

性能测试 PTS

操作指南

操作指南

场景配置

创建压测场景

创建压测场景是使用 PTS 压测的第一步，本文将帮助您了解什么是压测场景，以及如何创建一个压测场景。

什么是压测场景？

要发起一次性能压测，首先需要创建一个压测场景。一个压测场景包含一个或多个并行的业务（即串联链路），每个业务包含一个或多个串行的请求（即 API）。

API 是场景压测中的必需元素，用来定义串联链路中每个阶段 URL 的具体信息。API 是由用户行为触发的一条端上请求。例如，电商网站的登录、查询商品详情、提交订单等，分别对应一次用户行为中的多个请求 API。

串联链路是一组压测 API 的有序集合（类似于事务），具有业务含义。

示例

以下示例说明了串联链路和 API 之间的关系。

淘宝网需要压测两个业务，要求两个业务同时进行：

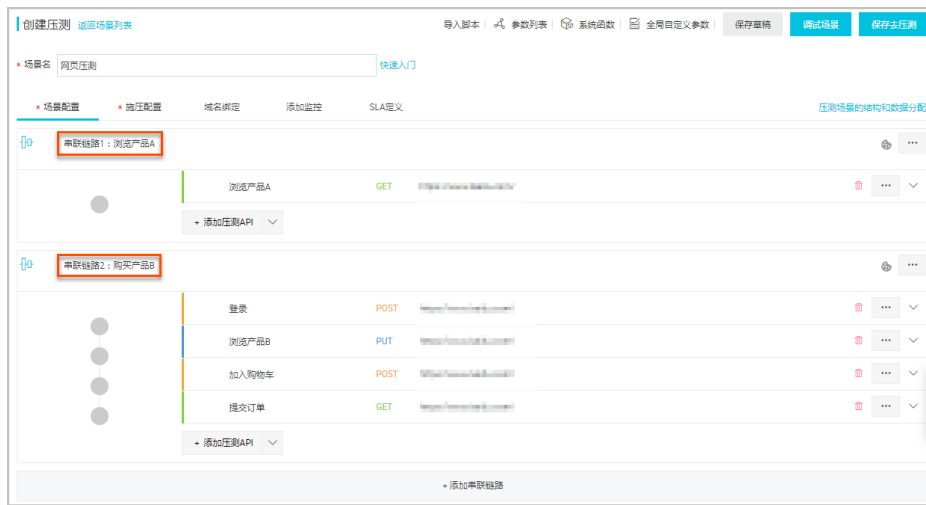
- 业务 A：浏览产品 A
- 业务 B：购买产品 B

业务 B 中包含四个请求，要求四个请求按照先后顺序发起：

- 请求 1：登录
- 请求 2：浏览产品 B

- 请求 3：加入购物车
- 请求 4：提交订单

那么在压测场景中的设置如下。



操作步骤

登录 PTS 控制台，在概览页单击**快速压测**，或者在左侧导航栏单击**创建压测**。

- 选择**快速压测**：PTS 最核心的压测方式，所有资源包均可使用。请前往步骤二，编辑压测场景。

选择**JMeter 压测**：您可以使用 JMeter 原生引擎发起压测，详见 JMeter 原生压测。

说明：如您有 JMeter 压测脚本，也可通过**导入脚本**功能，将脚本中 HTTP 请求接口的基本信息（如 Header、Body 等）导入到 PTS 场景中。详见 JMeter 导入 API 压测。

录制器：使用 PTS 的云端录制器录制请求，自动导入 PTS 发起压测。详见云端录制器。

在**创建压测**页面，输入场景名称，配置每个页签下的选项。

- 场景配置，参见自定义 API/请求。
- 施压配置，参见如何进行施压配置。
- 域名绑定，参见绑定域名。
- 添加监控，参见云监控。
- SLA 定义，参见配置场景 SLA。

单击 **+添加压测 API**，为某串联链路添加多个 API。启动压测后，多个 API 会按照顺序发起请求。

（可选）单击 **+添加串联链路**，为压测场景添加多个串联链路。启动压测后，多个串联链路会同时发

起请求。

在场景配置过程中，您还可以进行以下操作：

- 若需中途退出场景配置，单击**场景配置**页签右上角的**保存草稿**。保存草稿后，您可以在**压测场景**页面找到您的草稿态场景，继续编辑。
- 若需进行场景调试，单击**场景配置**页签右上角的**调试场景**，进行场景调试，详情请参见场景调试。
- 完成所有编辑后，单击**场景配置**页签右上角的**保存去压测**，启动压测。更多操作，详情请参见启动压测并调速。

