蚂蚁花呗
ANT CHECK LATER

# 花呗 - 亿级金融业务架构演进

蚂蚁金服 - 阿喆

蚂蚁金服
ANT FINANCIAL

# 目录

蚂蚁花呗
ANT CHECK LATER

蚂蚁花呗
ANT CHECK LATER

# 产品介绍

# 产品介绍

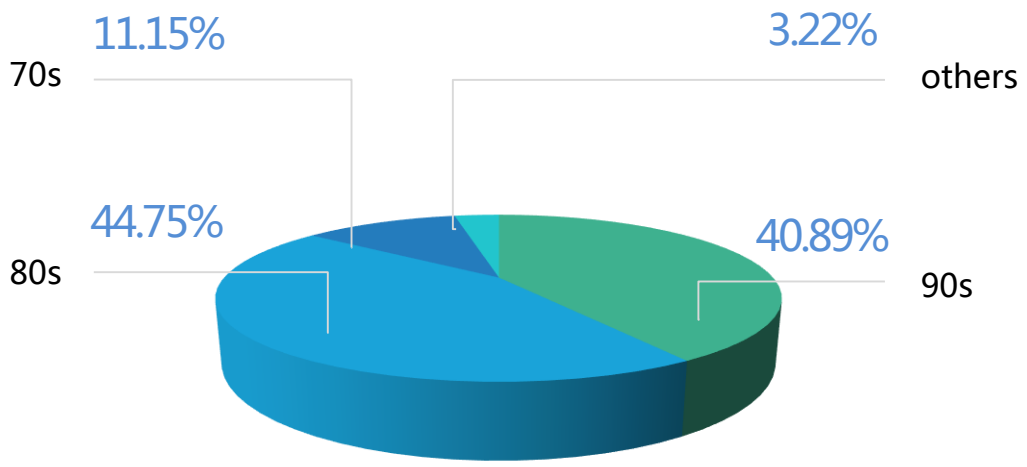
## 人群

**让消费金融触手可及**

### [ 服务年轻群体 ]

花呗活跃用户8成以上是8090后；

90后活跃用户占4成；

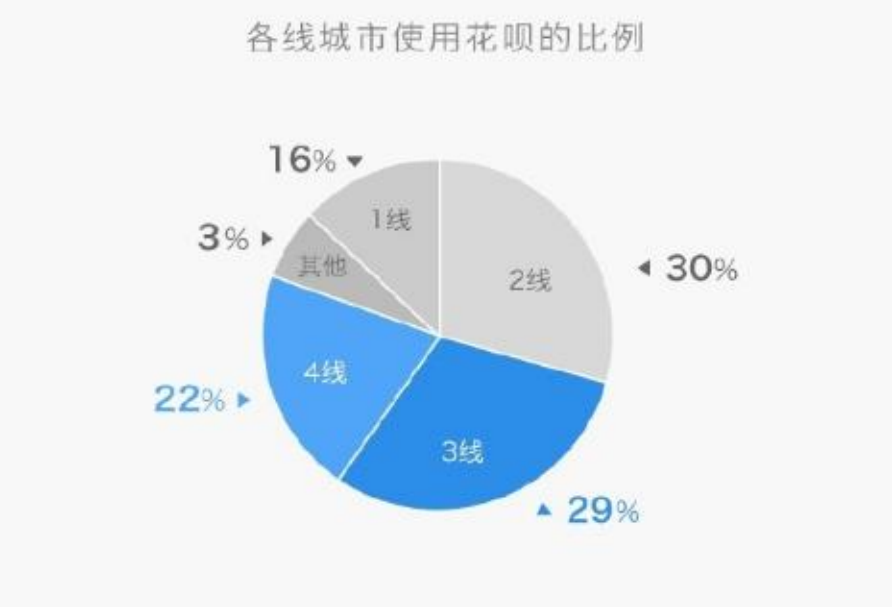



2015年双11消费用户年龄分布

### [下探至三四线城市]

三四线城市用户在购买数码商品时，

更青睐花呗




节选自《新数码消费趋势报告》

# 产品介绍

## 双十一表现


蚂蚁花呗
ANT CREDIT PAY
蚂蚁花呗
11.11的一天
双11
00:00 - 03:00
进入双11狂欢
1小时26分
支付笔数:
6048万笔
超过去年全天笔数
花呗付天猫国际
支付166万笔,
用全世界的东西，来满足你的挑剔。


07:00 - 09:00
便利店
起床犒劳熬夜奋斗的自己。
花呗付O2O便利店
支付13.6万笔
早起补充能量，
和千千万万剁手党共赴下个战场。
08:00 - 10:00
我和工作有个约会！
花呗付滴滴打车
支付16万笔
花呗付手机淘宝
支付1506万笔
上班路上最适合买买买，
合理利用所有空隙时间。
11:00 - 14:00
吃厌了食堂，今天中午换点新花样。
花呗买美食
支付2459万笔
中式西式随意切换，
每天都要吃得有新意。

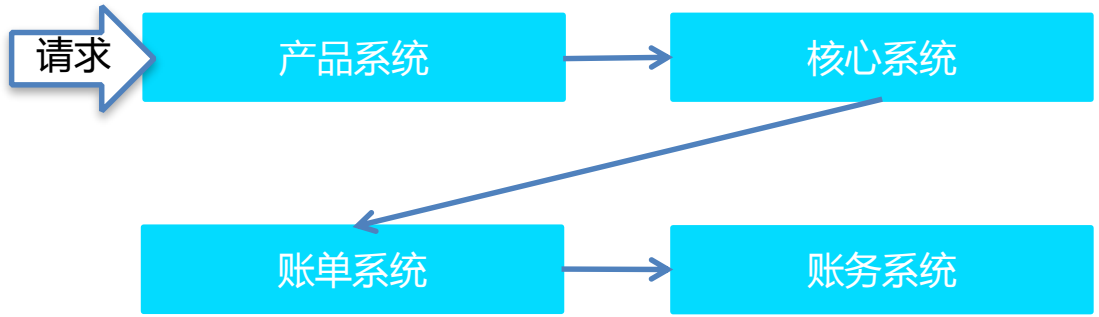

16:00 - 18:00
辛苦一天,是时候想想远方。
花呗付飞猪旅行
支付1万笔
花呗付12306
支付7万笔
远方已经不远，打足鸡血再奋斗。
19:00 - 21:00
奋战一天发现钱包已扁?分期无压力!
花呗分期
支付12万笔
锤子、大疆趁着低价赶紧入手，
分期还款心不慌。
22:00 - 24:00
睡前看下手机，发现返场有小惊喜。
花呗买零食
支付178万笔
花呗买情调
支付7万笔
potato chips

