



腾讯企点营销云 AB 实验分析 产品操作手册



腾讯云计算（北京）有限责任公司

版权声明

本文档著作权归腾讯云计算（北京）有限责任公司（以下简称“腾讯云”）单独所有，未经腾讯云事先书面许可，任何主体不得以任何方式或理由使用本文档，包括但不限于复制、修改、传播、公开、剽窃全部或部分本文档内容。

本文档及其所含内容均属腾讯云内部资料，并且仅供腾讯云指定的主体查看。如果您非经腾讯云授权而获得本文档的全部或部分内容，敬请予以删除，切勿以复制、披露、传播等任何方式使用本文档或其任何内容，亦请切勿依本文档或其任何内容而采取任何行动。

商标声明



“腾讯”、“腾讯云”及其它腾讯云服务相关的商标、标识等均为腾讯云及其关联公司各自所有。若本文档涉及第三方主体的商标，则应依法由其权利人所有。

免责声明

本文档旨在向客户介绍本文档撰写时，腾讯云相关产品、服务的当时的整体概况，部分产品或服务在后续可能因技术调整或项目设计等任何原因，导致其服务内容、标准等有所调整。因此，本文档仅供参考，腾讯云不对其准确性、适用性或完整性等做任何保证。您所购买、使用的腾讯云产品、服务的种类、内容、服务标准等，应以您和腾讯云之间签署的合同约定为准，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证

一、入门指引

1. 什么是 TAB？

TAB（Tencent AB Test）是一站式的 AB 实验平台，覆盖接入、分流、指标计算、数据报告整个实验流程，帮助我们更高效更科学地通过 AB 实验进行产品迭代及决策。

2. TAB 的目标？

➤ More

支持更多的实验同时运行；

在整体流量固定的情况下，可支持更多实验。

➤ Better

更好的易用性；

能识别更好的实验，不合理的实验不允许在线上流量进行。合理的但效果很差的实验，比如有 bug 的实验或是产生很差用户体验的实验，能很快的被识别并且下线。

➤ Faster

支持快捷创建实验，同时支持不具备工程师技能的同学可创建实验；

可以快速产出实验指标，便于分析实验结果。

3. 什么是 AB 实验？

AB 实验源自于生物医学里的双盲测试。在双盲测试中，病人被随机分成两组，在病人不知情的情况下分别给予安慰剂和测试用药进行服用。经过一段时间的观察，实验人员比较这两

组病人的病情变化是否具有差异，并以此来判断测试用药是否有效。

互联网公司的 AB 实验也延用了该概念。假设目前产品迭代中有 A、B 两个方案（例如方案 A 使用红色 banner，方案 B 使用蓝色 banner），业务可以通过对两个或多个属性相似的用户群体，分别在测试 A 方案和 B 方案带来的效果，收集相关数据，然后基于数据分析，给出最好产品方案并进行上线使用。

4. 为什么要进行 AB 实验？

➤ 获取真实的线上反馈

通过 AB 实验可以明确方案 A 和方案 B 在线上带来的效果差异，并从统计学的角度证实该上线方案与效果间的因果关系，同时较为准确与有效的量化上线策略所带来的增量。

➤ 小成本、低风险

大型互联网公司，每年每个季度要上线各种优化策略（包含：UI 设计、算法优化等等），AB 实验可以帮助业务做决策时在小成本和低风险的情况下，完成对最优解决方案的判断，并且使其决策的收益最大化。

5. 关键术语

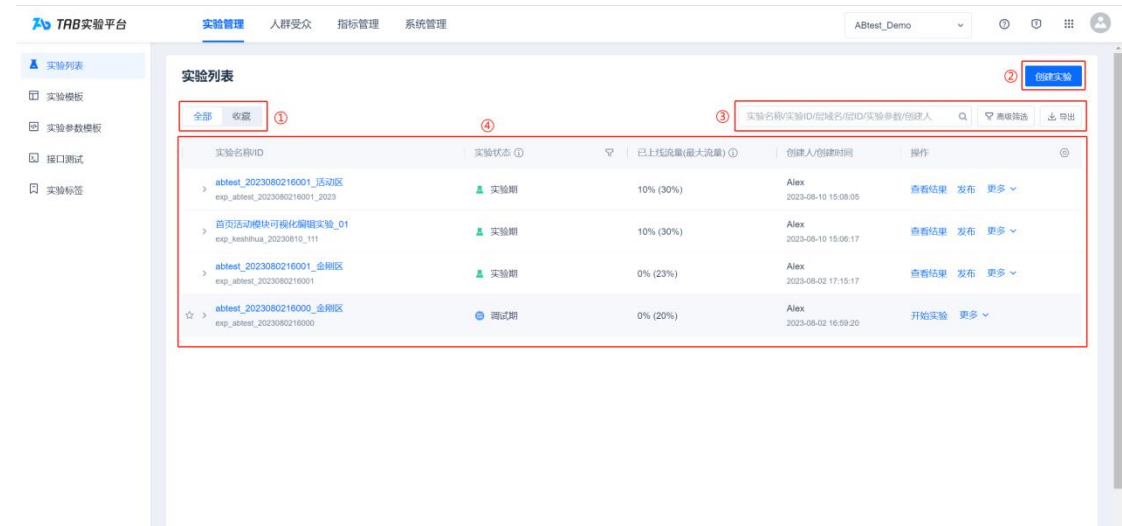
名词	解释
实验组、实验版本	一个实验组包含两个或两个以上的实验版本，针对每个版本下发不同的策略，并观察它们的效果。
实验参数	是对实验业务策略、功能的配置化映射，即不同的实验参数对应的不同的业务策略、功能。
受众规则	按照特定的规则下发实验版本。目前支持 2 种规则设定：

	<ul style="list-style-type: none"> • 按比例下发：按照流量比例随机下发； • 按标签下发：针对具有某种特性的用户群下发。
白名单用户	白名单用户是一群特殊的实验受众，在实验分流时被忽略，直接被配置实验版本，用以测试实验策略是否正常生效，是否存在 bug 需要修复。白名单用户数据不会计入实验结果。
层域管理	<p>对于一个业务产品而言，用户的流量是一定的。在 AB 实验中通过层域对流量进行合理划分，保证流量高可用以及分流的灵活性和科学性。</p> <ul style="list-style-type: none"> • 域：指垂直划分流量的分区。并列域之间是互斥的，不重叠的，可以更好地进行互不干扰的实验，所有的流量域加起来共享 100% 流量。 • 层：域流量的子集，用户流量被划分到不同桶中。同个域下不同层的流量是重叠的，通过 hash 对流量进行打散，使得多个层之间的用户流量是正交均匀的。 • 桶：用户流量的集合，通过 hash 函数将层的流量均匀划分到一定数目的桶内。
流量分桶	分为“百分桶”、“千分桶”、“万分桶”、“十万分桶”，用于支持流量动态伸缩，实现流量的完全正交。
实验状态	<ul style="list-style-type: none"> • 调试期：实验创建完毕进入调试期，表示实验还未正式开始。 • 空跑期：实验开始，编辑流量/设置上线流量后可进行 AA 空跑验证分流均衡性。 • 实验期：实验开始，编辑流量/设置上线流量后真正的 AB 策略

	<p>会分别生效，会进行实验数据累积。</p> <ul style="list-style-type: none">• 已全量：这个业务的所有线上用户，已生效实验的某个策略。• 已下线：AB 实验已下线。
--	---

二、实验管理

1. 实验列表



① 切换实验列表中罗列的 AB 实验——

- **全部**：当前业务创建的所有 AB 实验
- **收藏**：当前用户收藏的 AB 实验

② 点击【创建实验】按钮，进入实验创建流程

③ 搜索、筛选、导出 AB 实验——

- **搜索**：支持 **实验名称、实验 ID、层域名称、层 ID、实验参数、创建人** 的关键词模糊搜索 AB 实验
- **筛选**：支持按 **层域名称、层 ID、实验标签、创建人、创建时间范围、实验场景** 维度筛选 AB 实验

全部收藏

实验名称/实验ID/层域名/层ID/实验参数/创建人

收起筛选

导出

层(层域)名称/ID

实验标签

创建人

创建时间范围

请选择

请选择

全部

开始日期

至

结束日期

实验场景

重置

请选择

- **导出：**支持以 **CSV 格式**导出当前的实验列表



④ 浏览、筛选、以及操作 AB 实验——

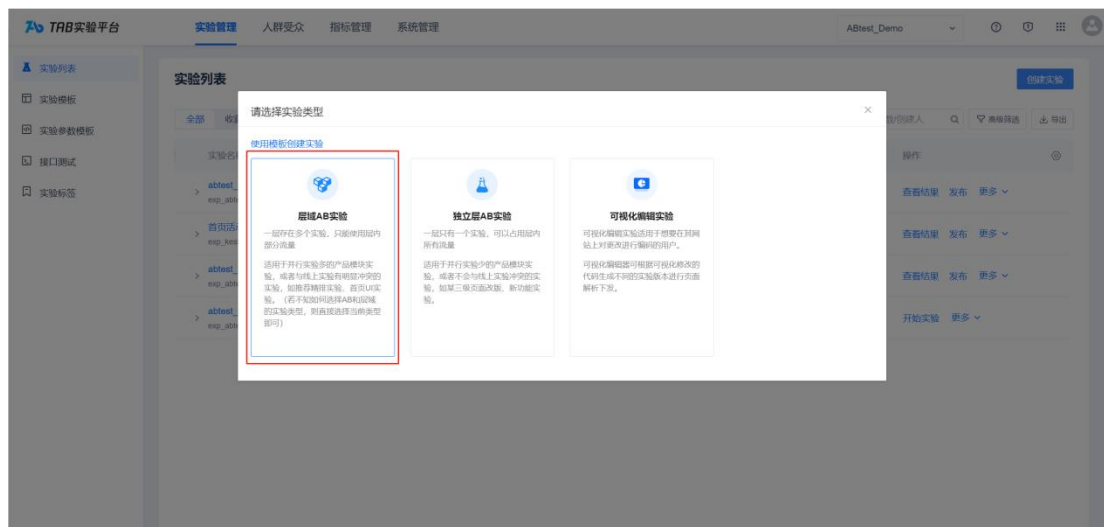
- 可操作项：
 - **开始实验**：开始进行实验，需要判断是先进入空跑期，还是直接进入实验期；
 - **进入实验期**：从空跑期进入实验期；
 - **管理白名单**：对当前实验各个实验版本的白名单组、临时白名单进行管理；
 - **接口测试**：输入分流 ID 以及一些条件，来查看该 ID 命中哪个实验版本、实验版本参数是什么、等等；
 - **编辑流量/设置线上流量**：编辑当前层线上的真实流量；
 - **放量/设置最大流量**：在空跑期、实验期，当线上流量不够用时申请设置最大流量；
 - **查看结果**：查看实验结果，包含实验报告、实验监控等页面信息；
 - **复制**：对实验所有配置信息进行复制，创建一个新的实验；
 - **发布**：实验已经获得预期效果，比如某个正向显著，将该版本进行全量发布；
 - **下线**：实验结束下线，实验下线后不可再恢复；
 - **删除**：实验数据删除。
- 支持修改表头展示字段；



2. 创建实验

2.1 层域 AB 实验

点击【创建实验】按钮，选择“层域 AB 实验”类型，进入实验创建流程：



① 填写基本信息——

- **实验 ID：**必填项，ID 全局唯一，需字母开头，且仅支持字母、数字、及下划线组成。
例如：exp_产品_子产品_业务_实验名称。
- **实验名称：**必填项，建议取与实验内容相关的名称，如有实验版本迭代可以增加版本号后缀，方便用户快速理解实验内容、是在哪个迭代版本的。
- **实验标签：**必选项，给实验打上的能标识实验特性的信息。实验标签中带“*”的标签

选项，都是必填的。

- **实验负责人：**必选项，默认是实验的创建者，也支持自定义实验负责人。实验负责人对实验的操作权限，同实验创建者。
- **实验设计：**可以填写实验方案或背景信息，例如提出问题、建立实验假设，等等。
 - 提出问题——例如为什么注册页面转化率不高？为什么推荐系统点击率不理想？
 - 建立实验假设——例如让注册按钮文案更有吸引力将提升点击率,或者更换一套推荐模型将提升点击率。

② 选择关注指标——

- **实验强相关指标：**从指标列表中挑选与当前实验强相关的指标
 - **T1-期望提升指标：**当前实验要验证的核心提升目标。
 - ◆ 添加数量不超过 50 个
 - ◆ 只可选择“均值类、比例类、比率类”3种类型的指标
 - **T2-其他普通指标：**不属于实验设计范围中的指标，可作为实验额外观察的指标。
- **业务默认添加指标：**无需用户操作，默认会将与业务强相关的指标加入当前实验
 - **T0-北极星指标：**业务的终极目标，是公司级别的战略性指标。
 - **T1-护栏指标：**测量不良用户体验的指标，是实验中不应该显著负向的指标。旨在

保护业务不出现异常。

- **实验结论设置：**配置规则，使后续实验上线后能够按照规则自动生成结论



③ 设置实验版本——

不同的实验参数（key、value）对应的不同的业务策略、功能，配置参数需与代码中约定的配置保持一致，主要用来支持此实验参数的下发。建议联系研发共同填写。

- **版本名称、版本 key：**实验版本和对照版本是一组相对概念。假设在实验中，参与的用户被随机地分配到 A 组和 B 组中，A 组用户在产品中体验到新策略，B 组用户在实验中体验的仍旧是旧策略。在这一实验过程中，A 组便为实验组，B 组则为对照组。
 - 实验中至少有 1 个实验组和 1 个对照组；
 - 输入详细、有辨识度的版本名称、版本 key，方便后续查看时更好的了解版本信息；
 - 点击【+添加版本】按钮，为当前实验创建更多实验版本。
- **配置实验参数：**实验参数是对实验业务策略、功能的配置化映射，即不同的实验参数对应的不同的业务策略、功能。配置参数需与代码中约定的配置保持一致。
 - 点击【+添加参数】按钮，添加更多参数；
 - 点击【使用模板】按钮，可以选择创建过的参数模板。

- **复制：**复制某个实验版本的实验参数信息
- **删除：**删除某个实验版本



④ 设置受众用户——

- **选择层域：**选择实验层，支持查看当前实验层的流量被哪些实验占用，剩余多少流量
- **选择受众规则：**
 - **实验层流量比例：**选择实验层后，该实验层剩余的流量比例
 - **按人群标签：**从人群标签清单中选择标签并配置标签之间的规则，从而实现人群圈选。只有被圈选的人群才可参与当前实验。
- **实验总流量：**当前实验各个版本在当前层预计最大可上线的流量比例。可以认为是“占坑流量”，后续编辑/设置线上流量需要在占坑的范围内进行设置。例如”实验版本 1“预设最大流量为 20%，那么后续编辑/设置线上流量只可 $\leq 20\%$ 。一般情况，不同版本的流量占比是平均的，尽可能保障实验置信度；各实验版本的最大流量占比的总和 $\leq 100\%$ 。
- **版本流量分配：**默认均匀分配，也可选择自定义流量分配。

以上，所有信息填写完毕，点击【进入调试期】按钮，实验即可创建成功。

2.2 独立层 AB 实验

点击【创建实验】按钮，选择“独立层 AB 实验”类型，进入实验创建流程：

① 填写基本信息——

- **实验 ID：必填项**，ID 全局唯一，需字母开头，且仅支持字母、数字、及下划线组成。

例如：exp_产品_子产品_业务_实验名称。

- **实验名称：必填项**，建议取与实验内容相关的名称，如有实验版本迭代可以增加版本号后缀，方便用户快速理解实验内容、是在哪个迭代版本的。

- **实验标签：必选项**，给实验打上的能标识实验特性的信息。实验标签中带“*”的标签选项，都是必填的。
- **实验负责人：必选项**，默认是实验的创建者，也支持自定义实验负责人。实验负责人对实验的操作权限，同实验创建者。
- **实验设计**：可以填写实验方案或背景信息，例如提出问题、建立实验假设，等等。
 - 提出问题——例如为什么注册页面转化率不高？为什么推荐系统点击率不理想？
 - 建立实验假设——例如让注册按钮文案更有吸引力将提升点击率,或者更换一套推荐模型将提升点击率。

The screenshot shows the 'TAB实验平台' (TAB Experimental Platform) interface. The top navigation bar includes '实验管理' (Experiment Management), '人群受众' (Audience), '指标管理' (Indicator Management), and '系统管理' (System Management). The main content area is titled '实验列表 / 创建实验' (Experiment List / Create Experiment). The form is divided into four steps: 1. 填写基本信息 (Fill in basic information), 2. 选择关注指标 (Select key indicators), 3. 设置实验版本 (Set experiment version), and 4. 设置受众用户 (Set audience users). The first step is active, showing fields for '实验ID' (Experiment ID) with value 'exp_234234', '实验名称' (Experiment Name) with value '4234234', '实验标签' (Experiment Label) with value '标准实验', '实验负责人' (Experiment Owner) with value 'Alice', and '实验设计' (Experiment Design) with a text area containing a placeholder about improving registration conversion. Navigation buttons '取消' (Cancel) and '下一步' (Next Step) are at the bottom right.

② 选择关注指标——

- **实验强相关指标**：从指标列表中挑选与当前实验强相关的指标
 - **T1-期望提升指标**：当前实验要验证的核心提升目标。
 - ◆ 添加数量不超过 50 个
 - ◆ 只可选择“均值类、比例类、比率类”3种类型的指标
 - **T2-其他普通指标**：不属于实验设计范围中的指标，可作为实验额外观察的指标。
- **业务默认添加指标**：无需用户操作，默认会将与业务强相关的指标加入当前实验
 - **T0-北极星指标**：业务的终极目标，是公司级别的战略性指标。

- **T1-护栏指标**：测量不良用户体验的指标，是实验中不应该显著负向的指标。旨在保护业务不出现异常。

- **实验结论设置**：配置规则，使后续实验上线后能够按照规则自动生成结论



③ 设置实验版本——

不同的实验参数（key、value）对应的不同的业务策略、功能，配置参数需与代码中约定的配置保持一致，主要用来支持此实验参数的下发。建议联系研发共同填写。

- **版本名称、版本 key**：实验版本和对照版本是一组相对概念。假设在实验中，参与的用户被随机地分配到 A 组和 B 组中，A 组用户在产品中体验到新策略，B 组用户在实验中体验的仍旧是旧策略。在这一实验过程中，A 组便为实验组，B 组则为对照组。

- 实验中至少有 1 个实验组和 1 个对照组；
- 输入详细、有辨识度的版本名称、版本 key，方便后续查看时更好的了解版本信息；
- 点击【+添加版本】按钮，为当前实验创建更多实验版本。

- **配置实验参数**：实验参数是对实验业务策略、功能的配置化映射，即不同的实验参数对应的不同的业务策略、功能。配置参数需与代码中约定的配置保持一致。

- 点击【+添加参数】按钮，添加更多参数；
- 点击【使用模板】按钮，可以选择创建过的参数模板。

- **复制**：复制某个实验版本的实验参数信息
- **删除**：删除某个实验版本



④ 设置受众用户——

- **选择层域**：选择层域，在该层域下新建一个实验层，可以占满 100% 的流量
- **关联实验场景**：标记当前实验归属哪个场景
- **选择受众规则**：
 - **实验层流量比例**：选择实验层后，该实验层剩余的流量比例
 - **按人群标签**：从人群标签清单中选择标签并配置标签之间的规则，从而实现人群圈选。只有被圈选的人群才可参与当前实验。
- **实验总流量**：当前实验各个版本在当前层预计最大可上线的流量比例。可以认为是“占坑流量”，后续编辑/设置线上流量需要在占坑的范围内进行设置。例如”实验版本 1“预设最大流量为 20%，那么后续编辑/设置线上流量只可 $\leq 20\%$ 。一般情况，不同版本的流量占比是平均的，尽可能保障实验置信度；各实验版本的最大流量占比的总和 $\leq 100\%$ 。
- **版本流量分配**：默认均匀分配，也可选择自定义流量分配。

实验列表 / 创建实验

1 填写基本信息 2 选择关注指标 3 设置实验版本 4 设置受众用户

设置目标受众

基于设置的层域信息，选择你希望参与此实验的受众群体以及流量比例。

* 选择层域 默认多层域

* 选择受众规则 ☒ 实验层流量比例 ☐ 按人群标签

* 实验的流量 100 % 该实验层可用流量: 0%~100%

* 版本流量分配 ☒ 均匀分配 ☐ 自定义流量分配

版本名称	最大可上线流量
exp_234234_A	均匀分配50%
exp_234234_B	均匀分配50%

☒ 实验创建完毕进入调试期，您可在实验列表页点击「更多-进入实验调试期」来开启实验

☐ 保存实验配置为模板

以上，所有信息填写完毕，点击【进入调试期】按钮，实验即可创建成功。

2.3 可视化编辑实验

点击【创建实验】按钮，选择“可视化编辑实验”类型，进入实验创建流程：

1. 填写基本信息——

- 实验 ID: **必填项**，ID 全局唯一，需字母开头，且仅支持字母、数字、及下划线组成。

例如：exp_产品_子产品_业务_实验名称。

- 实验名称: **必填项**，建议取与实验内容相关的名称，如有实验版本迭代可以增加版本号后缀，方便用户快速理解实验内容、是在哪个迭代版本的。

- **实验标签：必选项**，给实验打上的能标识实验特性的信息。实验标签中带“*”的标签选项，都是必填的。
- **实验负责人：必选项**，默认是实验的创建者，也支持自定义实验负责人。实验负责人对实验的操作权限，同实验创建者。
- **实验设计**：可以填写实验方案或背景信息，例如提出问题、建立实验假设，等等。
 - 提出问题——例如为什么注册页面转化率不高？为什么推荐系统点击率不理想？
 - 建立实验假设——例如让注册按钮文案更有吸引力将提升点击率,或者更换一套推荐模型将提升点击率。

The screenshot shows the 'TAB实验平台' (TAB Experimental Platform) interface. The main navigation bar includes '实验管理' (Experiment Management), '人群受众' (Audience), '指标管理' (Metrics), and '系统管理' (System Management). The left sidebar lists '实验列表' (Experiment List), '实验模板' (Experiment Template), '实验参数模板' (Experiment Parameter Template), '接口测试' (API Test), and '实验标签' (Experiment Label). The main content area is titled '实验列表 / 创建实验' (Experiment List / Create Experiment). It features a progress bar with four steps: 1. 填写基本信息 (Fill in basic information), 2. 选择关注指标 (Select metrics to track), 3. 设置实验版本 (Set experiment version), and 4. 设置受众用户 (Set audience users). The first step is active. Below the progress bar, the '基本信息' (Basic Information) section contains the following fields:

- * 实验ID (Experiment ID): exp_ (Format: starts with 'exp_', supports numbers and underscores, example: exp_产品_子产品_业务_实验名称)
- * 实验名称 (Experiment Name): 请输入实验名称, 长度不超过25
- * 实验标签 (Experiment Label): 请选择实验标签 (dropdown)
- 实验负责人 (Experiment Owner): Alex (dropdown)
- 实验设计 (Experiment Design): 基于【前置的XXX数据、XXX调研反馈】做出假设：如果【修改XXX功能】，可以达到【提高用户粘性】的效果，此效果可以使用【】来量化。

At the bottom right, there are '取消' (Cancel) and '下一步' (Next Step) buttons.