后续操作

场景创建完成后，单击左侧导航栏的**压测场景**，在场景列表中查看已创建的场景和草稿态的场景。

在场景对应的**操作栏**：

- 单击**启动**，发起压测。详见压测中调速。

单击**编辑**，进入场景编辑界面，可以编辑场景配置信息。PTS 提供了多种提升场景编排效率的功能，包括场景复制、串联链路的生效/失效、API 复制。参见效率相关功能。

说明：目前效率相关功能只对购买了 898 及以上资源包的用户开放。

单击**删除**，可删除该压测场景。之前生成的压测报告仍可查看。参见压测报告。

- 单击**复制**，可复制出一个新的场景。参见效率相关功能。

自定义 API/请求

在很多压测场景中，往往需要对 URL、Header、Body 进行自定义，或者将部分变量参数化，或者截取返回的 Response 中的内容提供给后续请求使用。您也可能希望能基于业务维度将某个出参的临界值作为异常判断标准，为是否继续压测提供参考（断言）。

URL、Header 和 Body 都可以通过参数化实现动态变化。参数化可使用“前置请求的出参”、“参数文件”、“系统函数”、字符串或者他们的组合，详细使用请参见参数化使用方法。

基本请求信息

压测场景创建完成后，可以直接编辑压测请求。基本配置项见下表。

配置项	说明	示例
-----	----	----

API 名称	一般用于标识该 API 的业务含义，最长支持 50 个字符。	获取商品列表
压测 URL	请求的 URL，可包含参数，如 <code>http://域名/路径?key1=value1&key2=value2</code> 。URL 输入框中最大支持 2083 个字符。	<code>http://www.example.com</code>
请求方式	HTTP 请求方式，可以选择 GET、POST、PUT 或 DELETE。	/
超时时间	请求发送之后该虚拟用户等待应答的最长时间，单位为毫秒。默认值 5000，即 5 秒。 对于在线业务，不建议超时时间超过 5000 毫秒。主要行业的响应时间建议如下： - 互联网企业：500 毫秒以下，例如淘宝业务 10 毫秒左右； - 金融企业：1000 毫秒以下为佳，部分复杂业务 3000 毫秒以下； - 保险企业：3000 毫秒以下为佳； - 制造业：5000 毫秒以下为佳。	500
允许 302 跳转	允许跳转即统计跳转终态页面的状态码（最大深度为 10），否则按照当前请求的返回结果统计。	/
数据轮询一次	当使用参数文件来组装压测请求参数时，勾选数据轮询一次，能够保证组装后的每条请求信息都不重复。	/

压测 URL 示例：

如果只需对相同的 URL 做简单的批量请求（比如打开 `http://www.example.com`）压测，那么直接在 URL 定义中填写需要压测的 URL 即可（此处最多可输入 2083 个字符）。



不同请求使用不同参数的情况，可以通过参数化来实现。

Header 定义

对 API Header 的定义，可以通过以下三种方式，使 Header 在不同范围内生效。您可以根据实际场景，选择最适合的方式设置 Header。

- 单个 API Header：该 Header 仅对当前 API 生效。
- 串联链路 Header：压测时，该 Header 将自动添加到当前串联链路的所有 API 中。
- 全局 Header：压测时，该 Header 将自动添加到压测全场景的所有 API 中。

说明：

- 若通过以上方式设置的 Header 的 Key 相同，则范围小的 Header 定义生效。
- 若通过以上方式设置的 Header 的 Key 均不相同，则三种 Header 定义均生效。

设置单个 API Header

为单个压测请求定义 Header 信息，您可以在每个 API 的 **Header 定义** 页签下，填写固定 Key 和 Value。

示例：

基本请求信息		Header 定义	出参/断言定义 ?	文本编辑
Key	Value	操作		
token	123abc	删除		
Cookie	session=12345	删除		
输入Header Key	输入Header Value			