# 挑战&应对

# 挑战&应对

## 花呗上线时技术架构


- ***部署架构：单机房部署，***
- ***应用架构：从上到下分为产品，核心，账单，账务4个系统***




- ***调用链路：几乎所有的业务（支付，确认收货，还款，账单查询）都需要经过4个系统完成***

# 挑战&应对

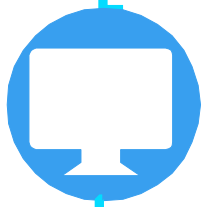
## 海量用户增长带来的挑战

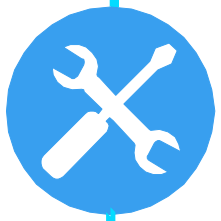
### 调度任务

数千万用户的定时业务在简单粗暴的执行方式下，常常难以控制并发和出现名单遗漏等质量问题

### 应用架构

系统处理链路很长，整体性能不高，任何一点系统抖动，都会影响这个业务的稳定性





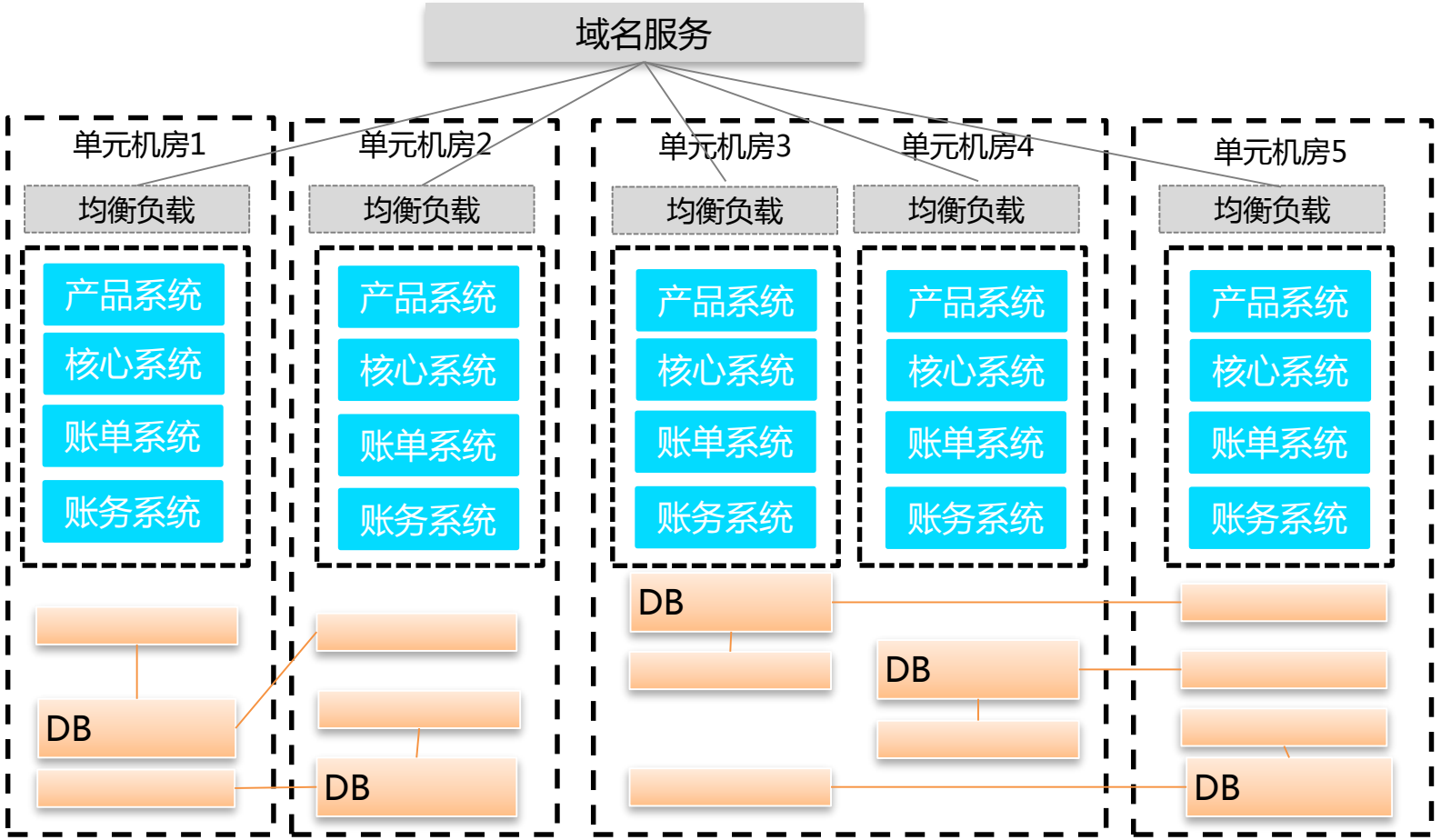
单机房部署出现业务瓶颈，从日益增长的用户规模，带来存储和应用部署瓶颈。

稳定性考虑，如果遇到机房级别故障，就会影响所有用户，这个也是不可接受的

### 部署架构

# 挑战&应对

## 升级异地多活的架构

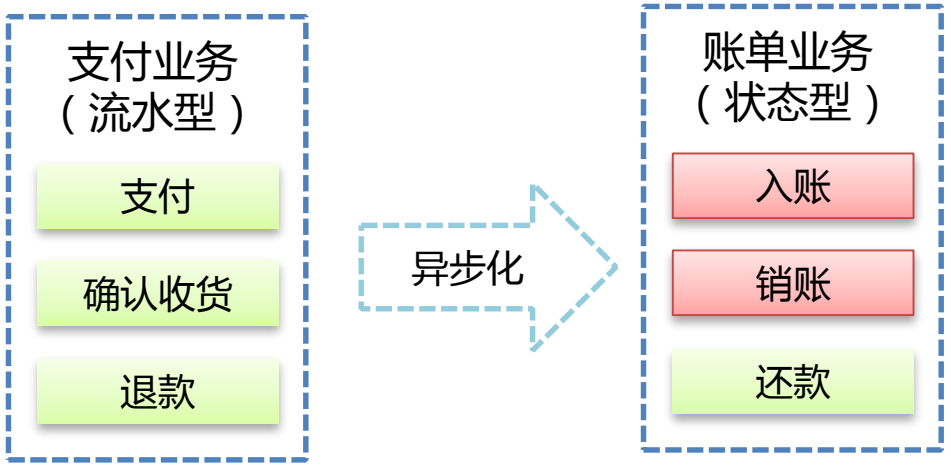


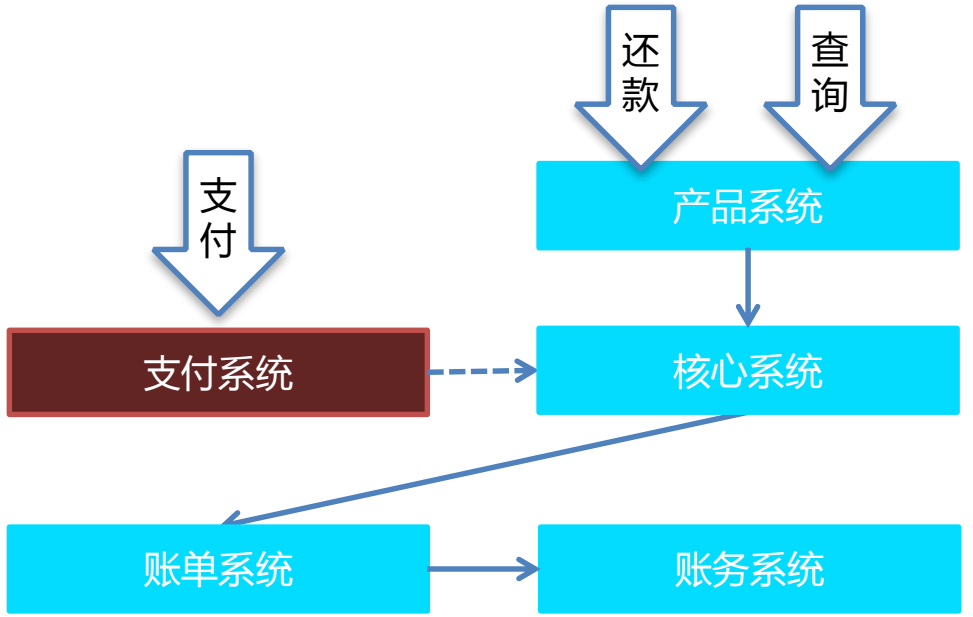

- ***升级成蚂蚁LDC架构***
- ***进入Rzone***
- ***整体数据基于用户维度拆分成5份，分别进入5个机房***
- ***数据做同机房和跨机房互备***

