

QCon[上海站]

全球软件开发大会 2016

百亿级通用推荐系统实践

吕慧伟

Tencent 腾讯

huiweilv@tencent.com



促进软件开发领域知识与创新的传播



关注InfoQ官方信息
及时获取QCon软件开发者
大会演讲视频信息



[北京站] 2016年12月2日-3日
咨询热线: 010-89880682



[北京站] 2017年4月16日-18日
咨询热线: 010-64738142

| 个人简介：吕慧伟

 腾讯

2015年10月至今

- 腾讯云布道师，SNG社交网络运营部高级工程师

- 神盾推荐系统性能优化和架构设计
- 腾讯云推荐引擎架构设计

 Argonne
NATIONAL LABORATORY

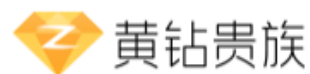
2013年 - 2015年

- 美国阿贡国家实验室博士后
- MPI开源版本MPICH核心开发者

 中国科学院计算技术研究所
Institute Of Computing Technology Chinese Academy Of Sciences

2007年 - 2013年

- 中国科学院计算技术研究所博士
- 专业方向计算机系统结构（并行计算）



实践中打磨而成的通用推荐系统

历时4年，12个业务，200多个在线场景，上百亿推荐请求/天

用户A



用户B



业务系统

推荐系统

数据接入

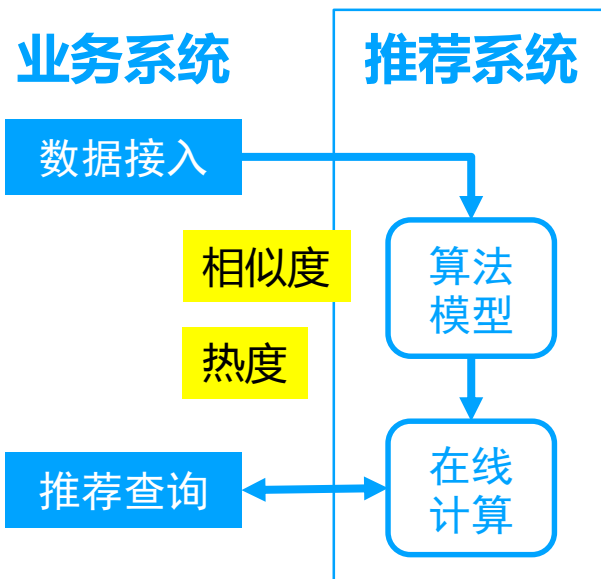
决策树

算法模型

推荐查询

在线计算

- 用户：**我喜欢**玩游戏 / 购物
- 产品：**预测**用户喜欢的手机应用



- 用户：我想看**相似+热门**的视频和资讯
 - 提升用户体验
 - 提升产品活跃度和内容质量



业务系统

推荐系统

数据接入

分类

算法模型

推荐查询

在线计算

- 用户：请**挑选**出我喜欢的游戏

| 推荐系统作用

- 提升用户体验
 - **个性化**：快速找到我感兴趣的信息
- 提高产品销售
 - **互联网+**：精准连接用户和产品
- 发掘长尾价值
 - **移动互联网**：主动推送用户兴趣

| 如何打造日均百亿级请求的推荐系统？

- 首先，要有用户，为此需要支持尽可能多的业务和场景
 - 通用算法库
- 其次，后台系统必须能够支持海量用户请求
 - 实时计算平台

1

通用化设计

推荐系统之模型训练

推荐系统：根据用户的历史行为和特点，推荐他喜欢的物品或内容，改进用户体验和提高产品销售。

• 模型

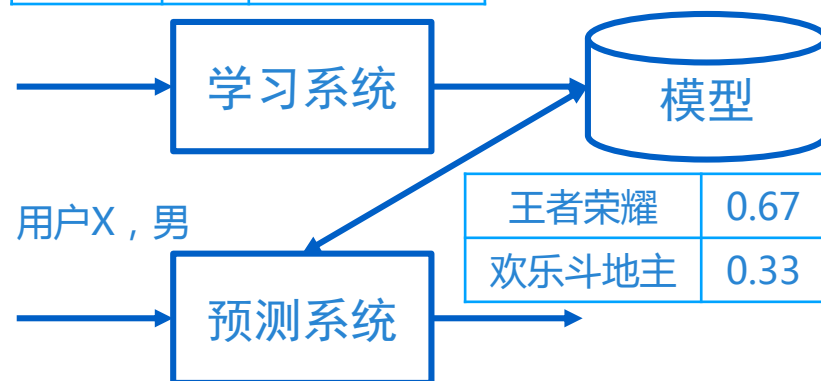
- **协同**：我的朋友喜欢什么
- **内容**：喜欢A所以喜欢B
- **知识**：根据我的需要推荐

• 学习系统

- 用户行为表
- 物品特征

用户A	男	王者荣耀
用户B	女	奇迹暖暖
用户C	男	王者荣耀
用户E	女	奇迹暖暖
用户F	男	欢乐斗地主

王者荣耀 | 男 = 0.67
奇迹暖暖 | 女 = 1
欢乐斗地主 | 男 = 0.33



通用化设计：模型训练

- 样本库

- 从流水日志中提取的用户行为和特征

用户A	男	王者荣耀
用户B	女	奇迹暖暖
...

样本库

- 特征库

- 用户和物品的属性等特征

王者荣耀	MOBA
奇迹暖暖	休闲
...	...

