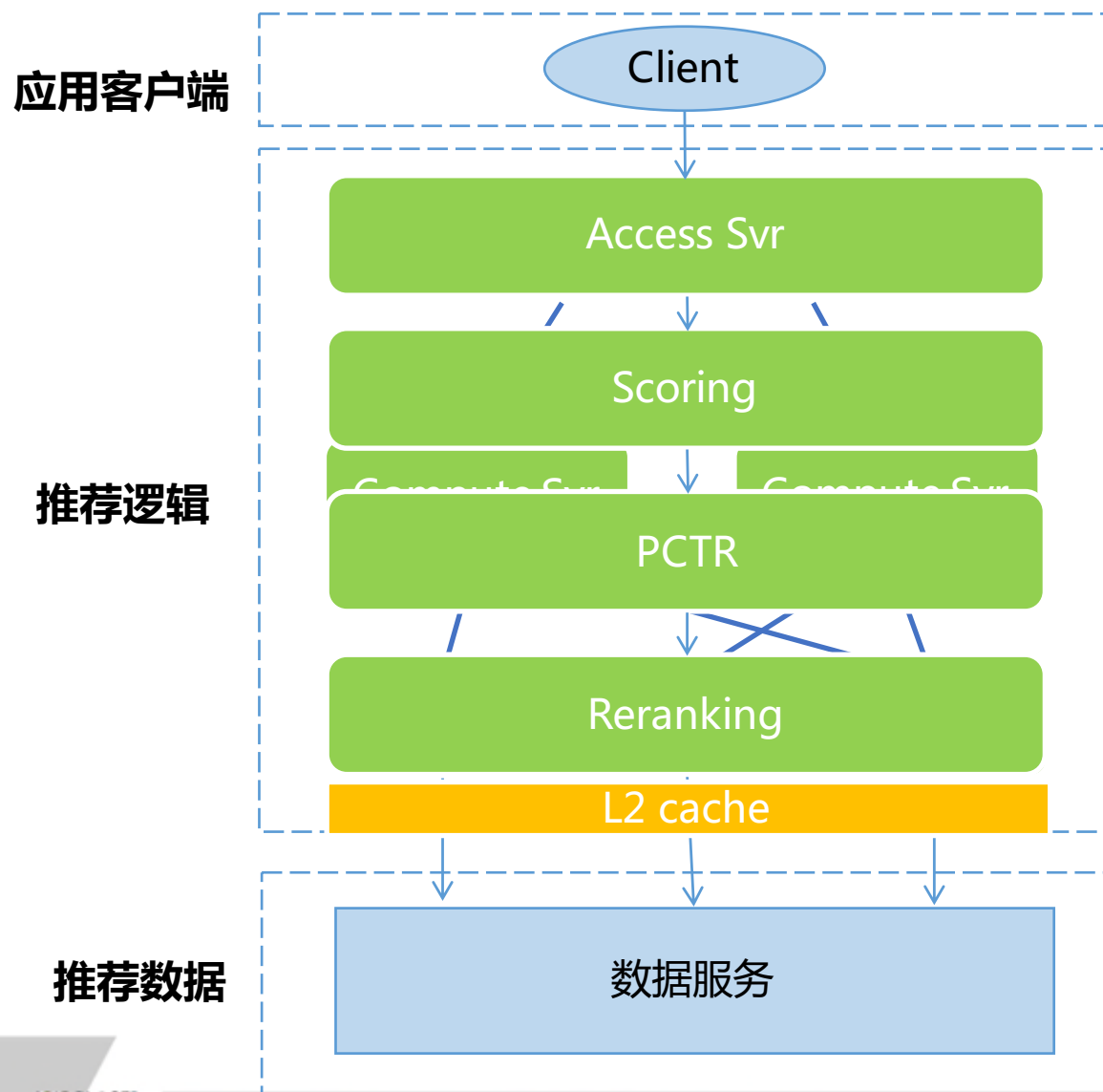
在线和实验算法100多个，15分钟模型数据迭代更新一次，每天更新9600多次

集群管理

上千台在线计算服务器的管理



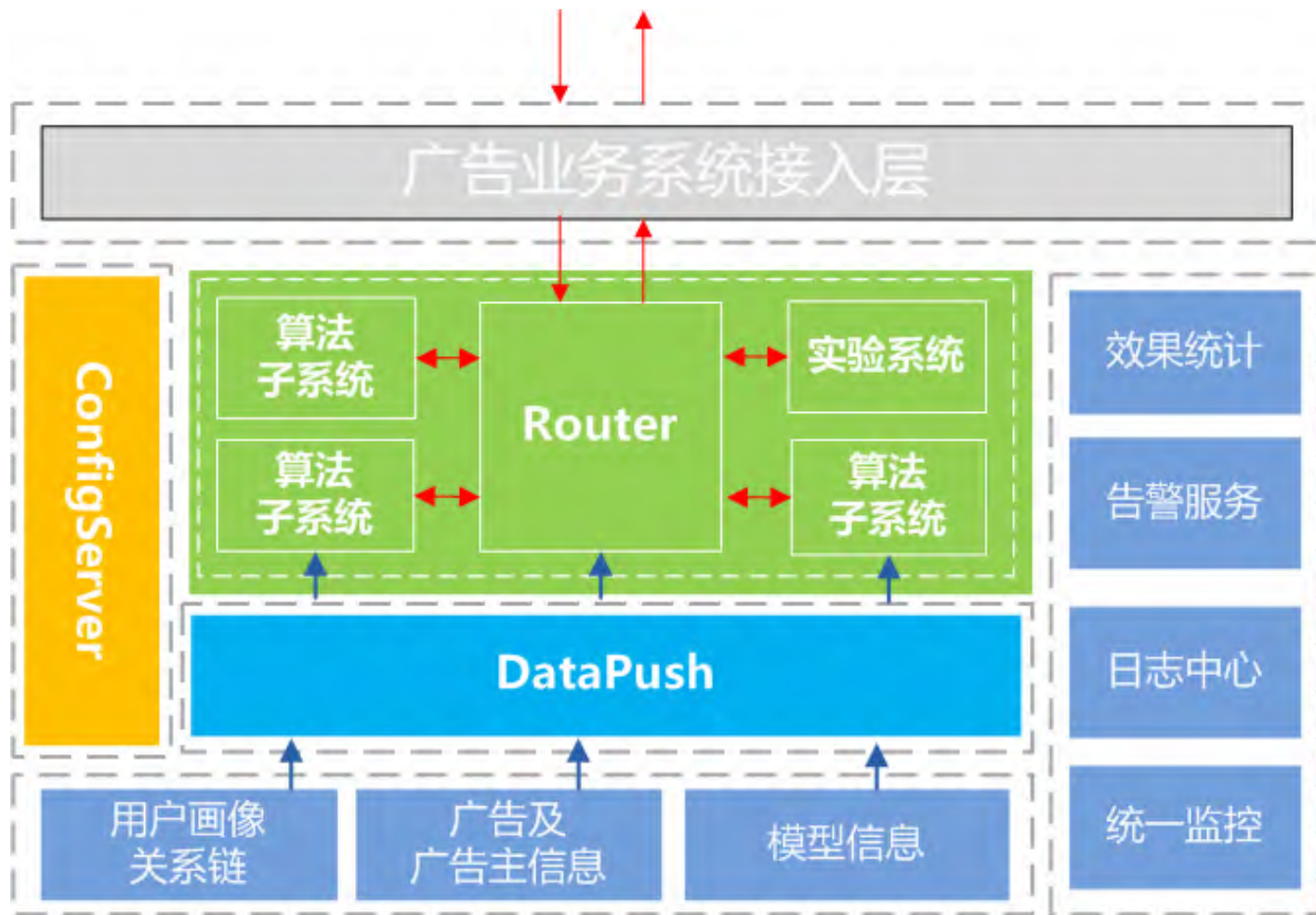
TRE一代架构



- **核心需求**
 - **海量+实时**：广告计算复杂响应时耗要求高
 - **大规模分布式系统流量、算法和数据管理**
- **主要痛点**
 - **扩展难**：业务定制化，分支版本众多，管理困难；数据层层传递，下游变更，所有上游计算层都跟着变；质量监控依赖外部系统 join，成本很高
 - **静态路由，关联系统各自为政**：路由事先静态生成，无法动态变化；系统模块增加，但相互独立；配置繁多，依赖性配置人工出错频繁
 - **可运营性差**：加载算法状态，缓存状态不透明；更新算法子系统配置需重启，集群无法动态扩缩容。



