



移动开发平台-AB TEST

技术白皮书

1.11.0 版本

文档版本：V20200701

蚂蚁科技技术文档

蚂蚁科技集团有限公司版权所有 © 2019，并保留一切权利。

未经蚂蚁科技事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。

商标声明



及其他蚂蚁科技服务相关的商标均为蚂蚁科技所有。

本文档涉及的第三方的注册商标，依法由权利人所有。

免责声明

由于产品版本升级、调整或其他原因，本文档内容有可能变更。蚂蚁科技保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在蚂蚁科技授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过蚂蚁科技授权渠道下载、获取最新版的用户文档。如因文档使用不当造成的直接或间接损失，本公司不承担任何责任。

目 录

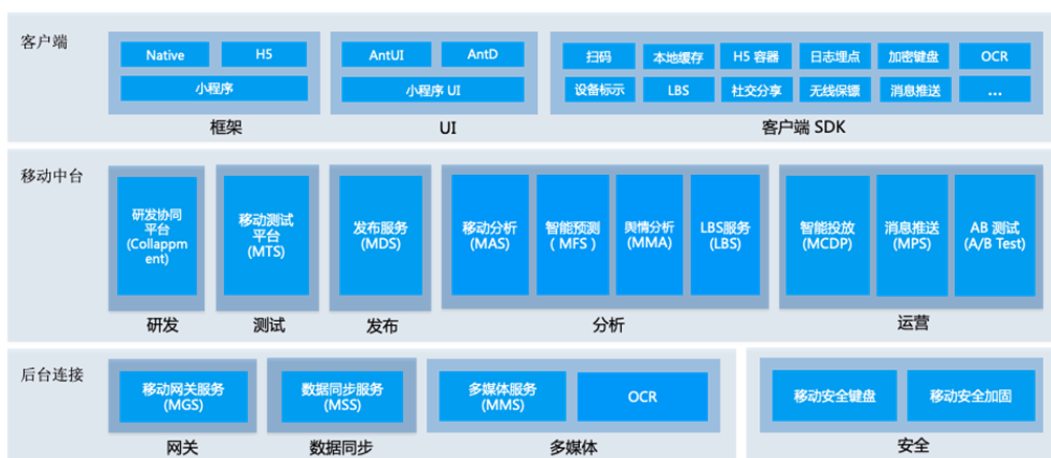
1 关于 mPaaS	1
1.1 什么是 mPaaS	1
1.2 系统架构	2
1.3 方案优势	2
2 ABTest ABTEST.....	6
2.1 什么是 ABTEST	6
2.1.1 产品背景	7
2.1.2 发展现状	7
2.2 产品优势	8
2.3 系统架构	8
2.4 性能指标	9
2.5 功能原理	9
2.5.1 ABTest 的统计科学原理	9
2.5.2 ABTest 的领域模型	10