# 挑战&应对

## 应用架构升级

- ***业务抽象：对于花呗绝大部分业务，都可用抽象成，流水型业务+状态型业务，考虑到业务的频繁度和稳定性将业务分成，支付业务和账单业务。***
- ***架构解耦：将支付业务和账单业务解耦，一笔支付成功后，异步化入账。***




- ***升级后优点：核心支付业务，链路缩短3/4，稳定性提升***
- ***将业务拆分成流水型和状态型，为后续容灾奠定基础***

# 容灾

容灾的关键

# FAILOVER能力

- 应用链路来看DB是在这个调用链路的最底层，同时也是故障的单点。
- 当出现机房级别或者地区级故障时，网络流量可以快速切换到另一个机房，但DB不行，尤其是金融级数据，不能容忍不一致的情况
- DB出现问题，恢复耗时相当高，尤其当出现机房级或者地区故障时，耗时更长
- Failover能力是为了保证核心业务功能在容灾恢复阶段仍然可用而提出来的

# 容灾

蚂蚁花呗 ANT CHECK LATER

支付类业务特点（流水型）

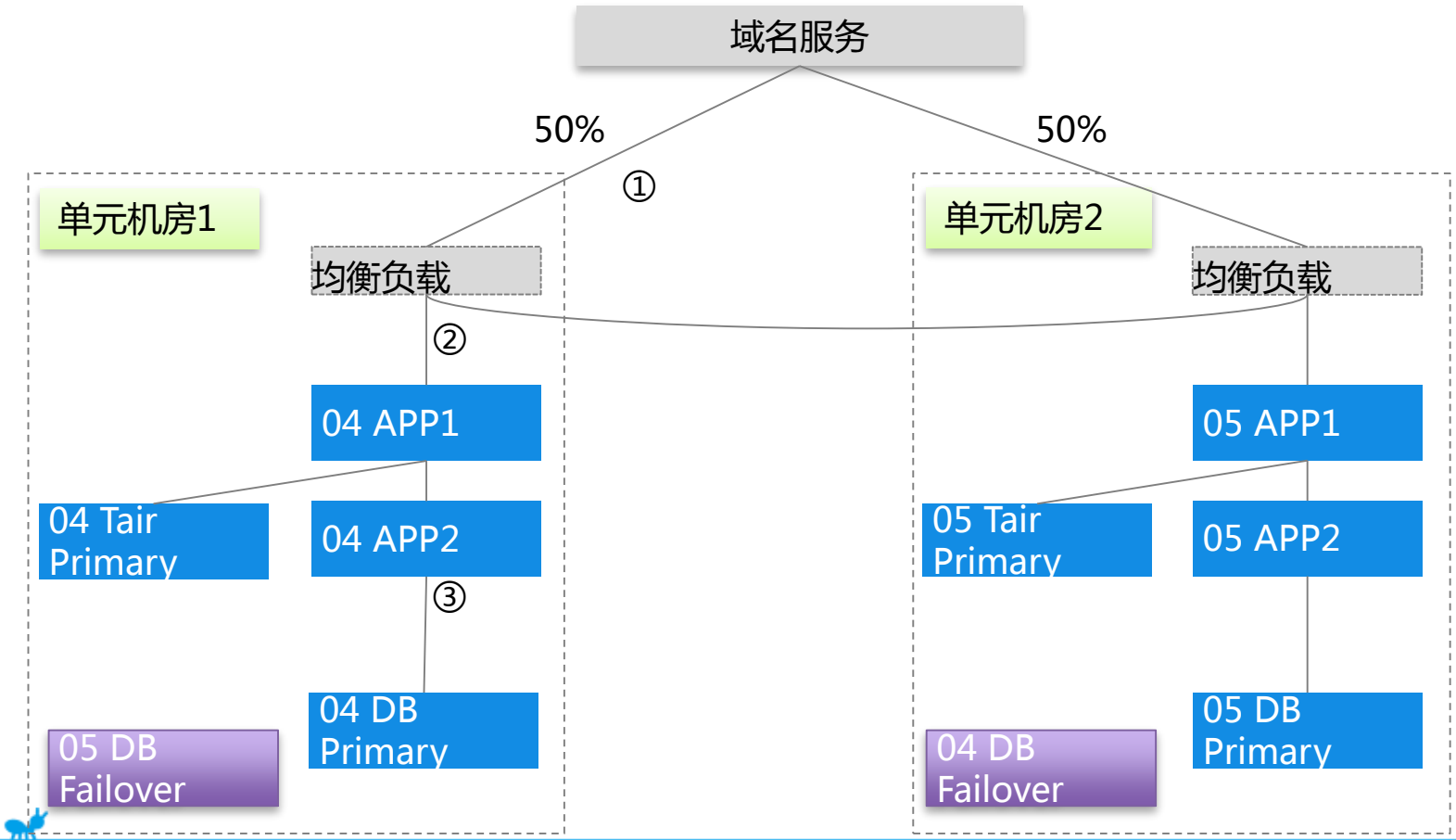
## 业务特点

- 多次业务操作之间没有关联性
- 该类业务操作对于数据库通常只做一条或者多条insert操作
- 该类业务操作不依赖数据库中已有的数据

**解决方案：**因为操作之间不依赖，都是插入操作，将流量切换到正常机房，提供可写的库保证业务正常

# 容灾

## 支付类业务容灾



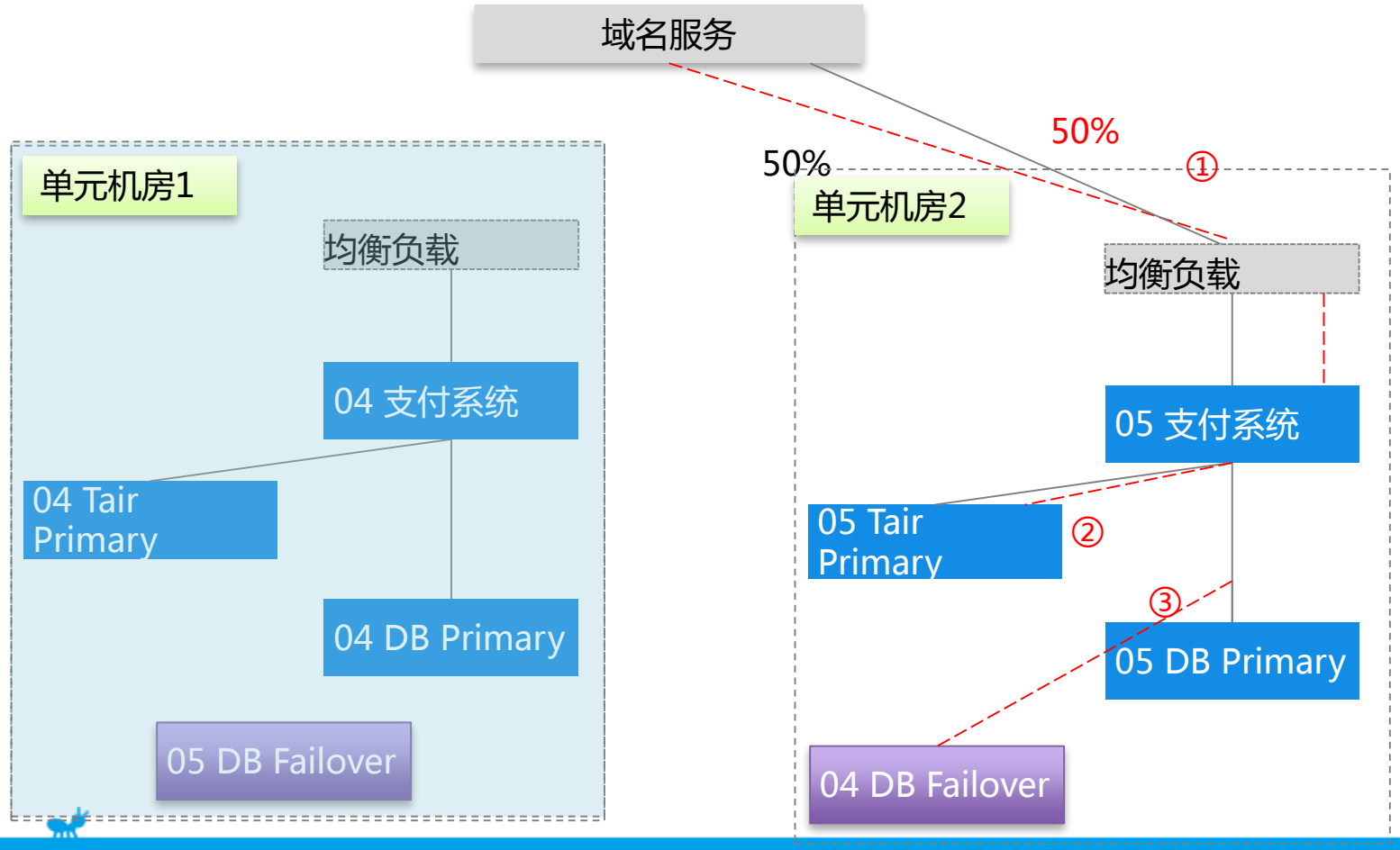