TRE二代架构

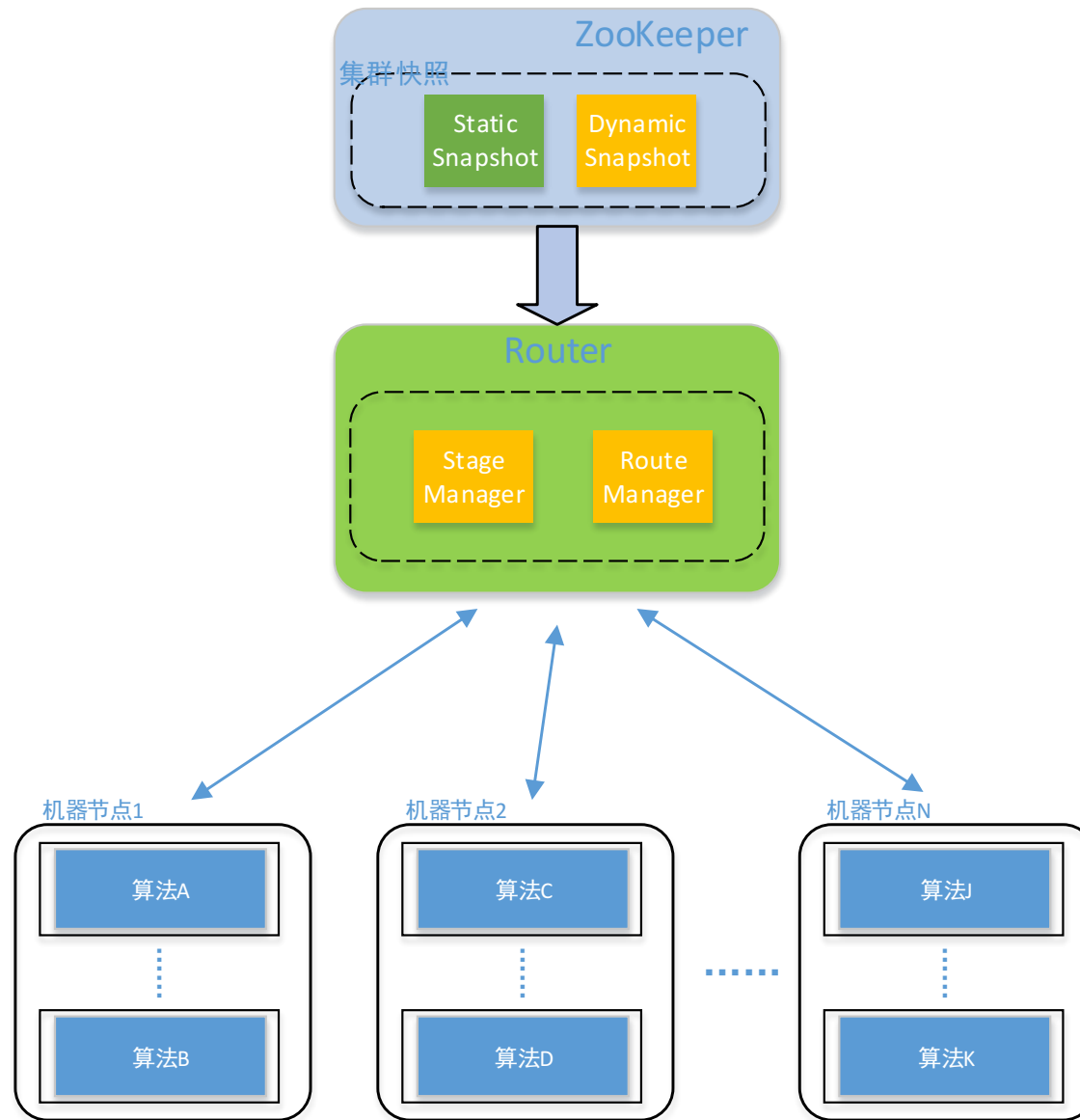


- **主要优点**
- **易扩展:** 通过中心节点 Router 进行不同的业务流程控制, 数据管理, 多模块日志join
- **动态智能路由:** 集群管理中心提供动态路由, 关联系统根据同一份路由数据进行数据和请求的联动变更
- **集群透明:** 集群管理中心收集节点元数据, 包括配置, 算法, 节点状态等, 前台可展示, 运维监控更方便