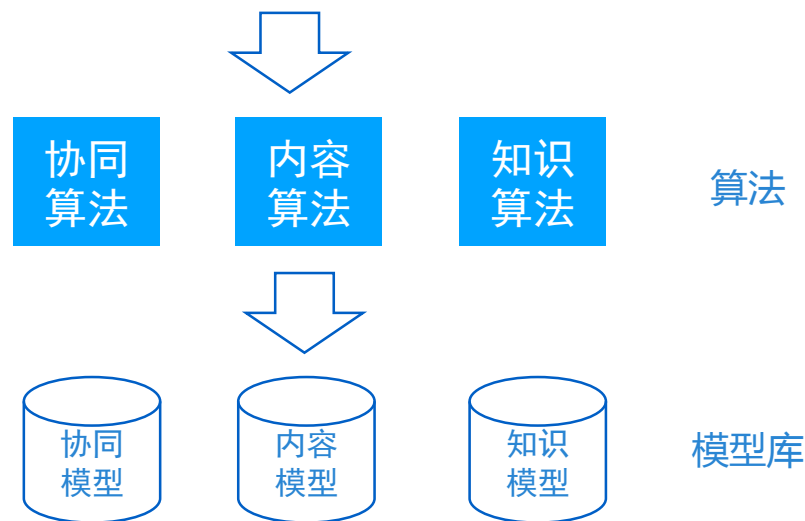
特征库

- 模型库

- 从样本和特征计算得到的训练模型

算法配置表

- 存储数据、算法和模型的映射关系，使得算法可以复用



我的模型 算法A 样本1 特征1 算法配置表

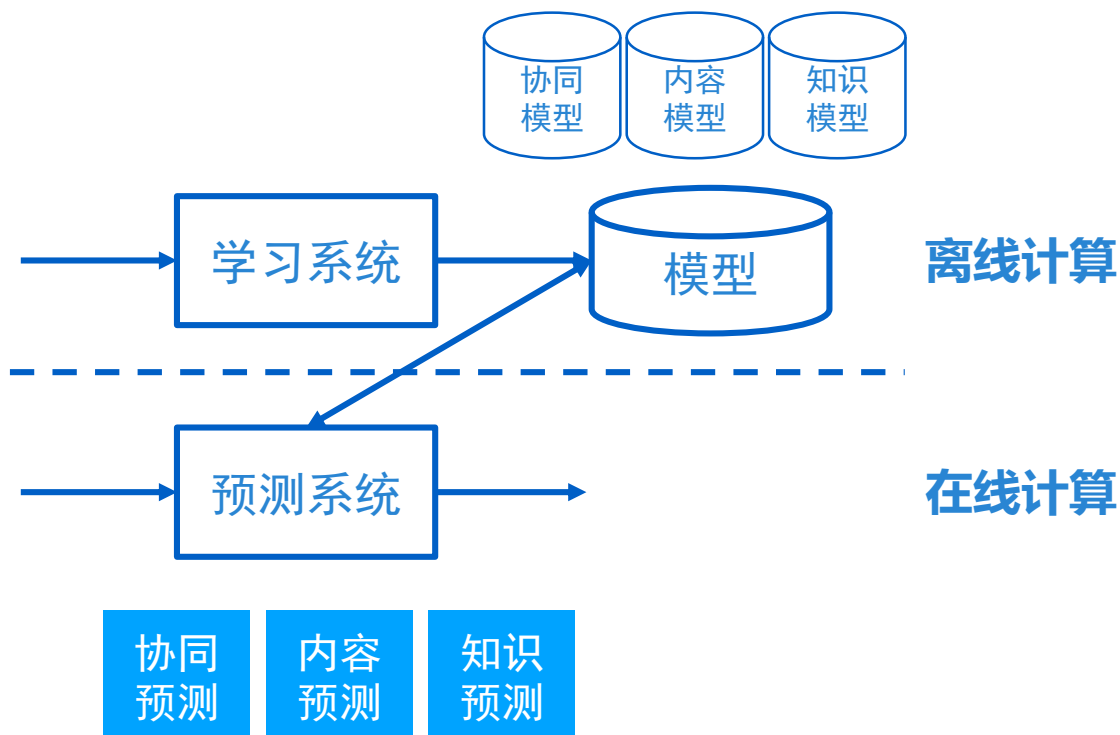
通用化设计：预测系统

- 预测系统

- 提供和训练模型相对应的预测计算

- 预测系统要求

- 低延时
- 海量请求
- 高可靠



2

百亿级用户请求

| R2：面向海量在线服务的实时计算平台

- 海量
 - 每天处理**上百亿**的个性化推荐请求
- 实时
 - 请求处理平均延时**18ms**
- 可靠
 - 系统稳定性**99.99%**

| R2设计初衷：一切为了线上服务

R2

- **面向线上服务**
 - 海量请求，低延迟
- **高可用**
 - 无单点故障
- **方便扩容**
 - 计算资源可插拔
- **动态调度**
 - 方便负载均衡
- **方便运维**
 - 紧密结合运维工具