**正常场景**：

跨机房增加一个数据库，我们称为failover库。容量可以比正式库小

正常流量是不会读写fo库，只有在failover状态下中才使用

# 容灾

## 支付类业务容灾




**故障切换场景**

①：单元机房1流量切换至单元机房2

②：单元机房1 tair集群切换到单元机房2集群

③：单元机房1 DB切换到单元机房2的failover库

# 容灾


**核心难点**：

如何保证网络重试等场景，

重复请求的幂等问题

**缺点**

- 多库读取，回迁成本

# 容灾

蚂蚁花呗 ANT CHECK LATER

账单业务特点（状态型）

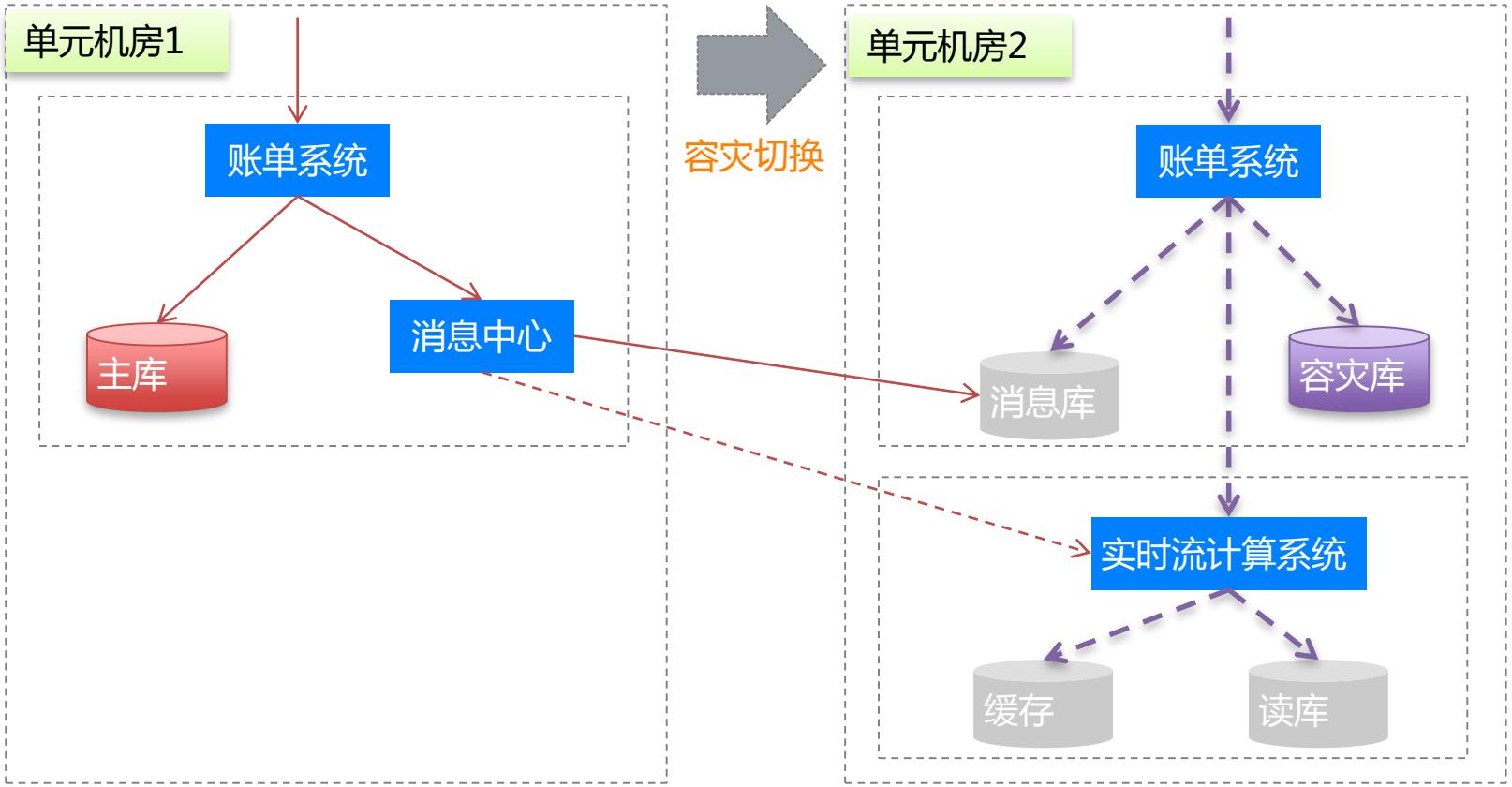
## 业务特点

- 业务操作之间相互关联，前后操作有影响
- 每次业务操作需要依赖最新的数据
- 每次业务通常对于数据库操作比较复杂，涉及操作记录比较多，有新增和修改
- 不能容忍数据不一致

**解决方案：**实时同步一份数据去异地，当出现故障时，将流量切换到异地机房，使用实时同步数据继续进行业务

# 容灾

## 账单类业务容灾（建设中）




**正常场景**

- 任何对于账单写操作，通过事务消息，同步到另一个机房中，数据落入缓存中和读库中

**故障场景**

- 流量切到单元机房2，先从消息中心DB中获取积压消息形成黑名单账户。黑名单账户中的数据是不准确的，不能进行业务处理。
- 写操作发现没有命中黑名单，则从读库中取出数据初始化容灾库，开始进行业务出来
- 读操作统一从读库中获取数据，提供查询服务