—→ 曝光请求流程

—→ 数据推送流程

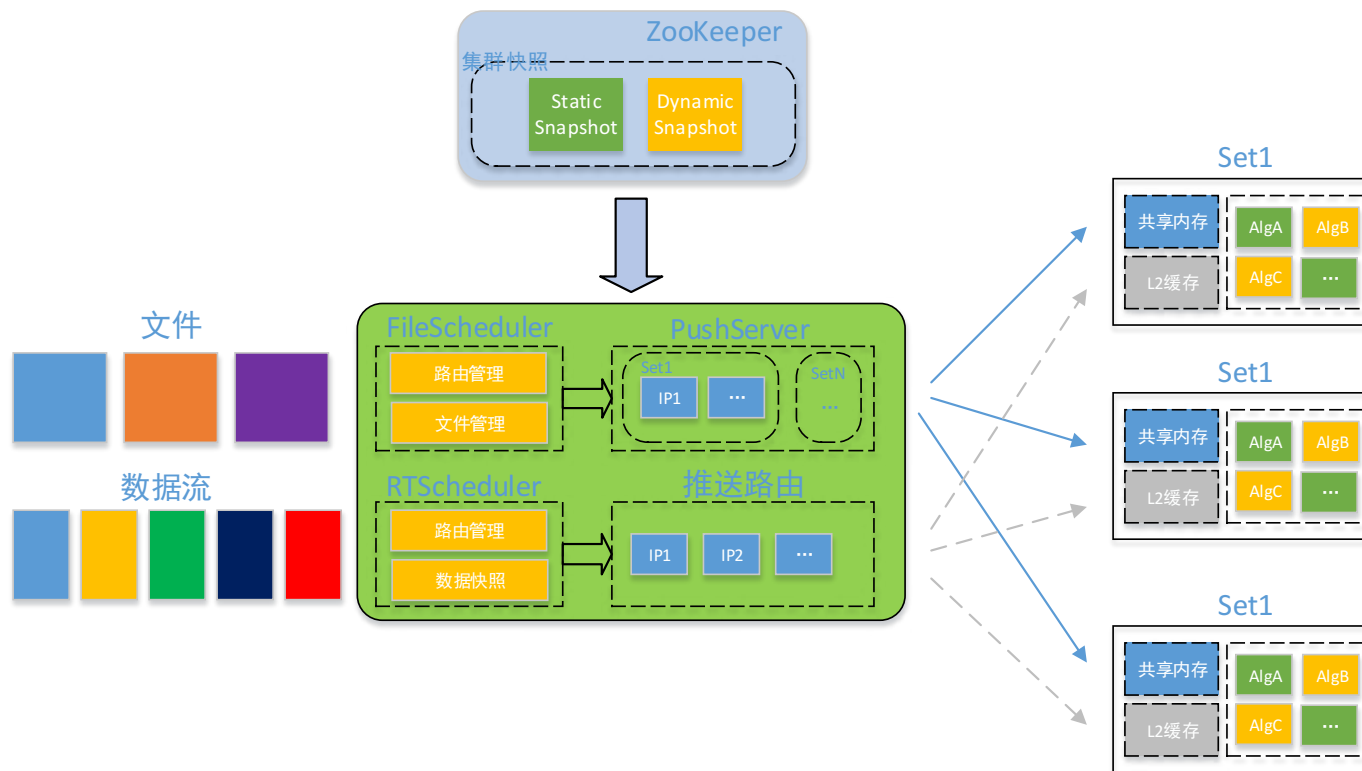
TRE算法管理



- 核心需求
 - 针对不同流量建模
 - 支持100+在线，实验算法效果调优，频繁更新，上下架
 - 支持多种业务流程
- 特色功能
 - 插件式管理
 - 支持动态上下架
 - 算法作为系统调度的路由依据
 - 接口解耦，状态机分离
 - 配置驱动流程调度