现有流计算系统



Baidu Dstream



Yahoo S4

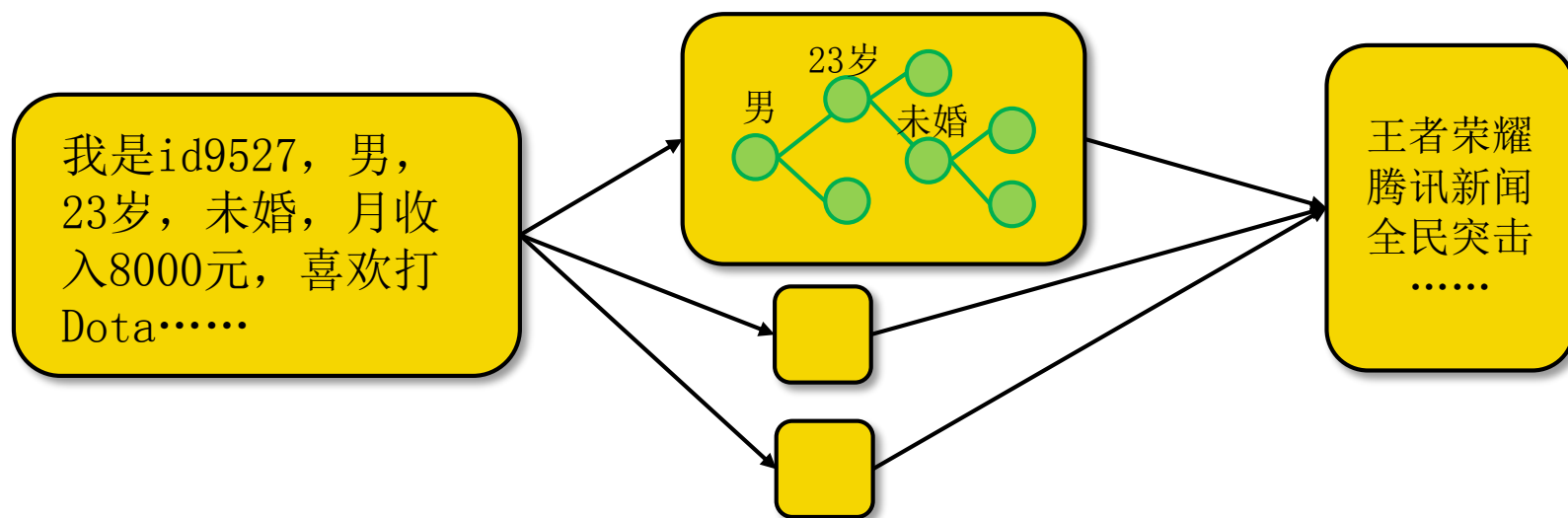


业务场景：猜你喜欢手机应用

用户特征

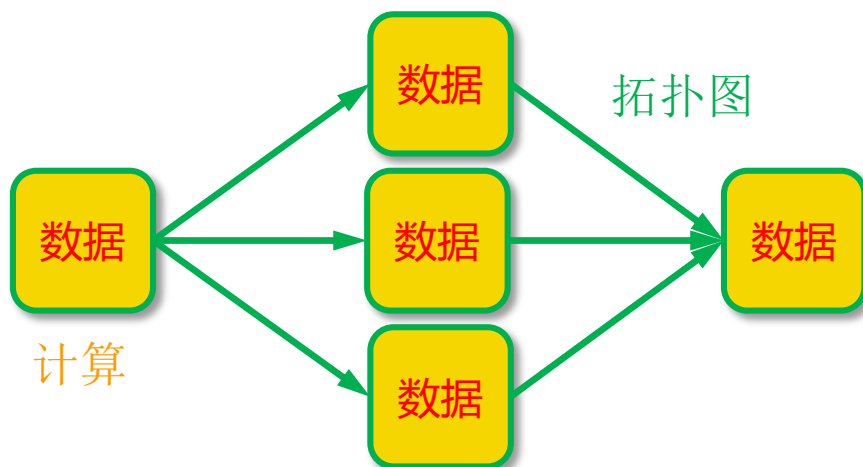
决策树

结果排序



- 1. 根据id得到用户特征
- 2. 使用决策树判断喜欢某个应用的概率
- 3. 对结果重新排序