2. 选择关注指标——

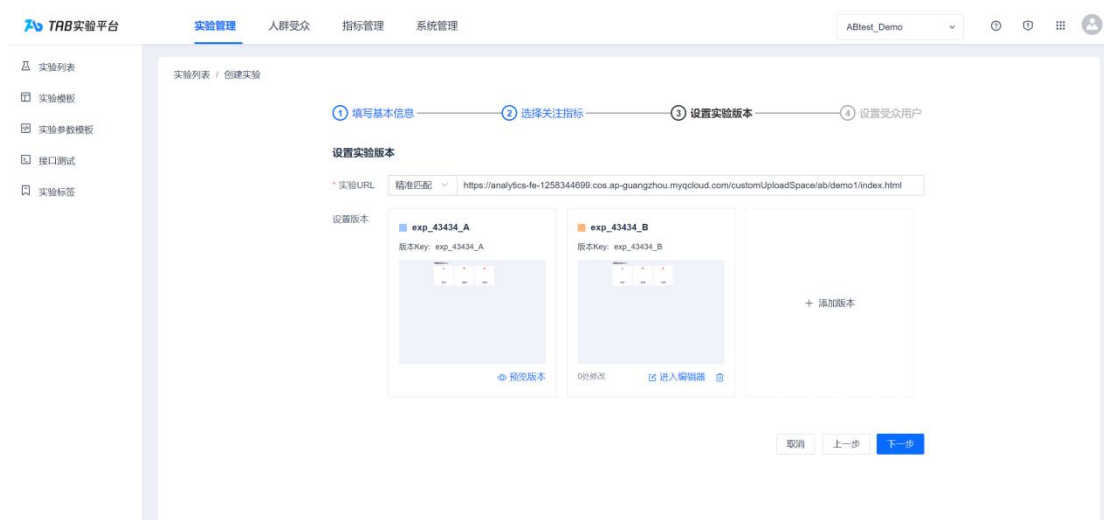
- **实验强相关指标**：从指标列表中挑选与当前实验强相关的指标
 - **T1-期望提升指标**：当前实验要验证的核心提升目标。
 - ◆ 添加数量不超过 50 个
 - ◆ 只可选择“均值类、比例类、比率类”3种类型的指标
 - **T2-其他普通指标**：不属于实验设计范围中的指标，可作为实验额外观察的指标。
- **业务默认添加指标**：无需用户操作，默认会将与业务强相关的指标加入当前实验
 - **T0-北极星指标**：业务的终极目标，是公司级别的战略性指标。

- **T1-护栏指标：**测量不良用户体验的指标，是实验中不应该显著负向的指标。旨在保护业务不出现异常。



3. 设置实验版本——

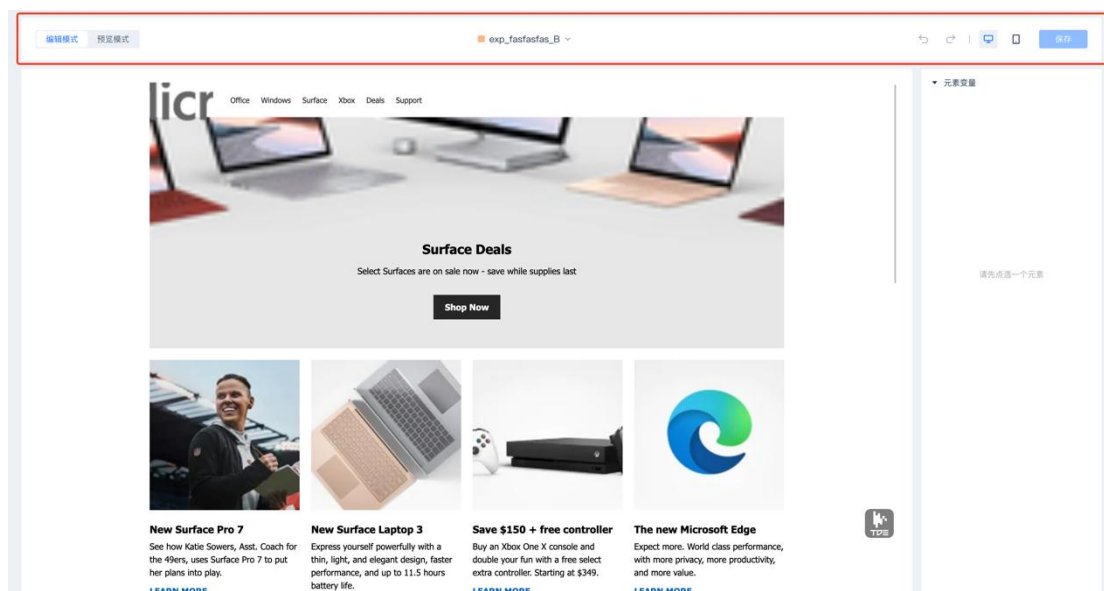
- **实验 URL：**作用于触发实验下发的 URL （实验 URL 需要以 http://或 https://协议作为开头）
 - **精准匹配：**默认的实验 URL 配置方式，用户访问的 URL 需要与实验 URL 完全一致，才能参与实验；
 - **模糊匹配：**忽略 URL 中查询的参数和锚点，用户访问的 URL 域名和路径匹配即可参与实验。
- **添加版本：**可添加多个实验版本
- **版本名称、版本 key：**默认实验版本 A 为对照组，不支持修改；其余实验版本均为实验组
- **预览版本：**可预览已填写的实验 URL 的对照版本页面内容
- **进入编辑器：**支持进入已填写的实验 URL 页面，并对其进行可视化编辑



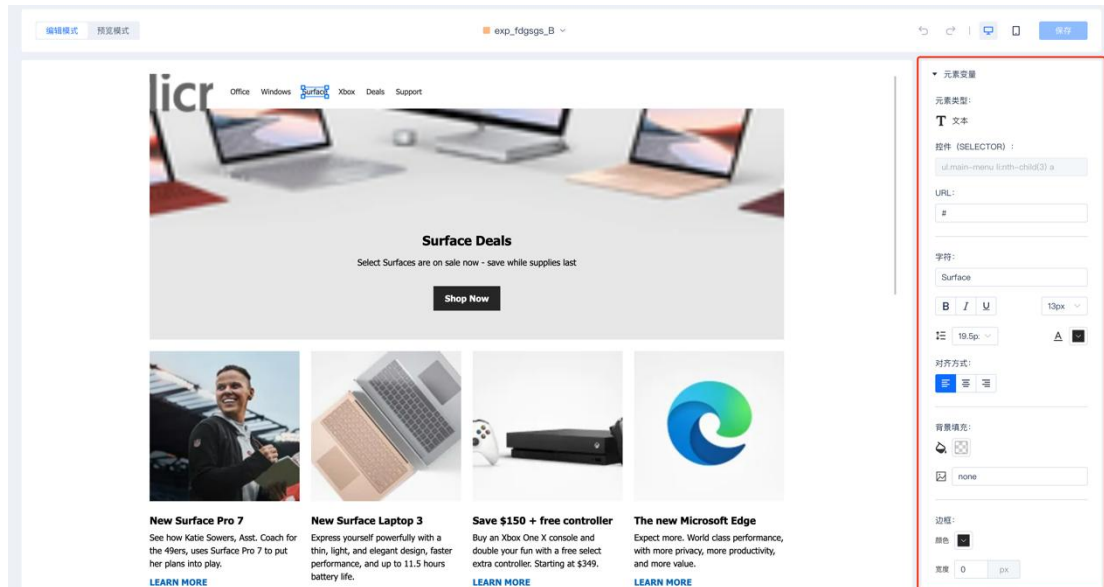
可视化编辑区域——

• 顶部功能区

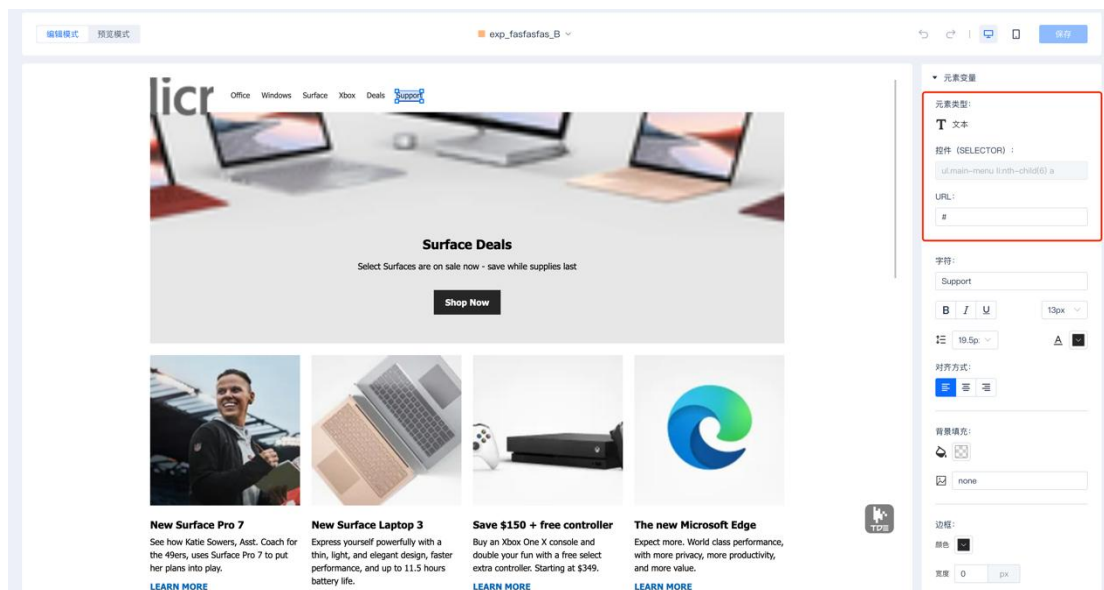
- 支持切换 “编辑模式、预览模式”
- 支持切换 “实验版本”
- 支持对编辑内容进行“上一步、下一步、保存”的操作
- 支持切换桌面端、移动端的视图模式



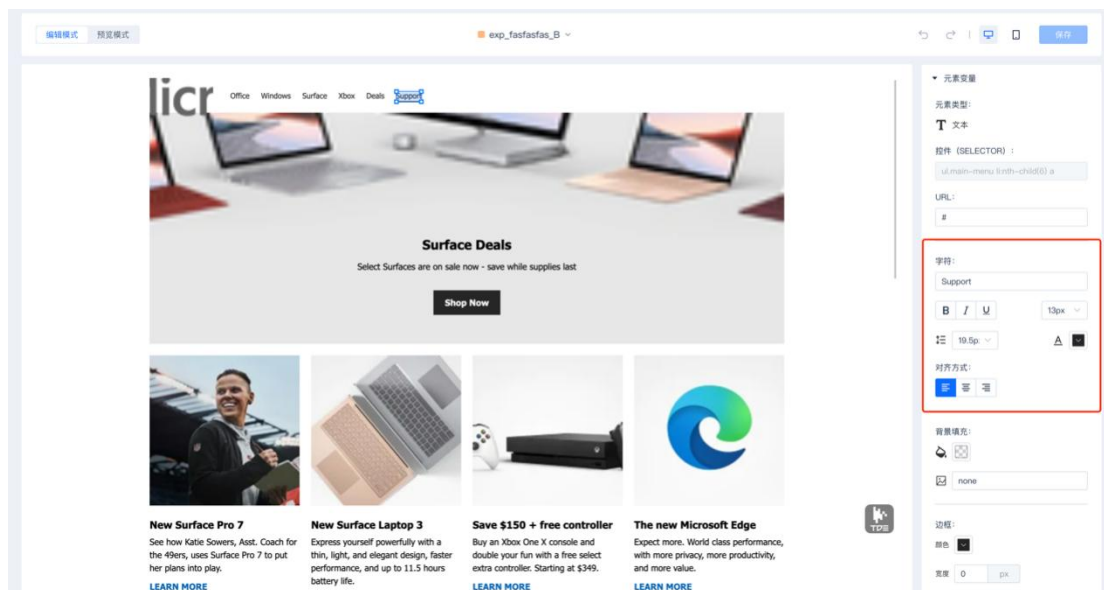
- **功能面板：**在“编辑模式”下，点击面中的任意元素即可触发功能面板



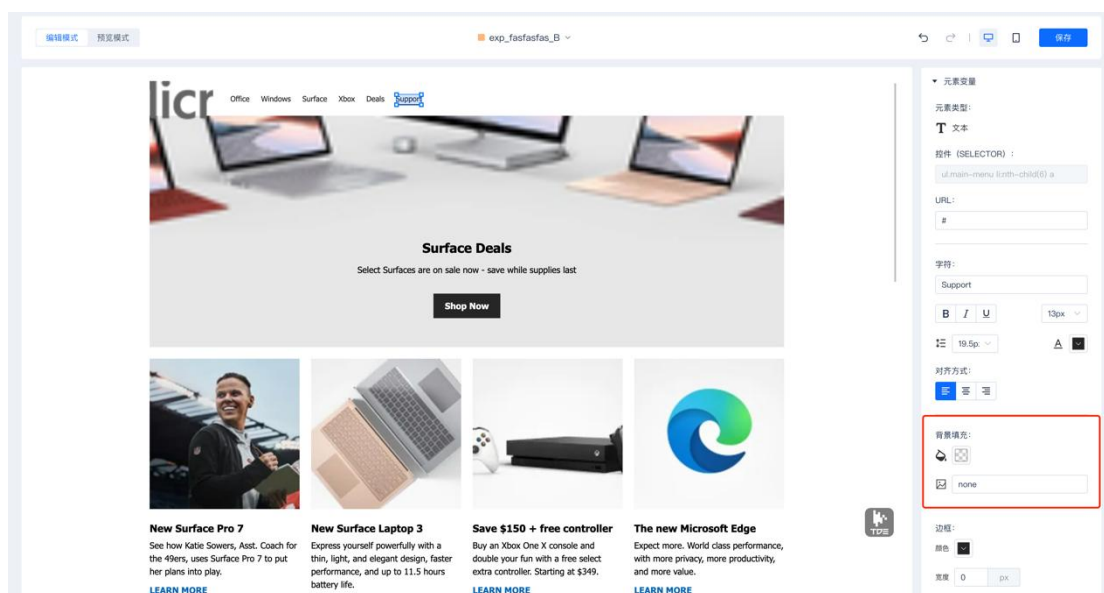
- 元素类型：元素基本信息展示；可新增或修改元素所关联的 URL



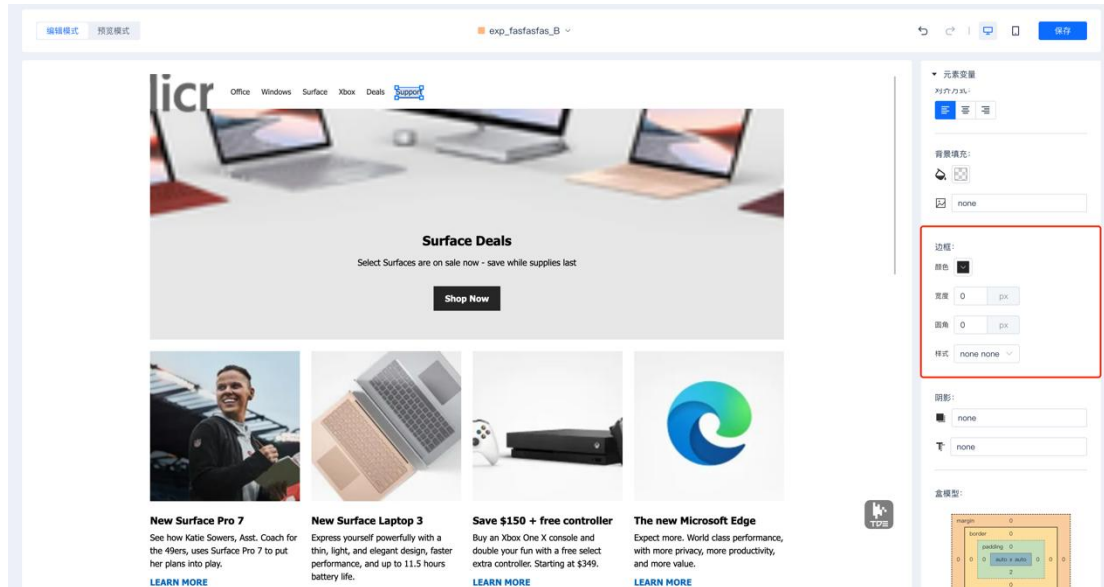
- 字体排版：支持对文本内容、加粗/斜体/下划线、字号、行高、颜色、对齐方式等属性进行编辑



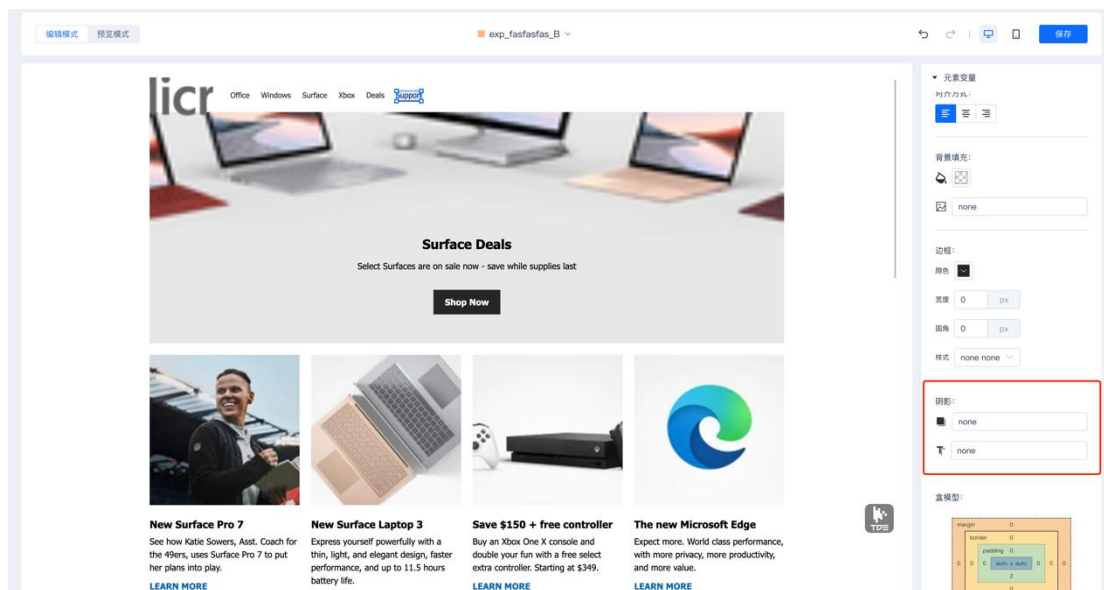
- **背景填充：**支持背景色和背景图两种填充方式



- **边框：**支持对元素的边框颜色、宽度、圆角、样式进行编辑

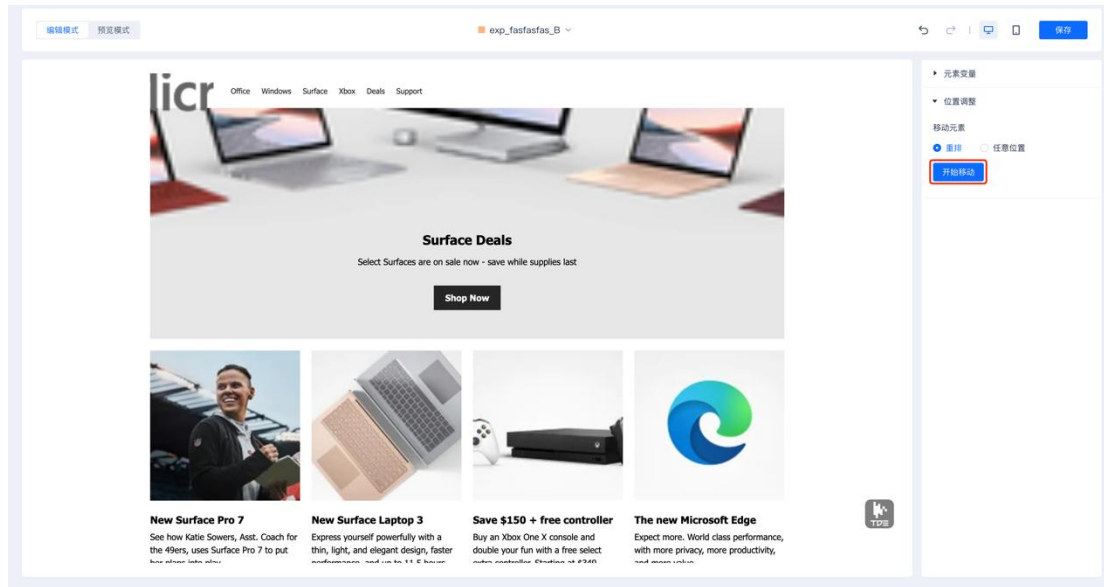


- **阴影：**支持对元素阴影和字体阴影进行编辑



- **移动元素：**支持对元素位置进行快速更改

- ◆ **重排：**点击【开始移动】按钮，选择需要重排的元素，对其位置进行重排
- ◆ **任意位置：**点击【开始移动】按钮，选择需要改变位置的元素，即可将其拖动到任意位置



4. 设置受众用户——

- **选择层域：**选择实验层，支持查看当前实验层的流量被哪些实验占用，剩余多少流量
- **选择受众规则：**
 - **实验层流量比例：**选择实验层后，该实验层剩余的流量比例
 - **按人群标签：**从人群标签清单中选择标签并配置标签之间的规则，从而实现人群圈选。只有被圈选的人群才可参与当前实验。
- **实验总流量：**当前实验各个版本在当前层预计最大可上线的流量比例。可以认为是“占坑流量”，后续编辑/设置线上流量需要在占坑的范围内进行设置。例如”实验版本 1“预设最大流量为 20%，那么后续编辑/设置线上流量只可 $\leq 20\%$ 。一般情况，不同版本的流量占比是平均的，尽可能保障实验置信度；各实验版本的最大流量占比的总和 $\leq 100\%$ 。
- **版本流量分配：**默认均匀分配，也可选择自定义流量分配。