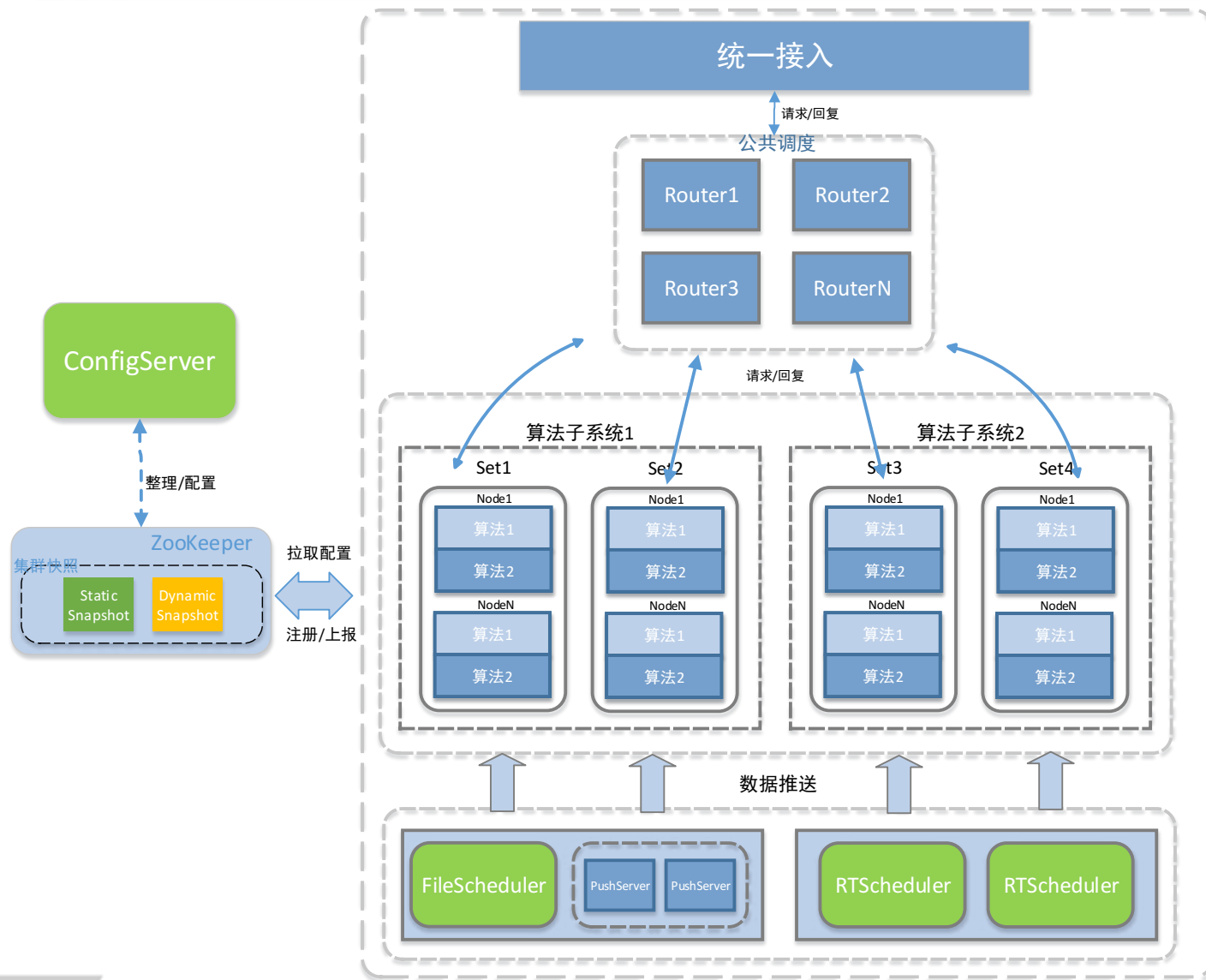
TRE实时数据更新



- 核心需求
 - 保证算法模型一致性
 - 支持100+在线实验算法模型数据推送
 - 每15分钟粒度定期推送到线上几百台服务器
 - 保证关键性数据查询命中
- 特色功能
 - 文件Pipeline任务调度
 - 100MB文件, 10S内发送到集群所有集群上
 - 实时流数据毫秒级更新到server缓存
 - 数据快照, 支持计算节点快速恢复