# 容灾


**难点**

**□如何将数据不一致的账户进行隔离**

**缺点**

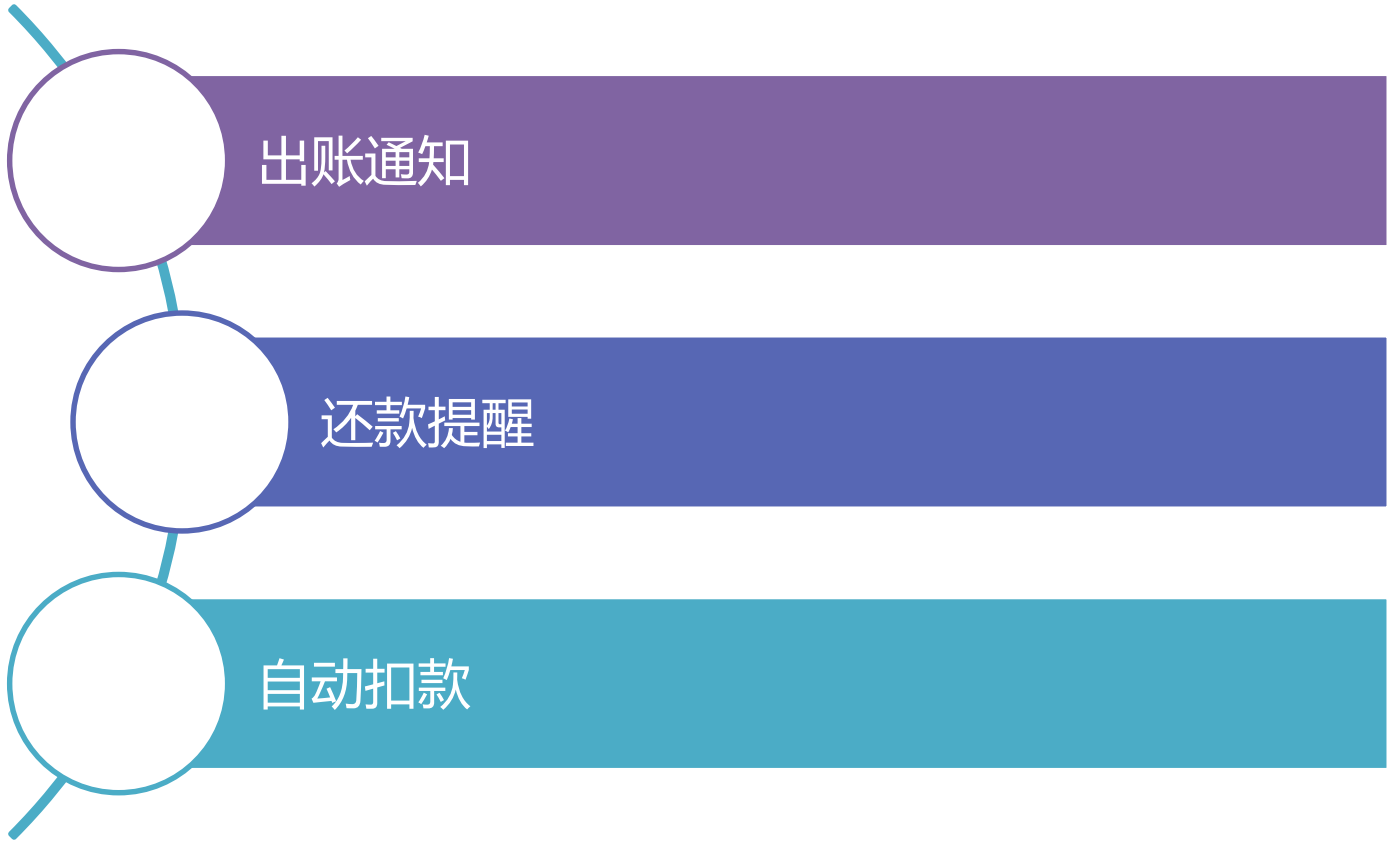
□操作复杂，增加RT，增加了依赖

# 分布式调度

# 分布式调度

蚂蚁花呗
ANT CHECK LATER

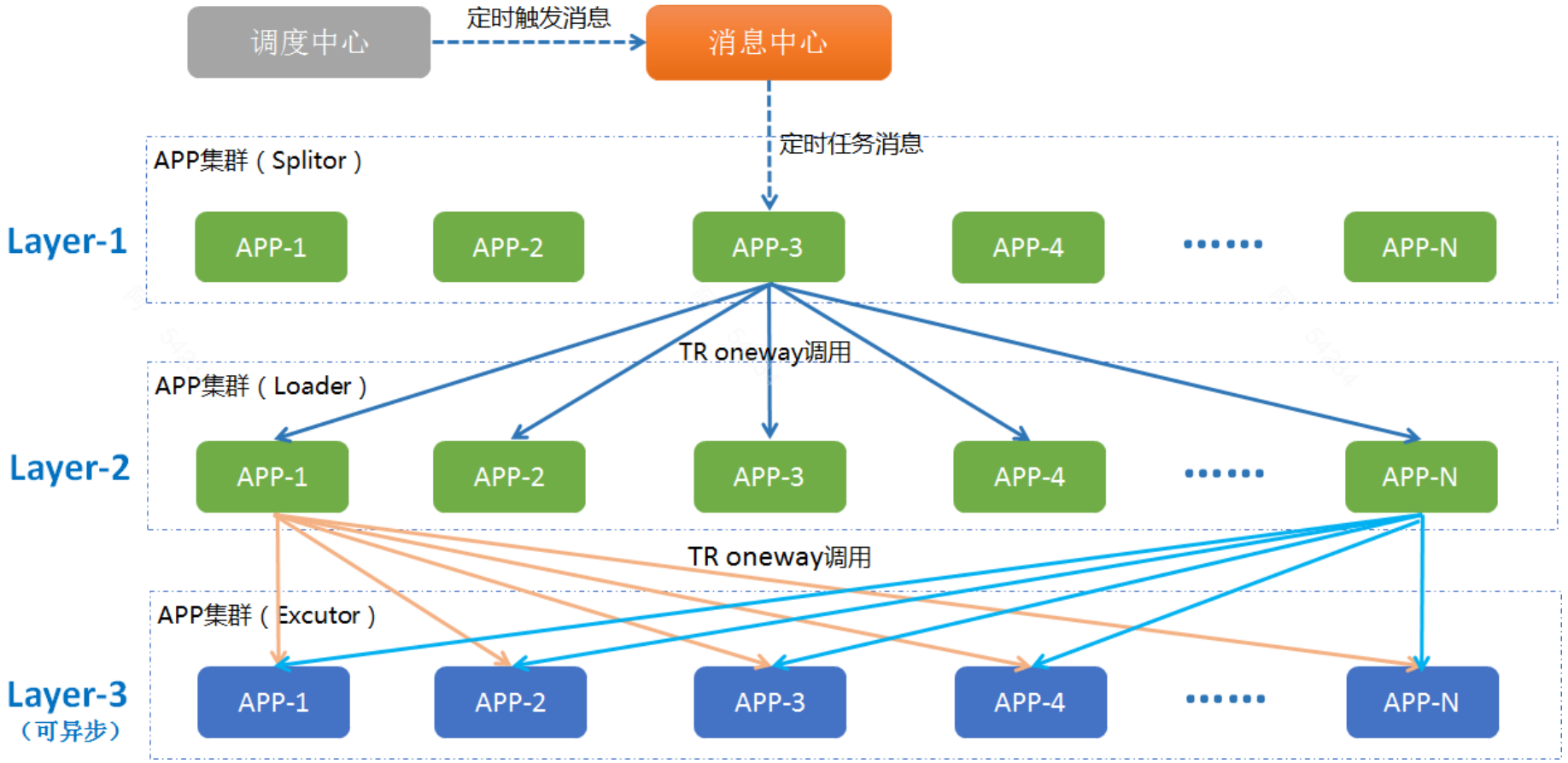
## 基于调度的业务




- 业务特点：花呗有许多基于固定时间点的业务，这些业务处理数据量大，而且不允许失败，耗时长，逻辑复杂

# 分布式调度

## 花呗调度架构




- 技术选型：基于蚂蚁金服分布式调度框架
- Layer-1：获取需要扫描的用户ID区间
- Layer-2：获取ID区间需要处理的用户名单，将用户名单包装成任务
- Layer-3：执行一个个任务