以上，所有信息填写完毕，点击【**进入调试期**】按钮，实验即可创建成功。

3. 调整实验状态

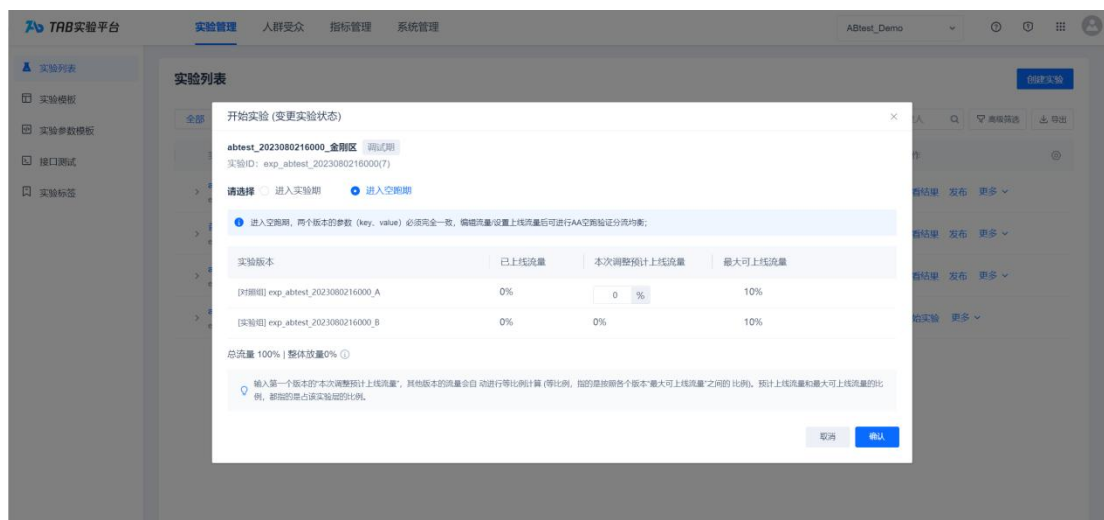
实验创建完成后，默认进入调试期。在调试期实验测试没问题后，即可选择直接进入空跑期或实验期。

点击【**开始实验**】按钮：

3.1 进入空跑期

进入空跑期，实验版本的参数（key、value）必须完全一致，编辑流量/设置上线流量后可进行 AA 空跑验证分流均衡。

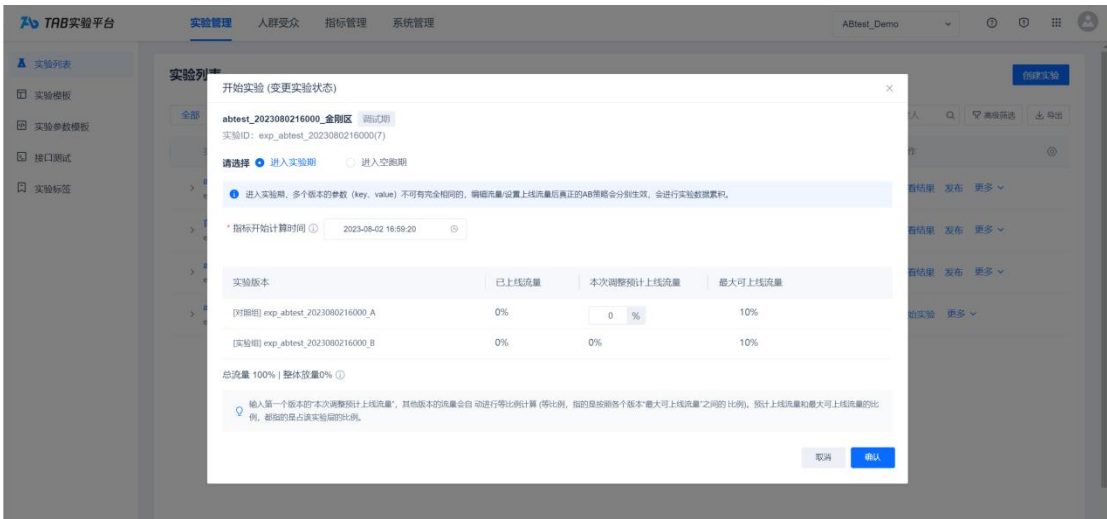
本次调整预计上线流量：输入第一个实验版本的“本次调整预计上线流量”，其他实验版本的流量会自动进行等比例计算（等比例，指的是按照各个版本“最大可上线流量”之间的比例）。预计上线流量和最大可上线流量的比例，都指的是占该实验层的比例。



3.2 进入实验期

调试期/空跑期状态的实验，可以进入实验期。进入实验期时，多个实验版本的参数（key、value）不可有完全相同的，编辑流量/设置上线流量后真正的 AB 策略会分别生效，并进行实验数据的累积。

- **实验开始时间：**实验策略下发开始的时间，会进行实验数据累积计算。实验期开始时间，必须和实验策略生效时间保持一致，且实验过程中不能随时修改策略。
- **本次调整预计上线流量：**输入第一个实验版本的“本次调整预计上线流量”，其他实验版本的流量会自动进行等比例计算（等比例，指的是按照各个版本“最大可上线流量”之间的比例）。预计上线流量和最大可上线流量的比例，都指的是占该实验层的比例。



3.3 不同实验状态对流量、参数的限制

• 变更实验状态

变更实验状态	限制条件	含义
调试期->调试期	不可以编辑流量，可以修改实验参数	表示实验未正式开始，可以调试白名单、测试接口。
调试期->空跑期	可以编辑流量、可以修改实验参数	进入空跑期，实验版本的参数（key、value）必须完全一致，编辑流量/设置上线流量后可进行 AA 空跑验证分流均衡。
调试期/空跑期->实验期	可以编辑流量、可以修改实验参数	进入实验期，实验版本的参数（key、value）不可有完全相同的，编辑流量/设置上线流量后真正的 AB 策略会分别生效，会进行实验数据累积。

• 当前实验状态

当前实验状态	限制条件
--------	------

调试期	不可以编辑流量，可以修改实验参数
空跑期	可以编辑流量、可以修改实验参数
实验期	可以编辑流量、不可以修改实验参数
已全量	不可以编辑流量、不可以修改实验参数
已结束	不可以编辑流量、不可以修改实验参数

4. 调整实验流量

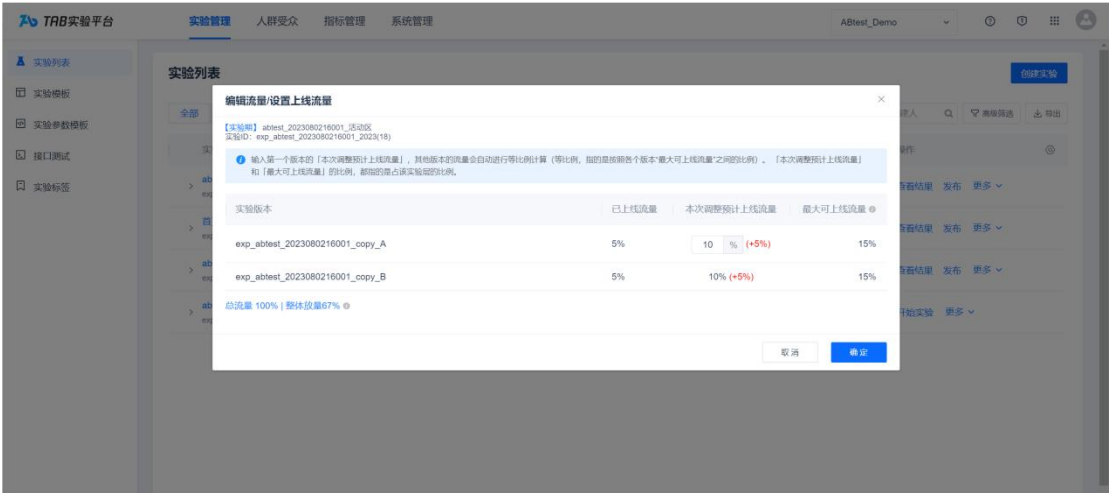
4.1 编辑流量/设置上线流量

在“最大可上线流量”这个范围内，去调整线上真正生效的流量。

在“更多”中，点击【编辑流量/设置上线流量】按钮：

实验名称/ID	实验状态	已上线流量(最大流量)	创建人/创建时间	操作
☆ > abtest_2023080216001_活动区 exp_abtest_2023080216001_2023	实验期	10% (30%)	Alex 2023-08-10 15:08:05	查看结果 发布 更多 ^
> 首页活动模块可配置编辑实验_01 exp_keshihua_20230810_111	实验期	10% (30%)	Alex 2023-08-10 15:06:17	编辑流量/设置上线流量 白名单管理 接口测试
> abtest_2023080216001_全刚区 exp_abtest_2023080216001	实验期	0% (23%)	Alex 2023-08-02 17:15:17	放量/设置最大流量 复制
> abtest_2023080216000_全刚区 exp_abtest_2023080216000	调试期	0% (20%)	Alex 2023-08-02 16:59:20	下线 权限配置

输入第一个实验版本的“本次调整预计上线流量”，其他实验版本的流量会自动进行等比例计算（等比例，指的是按照各个版本“最大可上线流量”之间的比例）。预计上线流量和最大可上线流量的比例，都指的是占该实验层的比例。



4.2 放量/设置最大流量

调整各个实验版本“最大可上线流量”。

在“更多”中，点击【放量/设置最大流量】按钮：

实验名称/ID	实验状态 ①	已上线流量(最大流量) ①	创建人/创建时间	操作
☆ > abtest_2023080216001_活动区 exp_abtest_2023080216001_2023	实验期	10% (30%)	Alex 2023-08-10 15:08:05	查看结果 发布 更多 ^
> 首页活动模块可视化编辑实验_01 exp_keshihua_20230810_111	实验期	10% (30%)	Alex 2023-08-10 15:06:17	编辑流量/设置上线流量 白名单管理 接口测试 放量/设置最大流量 复制 下线 权限配置
> abtest_2023080216001_金刚区 exp_abtest_2023080216001	实验期	0% (23%)	Alex 2023-08-02 17:15:17	
> abtest_2023080216000_金刚区 exp_abtest_2023080216000	调试期	0% (20%)	Alex 2023-08-02 16:59:20	

设置最大流量的百分比：可以选择占当前实验层/大盘的流量百分比。输入数值之后，即可

在下面放量详情查看本次放量后占当层/大盘的流量比例。

实验名称/ID	实验状态 ①	已上线流量
☆ > abtest_2023080216001_活动区 exp_abtest_2023080216001_2023	实验期	10% (30%)
> 首页活动模块可视化编辑实验_01 exp_keshihua_20230810_111	实验期	10% (30%)
> abtest_2023080216001_金刚区 exp_abtest_2023080216001	实验期	0% (23%)
> abtest_2023080216000_金刚区 exp_abtest_2023080216000	调试期	0% (20%)

设置最大可上线流量

【实验期】abtest_2023080216001_活动区
实验ID: exp_abtest_2023080216001_2023(18)
实验别名(实验ID): 多量02(duoceng02)

* 通知人
请输入名称搜索通知人或通知人

* 设置最大流量的百分比 ②
占当层 % 该实验层可用流量: 30%-70%

放量层
duoceng02(多量02)(当前层)

版本名称	原最大流量占当层	设置最大流量后占当层	设置最大流量后占大盘
exp_abtest_2023080216001_copy_A	15%	0%	0%
exp_abtest_2023080216001_copy_B	15%	0%	0%

4.3 发布实验

当发现实验组的新策略较对照组的策略，确实带来显著的正向收益，且业务侧共同认为新策略需要发布，则进行当前层实验发布。

当前层实验发布是无需研发修改改代码的，但是若涉及换层发布，则需要修改代码。

实验发布分为 2 种方式——

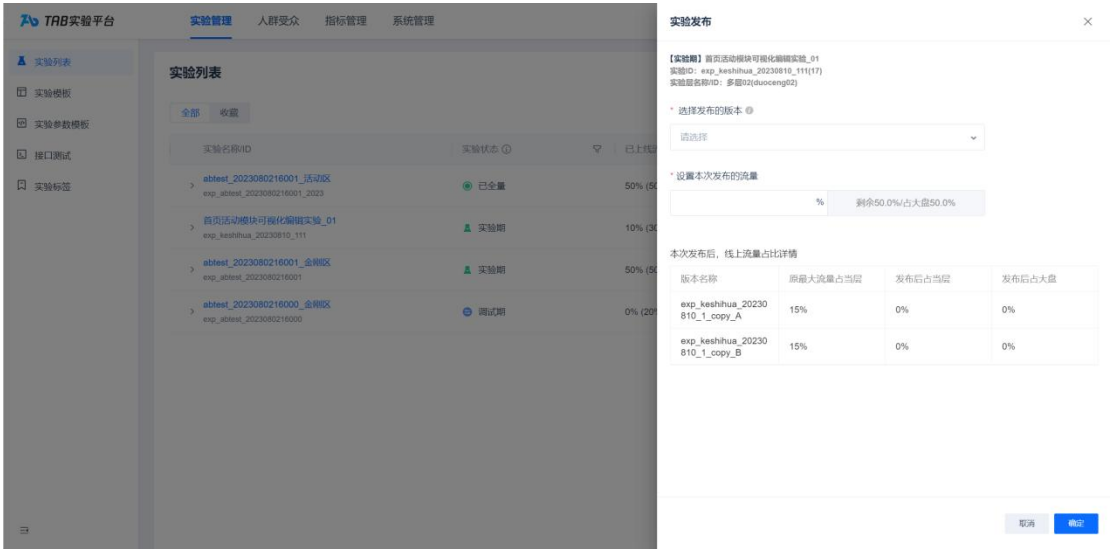
- **一次性全量发布：**实验已经获得预期效果，可将实验版本进行全量发布；
- **多次放量发布：**比如逐步发布 50%、70%、100%。举个例子，一个实验有 2 个版本，版本 A 有 50%流量，版本 B 有 50%，A 是效果更好的版本，需要全量发布版本 A；那么可以第一次发布版本 A 是 60%流量，第二次发布版本 A 是 70%流量，第三次发布版本 A 是 100%流量，第三次发布完成，即版本 A 全量。

点击【发布】按钮：

实验名称/ID	实验状态 ①	已上线流量(最大流量) ①	创建人/创建时间	操作
> abtest_2023080216001_活动区 exp_abtest_2023080216001_2023	已全量	50% (50%)	Alex 2023-08-10 15:08:05	查看结果 发布 下线 更多 ▾
> 首页活动模块可视化编辑实验_01 exp_keshihua_20230810_111	实验期	10% (30%)	Alex 2023-08-10 15:06:17	查看结果 发布 更多 ▾
> abtest_2023080216001_金刚区 exp_abtest_2023080216001	实验期	50% (50%)	Alex 2023-08-02 17:15:17	查看结果 发布 更多 ▾
☆ > abtest_2023080216000_金刚区 exp_abtest_2023080216000	调试期	0% (20%)	Alex 2023-08-02 16:59:20	开始实验 更多 ▾

- **选择发布的版本：**选择该实验的某个版本进行发布，自定义本次发布的流量。
 - 非发布版本的用户将会被打散至大盘（无实验策略），并不再受实验的任何策略影响（包括对照策略）
- **设置本次发布的流量：**选择某个实验层进行发布。发布到了 100%时，实验之前自带的人群标签筛选条件还会继续生效。不论是在当前层发布，还是换层发布，发布的流量只能代表是该层的流量占比，具体对应到线上流量占比，需要看这个层占全局比例是多少（如果该层是占全局的 100%，则表示线上 100%用户都会命中实验）。

- **当前层发布：**选择该实验所在实验层进行发布
- **换层发布：**选择某个层域后，在该层域下新建一个实验层来发布。当前实验换层发布后，TAB SDK 会同时返回两个正交的实验信息，具体命中哪个实验，会根据业务侧的生效逻辑来定。换层之前，请确认对用户造成的影响再操作。