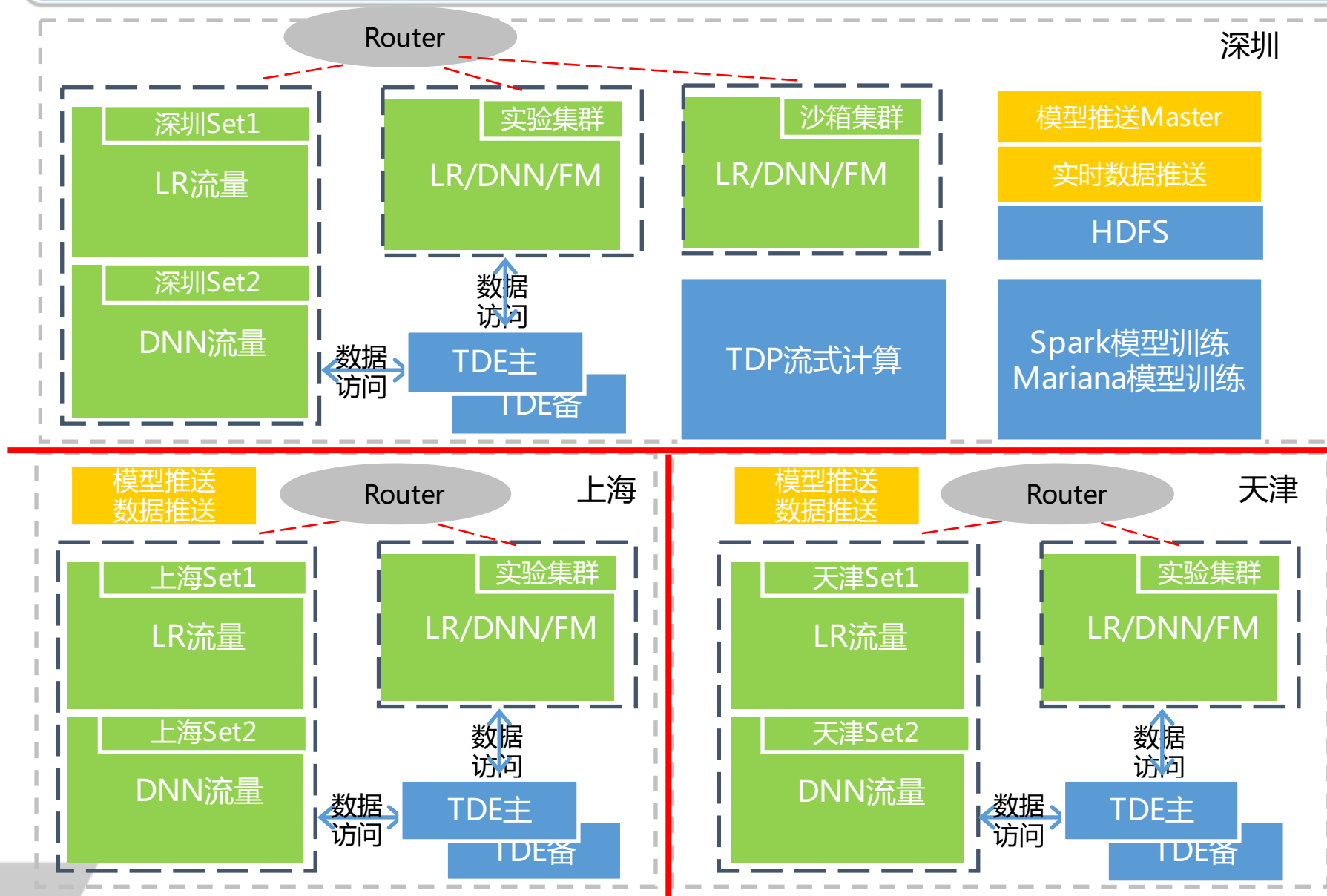
TRE集群管理



- 核心需求
 - 平行可扩容
 - 负载均衡管理
 - 路由管理
- 特色功能
 - 智能路由管理
 - set/ip级别的负载均衡
 - 动态扩缩容, 算法/模型/文件延迟删除机制
 - 状态上报, 收集整理