1 关于 mPaaS

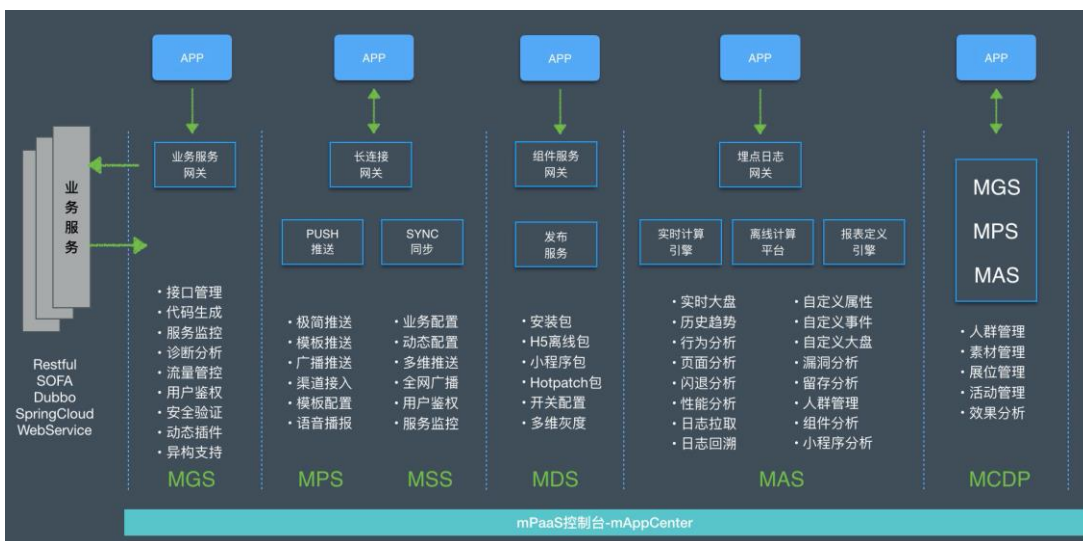
1.1 什么是 mPaaS

移动开发平台 (Mobile PaaS , 简称 mPaaS) 是源自于支付宝客户端 App 的移动开发平台, 为企业提供了移动开发、测试、运营及运维提供云到端的一站式解决方案, mPaaS 能有效降低技术门槛、减少研发成本、提升开发效率, 协助企业快速搭建稳定高质量的移动客户端 App。

mPaaS 能够提供 Native、H5、支付宝小程序三大开发框架; 100+ 的 UI 控件; 以及包括扫码, 本地缓存, 客户端埋点等 20+ 功能性 SDK, 可以让开发者快速接入搭建 App 所需要的基础能力。