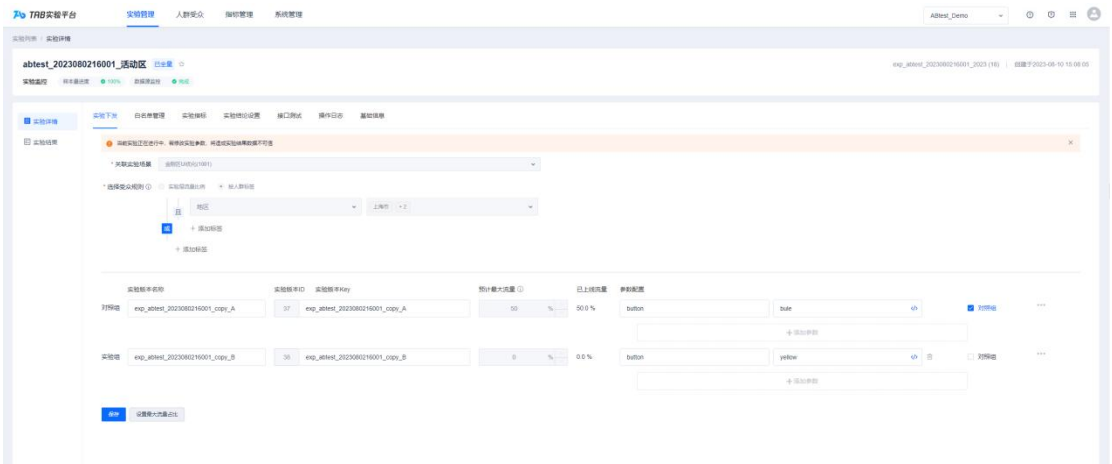


5. 实验分析

5.1 实验详情

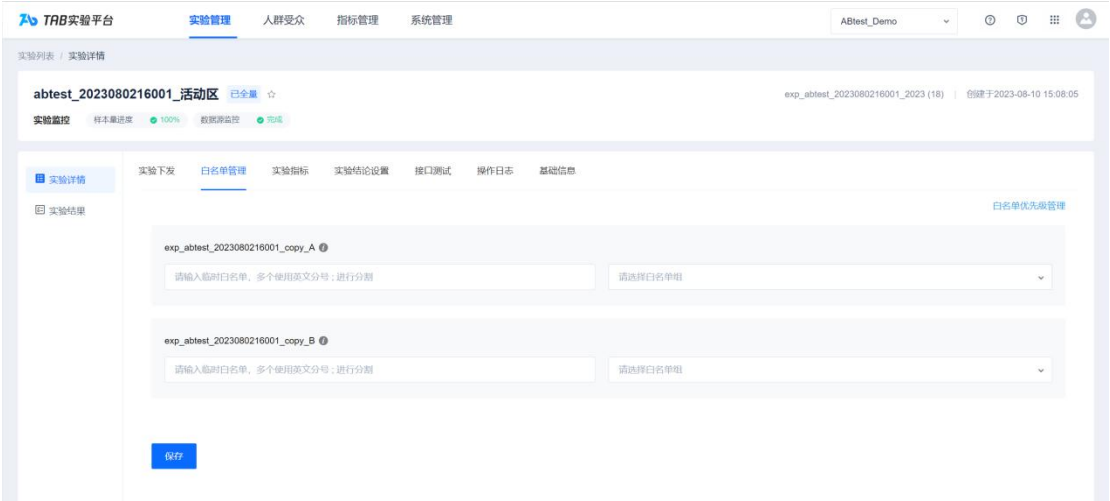
① 实验下发

支持查看和修改实验的版本信息，以及分流的比例。



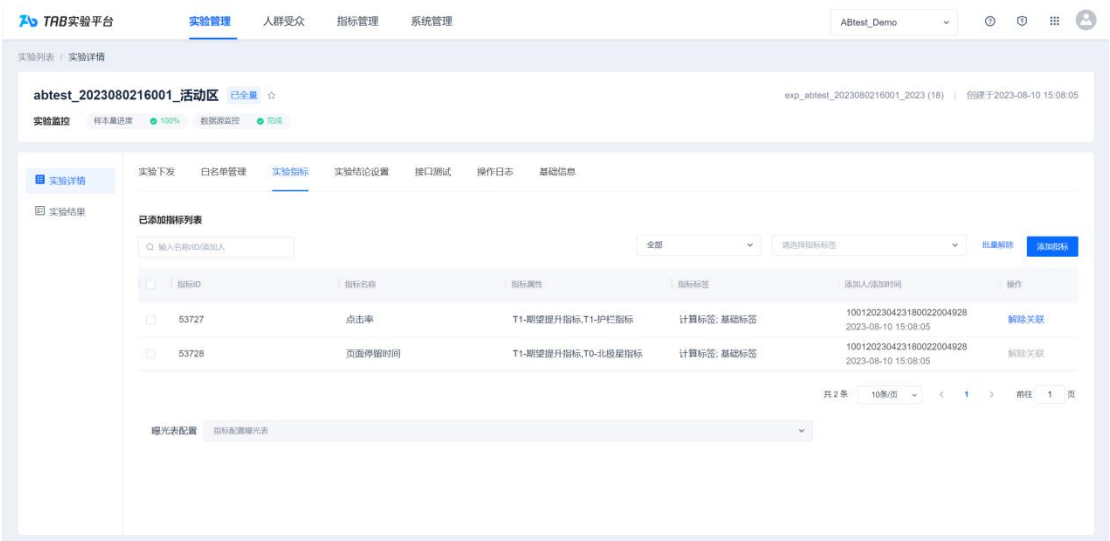
② 白名单管理

支持查看和编辑临时白名单，以及选择白名单组和设置不同白名单组生效的顺序。



③ 实验指标

支持添加或解除实验中的指标。



④ 实验结论设置

支持设置自动生成实验结论的规则。



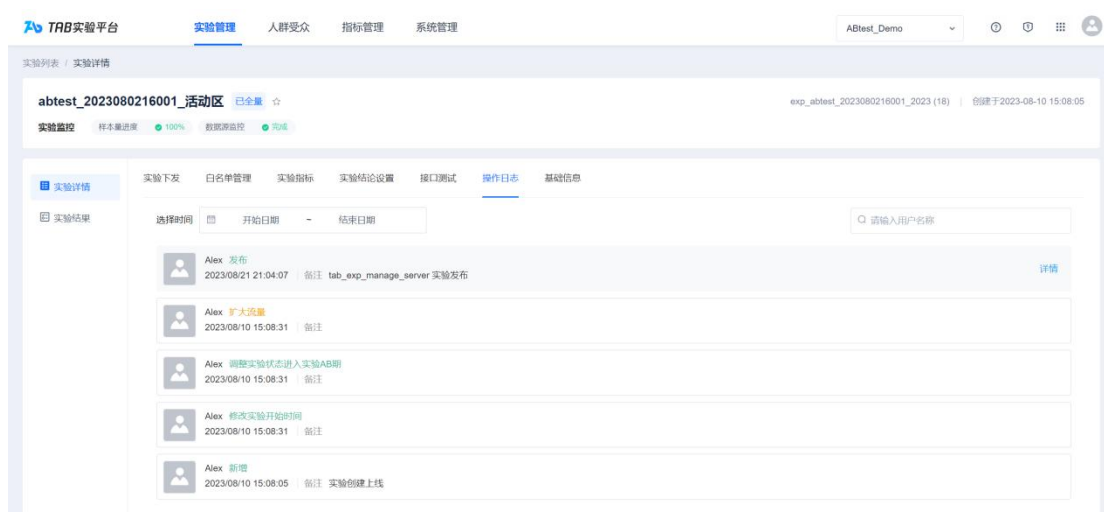
⑤ 接口测试

支持通过接口调试实验，需要传入调用分流的用户 ID。



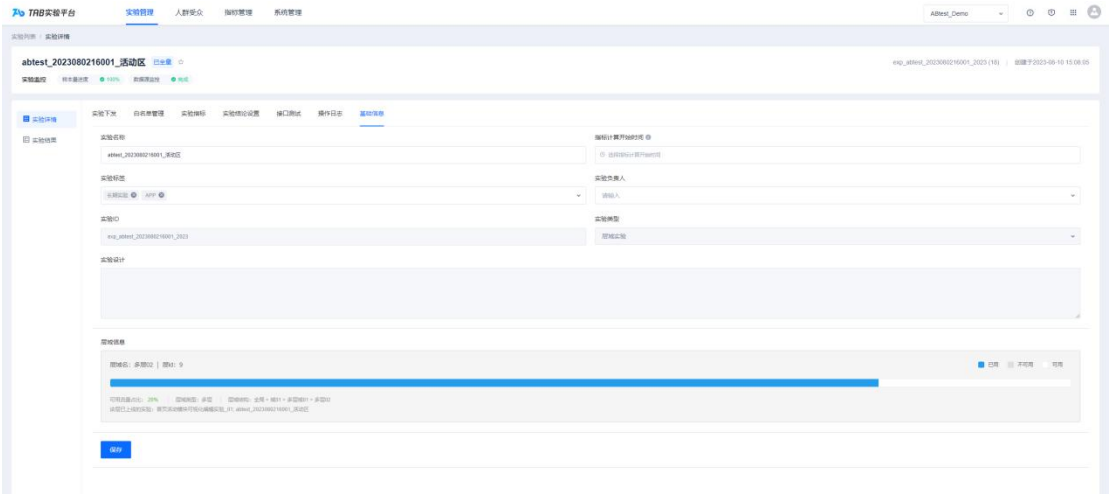
⑥ 操作日志

当前实验被用户调整过的记录。



⑦ 基础信息

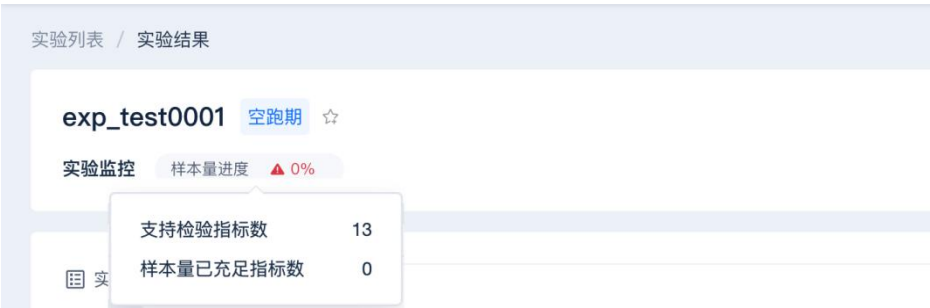
支持查看和修改实验的基本信息。



5.2 实验结果

一般而言，建议实验至少运行满一个自然周期（7 天）再来观测数据。

① 实验监控 - 样本量进度



- 实验指标只有在样本量累积足够的前提下，指标数据才能用作实验结论参考。此处样本量进度的检验方式采用 MDE 检验。MDE 也叫最小检测变化，是实验在当前条件下能有效检测的指标 diff 幅度。有效检测，指检出概率大于等于统计功效（默认为 80%）。只有在 MDE 小于等于实验设计初期设定的最小变化阈值后，实验样本量才是足够的，这时进行假设检验得出的显著性结果是最可靠的，建议以此时的显著性结果为依据进行实验效果决策。
- 目前并不是所有指标都支持 MDE 检验，因此本监控主要提示信息是：①支持检验的指

标数量有多少个；②在支持检验的指标里，MDE 样本量累积已达标的指标有多少个；

③样本量已累积达标的指标占有支持检验的指标的比例是多少。

- 本监控主要是提示实验指标样本量累积情况，不影响实验结论科学性判断。

② 实验结论



实验结论模块是为了降低实验分析门槛给出的平台实验结论建议，方便了解当前实验状况。

实验对于业务具体的影响仍需要业务具体情况具体分析。

实验结论基于实验的结论设置产生，规则如下：

标题	判断规则	提示文案
等待开始实验	实验在空跑期时	当前实验状态为空跑期，暂无实验结论。 建议进入实验期开始实验。
已产生优胜组	实验组结果满足分析模板设置的条件（满足条件的可能有多个）	[XX 实验组 1de]、[XX 实验组 2]的表现 优于[xx 基准组]，已产生 x 个正向显著指标，x 个负向显著指标。建议发布实验。
未产生优胜组	实验组结果不满足分析模板设置的条件	各实验组和[xx 基准组]对比，未产生 优胜组。已产生 x 个正向显著指标，x 个 负向显著指标。
暂无实验结论	没有配置结论规则	本实验尚未配置实验结论规则，因此暂无

多组对比 | 详细对比

假设检验参数

下载数据

添加指标

指标	基准 A	B 对比	C 对比	D 对比
指标名称 (天)	4.47%	+0.5737%	+1.3496%	+0.1099%
T0-北极星指标		4.52%	4.54%	4.48%
指标名称 (分)	45.96	+0.1264%	+0.2210%	-0.0979%
T0-北极星指标		46.13	46.07	45.7
指标名称 (天)	2.1269	+0.1400%	+0.1266%	+0.0615%
T0-北极星指标		2.1299	2.1296	2.1283
指标名称 (天)	12.4547	+0.1059%	-0.1933%	-0.3132%
T0-北极星指标		12.4888	12.443	12.3703
指标名称 (天)	5.1434	+0.1438%	+0.1948%	+0.2470%
T0-北极星指标		5.1799	5.1555	5.1228
指标名称 (天)	0.3157	+0.3371%	+0.7568%	-0.1600%
T0-北极星指标		0.3174	0.3184	0.3148
指标名称 (天)	20.3437	+0.1232%	-0.0235%	+0.0983%
T0-北极星指标		20.4567	20.304	20.2302
指标名称 (天)	11.48%	-0.3532%	-0.1161%	-0.2124%
T0-北极星指标		11.48%	11.52%	11.43%

适合多个实验组同一个对照组进行多组比较，方便查看各实验组表现情况。

多组对比 | 详细对比

假设检验参数

下载数据

添加指标

指标	基准 A	C 对比	B 对比	D 对比
指标名称 (天)	5.72%	+1.0303%	+0.4733%	-0.2113%
T0-北极星指标		5.78%	5.75%	5.71%
指标名称 (分)	71.45	+0.1505%	+0.1493%	-0.6847%
T0-北极星指标		71.56	71.56	70.96

- 多组对比模式下，第一列展示指标名，第二列基准列展示基准版本（一般是第一个对照组）的指标值，后续列展示剩余对比版本的指标值（灰色字体）、相对差异（红绿色字体）、显著性（色块加箭头）；
- 图中用色块区分显著性，如图绿色色块表示正向显著，红色色块表示负向显著，无色色块绿色字表示指标值有提升但不显著，无色色块红色字表示指标值有下降但不显著；
- 相对差异计算公式 = (对比组指标值 - 基准组指标值) / 基准组指标值

1-2) 详细对比

多组对比 | 详细对比

假设检验参数

下载数据

添加指标

指标	基准 A	对比 C	相对差异	相对差异与趋势(天)	P值
指标名称 (天)	5.72% 101.57K/1.77M	5.78% 102.64K/1.77M	+1.0439% (±0.79%)		
T0-北极星指标					
指标名称 (分)	71.45 126.77M/1.77M	71.56 126.99M/1.77M	+0.1295% (±0.32%)		
T0-北极星指标					
指标名称 (天)	3.3577 5.96M/1.77M	3.3617 5.97M/1.77M	+0.1189% (±0.1719%)		
T0-北极星指标					
指标名称 (天)	19.2987 34.24M/1.77M	19.2824 34.22M/1.77M	-0.2173% (±0.39%)		
T0-北极星指标					
指标名称 (天)	7.9652 14.13M/1.77M	7.9814 14.16M/1.77M	+0.1780% (±0.34%)		
T0-北极星指标					
指标名称 (天)	0.4954 878.85K/1.77M	0.4976 883.03K/1.77M	+0.3427% (±0.87%)		
T0-北极星指标					
指标名称 (天)	30.8194 54.68M/1.77M	30.77 54.61M/1.77M	+0.0265% (±0.31%)		
T0-北极星指标					
指标名称 (天)	11.1% 196.98K/1.77M	11.16% 197.97K/1.77M	+0.0139% (±0.34%)		
T0-北极星指标					

适合一个实验组同一个对照组进行两两比较，方便进行两组详细分析对比。

详细对比模式下，基准组默认为第一个对照组，对比组默认为第一个实验组，可下拉切换其他组。基准和对比列显示该组指标值。

同多组对比一样，绿色色块表示正向显著，红色色块表示负向显著，无色色块绿色字表示指标值有提升但不显著，无色色块红色字表示指标值有下降但不显著。

相对差异下括号中的数字代表置信区间，如下图中的+1.0303%(±0.8508%)表示相对差异在设定置信度（默认为 0.05）的条件下的置信区间为[0.1795%,1.8811%] 。



简单解读置信区间：假设策略上线，有 95%的把握会看到目标用户的真实指标提升在 [0.1795%,1.8811%]这个范围内。置信区间不包含 0 和统计显著是等价的。

2) 相对差异与趋势（天）

在基础分析列表可以看到相对差趋势图的缩略柱状图,点击缩略图可查看详细指标卡片如下图：



实验平台提供累计口径与非累计口径的统计数据，对应的统计口径如下：

非累计数据指标值 = 指标值按天累加 / 用户按天去重，可简单理解为每日的人均时长。

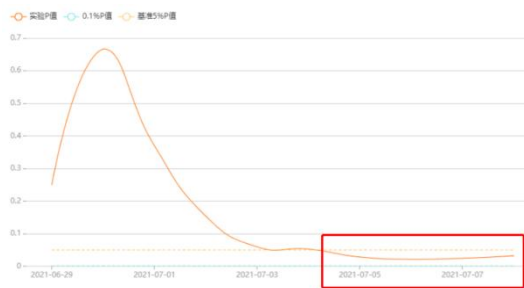
累计数据指标值 = 指标值多天累加 / 用户多天去重

一般情况下推荐使用累计口径，因为累计口径可以避免样本量不均衡，从而便于归因。

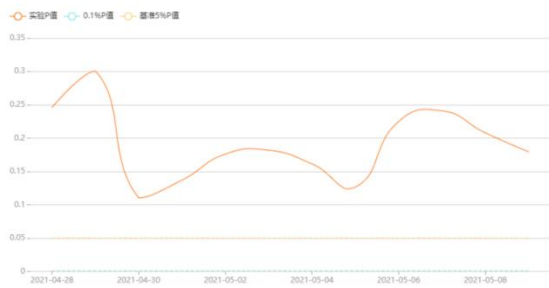
3) P 值

当 P 值小于**显著性水平**（默认为 0.05）时，认为实验组和对照组的该指标在统计上显著不同。在这种情况下，对应的指标值会在相对差异一格显示色块，**绿色**色块表示正向显著，**红色**色块表示负向显著

除了看 P 值大小，还可以点击 P 值左侧的黄色趋势线观察 P 值变化趋势。预期是观测的 P 值趋于稳定（下左图），不存在剧烈波动：



左图（P 值稳定小于置信度）



右图（P 值存在波动）

比较相对差异和预先设定变化阈值：

- 需要结合业务上的众多因素来考量实验策略带来的收益是否值得推广。P 值只说明实验组与对照组显著不同，并不能说明究竟提升了多少。在很多情况下存在 P 值很小，但是相对差异也很小，需要特别注意。
- 只靠 P 值不足以得到量化的结论，通过相对差异一栏，可以看出指标的变化是否达到预期以及相对差异的置信区间。

4) 最小检测变化（MDE）

最小检测变化是实验在当前条件下能有效检测的指标 diff 幅度。有效检测，指检出概率大

于等于统计功效（默认为 80%）。只有在 MDE 小于等于实验设计初期设定的最小变化阈值后，实验样本量才是足够的，这时进行假设检验得出的显著性结果是最可靠的，建议以此时的显著性结果为依据进行实验效果决策。

建议在业务关心的所有指标 MDE 右侧出现蓝底勾号时再进行分析。此时 MDE 小于等于实验设计初期设定的最小变化阈值或小于等于 1%。

5) 非累积平均 (天): 基准 | 对比



此处显示非累积平均天口径下的基准组平均指标值、对比组平均指标值、平均相对差异：

对应的统计口径如下：

- count 和 sum 类：每天的非累积数据的平均
- mean、ratio、proportion 类：N 天的非累积数据的分子总和，除以 N 天的非累积数据的分母总和

⑤ 其它设置

1) 假设检验参数



- **显著性水平(α)**: 显著性水平是用来判断实验是否显著的依据, 若 P 值 $<$ 显著性水平, 则认为统计显著。显著性水平默认为 5%。
- **统计功效($1-\beta$)**: 统计功效是在实验策略有效的情况下, 检验出来有区别的概率。统计功效默认为 80%。
- **多对比修正**: 当一次实验中要进行多次 ($n \geq 2$) 假设检验时, 就需要做多重假设检验校正, 进行修正的目的是为了降低第一类错误的概率。
- **求和类指标检验**: 默认不使用。若启用, 会将指标退化为均值类指标进行假设检验。

2) 下载数据



点击后可选按照累计口径或者非累计口径分别下载数据。

3) 添加指标



若需要在实验开始后补充指标，可点击“添加指标”按钮添加。添加指标后，默认在第二天产出这些新增指标的数据结果。

6. 实验模板

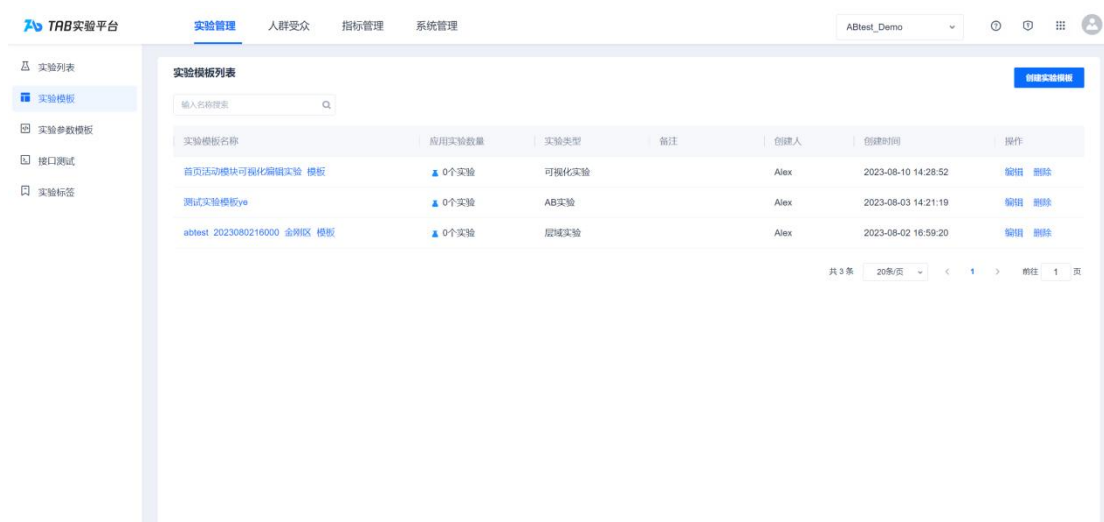
6.1 实验模板的价值

实验模板可帮助业务解决如下问题：

- 业务自主进行实验、但又不知道如何进行实验设计，以及过去业务中有不合理的实验设计习惯等；
- 业务每次配置实验平台时，多次进行指标选择、实验版本配置等重复性劳动。

6.2 实验模板列表

- **搜索：**输入实验模板名称，进行模糊搜索
- **创建实验模板：**支持创建层域 AB 实验、独立层 AB 实验和可视化编辑实验
- 点击“**实验模板名称**”：查看实验模板详情信息
- 点击“**应用实验数量**”：查看该实验模板用于哪些实验
- **编辑：**只有模板创建者，可以编辑模板信息
- **删除：**只有模板创建者，可以删除模板



6.3 创建实验模板

点击【创建实验】按钮，进入创建流程：



实验模板创建流程基本完全复制了实验创建的流程, 额外需要填入的信息有: 实验模板名称、备注、实验类型。

另外, 在配置“选择关注指标”时, 还可以对实验强相关指标, 后续是否能 2 次修改选择的指标进行设置:

- T1-期望提升指标

- 可移除模板的指标, 可添加新的指标: 例如, 模板有 N 个指标, 可以移除模板的

N 个指标；还可以添加“非模板”的新的 N 个指标。

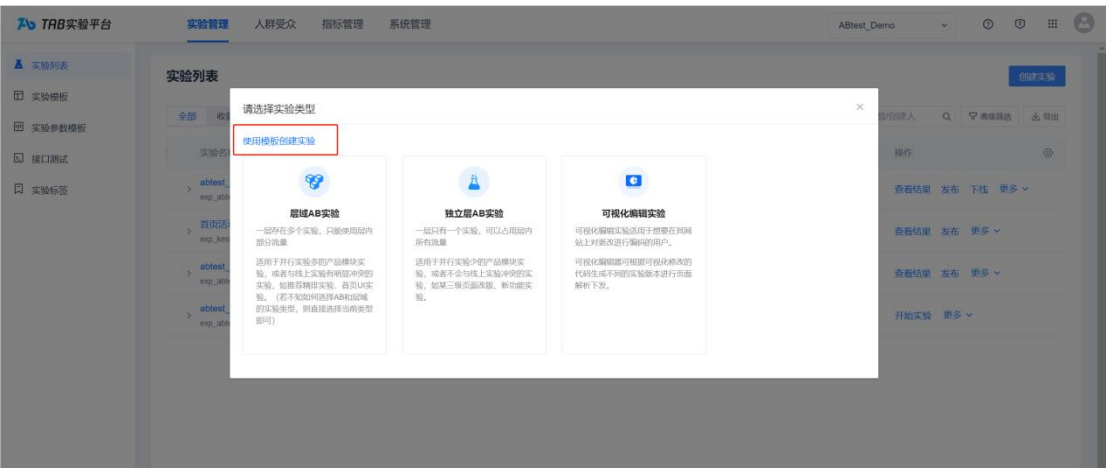
- 不可移除模板的指标，不可添加新的指标：例如，模板有 N 个指标，不能移除，也不可以添加新的指标。

• T2-其他普通指标

- 可移除模板的指标，可添加新的指标
- 不可移除模板的指标，可添加新的指标



后续在创建实验时，可以直接使用模板：



7. 实验参数模板

7.1 实验参数模板列表

- **创建实验参数模板：**进入实验参数模板创建流程
- **编辑：**编辑模板信息
- **删除：**删除模板

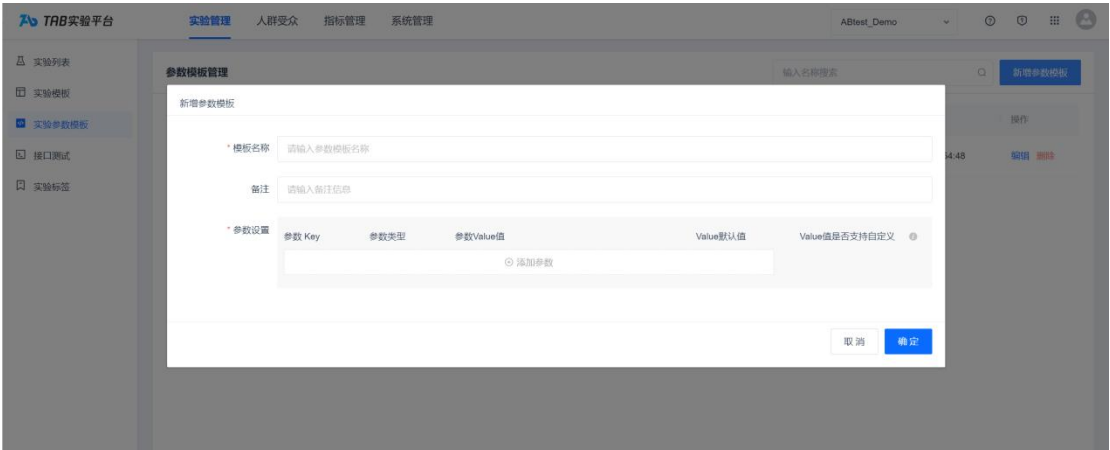


7.2 创建实验参数模板

点击【**新增参数模板**】按钮，进入创建流程：

- **参数 key：**在一个参数模板下，key 值不能重复
- **参数类型：**支持 int、string、boolean
 - **int：**支持整数
 - **string：**支持字符串，由字母、数字、下划线等组成（不建议使用中文）
 - **boolean：**只允许选择 FALSE 和 TRUE，默认 FALSE
- **参数 value 值：**多个 value 可使用英文“;”隔开
- **value 默认值：**默认的 value 值
- **value 值是否支持自定义：**默认关闭，表示该 value 不可自定义