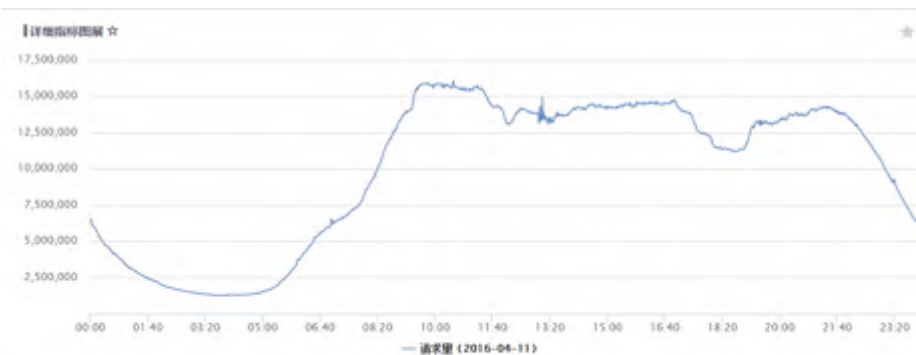
TRE多地容灾



TRE集群监控



- 集群状态监控
- 节点状态
- 算法版本
- 路由版本
- 配置版本

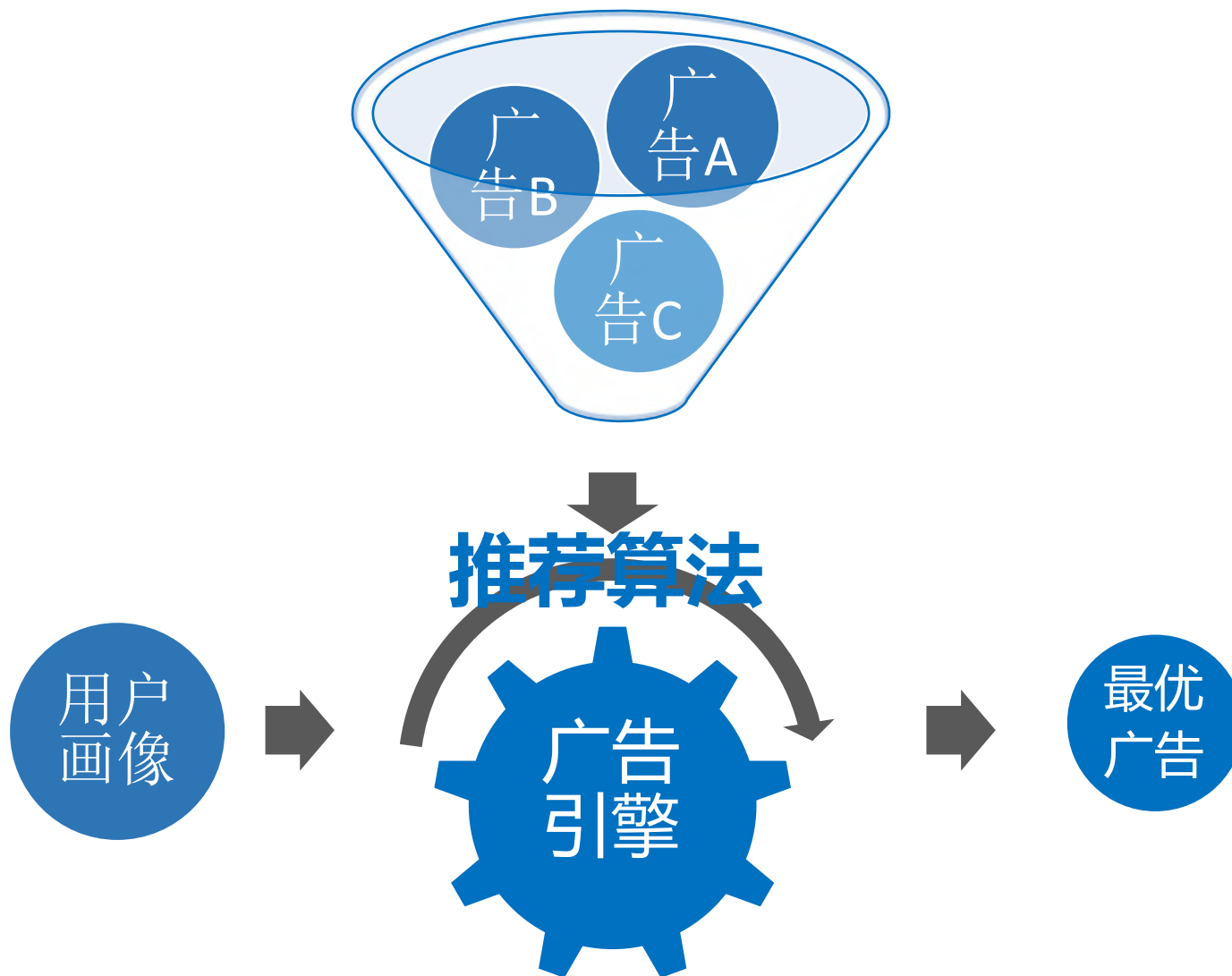


- 秒级指标监控
- 请求执行状况监控
- 算法效果监控
- 收入监控

时间	数量				比率			均值
	请求量	成功量	业务失败量	系统失败量	成功率	业务失败率	系统失败率	平均耗时
总计	14,345,666,438	14,344,647,454	0	918,984	99.9936%	0%	0.0064%	3857654336.250072
23:59	6,391,677	6,391,372	0	305	99.9952%	0%	0.0048%	3.486094
23:58	6,401,690	6,401,566	0	132	99.9979%	0%	0.0021%	3.486552
23:57	6,408,830	6,408,719	0	120	99.9981%	0%	0.0019%	3.506201
23:56	6,488,574	6,488,358	0	216	99.9967%	0%	0.0033%	3.503416



总结



未来展望

□更加实时

- 在线算法

- 实时用户画像

□更加精准

- 更多特征

- 深度学习

□更加通用

- 多业务支撑

- 第三方开放



Join us !

WELCOME TO JOIN US

加入我们



DTCC

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

SequeMedia
数据媒体

IT168

ChinaUnix

ITPUB



THANKS

SequeMedia
威拓传媒

IT168.com

ChinaUnix

ITPUB

THANKS