mPaaS 专注于提供移动端所需功能组件, 其最核心的为 mAppCenter 控制台和 6 大功能组件：

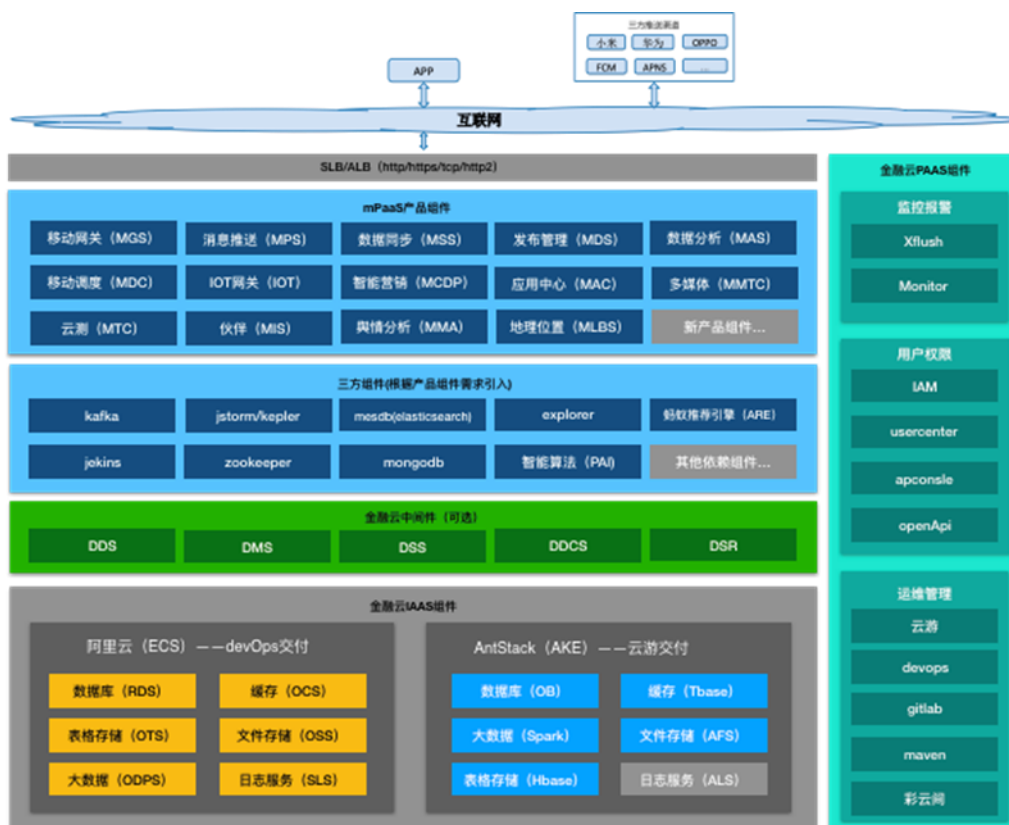


- **移动网关 MGS**：为连接移动客户端与服务端的组件产品, 负责客户端与服务端各业务系统之间的通讯、数据交换、配置管控。

- 移动推送 MPS：是 mPaaS 中的消息推送组件，针对不同的场景推出多种推送类型，满足个性化的推送需求。
- 移动同步 MAS：是 mPaaS 平台的一个核心基础服务组件，是一项方便您轻松地大规模收集、可视化并理解应用程序使用数据的服务。
- 移动同步 MSS：提供基于传输控制协议（Transmission Control Protocol，简称 TCP）和安全套接层（Secure Sockets Layer，简称 SSL）的安全数据通道，能够及时、准确、有序地将服务器端业务数据主动地同步到客户端 App。
- 实时发布 MDS：提供例如版本升级包、热修复包的管理和发布服务。支持通过白名单进行灰度发布，支持高级过滤规则，比如指定机型，来进行更精准的灰度发布。

1.2 系统架构

客户端开发和移动中台能力都是针对 App 本身，一个完整的 App 需要通过服务端来获取更高阶的能力。除了客户端框架和基础组件之外，云端基础服务（如 API 网关、SYNC 数据同步、PUSH 通知等）提供了接口管理、流量管控、用户鉴权、H5 离线包、热修复包、性能分析等运营运维能力，构建了一个高稳定、高可靠以及高效率的后台连接服务。



1.3 方案优势

- 高可用保障体系

mPaaS 线上的高可用保障体系由灰度发布、实时监控、诊断定位和容灾处理四大部分组成。

➤ 灰度发布

灰度发布是预防风险的有效手段之一，对于客户端来说，线下测试的再怎么完备也无法保证在用户环境下一切正常，直接发布至全量用户是非常危险的操作，在支付宝内部属于严重违规。我们的发布平台提供多种灰度策略，包括白名单灰度、时间窗灰度、百分比灰度、基于机型地域系统等维度的灰度。新版发布前，优先选取活跃用户和问题高发的机型进行灰度，灰度期间发现并修复问题，不断扩大灰度范围，直到闪退率、卡死率等指标符合发布标准后再进行全量的发布。

➤ 实时监控

首先，制定各种线上监控指标，包括闪退、卡死、流畅度、流量、内存、存储以及业务不可用埋点等。

其次，闪退、卡死、业务不可用埋点等高优先指标实时上报，第一时间发现异常；数据上报使用独立的进程，确保不影响主业务逻辑；当处于业务高峰时期，比如春节红包、双 11 等大型活动时，我们能够动态调整上报策略以缓解日志服务端的压力。除了自动上传和周期性上传策略外，我们通过下发诊断指令至客户端去获取平时用不到但驻留在客户端的日志，比如 logcat 日志。

➤ 诊断定位

mPaaS 能够根据客户端上报的各种埋点日志，完整描绘出用户的操作路径，根据这些信息，尝试重现用户的问题，数据的真实性相对用户提供的信息更加可靠，能减少错误信息产生的干扰。另外，通过诊断指令上传的 logcat 日志，获取更加完整的信息，能够更加清晰的定位问题。

➤ 容灾处理

预置开关到业务逻辑中，当出现业务大面积异常或资损时，后台推送业务开关至客户端中，将业务临时屏蔽下线。

客户端在启动阶段发生的死锁、闪退或者主页异常超过一定阈值时，将会自动清理应用数据，还原应用至初始状态，能够解决一部分由于数据异常导致的启动问题。

使用 hotpatch 技术可修复原生代码，同样 hotpatch 技术本身存在风险，因此也要经历灰度发布的阶段来逐步验证线上稳定性，一旦发生因 patch 引起的问题，需要立刻回滚 patch。

● 动态研发模式

相较于原生应用开发，Web 应用的优势非常明显：

- 只需要一套代码，Web 应用即可在 iOS 和 Android 客户端中运行，可相对减少开发人员投入
 - 每个用户日常使用的功能仅仅是支付宝庞大平台中的一小部分，H5 应用可实现动态下发，因此可以消除冗余的存储，降低包大小。
 - 采用 Web 技术作为动态研发模式的基础。
 - Web 应用迭代摆脱了客户端集中时间点发布的束缚，各业务线迭代计划变得自主可控。
- 打磨 Web 体验