- key 值是否支持自定义：默认关闭，表示 key 不可自定义

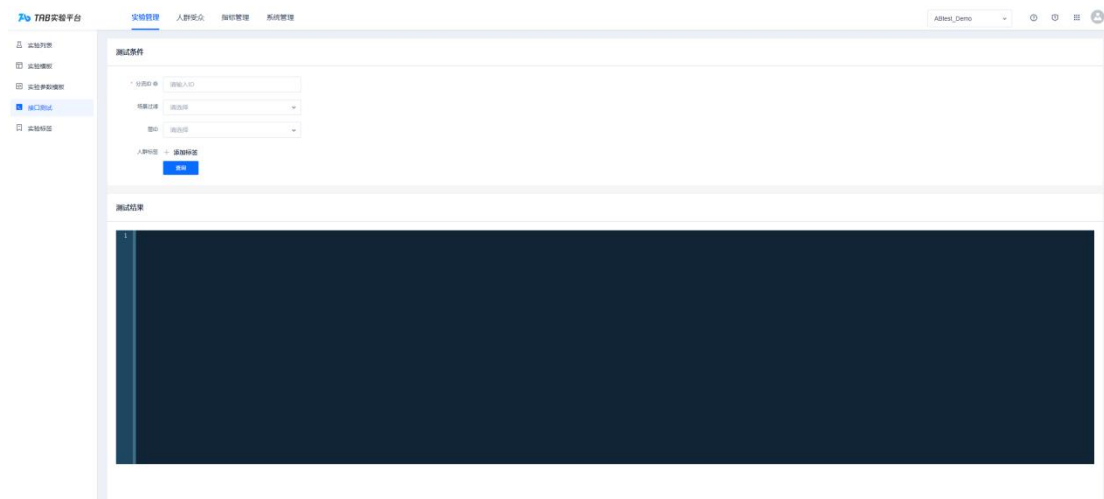


后续在配置参数时，可以直接使用模板：

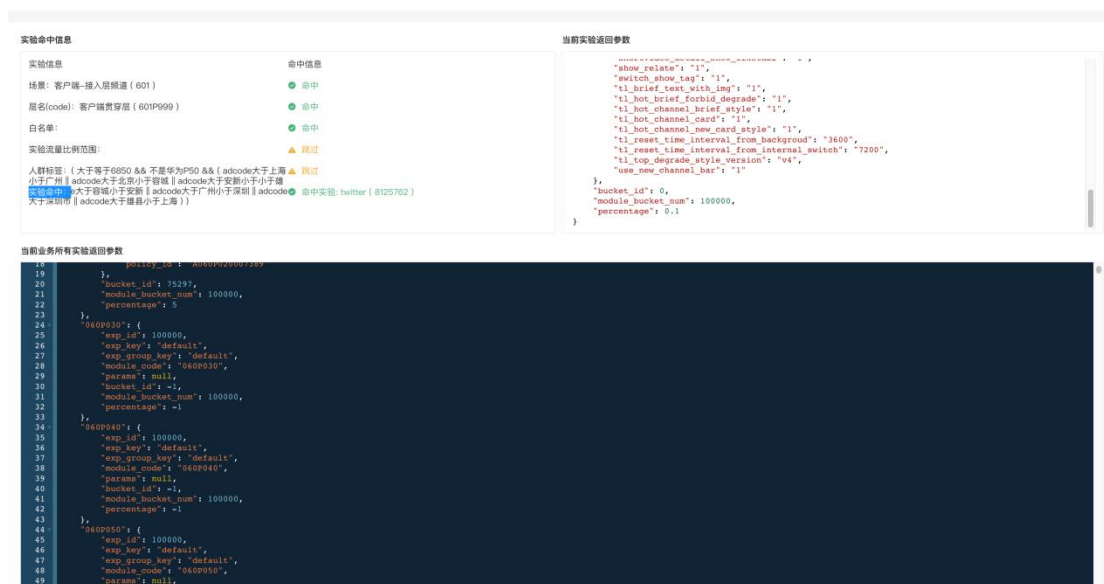


8. 接口测试

- 分流 ID：必填项，可以理解为用户的 ID，业务在接入 TAB 实验平台时，会取用户 ID 前 28 位字符作为分流 ID
- 业务 ID：必填项，默认是当前业务，可选择您其它业务
- 场景过滤：非必填项，可通过选择实验场景，来过滤目标实验
- 层 ID：非必填项，从当前业务中选择层来过滤结果
- 人群标签：非必填项，选择在标签管理中的标签组及标签



点击【**查询**】按钮，即可获取测试结果：



9. 实验标签

9.1 实验标签列表

- **创建实验标签：**进入实验标签创建流程
- **编辑：**编辑标签信息
 - **注意：**平台默认标签不支持编辑
- **删除：**删除标签

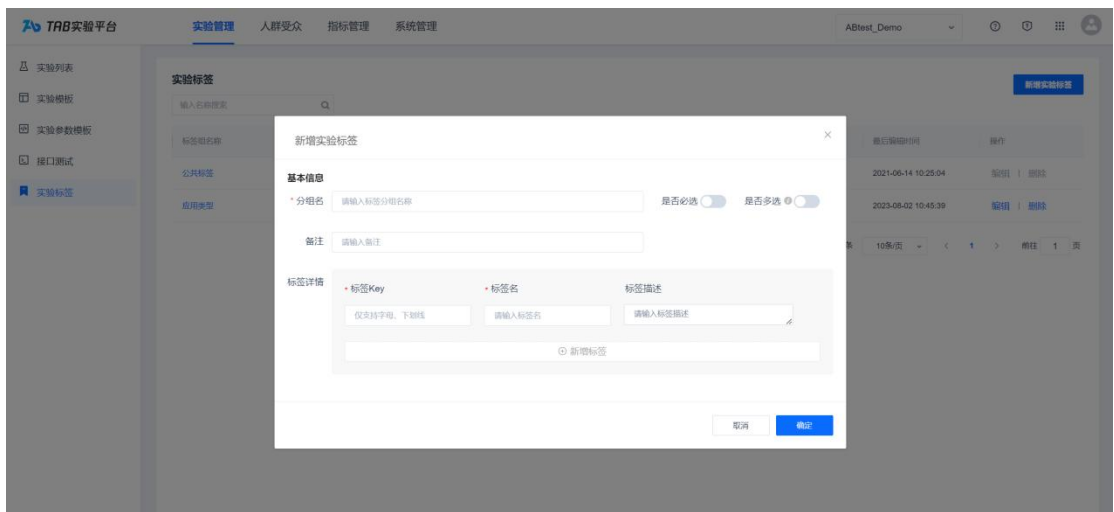
■ 注意：平台默认标签不支持删除



9.2 创建实验标签

点击【新增实验标签】按钮，进入创建流程：

- **分组名**：必填项，标签组的名称
- **是否必选**：在创建实验流程中，是否是必选项
- **是否多选**：标签组下的标签是单选或多选
- **标签 key**：必填项，标签的 key 值
- **标签名**：必填项，标签的名称



后续在选择实验标签时使用：

* 实验标签

请选择实验标签

实验负责人

jiankiiTest02

实验设计

实验标签

* 所属分行

guangmingTAG

* 实验类型

* 所属部门

广州分行

上海分行

深圳分行

总行

北京分行

果【修改XXX功能】，可以达到【提高用户粘性】的效果，此效果可以使

取消

下一步

三、人群受众

业务可以通过标签将人群属性管理起来。在创建实验的时候通过使用标签组合，来圈定实验的目标人群。

1. 标签组列表

- **新增标签组：**进入标签组创建流程
- **编辑：**编辑标签组信息
- **删除：**删除标签组
 - 删除标签组，后续新建/修改实验时均无法再添加该人群标签组
 - 已在实验中添加的人群标签不受影响

TAB实验平台

实验管理 人群受众 指标管理 系统管理

ABTest_Demo

人群标签

人群标签

标签组管理

搜索标签名字

新增标签组

标签组名称	标签组 ID	备注	创建人	创建时间	操作
地区	area	地区属性	Alex	2023-05-02 17:07:51	编辑 删除

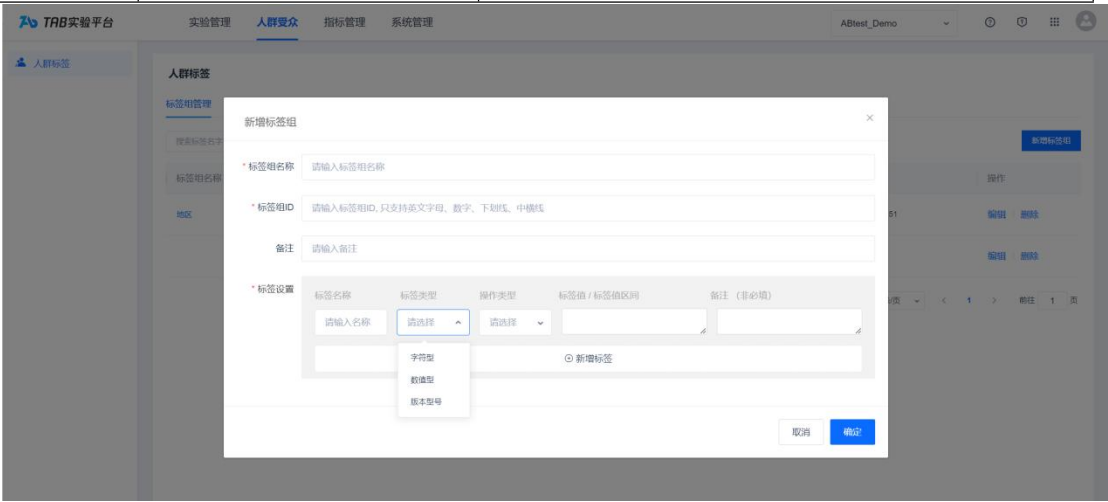
共 1 条 10条/页 < 1 > 前往 1 页

2. 新增标签组

点击【新增标签组】按钮，进入创建流程：

- **标签组名称：**必填项，支持中文、字母、数字等
- **标签组 ID：**必填项，只支持英文字母、数字、下划线、中横线
- **标签类型：**必填项，标签类型包括 3 种——

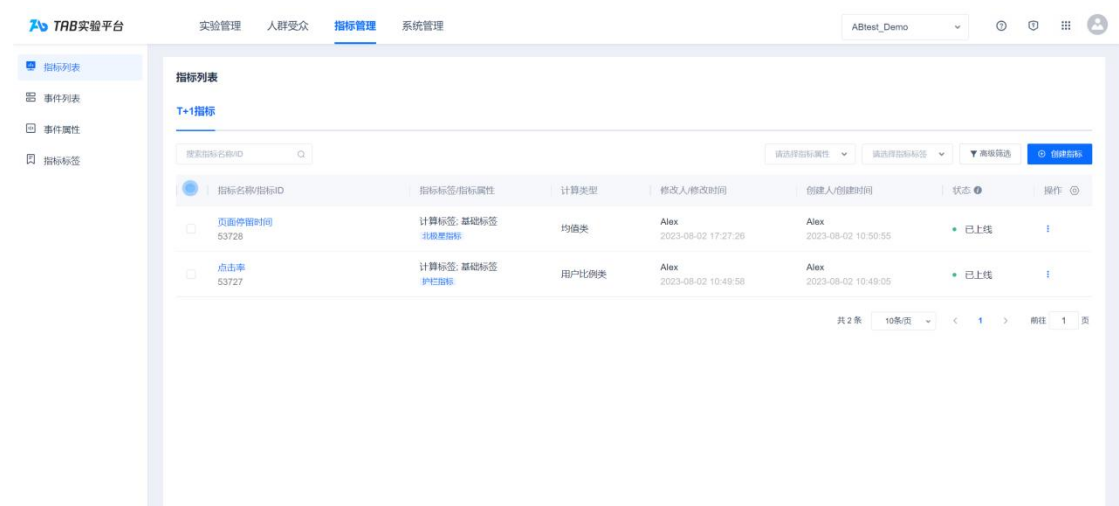
标签类型	操作符	注意事项
字符型	等于、不等于、包含	
数值型	小于、左闭右开、闭区间、 左开右闭、大于、等于	数字型标签是本身自带比较操作符的，因此只要传入的数字是符合该标签的约束，即满足条件。 注意： 不要给实验配置自相矛盾的标签值。比如，配置了>3 的标签值的同时又配置了<1 的标签值。
版本号型	小于、左闭右开、闭区间、 左开右闭、大于、等于	注意： 用标签来管理版本号时，最好保持版本号位置一致来设置。



四、指标管理

1. 指标列表

点击【创建指标】按钮，进入指标创建流程：



1.1 指标信息填写

- 指标名称：必填项，指标中文名称
- 指标标签：必填项，给指标做标记
- 指标口径：必填项，指标的计算方式
- 指标负责人：必填项，指标由谁负责
- 指标属性：必填项，用于指标分类。包括 4 种——

指标属性	定义	如何关联实验
北极星指标	平台或产品的终极目标，是公司级别的战略 略性指标，如人均停留时长。	实验创建时自动关联到实验 中，建议不超过 10 个。
期望提升指标	当前实验要验证的核心提升目标。与北极	实验创建时需手动关联到实

	星指标相关，且可以引导我们提升北极星指标的。北极星指标往往比较难直接影响，我们用期望提升指标作为提升北极星指标的间接证据，且一般是短期的较直接的衡量。如 CTR，推送接收。	验中，总共不超过 3 个。
护栏指标	测量不良用户体验的指标，实验中不应该显著负向的指标；旨在保护我们产品/功能不出现异常，如屏蔽推送，卸载数，延迟，安全性等。	实验创建时自动关联到实验中，建议不超过 10 个。
无固定属性	不属于实验设计范围的指标，当前实验额外观察的指标。	

- 是否 DS 认证指标：必填项，是否为业务方 DS 认可的指标

TAB实验平台

实验管理 人群受众 指标管理 系统管理

ABTest_Demo

指标列表

事件列表

事件属性

指标标签

指标列表 > 新增指标

1 指标信息填写

* 指标名称

请输入指标中文名

0/40

* 指标标签

请选择

* 指标口径

请输入清晰易懂的指标口径，方便后续实验中的指标应用和结果数据解读。例如：人均播放时长=播放时长的求和 / 命中实验的用户数

0/100

* 指标负责人

Alex

* 指标属性

请选择

* 是否DS认证指标

DS团队存在DS团队和业务方自主实验的数据，在此基础上对指标是否由DS团队进行标识

是

否

下一步

1.2 指标配置



① 数据表配置

- 选择实验曝光表（该表专门用于记录实验曝光信息）
- 选择指标数据源（该表专门用于记录上报的各个事件）
 - 指标数据源支持添加多个
 - 单指标数据源时：实验曝光表 LEFT JOIN 指标数据源
 - 多指标数据源时：曝光表 LEFT JOIN （指标数据源表 A INNER / OUTER JOIN 指标数据源表 B）
 - ◆ 指标类型为“留存类”时：曝光表 LEFT JOIN （指标数据源表 A UNION 指标数据源表 B）

② 计算方式配置

- 非留存类指标
 - 均值类：也叫 uv 均值类，累计指标求和/用户去重计数，形如 SUM(A) / COUNT(uin)，例如人均播放时长、人均时长、人均点击。

计算方式配置

* 指标计算类型

均值类

事件加工 输出按天聚合的表: guid、grayid、ds、变量X

A	tab_coconu...midtable_day	category_name / 计数	Count(category_name)
B	nbase_t_sd...event_to_tdw	event_value / 计数	Count(event_value)

⊕ 新增事件

基于加工结果筛选用户 ☐

* 组合规则

分子 A 的求和 / 分母 B>0 的用户数

- **用户比例类**: 也叫 uv 比例类, 某条件下用户去重计数 / 用户去重计数, 形如 $COUNT(uin) / COUNT(uin)$, 例如点击用户占比、UV 转化率。