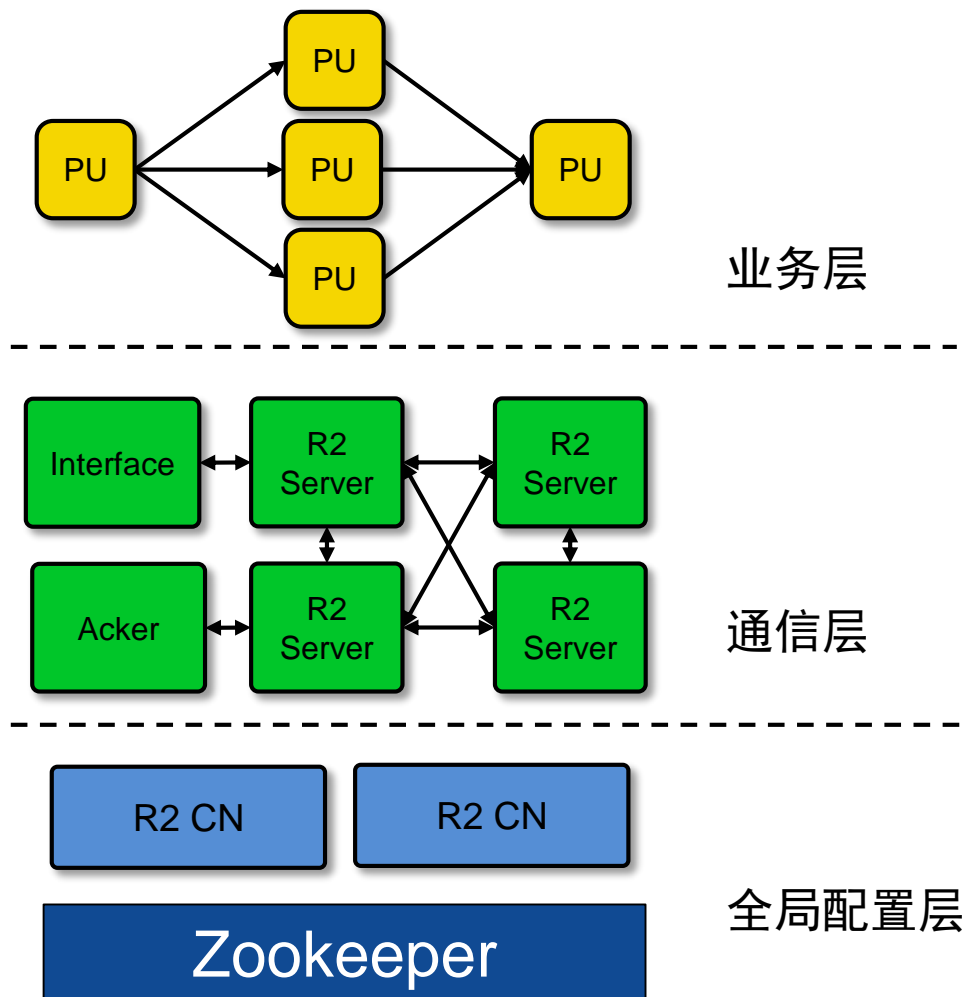
| R2计算模型



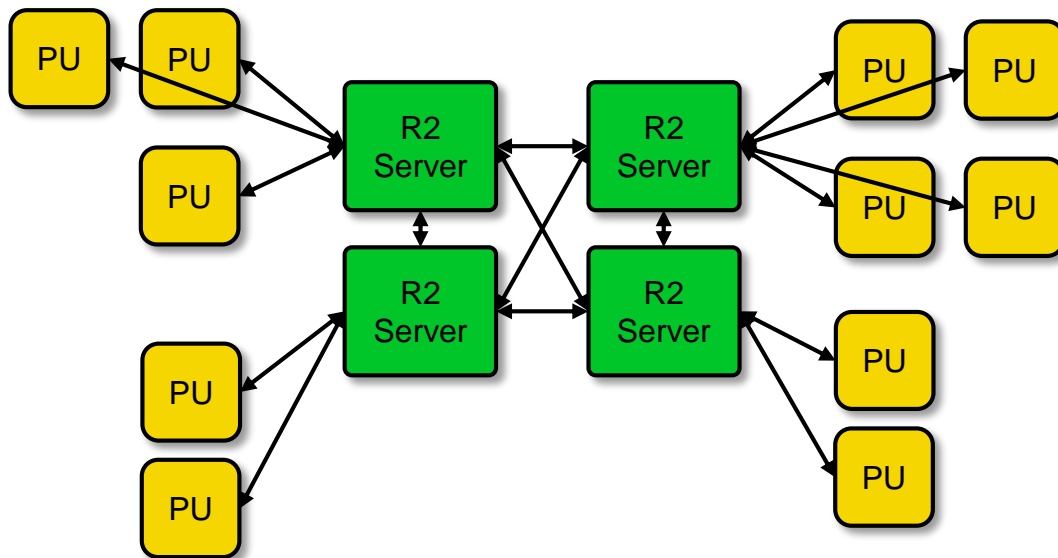
- 流式计算模型
 - 对连续的**数据流**在**拓扑图**上的**计算**

R2架构

- 业务层
 - 业务处理逻辑
- 通信层
 - 基于名字的通信
 - 分布式流处理
- 全局配置层
 - 可适应的拓扑配置
 - 动态扩缩容
 - 动态负载调度