针对 Web 应用在用户体验、性能方面与原生应用的差距，mPaaS 做了大量改进，主要在以下几个方面：

➤ 前后端分离

页面资源离线化，节省了资源请求消耗的时间，使页面打开速度提升明显，解决了在网络环境较差下容易出现白屏的问题。同时，数据请求使用 Native 网络通道，可优化的空间更大，安全性更高。

➤ 增量更新

客户端更新某业务应用版本时，无需下载完整的新版本资源包，仅需下载由发布平台根据客户端本地安装版本计算生成的体积更小的增量包，这样不仅能够节省带宽和流量，也提升了业务更新的速度。

➤ 推拉结合

解决业务最新版本覆盖率的问题，每次发布新版本时，业务可主动触发消息到客户端，客户端收到通知后会更新该业务应用版本。同时，客户端会定时检查服务端是否有版本发布，这样能够保证版本发布后大多数用户在短时间内获得最新的应用。

➤ 容错补偿

发布离线资源时，发布平台会自动生成对应的在线 URL 并配置到应用信息中，当客户端加载 Web 应用时发现离线包不可用，会立刻启用该 URL 加载内容，能够最大程度保证业务可用性。解决由于网络、安全或者存储权限等原因，导致客户端不能使用或者及时获得离线包的问题。

➤ Android 独立浏览器内核

在 Android 系统中，不同系统、不同厂商中的浏览器内核存在差异，这导致层出不穷的兼容性问题，为了解决并掌控这些问题，mPaaS 引入了独立的 UC 浏览器内核并集成在应用中。根据数据统计，使用 UC 浏览器内核后浏览器相关的闪退和

ANR 有明显的下降。同时，针对安全上出现的漏洞，可以在第一时间修复并发布，远比厂商升级更有效率。

➤ Web 应用全方位监控

资源加载异常、JS 执行异常、白屏、加载耗时等性能数据会被收集上报至后台，帮助不您及时发现异常。

● 节约开发测试成本，提升研发效率

➤ 统一的 App 开发标准

借助统一的客户端框架和开发标准，满足大团队协同开发需求，降低系统之间的耦合性。开发框架采用模块化的架构思路，具备高度可扩展性。IDE 插件提供工程化能力，可以快速上手，提高开发效率。