# 分布式调度

蚂蚁花呗 ANT CHECK LATER

## 面临的问题

### 并发压力大&难以度量

下游系统压力大，银行通道
不同时间段压力不同，同时无法
很好度量多少并发是极限性能

### 质量隐忧

涉及都是对客业务，短信，扣款，处理
数据量大，一旦出现问题影响面非常大

框架层面

### 批量失败

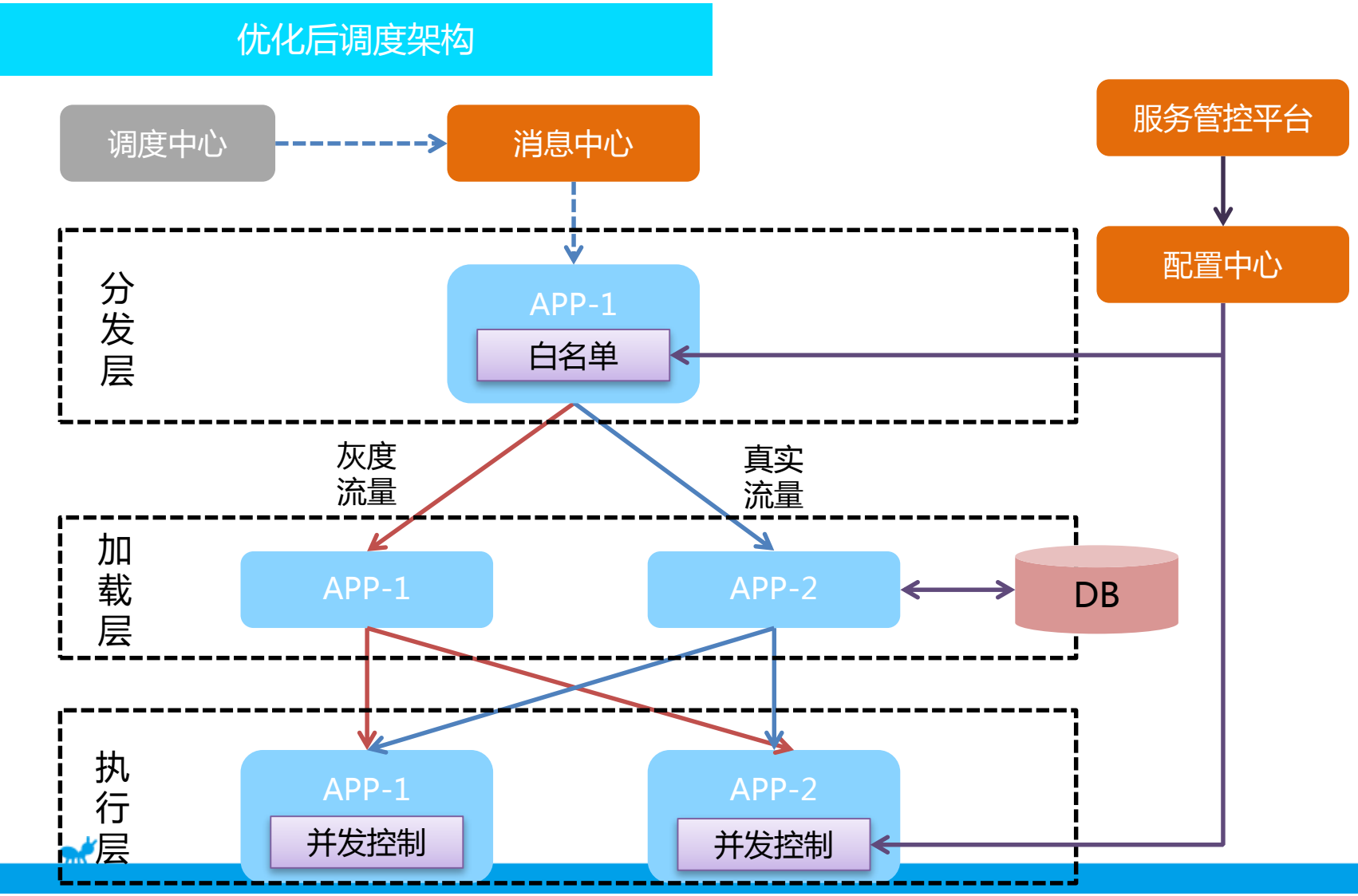
涉及用户量大，在整个链路中
一旦有抖动，就会有批量失败。

### 名单遗漏

数据库抖动导致名单获取遗漏

业务层面

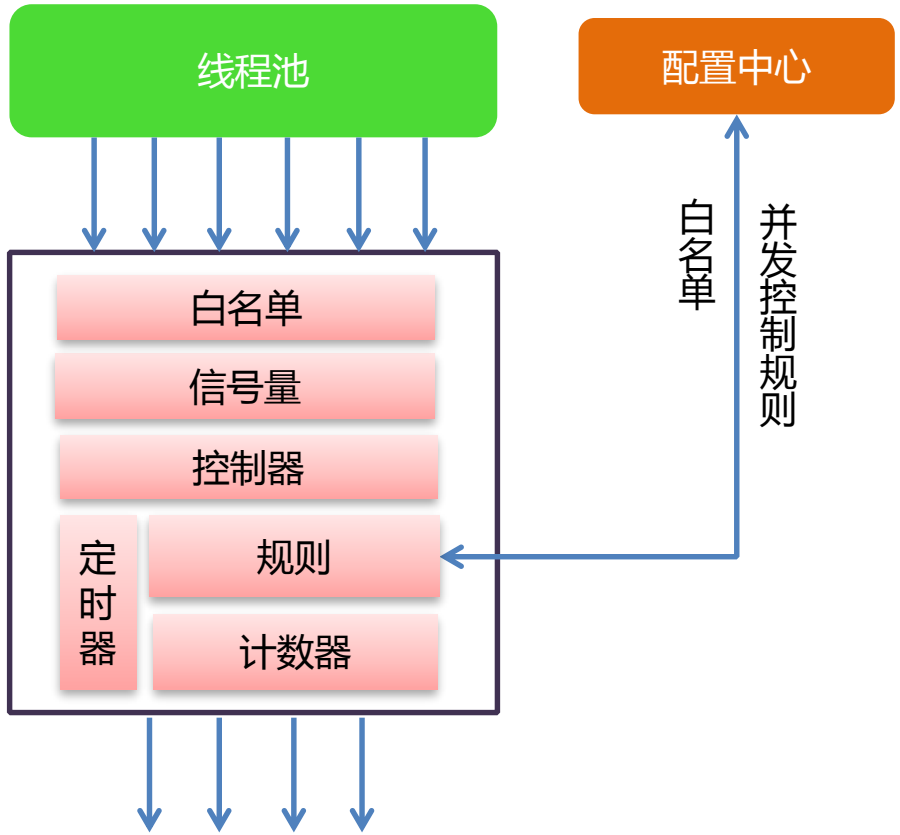
# 分布式调度




- 任务持久化：解决执行任务失败或者机器重启任务丢失问题
- 任务解耦：将生成任务名单和执行任务变成两个调度，一个调度只做生成任务，一个任务只做执行任务。当任务生成失败，会重试
- 灰度&并发组件：自动决策是否灰度，分时间段控制并发度