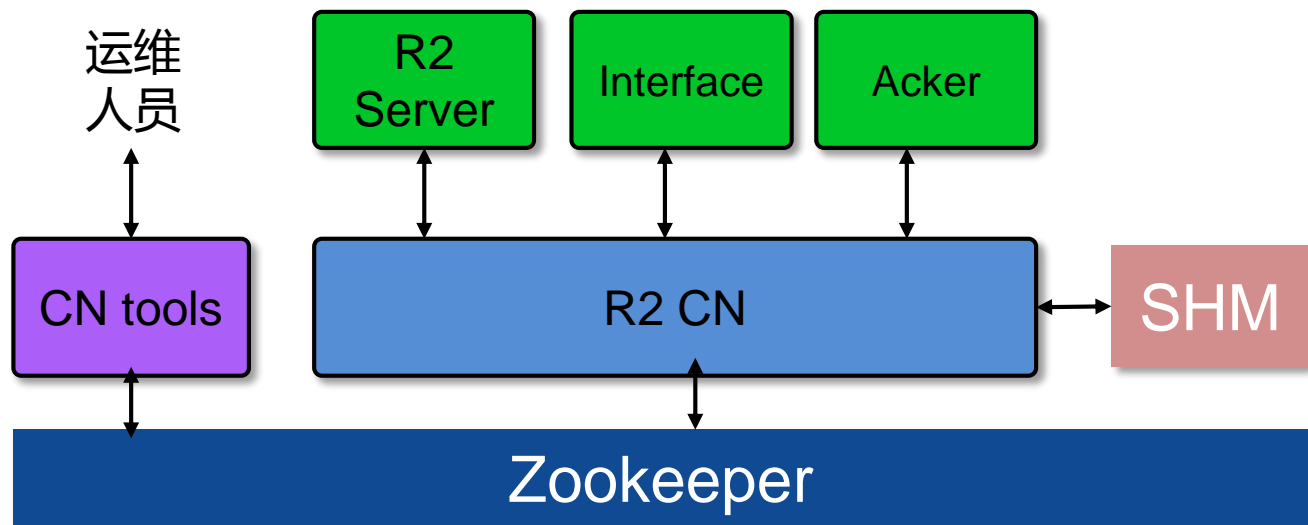


| 通信层架构



- 负载启动PU（计算单元）和监控心跳
- 数据包转发
 - PU之间的通信通过R2Server转发

| 全局配置层架构



- 节点注册
- 名字服务
 - 拓扑管理
 - 动态扩缩容
 - 负载调度

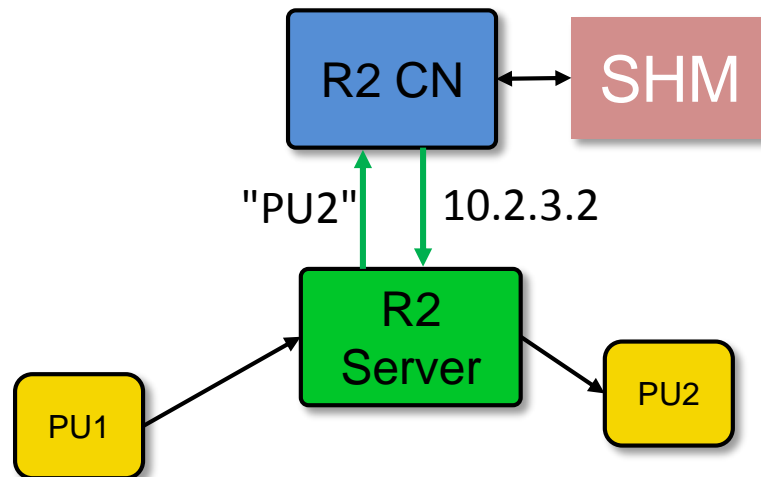
| 基于名字的消息通信

- PU编程接口

- `send_pack(PU_NAME, key, value)`
- 应用程序指定下一个PU的名字

- 名字注册

- R2 CN来决定具体是由PU组中具体某一个PU来处理

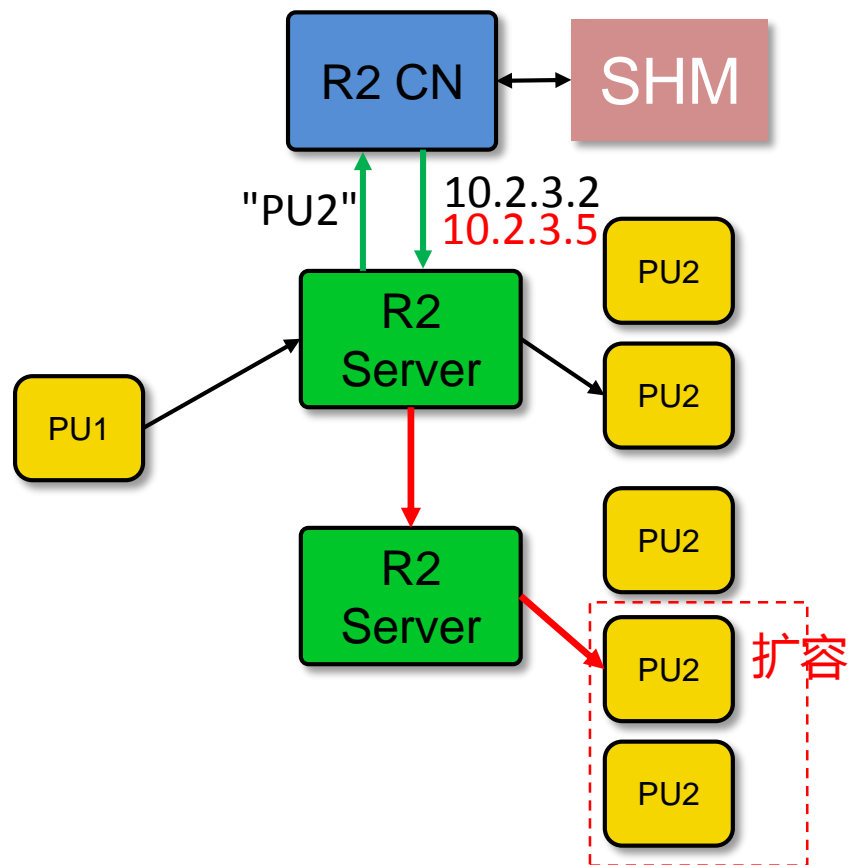


```
send_pack("PU2", key, value)
```


动态扩容

- 增加PU的实例数
 - 每个PU有一个index，目的PU的index在路由时由框架根据key和PU_CNT决定
 - key来自用户请求
 - PU_CNT保存在R2 CN中，实时更新