* 指标计算类型

4. 用户比例类

指标数据源处理 输出按天聚合的表: guid、grayid、ds、变量X

5. A dws_app_model_general_omg in_vstart_cnt / 求和 Sum(in_vstart_cnt)

添加变量

实验指标定义

6. 某条件下的用户去重计数

* 组合规则

分子 A>0 的用户数 / 分母 命中实验 的用户数

A>0 命中实验

7. 用户去重计数

- **比率类**: 累计指标求和 / 累计指标求和, 形如 $SUM(A) / SUM(B)$, A、B 一般都为随机变量, 例如 CTR、单 PV 时长。

* 指标计算类型

4. 比率类

指标数据源处理 输出按天聚合的表: guid、grayid、ds、变量X

5. A dws_app_model_general_omg in_poster_click_cnt / 求和 Sum(in_poster_click_cnt)

6. B dws_app_model_general_omg in_poster_imp_cnt / 求和 Sum(in_poster_imp_cnt)

添加变量

实验指标定义

* 组合规则

7. 累计指标 A 求和 的求和 / 8. 累计指标 B 求和 的求和

- **用户数据求和类**: 也叫 SUM 类指标, 累计指标求和, 形如 $SUM(A)$, 例如总时长、

总点击。

* 指标计算类型

4. 用户数据求和类

指标数据源处理 输出按天聚合的表: guid、grayid、ds、变量X

5. A dws_all_xsadexp_ul_tab_incor in_xsad_ul_income_with_co... Sum(in_xsad_ul_income_with_contract) </>

+ 添加变量

实验指标定义

* 组合规则 6. 累计指标求和

A 的求和

- **用户计数类：**也叫去重计数或者 COUNT 类，形如 COUNT(uin)，例如点击人数、累计播放用户数。

* 指标计算类型

3. 用户计数类

指标数据源处理 输出按天聚合的表: guid、grayid、ds、变量X 4. 由于是计算去重用户计数，所以可以不选择用户 id

A 请选择数据源表 请选择字段及函数 </>

+ 添加变量

实验指标定义

* 组合规则 5. 去重用户计数

命中实验 的用户数

- **人均活跃天数类：**满足某条件下天数 / 用户去重计数，形如 SUM(user_day) / COUNT(uin)，用于评估实验对 DAU 的影响。

* 指标计算类型

3. 人均活跃天数类

指标数据源处理 输出按天聚合的表: guid、grayid、ds、变量X

4. A dws_app_model_general_omg imp_date / 计数 Count(imp_date) </>

+ 添加变量

实验指标定义

* 组合规则 5. 满足某条件下天数 / 用户去重计数，即 A > 0 的用户的实验期间人均天数

满足 A>0 的用户的实验期间人均天数

留存类指标

- **去重口径留存：**实验期内用户首次命中实验后，N 日后又活跃的用户比例（重点

在于去重人数)。只统计用户首次命中实验的数据，属于延迟类指标_UV 比例。

3. * 指标计算类型

去重口径留存

指标数据源处理 输出按天聚合的表: guid、grayid、ds、变量X

A 请选择数据源表 请选择字段及函数

添加变量

实验指标定义

* 组合规则 5.首次命中实验的用户 6.次日也命中实验的用户 * 间隔天数 7. 1天表示次日留存

初始行为 命中实验 的用户数 结束行为 命中实验 的用户数 1 天

- **非去重留存：**实验期内每次用户命中实验后，N 日后又活跃的 pv 比例（重点在于非去重次数），属于延迟类指标_比率类。

4. * 指标计算类型

非去重留存类

指标数据源处理 输出按天聚合的表: guid、grayid、ds、变量X

A 请选择数据源表 请选择字段及函数

添加变量

实验指标定义

* 组合规则 5.初始时刻命中实验用户数 6.初始时刻 +n 命中实验用户数 * 间隔天数 7. n = 1

初始行为 命中实验 的用户数 结束行为 命中实验 的用户数 1 天

③ 异常值排除

- **异常值处理：**从数据源中选择“异常用户表”，作为该指标不计入统计的用户

异常值处理

☒ 异常用户名单 按黑名单过滤，指剔除异常用户ID

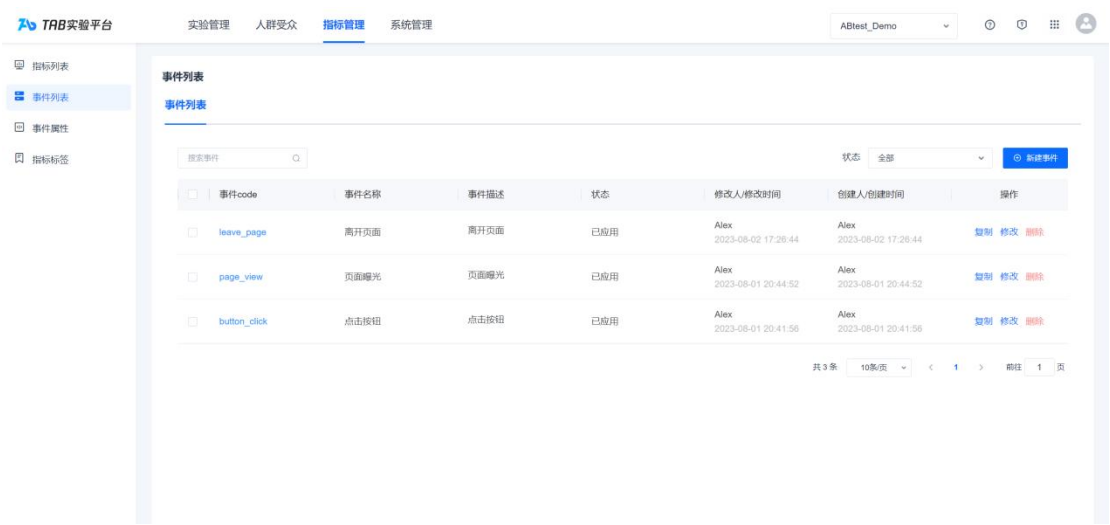
dim_usr_qimei_tab_abnormal_di

dim_usr_qimei_tab_abnormal_di

☐ 离群值替换

☐ 离群值剔除

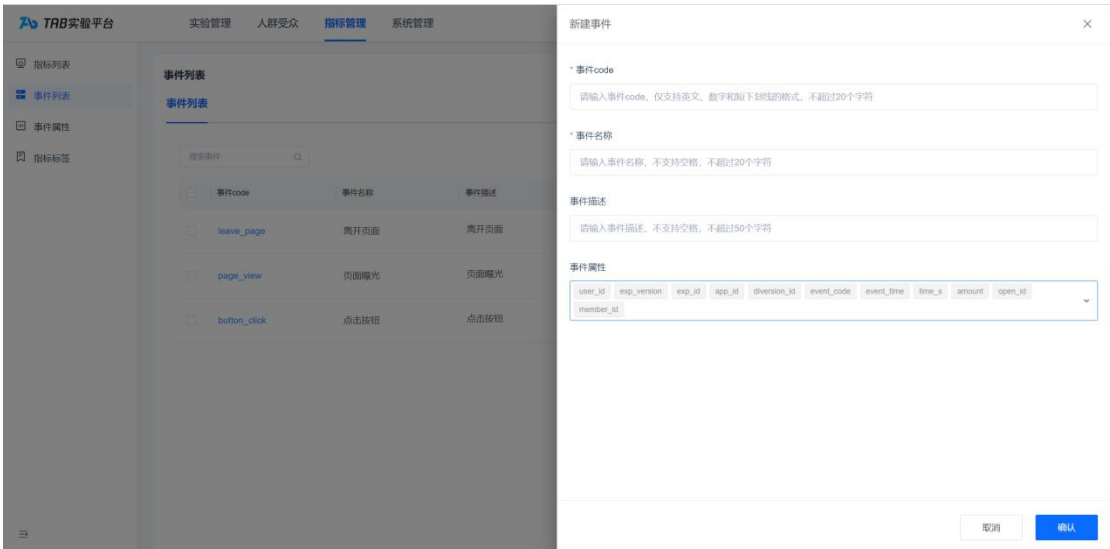
2. 事件列表



- **新建事件：**进入事件新建流程
- **复制：**复制当前事件数据进行创建
- **修改：**进入事件编辑流程
- **删除：**删除当前事件数据

2.1 新增事件

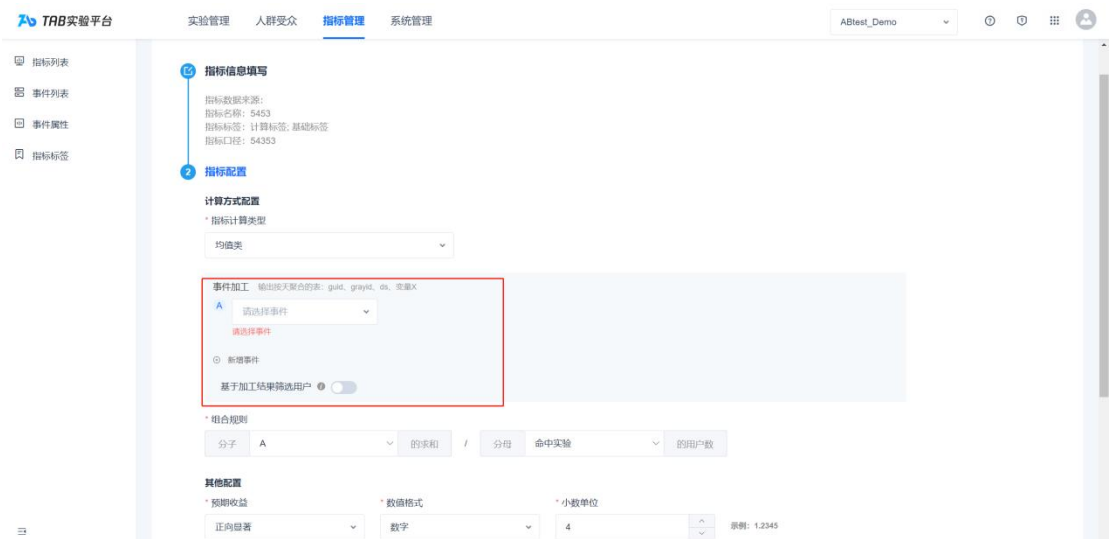
点击【新建事件】按钮，进入事件创建流程：



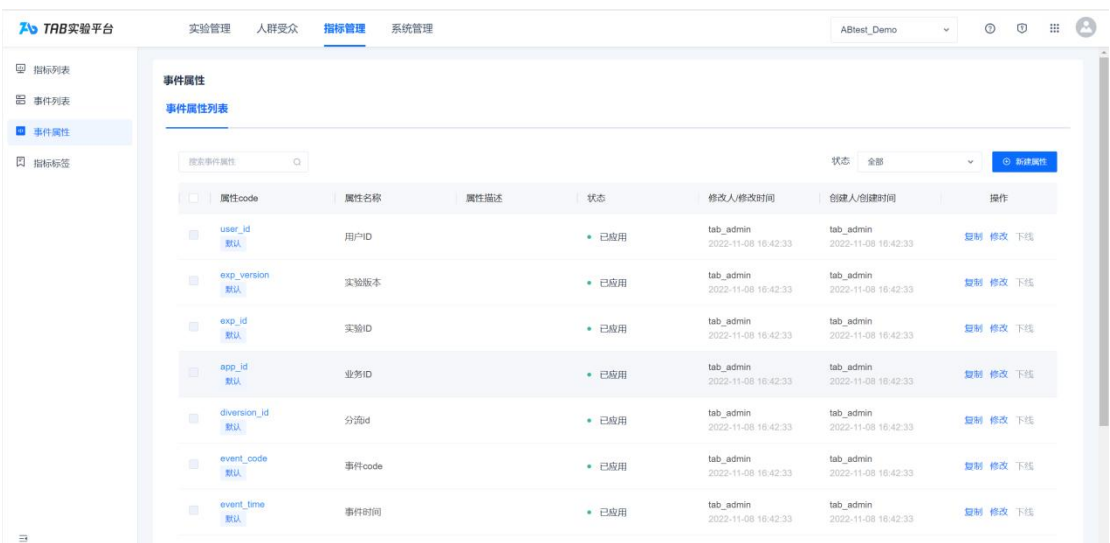
- **事件 code：**必填项，仅支持英文、数字和短下划线的格式，不超过 20 个字符
- **事件名称：**必填项，不支持空格，不超过 20 个字符
- **事件描述：**不支持空格，不超过 50 个字符
- **事件属性：**给当前事件添加或删除属性，可选择系统默认属性和自定义属性

2.2 使用事件

在新建指标时，配置指标计算类型中需要选择已经配置好的事件



3. 事件属性



- **新建属性：**进入属性新建流程
- **复制：**复制当前属性数据进行创建
- **修改：**进入属性编辑流程
- **下线：**下线当前属性数据，下线后的属性在事件创建中无法选到

2.1 新增属性

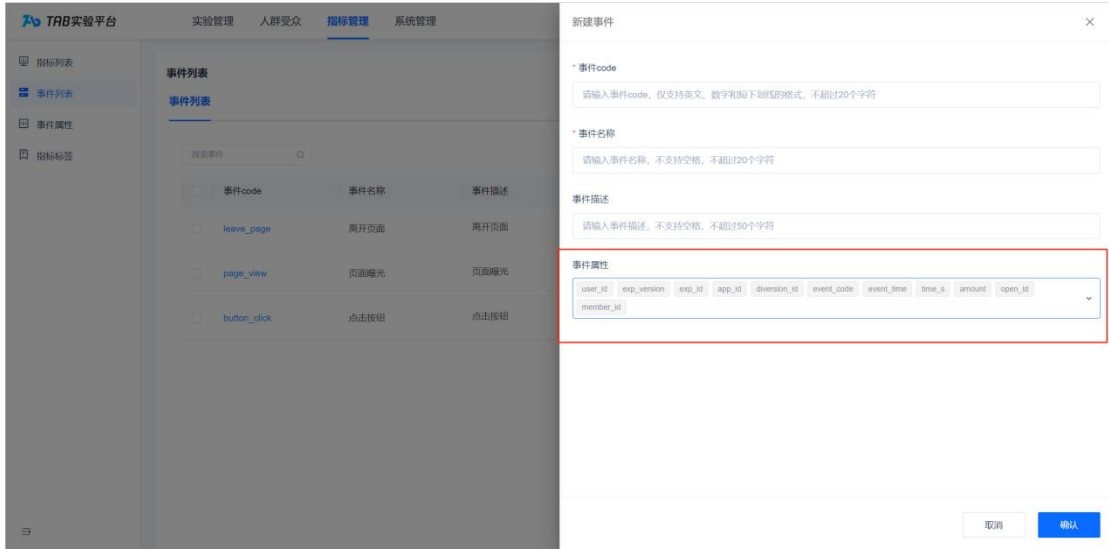
点击【新建属性】按钮，进入属性创建流程：

属性code	属性名称	属性描述
user_id 默认	用户ID	
exp_version 默认	实验版本	
exp_id 默认	实验ID	
exp_id 默认	业务ID	
diversion_id 默认	分流id	
event_code 默认	事件code	
event_time 默认	事件时间	

- **属性 code：**必填项，仅支持英文、数字和短下划线的格式，不超过 20 个字符
- **属性名称：**必填项，不支持空格，不超过 20 个字符
- **属性类型：**必填项，枚举值单选【int】【float】【string】【boolean】【datetime】
- **属性描述：**不支持空格，不超过 50 个字符

2.2 使用属性

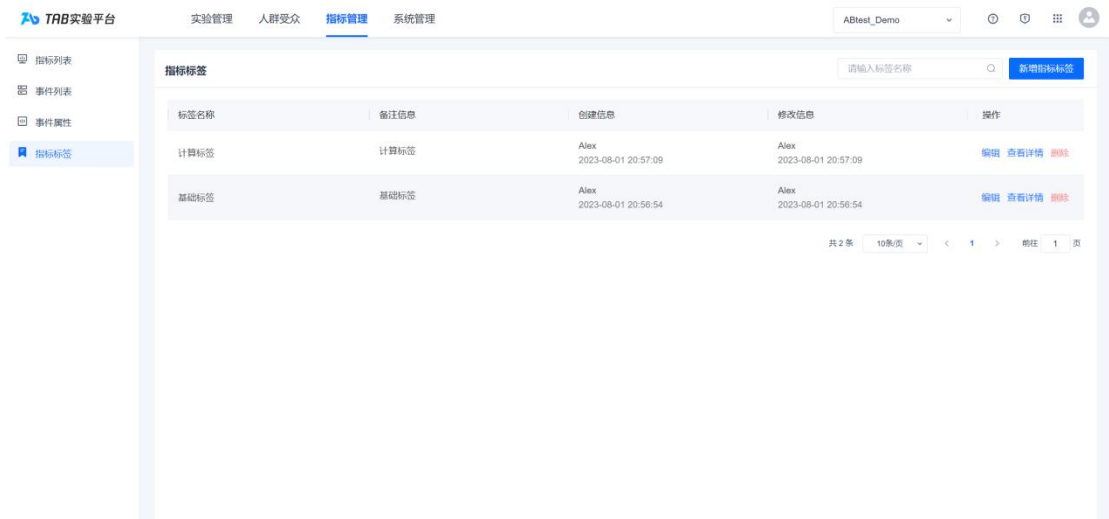
在新建事件时，配置事件属性中可选择已经创建且非下线状态的事件属性



4. 指标标签

3.1 指标标签列表

- **新增指标标签：**进入指标标签创建流程
- **编辑：**编辑标签信息
- **删除：**删除标签



3.2 创建指标标签

点击【**新增指标标签**】按钮，进入创建流程：

- **标签名称：**必填项，指标标签的名称
- **父级标签：**该指标标签的上一级
- **标签描述：**指标标签的描述信息
- **关联指标：**该指标标签与哪些指标挂钩

五、系统管理

1. 业务参数设置

1.1 实验治理设置

① 指标标签设置

设置指标标签层级后，可对指标按层级进行分类。在新建标签时，可以选择父级标签。最多可设置五级层级分类。

指标标签设置

设置指标标签层级后，可对impala指标按层级进行分类

最多可设置层级

3

② 期望提升指标 - 必填

打开开关后，在创建实验的时候，期望提升指标是必填项。

期望提升指标，必填 

打开开关后，在创建实验的时候，期望提升指标是必填项。

③ 实验关联场景

- 实验关联场景 - 开关打开：在创建实验的流程中，需要填写实验关联的场景
- 必填开关打开：由非必填→ 必填选项

实验关联场景 

开启高级上报/关联场景，AB实验类型，开关实验类型在创建的时候关联特定场景，场景信息从实验管理-场景管理中添加

场景默认值

name2 

实验场景必填 

开启后新增场景必填，AB实验场景必填

④ 实验开始禁止参数修改

开启这个功能后，仅实验状态在“调试期”可更改实验参数。空跑期、实验期、已全量、已下线，则不可修改实验参数。

实验开始禁止参数修改 

开启这个功能后，仅实验状态在“调试期”可更改实验参数

⑤ 创建实验配置实验参数时，强制只可使用参数模板

开启该功能后，“创建实验”在配置实验参数（key、value）的时候，只可使用实验参数模板的信息；创建实验完毕，后续更改实验参数不受此开关的限制。

创建实验配置实验参数时，强制只可使用参数模板 

开启该功能后，“创建实验”在配置实验参数（key、value）的时候，只可使用实验参数模板的信息；创建实验完毕，后续更改实验参数不受此开关的限制

⑥ 创建实验，检验是否与其他实验层 Key 相同

创建实验，检验是否与其他实验层 Key 相同。开启该功能后，表示需要进行实验参数校验。

创建实验，检验是否与其他实验层Key相同 

开启该功能后，表示需要进行实验参数校验

⑦ 实验流量分桶

分桶区别：

- **百分桶**：实验下线后流量不能再使用，即流量是不可再用状态，因此慎重选择百分桶。
- **千分桶/万分桶/十万分桶**：即一层由 1K/1W/10W 个桶组成，理论上桶数越多，波动越少。每层有 12%的不可用流量用于流量打散,减少实验波动。即每层可用流量为 88%。

实验下线后流量可复用。（放量或发布后可将实验流量占满 100%）

实验流量分桶

可根据DAU评估实验流量分桶，勾选后可在创建层域时选择对应流量分桶

☐ 百分桶 ☒ 千分桶 ☒ 万分桶 ☒ 十万分桶

2. 白名单组管理

2.1 背景信息

① 什么是白名单用户？

在实验正式开启之前，通常需要先选择几名用户进入测试阶段，观察实验过程是否能正常运行，以及实验策略是否生效等。参与这一步的用户被称为“白名单用户”。（白名单用户数据不会被计入实验结果）

② 白名单用户分类

- **临时白名单**：在创建/管理实验时，可以通过手动输入用户 ID，添加只针对此实验的一次性白名单
- **长期白名单**：为了在之后实验中重复调用同一组白名单，可以通过手动输入/文件上传

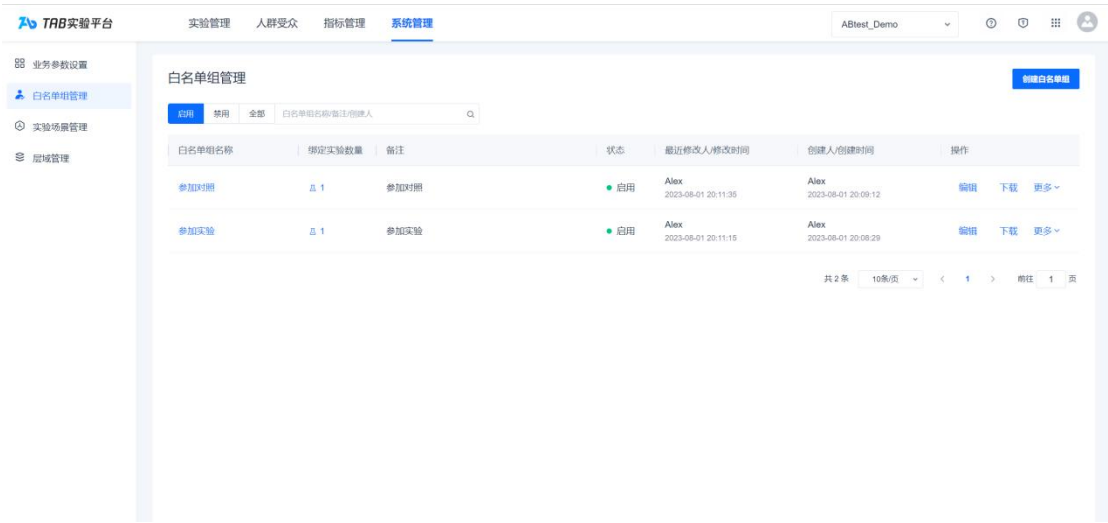
的方式，保存常用白名单，避免每次管理实验时重复性输入用户 ID 的冗杂和易错。

③ 什么是白名单组？

多个白名单用户，在一个组内，即被称为白名单组。

2.2 白名单组列表

- **搜索：**输入“白名单名称”即可搜索
- **创建白名单组：**点击【创建白名单组】按钮，填写信息，即可创建
- **编辑：**可修改白名单组名称、备注，可添加新的白名单用户 ID，可删除现有的白名单用户 ID
- **查看：**可查看白名单内容、别名和创建时间、所绑定实验的 ID、版本和绑定时间
- **更多：**可下载白名单、禁用/上线白名单

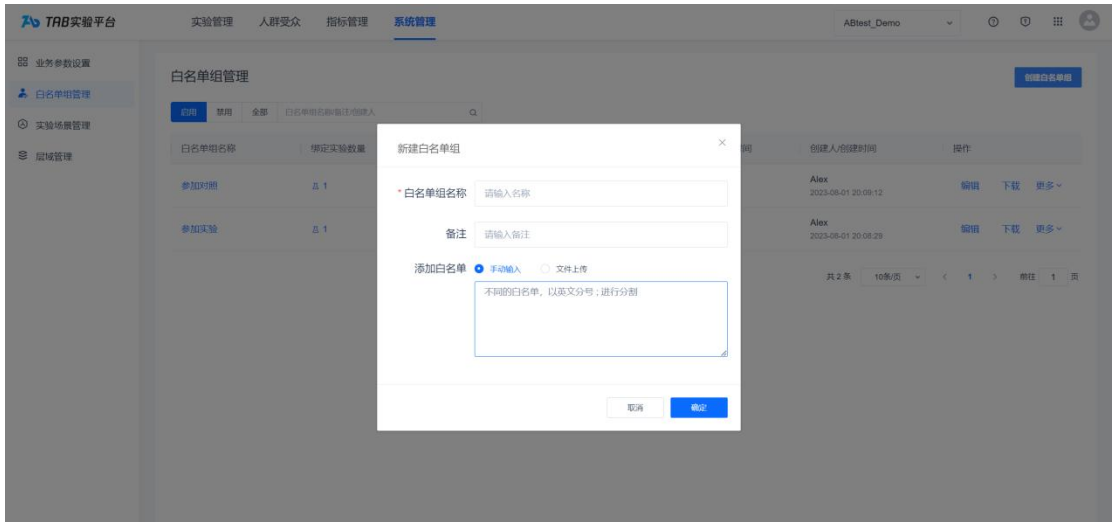


2.3 新增白名单组

点击【创建白名单组】按钮，进入创建流程：

- **手动输入：**手动输入白名单 ID，还可在该页面删除已添加的白名单组
- **文件上传：**通过文件上传的方式，添加白名单。仅支持 txt 文件，每一行代表一个白

名单。(单个实验版本绑定的白名单总数不得超过 5000 个，一个白名单组内白名单数量不得超过 5000 个)



2.4 使用白名单

在实验列表页，在操作中点击【白名单管理】按钮：



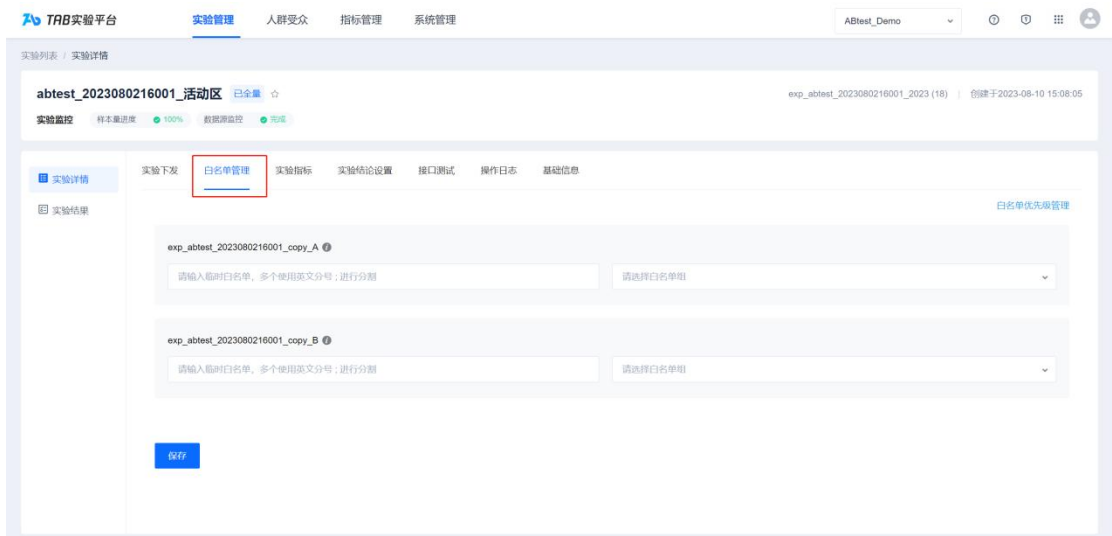
进入实验详情页：

在当前弹窗页面，可给该实验的各个版本添加白名单。

- **添加临时白名单：**填写用户 ID，该 ID 只针对该实验
- **选择白名单组：**选择已经创建的长期白名单组
- **白名单优先级管理：**若多个白名单组内添加了同一个白名单，那么可进行白名单优先级管理。比如白名单组“测试 1”和“测试 2”内，有同一个白名单“AA”，那么选择“测

测试 1”的优先级高于“测试 2”，那么白名单“AA”会优先命中“测试 1”所在的实验版本。

- **其他：**如果一个实验有 2 个版本，“版本 1”的临时白名单和“版本 2”的白名单组下的某个白名单用户重复,那么不做重复校验,但重复的这个白名单用户默认优先命中“版本 1”