# 分布式调度：解决方案

## 灰度组件&并发组件

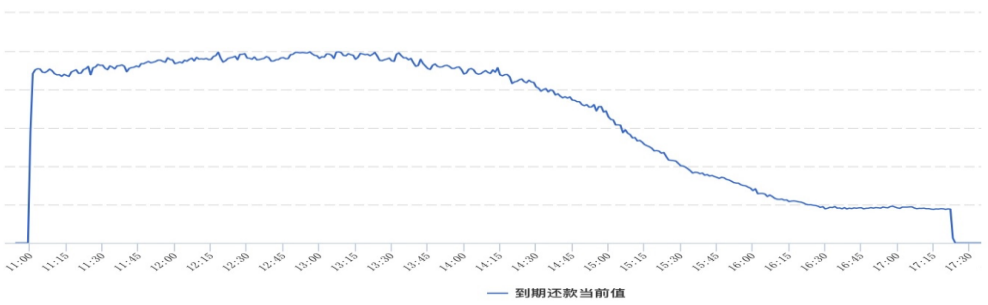


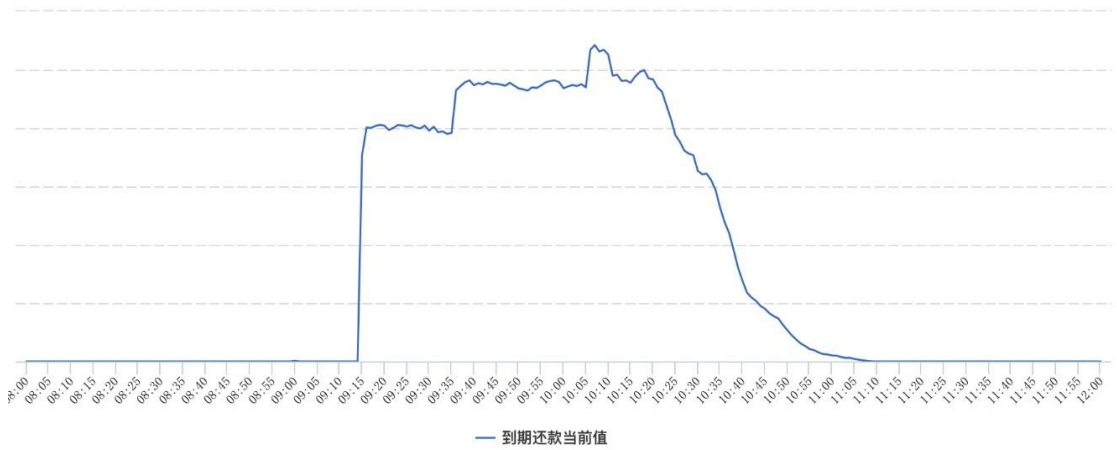

- 并发控制
  - 基于时间维度控制
  - 基于流量，失败量，处理时间控制
- 白名单
  - 配置中心配置白名单
  - 根据当前时间执行决策白名单是否激活