➤ 节约开发测试成本

- 自动构建打包，自动化测试，减少人力成本。
- 增强构建，发布管控。
- 标准自动化构建，打包，发布流程，增加效率，减少问题。

● 构建数字化运营能力，深入洞察用户行为

➤ 建立完善的客户端“数据采集、筛选、分析”链路

针对客户端业务数据进行统一采集、筛选和分析，通过标准格式，打通移动分析服务、消息推送等组件沉淀的业务数据，支持自定义事件设置。

➤ 用数据分析指导科学的业务决策

移动分析产生的数据可直接输出消息推送组件进行人群定向应用，用数据指导科学决策和精细化运营。

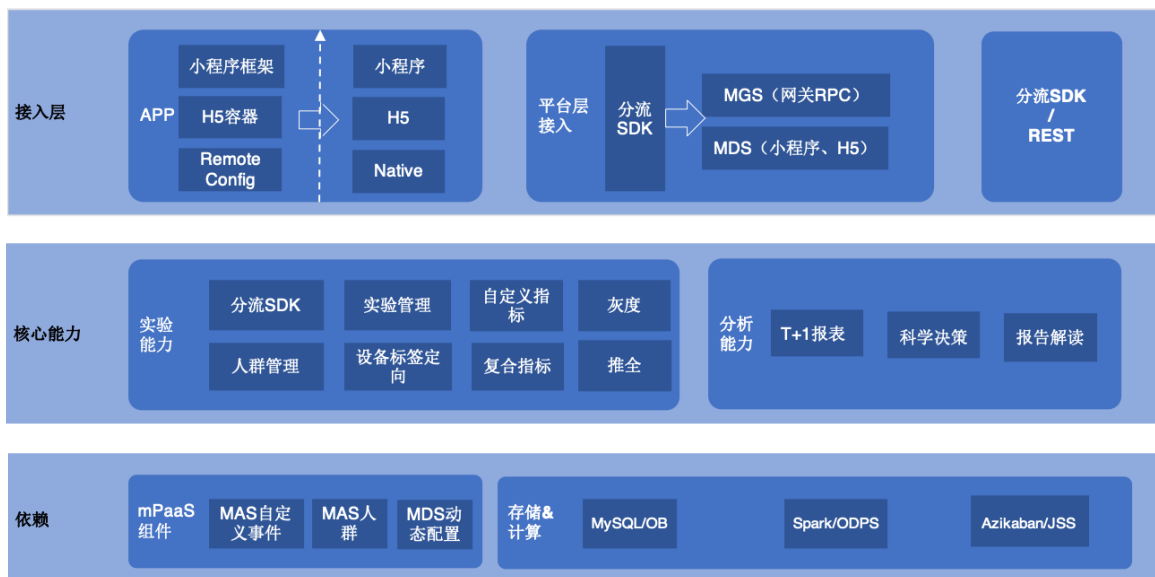
2 AB Test

2.1 什么是 AB Test

AB 实验是蚂蚁金服金融科技提供的一项用来验证产品新功能效果的科学在线测试方法，它能够通过对实际实验数据进行统计分析从而帮助用户做出对产品的不同功能方案进行取舍的决策。

在当前的业务发展环境下，每个公司都希望自己的业务能够健康稳定地发展，而合理地规划产品迭代的策略就显得尤为重要。AB 实验可以帮您找出这个更优秀的策略。无论您是想检测“不同 UI 或交互的产品效果”，还是“不同算法的优劣”或是“不同的广告投放方式的转化效果”，AB 实验都可以帮您实现。

AB 实验可与 mPaaS 框架中的其他组件相互关联，如移动分析、移动网关、实时发布等，形成一个完整的生态，便于您随时开展不同类型的 AB 实验。



产品功能如下：

- 实验室管理

创建实验室、管理实验室成员、在实验室废弃之后删除实验室。

在实验室下创建实验、调试实验检查实验方案对应的功能、运行实验立刻让实验生效、推全实验把表现好的实验方案推广到全部用户、结束实验快速下线实验。

实验运行过程中可以随时查看实验数据，关注核心指标，作出实验决策。

- 指标管理

系统内置了实验分流 PV、UV、7 日留存率三个通用指标。用户可以通过引用 MAS (Mobile Analysis Service) 的自定义事件埋点作为核心转化指标，即 MAS 指标。系统支持通过 MAS 指标的交集、差集、并集、除等 4 个计算方式创建复合指标。

- 变量管理

支持实验变量的创建和删除。通过引用 MDS (Mobile Delivery Service) 客户端配置模块的动态配置变量可以新建出 MDS 变量，新建自定义变量时支持指定默认值。

- 应用数据概览

应用数据概览页展示此 App 下实验的概览数据，包括：

1. 昨日全部流量
2. 本月累计流量
3. 运行中实验个数
4. 历史实验总个数
5. 流量走势图
6. 流量占用情况

2.1.1 产品背景

近年来，随着互联网的快速发展，Growth-Hacking 已经是一个很普遍的概念。Growth-Hacking 的目的就是使用更小更灵活的成本通过数据驱动来挖掘产品增长的奥秘。同时在 AARRR 这个模型中，打造成不断优良循环的流程，需要从数据分析中发现产品功能、运营策略与转化之间的相关性，思考他们之间因果关系。

如何衡量思考、创新想法的正确性？数据是最好的衡量标准，这就因此要求我们需要运用一些工具。而 AB 测试就是一个快速试错，用户影响尽可能小，通过数据科学决策的工具，它是 Growth-Hacking 最基础也是最重要的工具之一。

自 2000 年谷歌工程师将 ABTest 应用在互联网产品以来，A/B 测试在国内外越来越普及，逐渐成为互联网数据驱动产品的重要体现，通过 A/B 测试来驱动产品设计迭代优化。

2.1.2 发展现状

AB Test 源于蚂蚁金服集团下实验系统 Compaas，刚开始只服务于一些金融机构，目前服务于金融云公有云，政府、30 多家银行等私有云，支付宝国际合作伙伴等。未来会服务于更多行业和公司