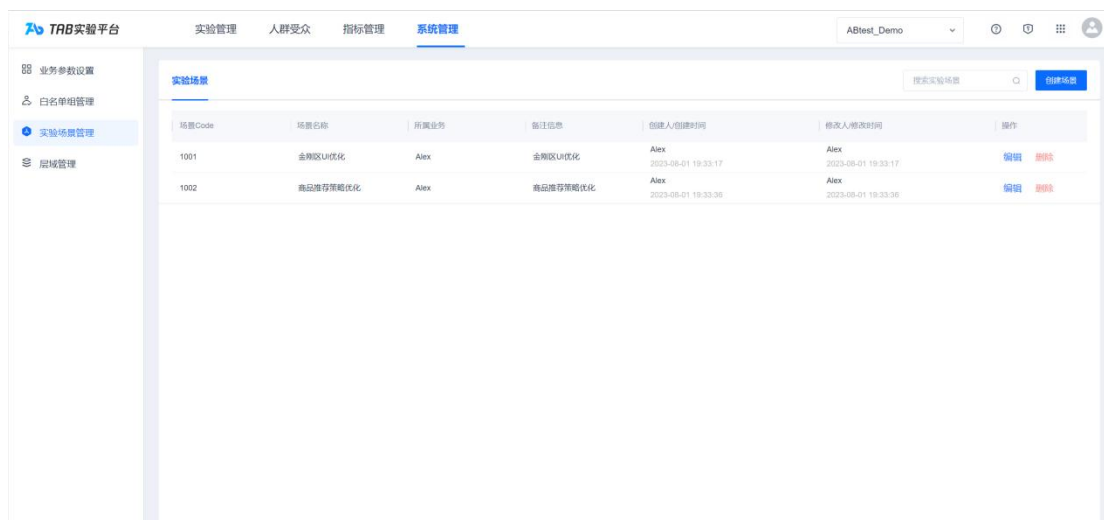


白名单命中规则：白名单同时添加在 2 个不同的实验层的实验，则 2 个实验都会命中；若添加在同层的 2 个实验，则后添加的会覆盖前面添加的。

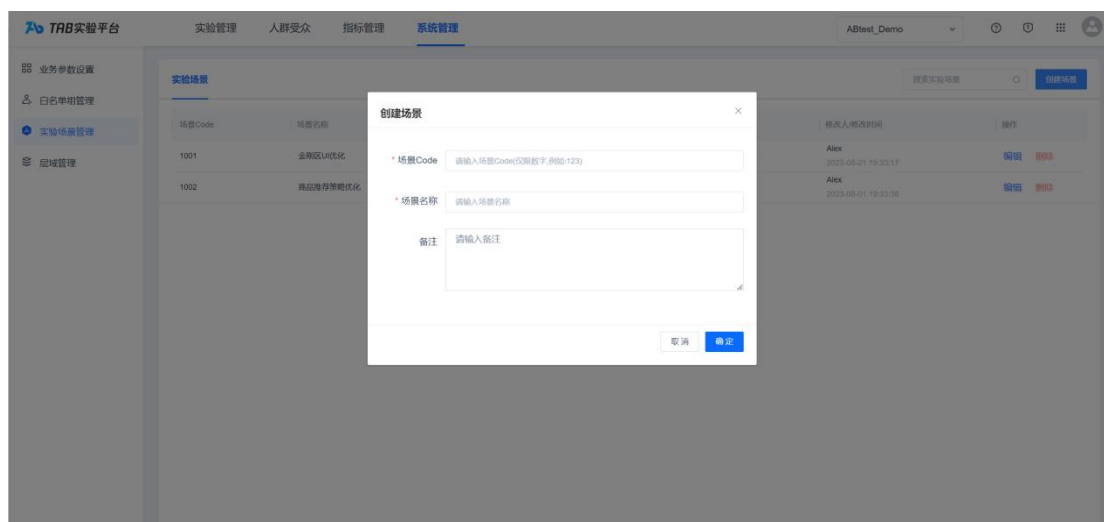
3.实验场景管理

实验关联场景，只针对“独立层 AB 实验”，其他类型的实验不需要手动关联。

创建实验的过程中添加“实验场景”，便于业务开发同学拉取该场景下的实验列表信息，过滤业务场景，避免过滤曝光问题。



点击【创建场景】按钮，填写场景 code 和场景名称：



4. 层域管理

4.1 背景信息

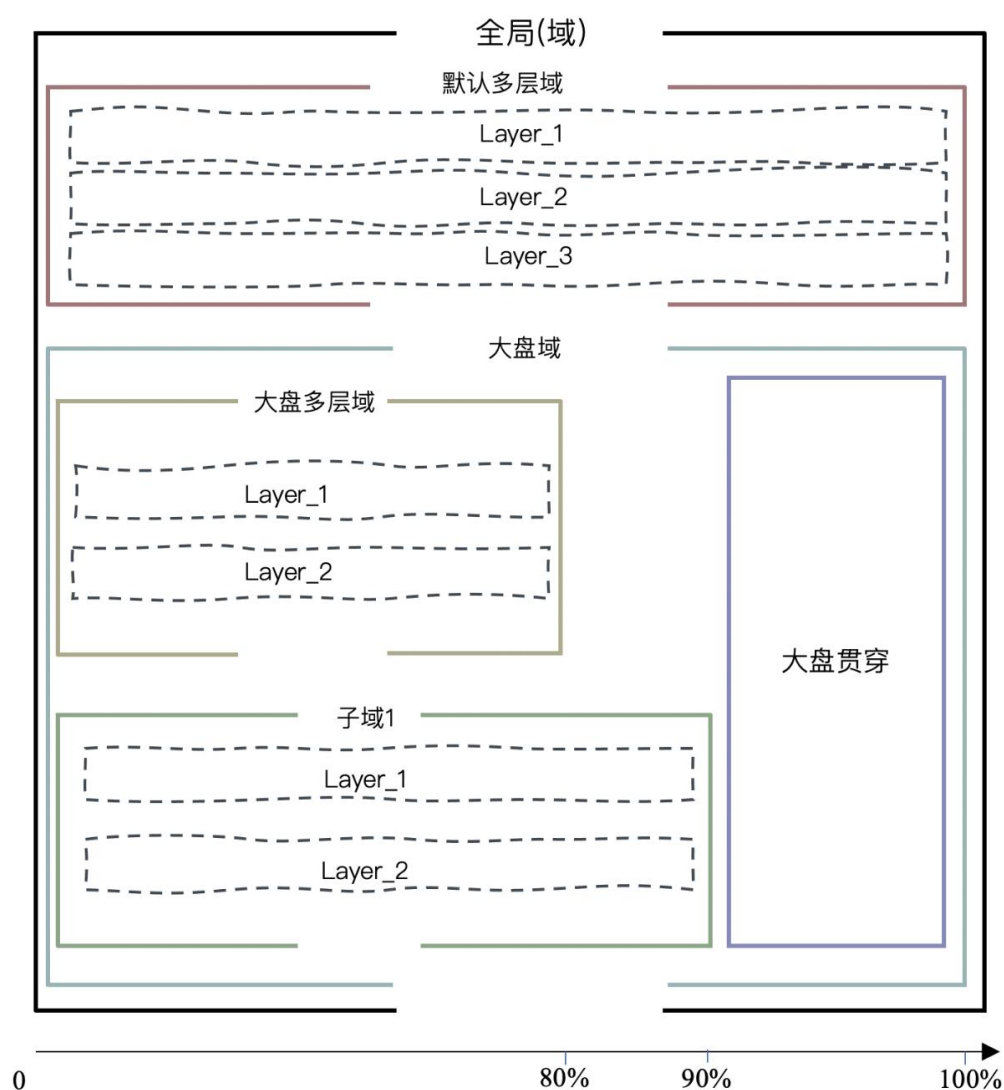
一个 AB 实验最大可占满大盘的百分百流量，各个实验版本之间是互斥的，多个实验之间是正交的。如果业务需要多个实验之间是互斥的，AB 实验是做不到的，因此引入层域实验来达到实验间互斥，流量贯穿等目的，适合大业务使用。

当业务只需要正交层，可使用默认多层域。

4.2 层域概念

- **多层：**实验是挂载多层或者单层上的，多层只能挂载在多层域上面。（多层其实还是一层，不是字面意思）
- **单层：**实验是挂载多层或者单层上的，单层只能挂载在域上面，流入流量可设定。
- **域：**子节点为单层、多层域、域，不能挂载实验，流入流量为 100% - 同级单层，最顶层的“全局”可以认为是一个域。
- **多层域：**子节点为多层，流入流量可设定。

实验只能挂载在单层或则多层下，其他节点不挂载实验。



4.3 层域列表

- **层域名**：可点击“层域名”列下的展开/收起 icon，即可展开收/起层域/实验层
- **占上级流量**：指的是“当前实验层”占“上一级层域”的流量占比
- **桶号**：流量的区分范围，如[98-100]，一共为 100 份流量，该层为 98、99、100 这三份流量
- **占全局流量**：由于层域划分分了很多层级，所以给出占全局流量字段，可以看到当前层占最上层全局下的比例
- **场景**：场景的作用是对层或者域做一个场景分类，方便在 SDK 上做实验信息获取的过滤，保证拉取的实验信息就是单独业务场景要用到的，避免过度曝光
- **流量分桶**：分为百分桶/千分桶/万分桶/十万分桶，除了百分桶，其他分桶粒度都有 12% 的预留流量
- **操作**：
 - **rehash**：该层流量 hash 重置，分桶分流处理（需要层上没有在线实验才能操作，否则会引起用户体验跳变）
 - **计算层域流量**：点开可以看到层上所挂在的实验是哪些

TAB实验平台

实验管理 人群受众 指标管理 系统管理

ABTest_Demo

🔍 ⚙️ 👤

业务参数设置

白名单管理

实验场景管理

层域管理

层域管理

该页面是对层、层域、域进行管理

层域信息 层操作记录

Alex

创建者 tab_admin

编辑

层域列表

输入层域名、层ID、用户名

层域名	层ID	层类型	占上级流量(桶号)	占全局流量	流量	流量分桶	修改者	创建者/创建时间	操作
全局	-1	域	100% ([1~100])	100%		无	tab_admin	tab_admin 0001-01-01 00:00:00	
默认多层域	-2	多层域	100% ([1~100])	100%		无	tab_admin	tab_admin 0001-01-01 00:00:00	
域01	module_yu01	域	100% ([1~100])	100%	0	百分桶	Alex	Alex 2023-08-01 19:50:55	



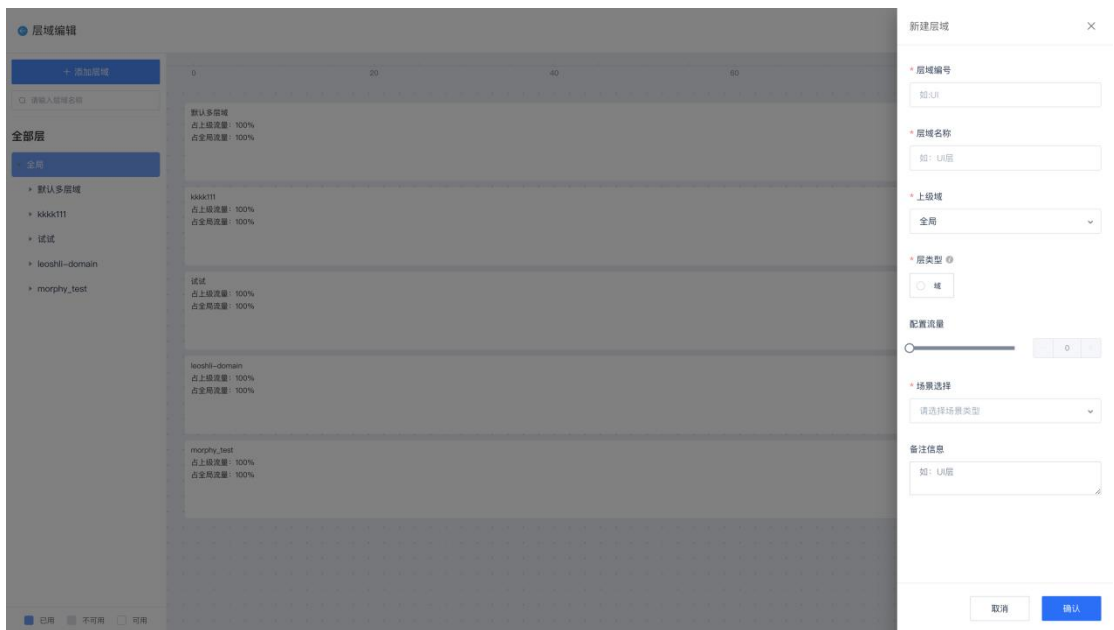
4.4 新建/编辑层域

点击【**编辑**】按钮，进入新建/编辑流程：

- **层域编号：**建议使用英文，这个 key 是 SDK 返回的结构中的 key，一个用户一层只会命中一个实验。
- **层域类型：**
 - **域：**是一个正交空间的划分，本质是域，上级也是域不可挂载实验。
 - **多层域：**是多层的载体，一般来说挂在域下，流量为 100%-单层。
 - **多层：**正交层，可以挂载实验。选择单层、多层的时候需要选定流量分桶，流量分桶默认为百分桶，建议为十万分桶。即一层由十万个桶组成，实验波动更少。
 - **单层：**是一个可挂载实验的实验层，一般用作贯穿层，因为域下流量处理为 100%-单层，这样单层会贯穿整个域。选择单层、多层的时候需要选定流量分桶，流量分桶默认为百分桶，建议为十万分桶。即一层由十万个桶组成，实验波动更少。
- **配置流量：**配置占上级
- **场景选择：**场景的作用是对层或者域做一个场景分类，方便在 SDK 上做实验信息获取的过滤，保证拉取的实验信息就是单独业务场景要用到的，避免过度曝光。部分大业务有开启场景过滤，如果使用，请先跟业务接入层同学了解清楚使用哪个场景，层域创建

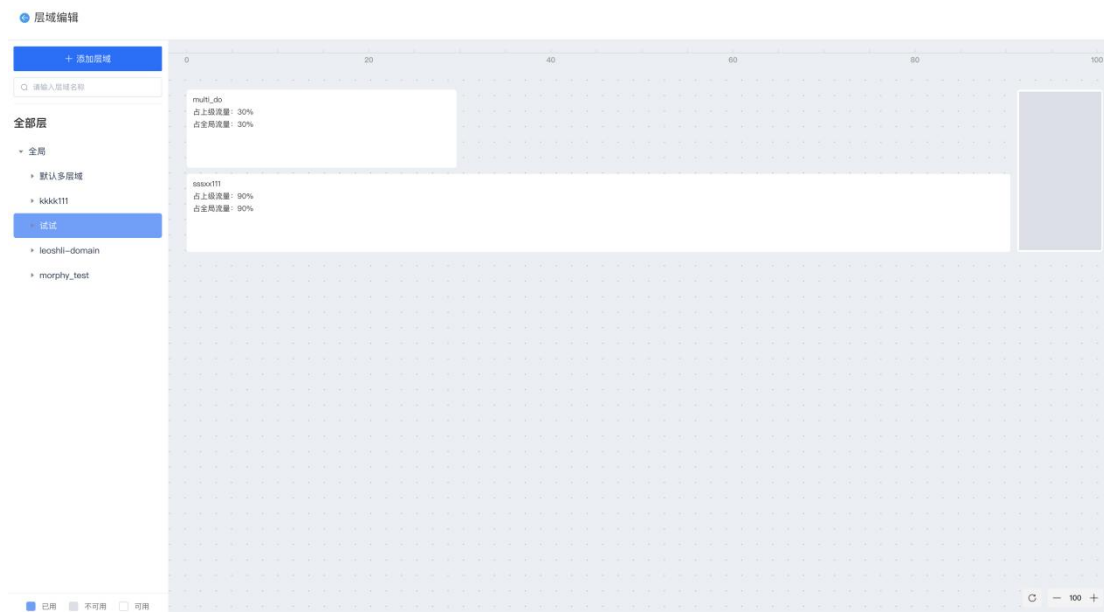
后，不支持修改。场景选择会影响到实验命中。

- **层桶数：**单层和多层需要设置层桶数，桶数选择尽量依据每个桶尽量大于 100 人，以防止某个桶没有用户的现象出现。例如 DAU 为 100W，则 $1000W/100=1W$ ，则可以使用万分桶。



4.5 查看层域

- **顶部数字：**顶部的 0 ~ 100 为用户进入这层的整体流量比例情况
- **已用：**该层上已有实验用到的流量范围
- **不可用：**不可用为 TAB 为层预留的 shuffle 流量，非百分桶层都会有 12%的预留流量
- **可用：**该层剩余多少流量可以使用



六、权限设置

点击页面顶部右上角【权限设置】按钮进入个人权限和空间权限设置页面



1. 我的申请列表

支持用户在【我的申请列表】中查看当前用户申请记录及结果



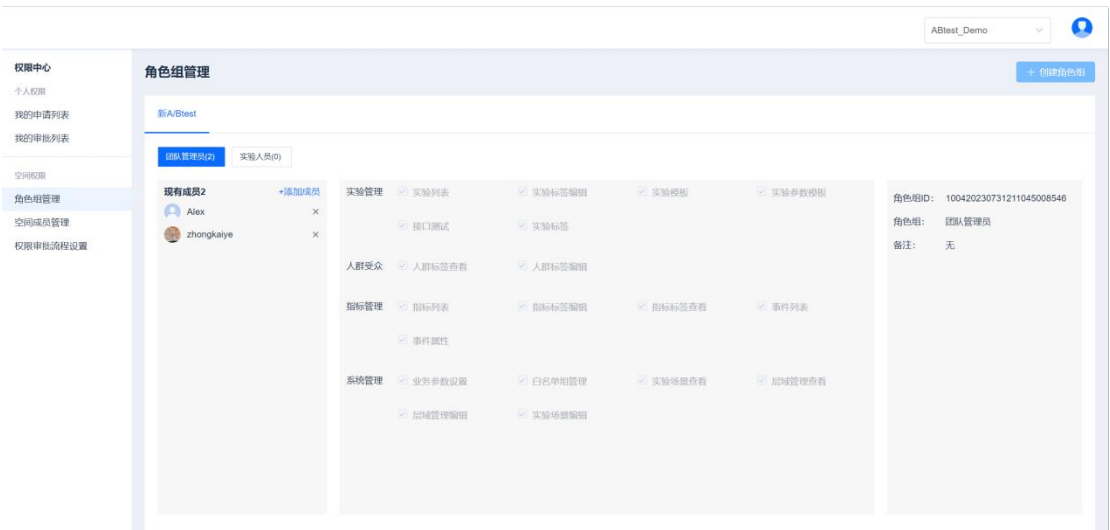
2. 我的审批列表

支持用户在【我的审批列表】中查看审批记录及结果，在列表中操作审批通过或拒绝

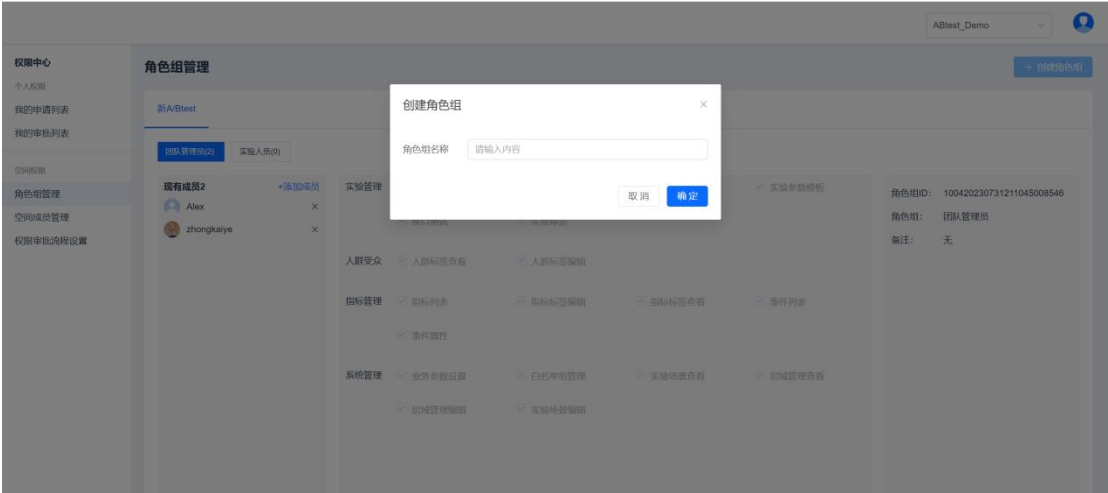


3. 角色组管理

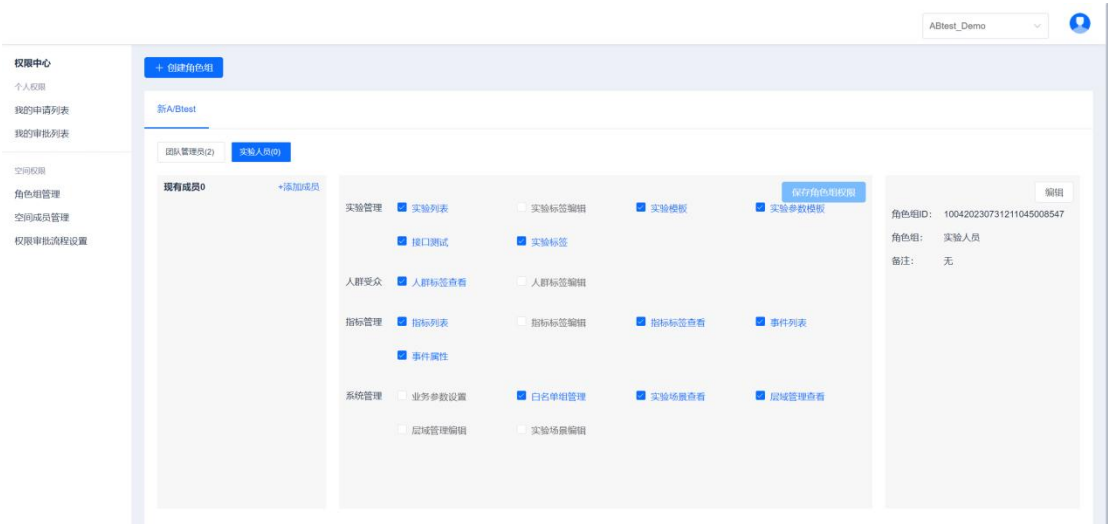
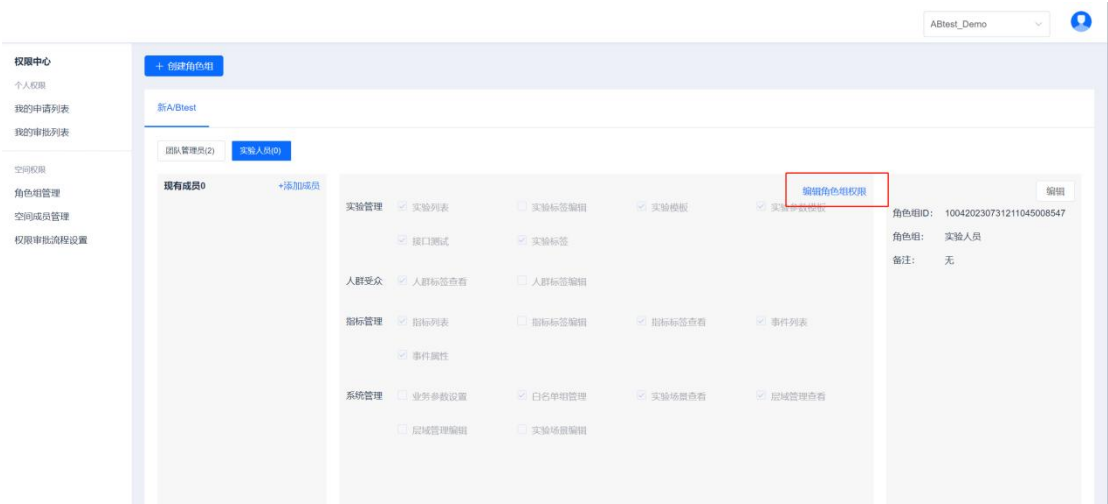
权限组管理中可以查看系统中所有的权限组和每个权限组中的成员,支持创建和编辑每个角色组中权限，添加和移除每个角色组中的成员