# 分布式调度：解决方案

蚂蚁花呗 ANT CHECK LATER

## 对比



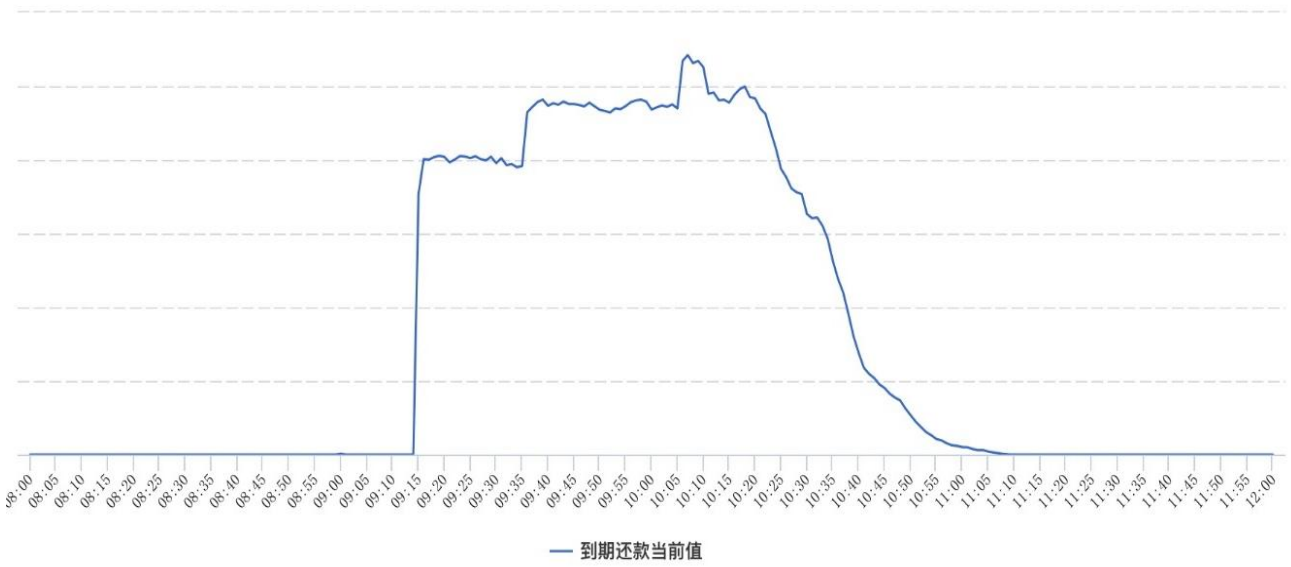

左图：改造前任务处理时间需要6小时

右图：改造后任务处理数据量，根据时间段自动进行并发控制，处理时间缩短为2小时

通过时间段并发控制，在线实时压测找到整个处理链路并发的极限，将处理速度提升到最大

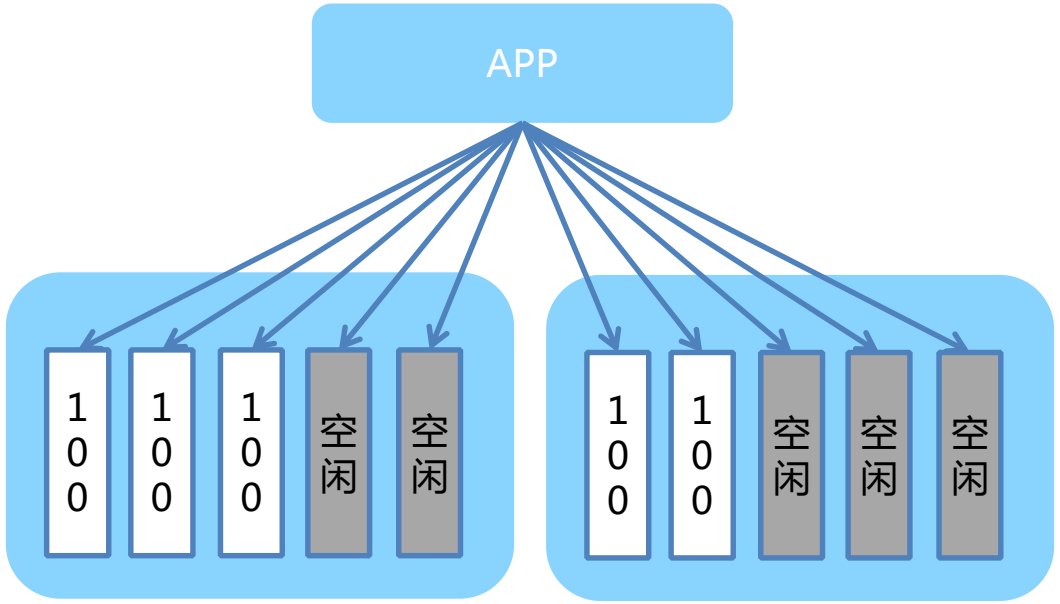
# 分布式调度：解决方案

## 拖尾问题




- 固定长度的任务切片导致长尾

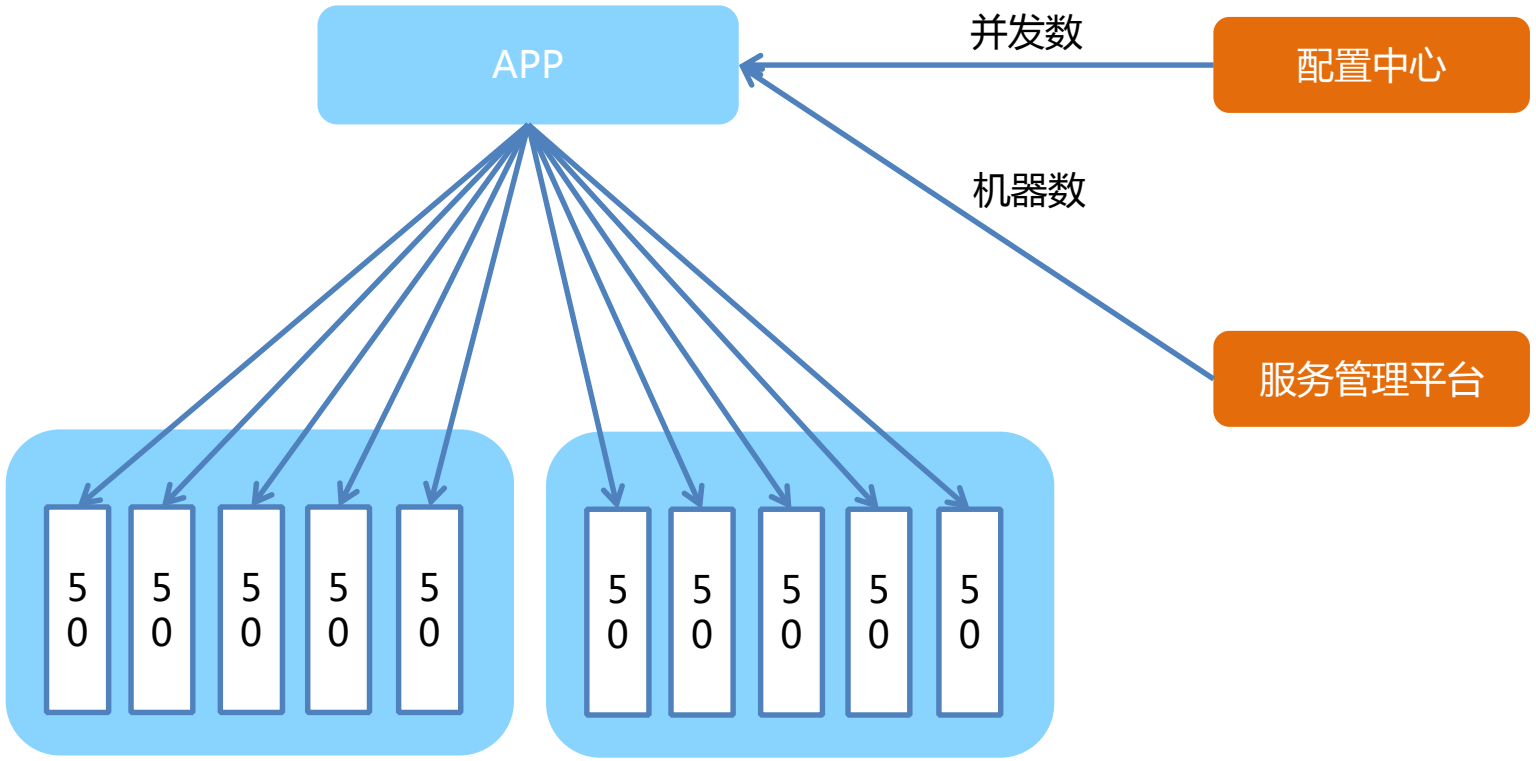
举例：500个用户场景




- 随着时间推移，需要处理的用户数减少后，固定用户数的任务切分方式，会造成任务生成过少，无法分发到足够的机器，造成整体集群并发度下降

# 分布式调度：解决方案

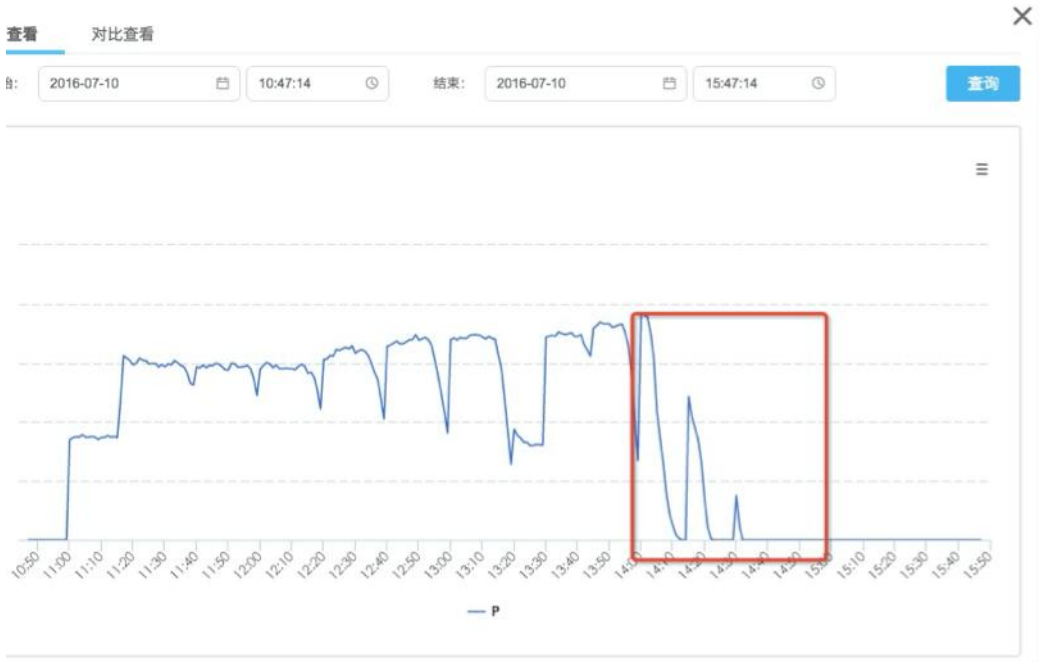
## 智能任务切片策略




- 任务切片策略
  - 基于在线机器数
  - 基于集群并发数
- 数据加载层逻辑
  - 获取集群并发数和机器数
  - 负责计算任务数据量
  - 执行切片
  - 分发到执行层

# 分布式调度：解决方案

## 对比


左图：改造前整个任务有持续1个小时拖尾时间

右图：改造后整体任务处理拖尾压缩到10分钟

而且整体效果会因为处理数据量越大，效果越明显

# 分布式调度

蚂蚁花呗
ANT CHECK LATER

展望

基于大数据个性化通知体系

基于大数据自动还款决策

# Q&A

蚂蚁花呗
ANT CHECK LATER

谢谢！