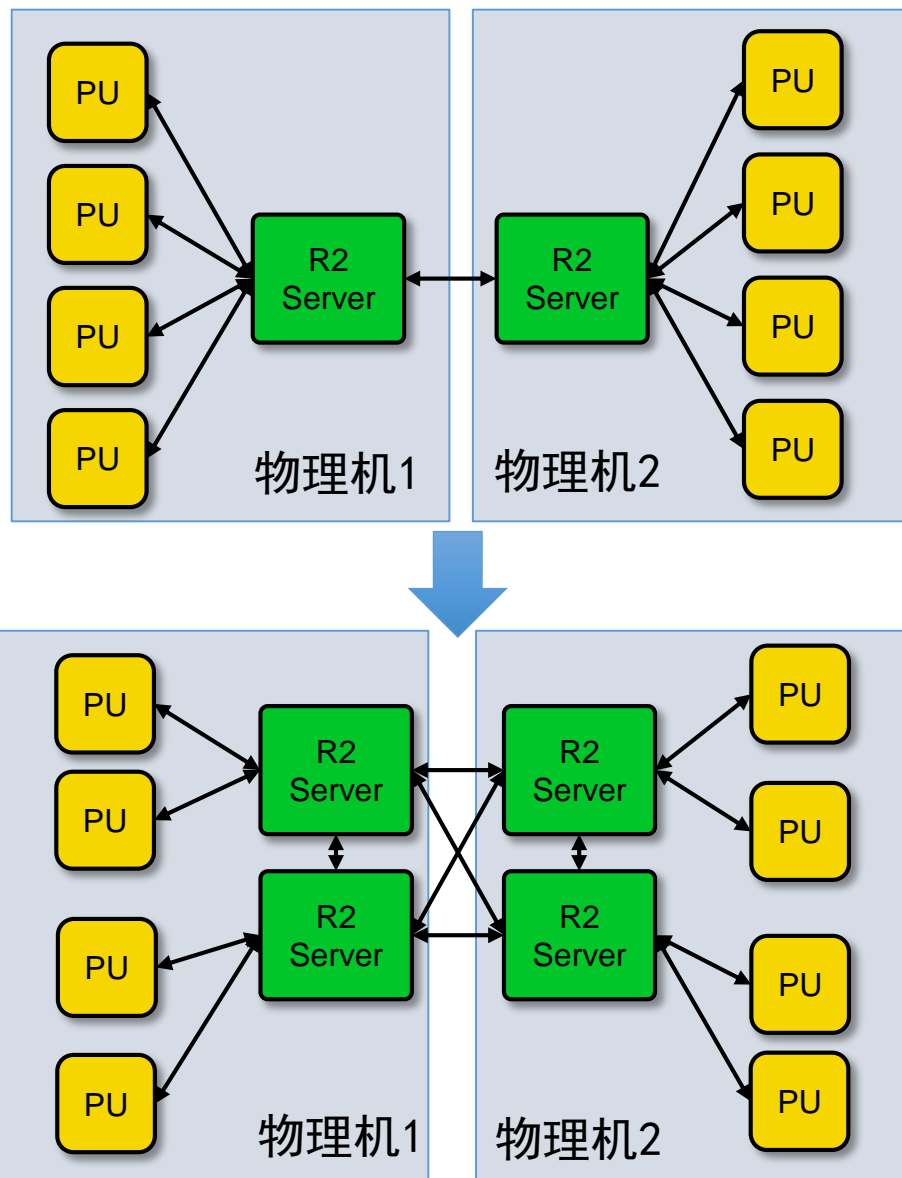
```
send_pack("PU2", key, value)
```



$$\text{Index} = \text{Partition}(\text{key}) \% \text{PU_CNT}$$

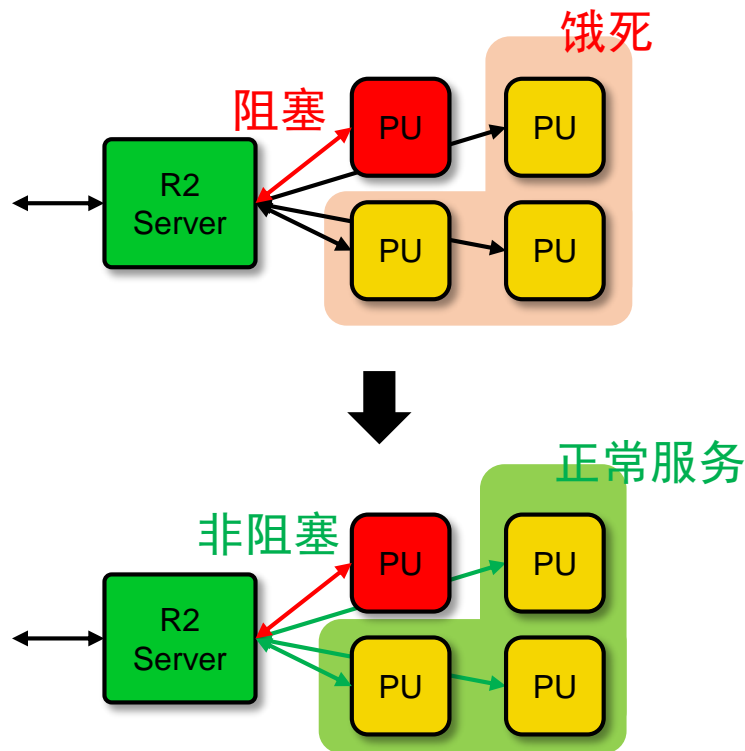
通信层重构

- R2Server负责通信层的信息转发
- 每个物理机从部署单个R2Server改为多个R2Server
 - 每个R2Server进程使用不同的端口
 - 注册不同的R2Server到全局配置层



提高系统整体可靠性

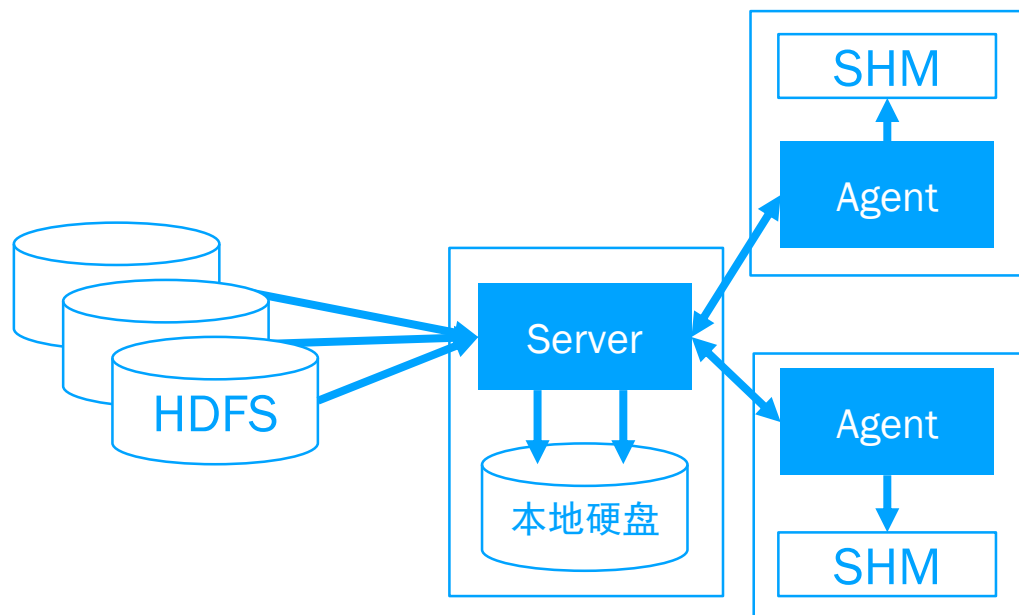
- 原有问题：阻塞写
 - R2Server写管道时为阻塞写，单个PU的阻塞会导致其他PU的心跳更新丢失
- 解决方案：非阻塞写
 - 管道写从阻塞改为非阻塞，在某个PU不可用时放弃这个PU继续服务其他PU
- 方法论：有损服务
 - 放弃绝对一致，追求速度极致
 - 少数服从多数



SSM：订阅式缓存系统

拉取模型数据到共享内存系统用于线上计算

- 高速缓存
 - 每秒千万级读写
- 数据秒级更新
- 客户端按需拉取
- 数据一致性
 - MD5+Version
- 通用数据格式
- 不间断服务
 - 主备切换，数据更新不影响服务