2.2 产品优势

- **自带客户端 SDK**

自带的客户端 SDK 可处理实验信息的拉取，分流信息、转化事件的埋点、上报。

- **支持多种不同类型的实验**

支持的实验类型有两种：客户端实验以及服务端实验。客户端实验的变量通过动态配置服务推送到客户端，快速生效；服务端实验通过 REST 接口开放分流服务，支持任意场景灵活对接。

- **支持编程模式**

您可通过客户端编程的方式获取变量值，以用于所需的实验。

- **支持 iOS/Android/H5/小程序/服务端**

您可在各种不同的系统、环境中开展 AB 实验，已达成所期望的实验效果。

- **支持多实验室**

一个应用下的不同实验室共享全部流量，同一个实验室下不同实验间流量互斥，合理分配帮助您科学有效使用实验流量。

- **支持定向实验**

通过设备标签、MAS 人群，您可对圈定的用户将不同的实验定向投放至指定人群，以获得更精准的实验效果。

- **支持创建复合指标**

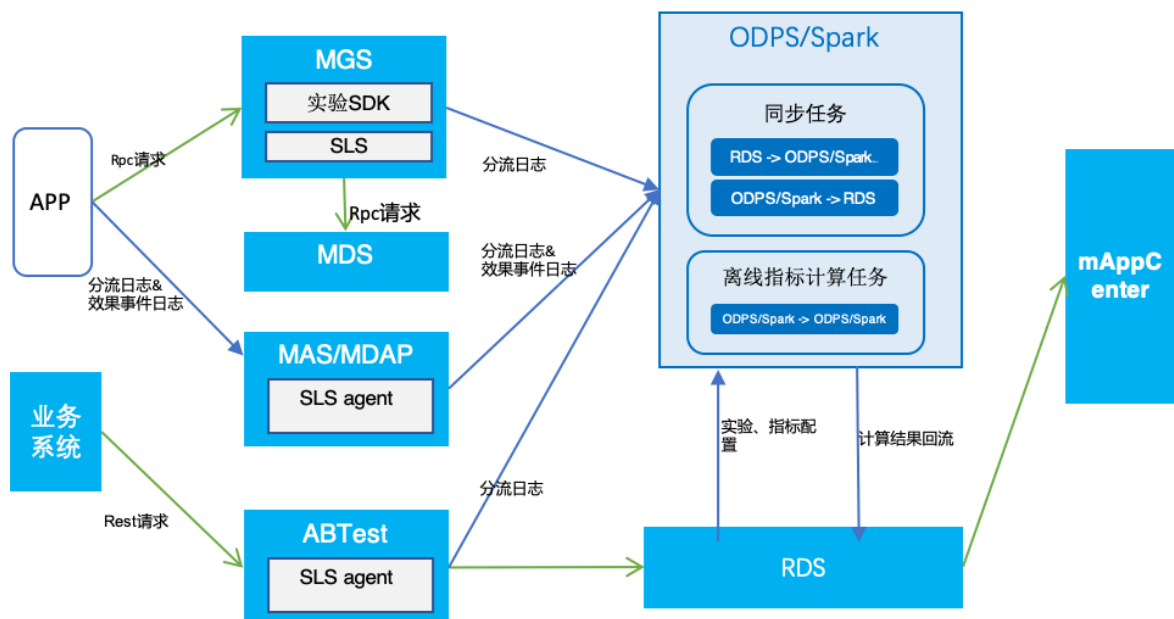
在已有指标不能满足您的需求时，您可以自定义复合指标。

- **支持权限管理**

同一个工作空间中，仅有获得权限的成员可以参与 AB 实验的操作，实现不同角色的权限分离。

2.3 系统架构

系统架构如下图所示。



- **MGS**：网关系统，作为连接移动客户端与服务端的组件产品，负责客户端与服务端各业务系统之间的通讯、数据交换。
- **MDS**：实时发布组件（Mobile Delivery Service，简称 MDS）提供例如版本升级包、热修复包的管理和发布服务，支持客户端 App 的动态配置发布更新。
- **MAS**：MAS 移动分析服务（Mobile Analysis Service，简称 MAS）是 mPaaS 平台的一个核心基础服务组件，是一项方便您轻松地大规模收集、可视化并理解应用程序使用数据的服务。
- **MDAP**：MAS 日志接口网关，接收移动端日志，dump 到本机磁盘。同机上有 flume 接收移动端日志，发送到 Kafka、HDFS 上 MAS 产品下的日志网关应用，
- **mAppCenter**：mPaaS 的统一控制台。