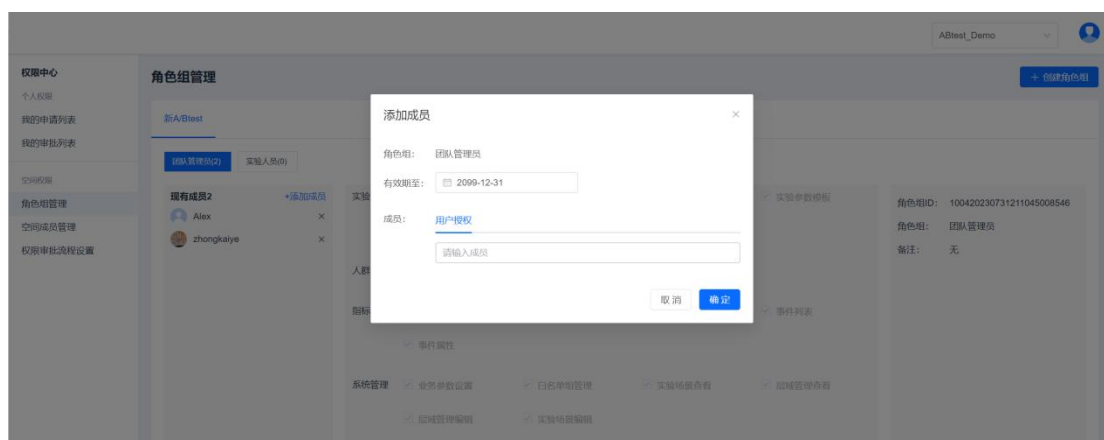
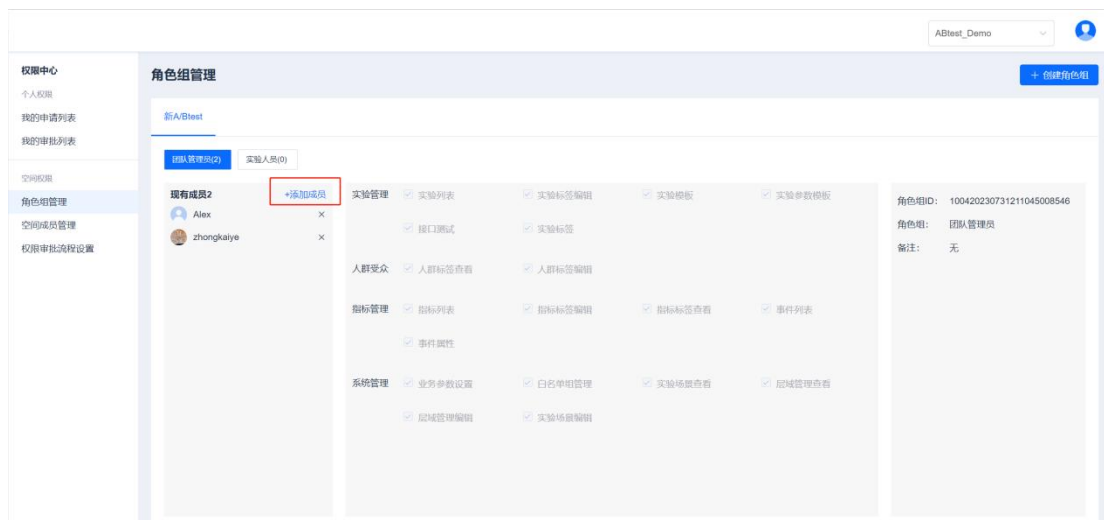
- 创建角色组



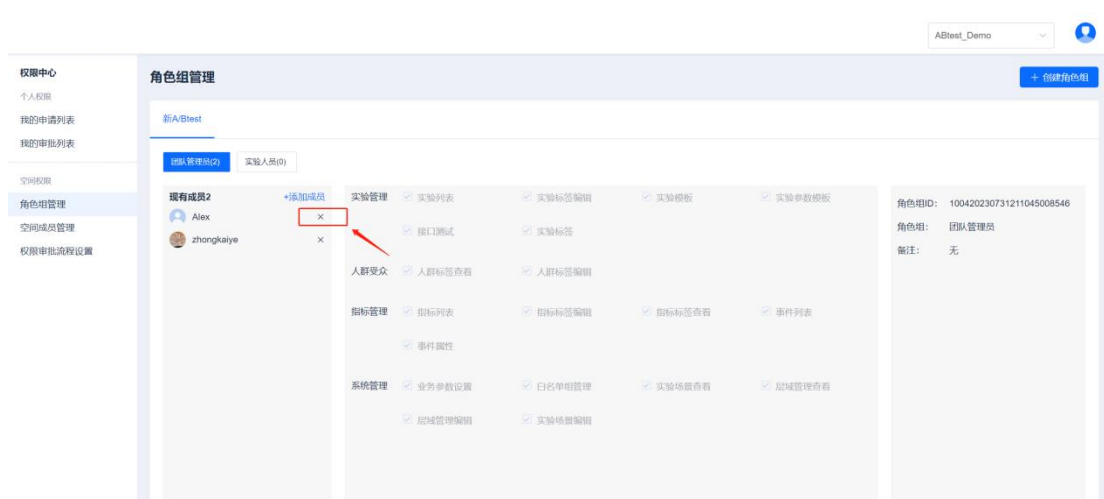
• 编辑角色组



• 添加成员



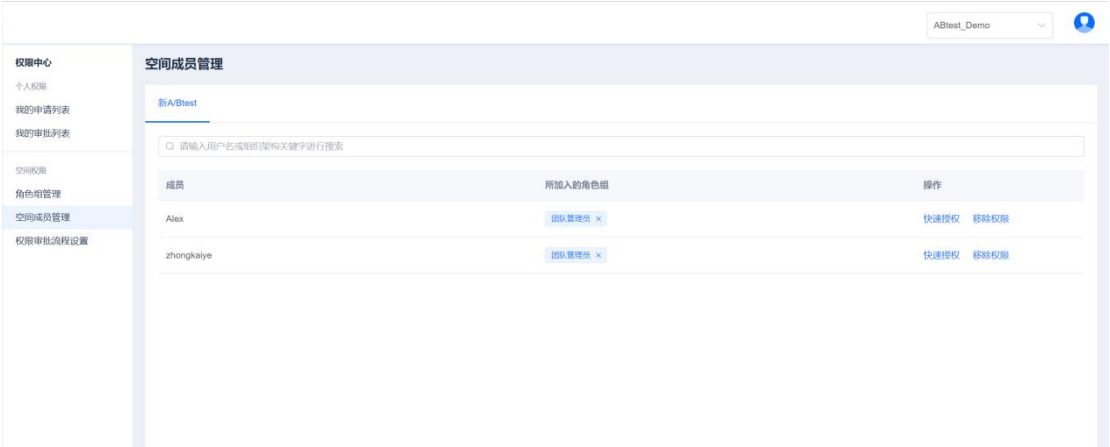
• 移除成员



3. 空间成员管理

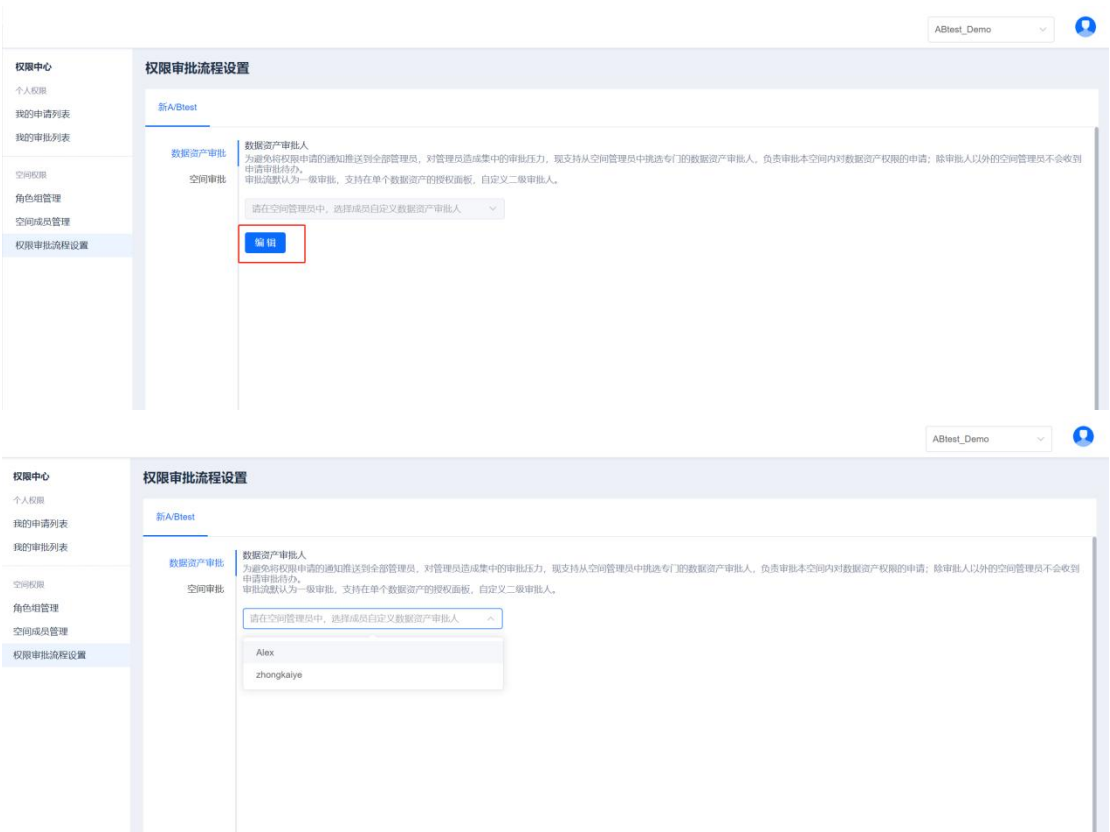
空间成员管理中可以查看系统中所有成员及所加入的角色组,支持快速给成员授权及移除权

限

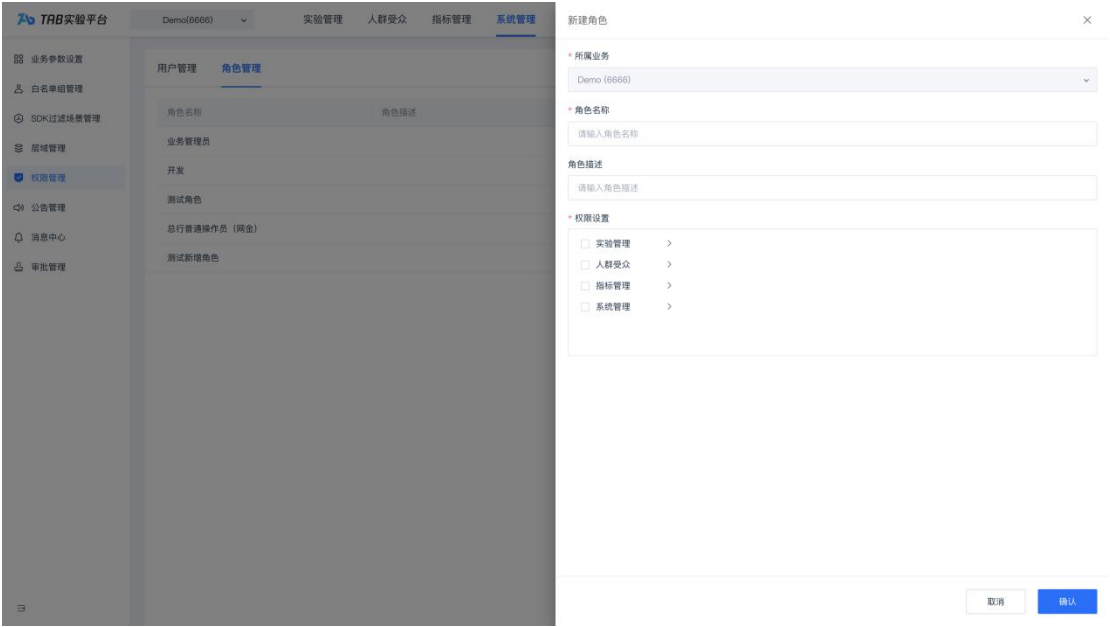
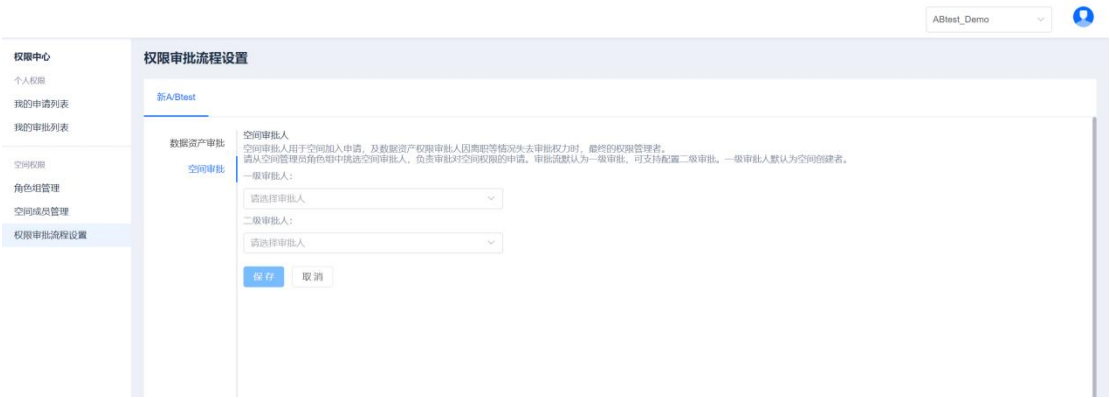


4. 权限审批流程设置

4.1 数据资产审批



4.2 空间审批



七、企业设置

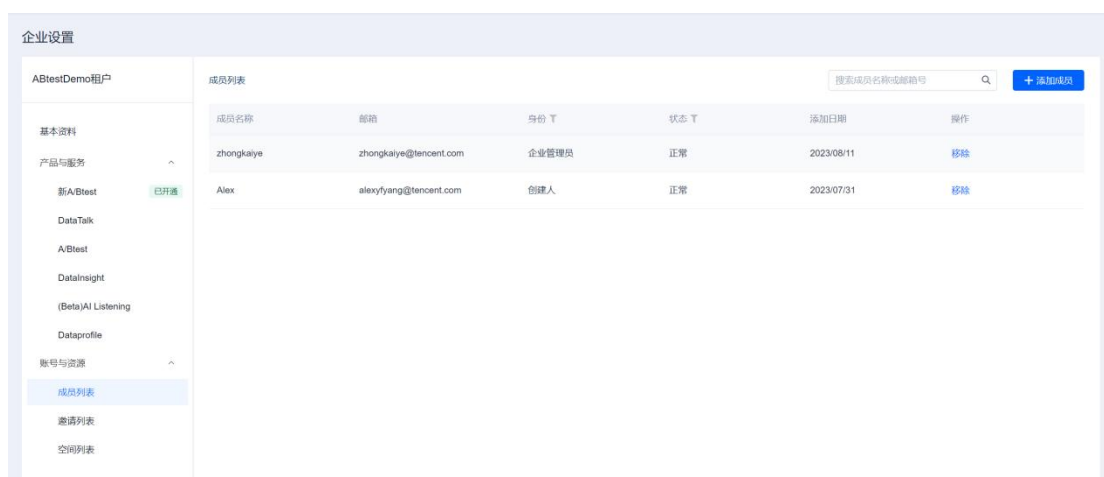
点击头像下拉中【企业设置】进入企业设置页面



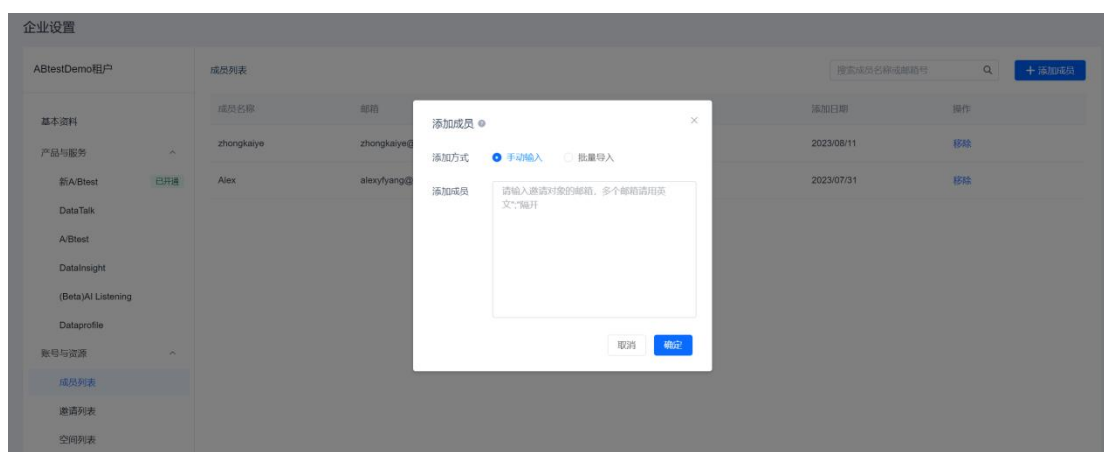
1. 账号与资源

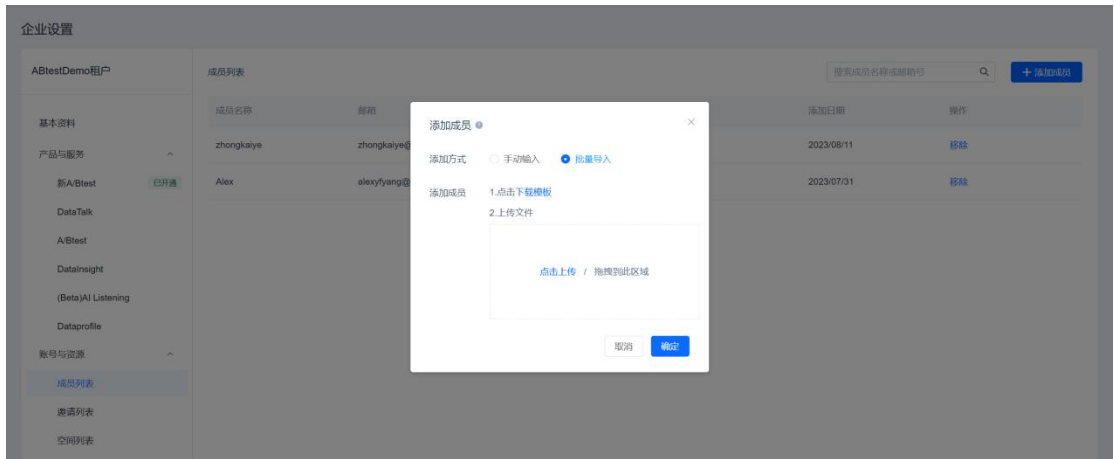
1.1 成员列表

支持添加、查看和移除系统中的成员



• 添加成员



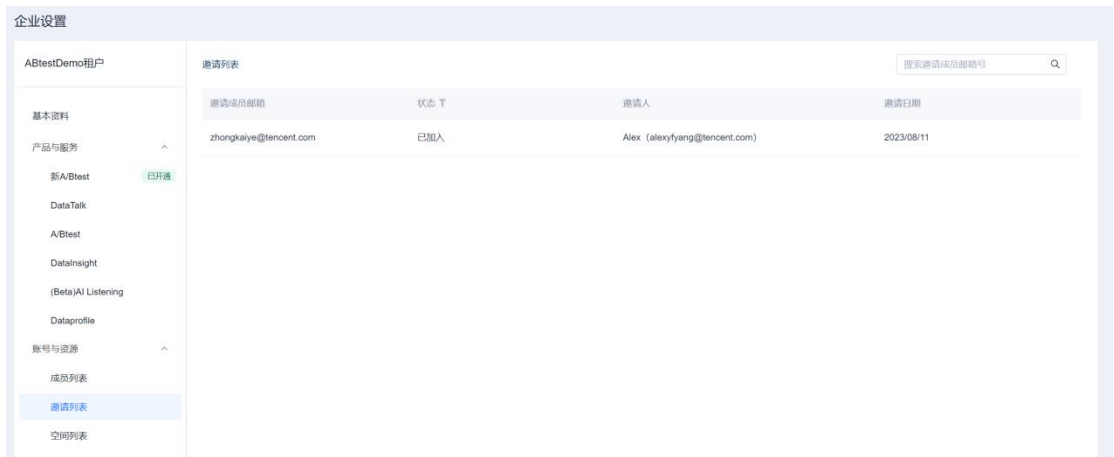


• 移除



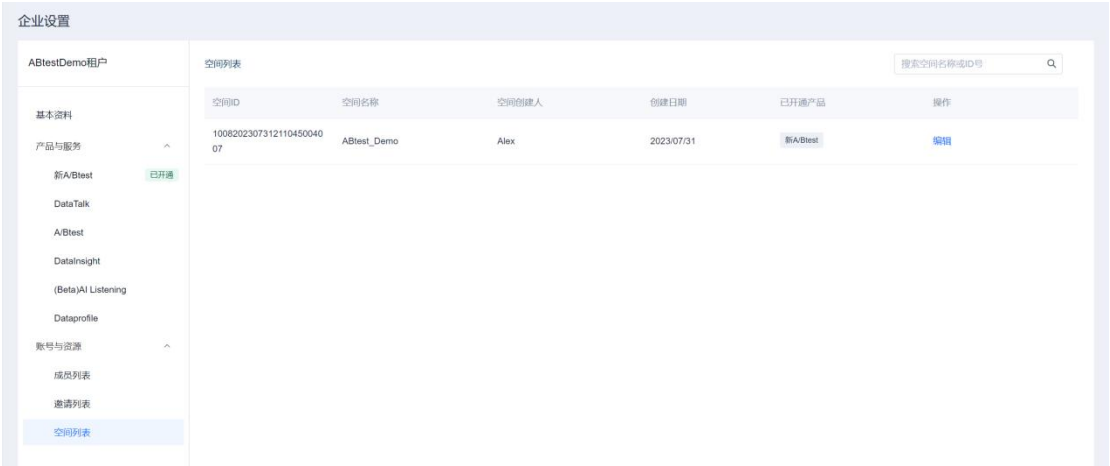
1.2 邀请列表

支持查看邀请加入的成员信息



1.3 空间列表

支持查看所有空间信息，支持编辑空间名称和空间管理员权限



- 编辑空间