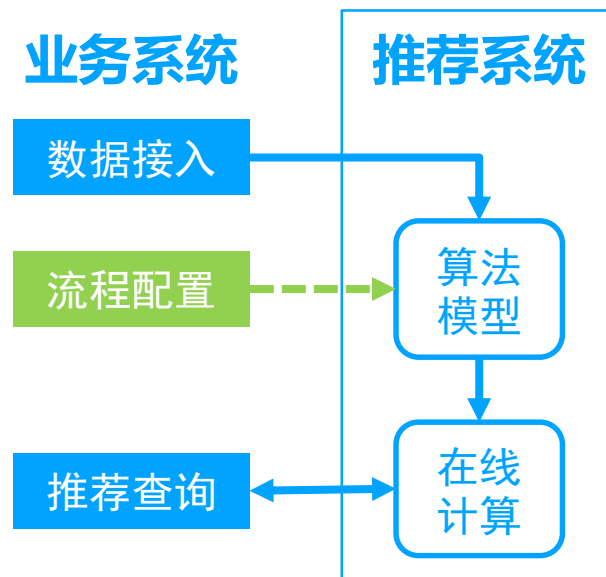
3

腾讯云推荐引擎

腾讯云推荐引擎

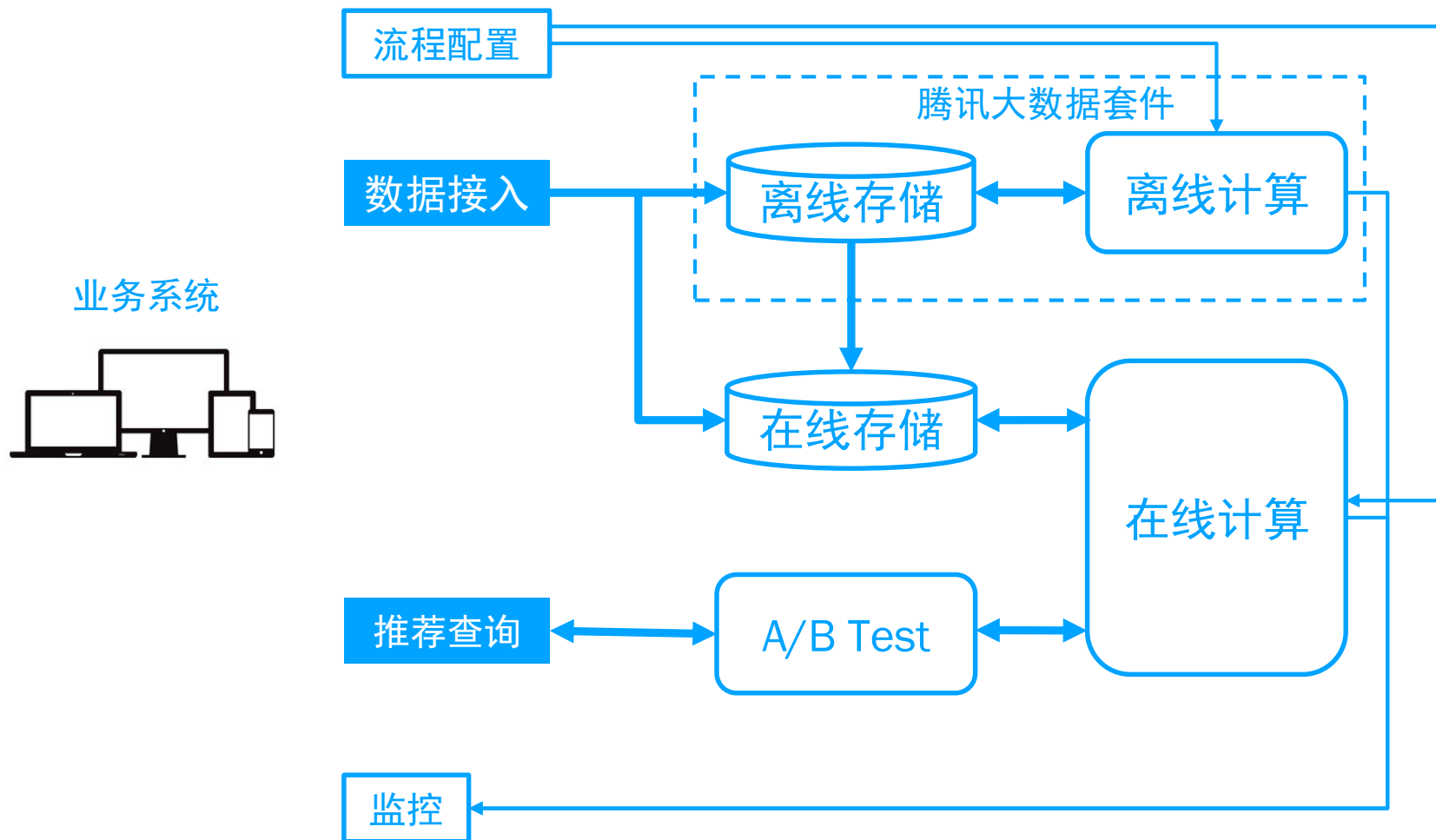
面向广大中小互联网企业打造的一站式云推荐引擎解决方案，提供安全、便捷、精准、可靠的推荐系统服务，提升其业务的点击转化率和用户体验。