2.4 性能指标

- MCDP 自带客户端 SDK，会按照配置策略缓存广告投放数据，减少到达客户端广告拉取量。
- MCDP 完全支持水平扩展，在请求突增的情况下，可轻松扩容。

2.5 功能原理

2.5.1 AB Test 的统计科学原理

A/B 测试的实验原理是统计学上假设检验（显著性检验）的一种形式：假设检验中的参数检

验是先对总体的参数提出某种假设，然后利用样本数据判断假设是否成立的过程。

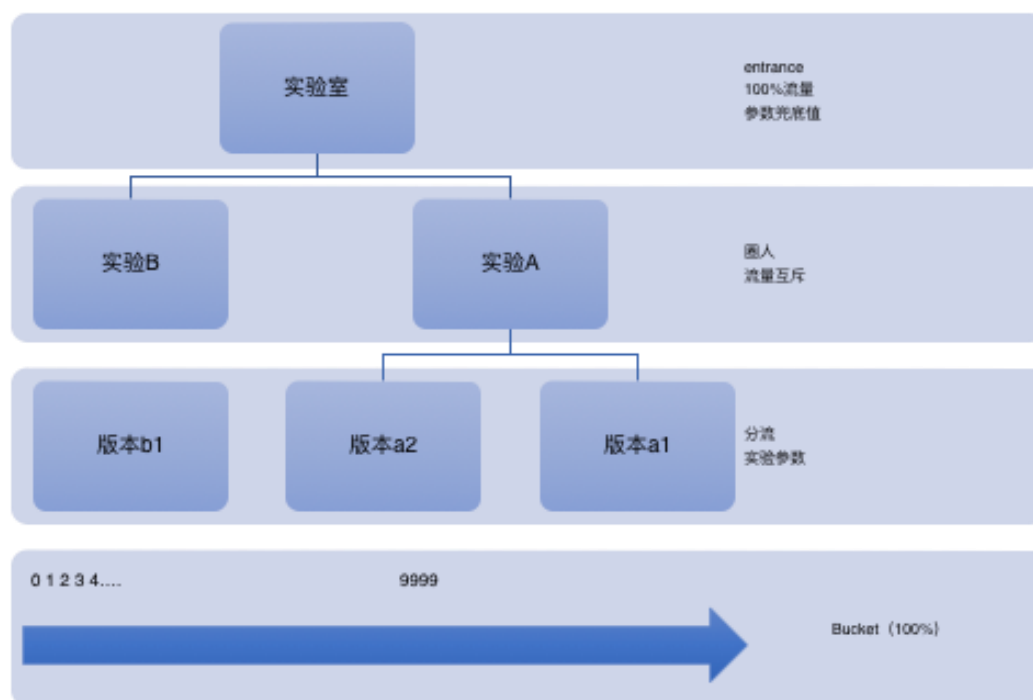
逻辑上运用反证法，统计上依据小概率思想：

小概率思想是指小概率事件（显著性水平 $p < 0.05$ ）在一次实验中基本上不会发生。

反证法是指先提出假设，再用适当的统计方法确定假设成立的可能性大小；如可能性小，则认为假设不成立。

具体到对比实验，就是假设测试版本的总体参数（优化指标均值）等于对照版本的总体参数，然后利用这两个版本的样本数据来判断这个假设是否成立。

2.5.2 AB Test 的领域模型



AB Test 通过“实验室”实现了一个分层，即每个“实验室”都拥有独立的 100% 流量。

- **实验室**

100% 的流量入口。支持设置参数的默认值兜底。业务上相互独立的每个业务应该拥有一个独立的实验室。

- **实验**

每个实验室下可以创建多个实验，实验间的流量是互斥的，达到流量隔离的目的。同时实验级别支持圈人定向条件，支持定向实验。

- **实验方案**

一个实验下可以创建多个实验方案 (版本), 实验方案上绑定实验参数的实验值。每个方案的流量分配通过和 1000 个 bucket 间的映射关系来分配。