- 一天接入，快速上线
- 模板化算法，节省99%代码
- 快速扩容，应对业务快速增长
- 稳定可靠，节省运维开销

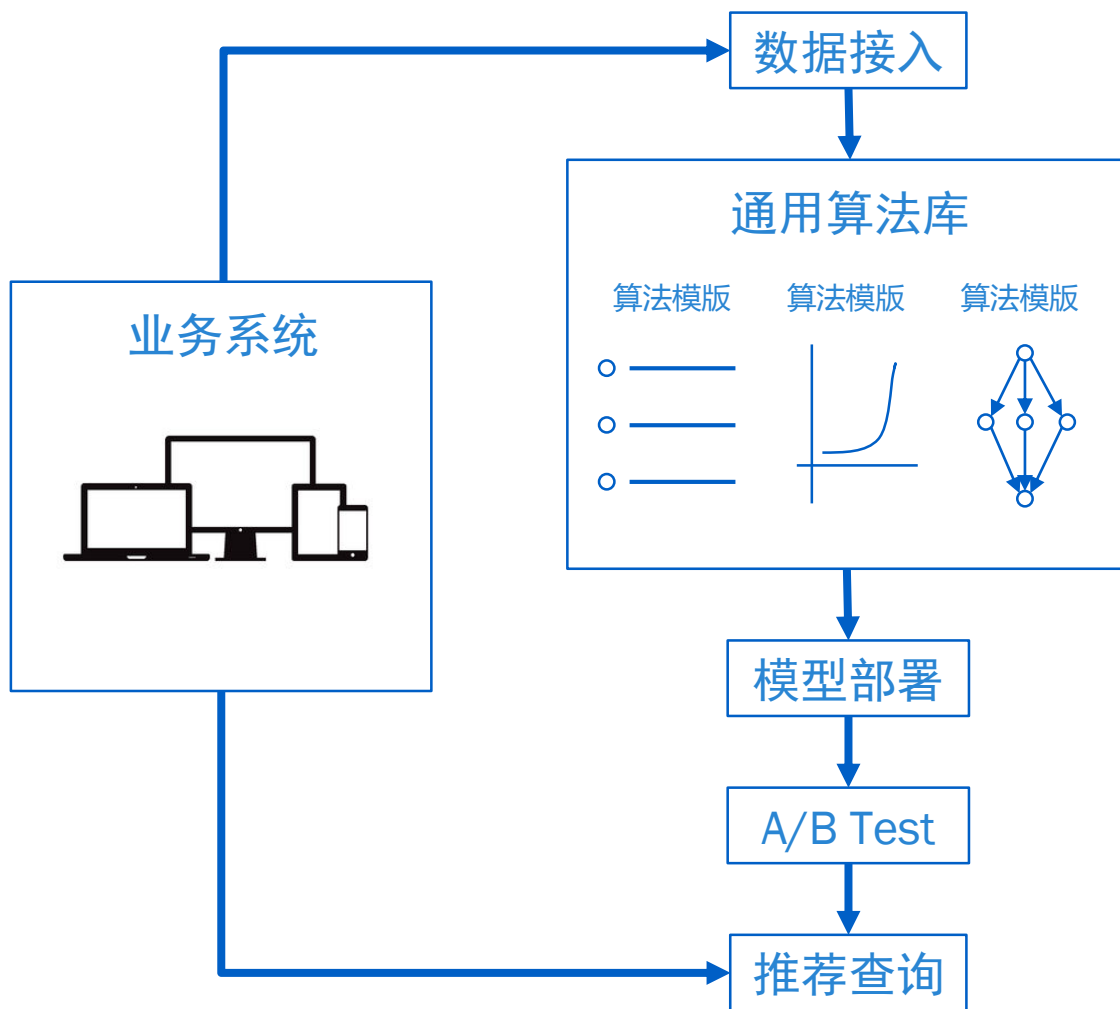


腾讯云
推荐引擎

云推荐引擎架构



解决方案



1

创建业务

规范的数据接入格式、完善的SDK、流程化业务场景创建

2

选择算法

通用的算法框架，模板化配置

3

一键上线

一键式模型部署上线、完善的Debug工具

4

效果评估

实时A/B Test效果评估系统

| 产品优势

- **安全**

- 尊重用户数据安全和隐私，支持专有云环境部署

- **便捷**

- 一键配置算法流程，提供一站式解决方案

- **精准**

- 提供丰富的算法模版，支持多种应用场景

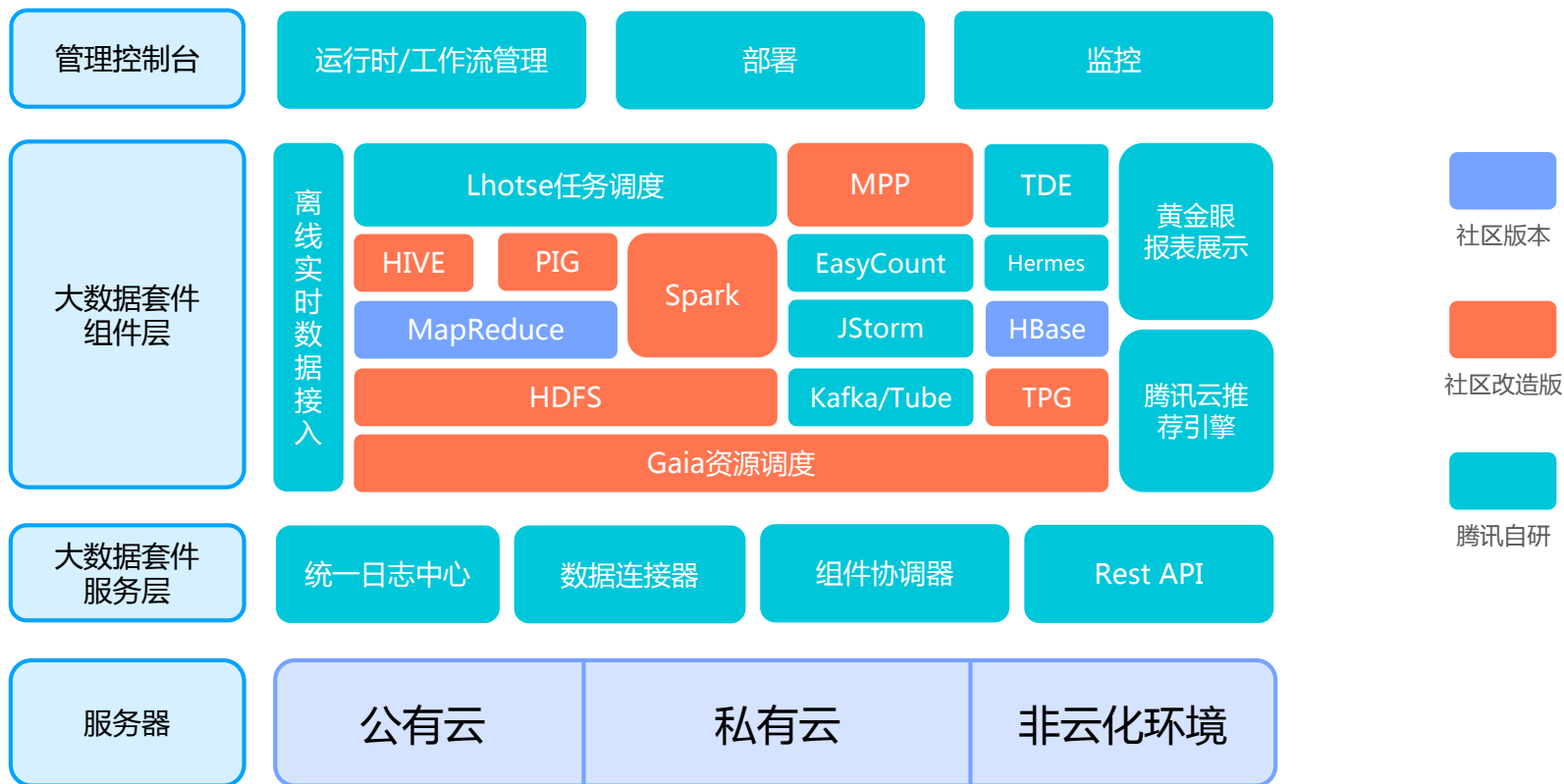
- **可靠**

- 经过多年实战经验打磨，值得信赖

业界方案

- 用户核心数据上传到公有云，数据安全不可控
- 通过开源软件搭建推荐系统，流程繁杂，费时费力
- 算法实现和优化过程冗长
- 很难应对突发事件，需要大量运维操作

腾讯大数据套件



1

通用化
设计

2

百亿级用
户请求

3

腾讯云推
荐引擎

联系我们

- huiweilv@tencent.com
- 微信交流群：“腾讯云推荐引擎”
- 欢迎技术交流 & 加入我们



腾讯云
推荐引擎