说明：

Value 中可以填写参数化的表达式。

函数、字符串的组合嵌套等更高级的参数编写方法，参见系统函数及字符串如何组合嵌套使用。

单击 **Header 定义** 页签右上角 **文本编辑**，可切换为文本形式，查看已输入 Key 和 Value。您也可以在此 **文本编辑** 形式下，批量写入多个 Header。示例如下：

```
token : 123abc
Cookie : session=12345
```

在 **Key** 输入框中，根据您填入的信息，会模糊匹配通用的 HTTP Request Header，您可以快速选择目标 Header。

- Content-type 需要在 **Body 定义** 页签进行设置。

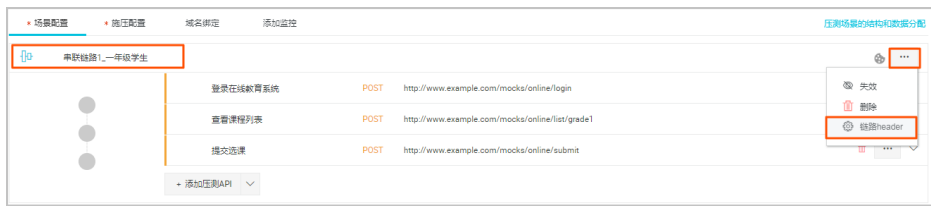
设置串联链路 Header

如果某串联链路中的多个压测请求需要同一个 Header，您可以通过设置串联链路 Header 的方式实现 Header 批量添加。

例如，在线选课业务中需要区分一年级学生、二年级学生，模拟这两种类型的用户同时发起请求的情况。您可以为一年级、二年级学生分别创建一个串联链路，在串联链路级别设置 Header，以更好地区分用户类型。

操作步骤如下：

在**创建场景**或**编辑场景**页面，单击串联链路最右侧的点号，在下拉列表中，选择**链路 Header**。



单击串联链路 Header 行，展开编写区域，填写 Header 的 **Key** 和 **Value**。示例如下：



说明：

Value 中可以填写参数化的表达式。

函数、字符串的组合嵌套等更高级的参数编写方法，参见系统函数及字符串如何组合嵌套使用。

单击 **Header 定义** 页签右上角**文本编辑**，可切换为文本形式，查看已输入 Key 和 Value。您也可以在**文本编辑**形式下，批量写入多个 Header。示例如下：

```
Grade : 1
Cookie : abc
```

在 **Key** 输入框中，根据您填入的信息，会模糊匹配通用的 HTTP Request Header，您可

以快速选择目标 Header。

Content-type 需要在 **Body** 定义页签进行设置。

设置全局 Header

如果整个压测场景中的所有请求都需要同一个 Header，您可以设置全局 Header 参数。

例如，页面压测中，为区分页面在 Chrome 和 Firefox 浏览器中的不同表现，可以创建两个压测场景，同时在每个压测场景中设置全局 Header，以更好地区分每个请求所属的场景。

操作步骤如下：

在**创建场景**或**编辑场景**页面，单击页面右上角的**全局自定义参数**。



在**全局自定义参数**对话框中，选择下方的**全局 Header**。

填写全局 Header 的 **Key** 和 **Value**。示例如下：

注意：全局 Header 的 Value 需填写确定值，**不支持**参数化表达式的解析。



说明：

单击 **Header 定义** 页签右上角**文本编辑**，可切换为文本形式，查看已输入 Key 和 Value。您也可以**在文本编辑形式下**，批量写入多个 Header。示例如下：

```
explorer_identification : chrome
Cookie : 123abc
```

- 在 **Key** 输入框中，根据您填入的信息，会模糊匹配通用的 HTTP Request Header，您可以快速选择目标 Header。

Body 定义

Body 定义的内容取决于请求参数类型。Content-Type，即请求参数类型，有三种：

- **x-www-form-urlencoded**：即 application/x-www-form-urlencoded，数据被编码为名称/值对；
- **raw**：
 - TEXT (text/plain)：即文本，纯文本格式的编码形式 (text/xml/html)；
 - JSON (application/json)：JSON 字符串。

Body 定义内容示例如下：

表单 (x-www-form-urlencoded) 示例：

以 Key/Value 形式输入:



单击 **Body 定义** 页签右上角**编辑文本框**，可切换为文本形式，最多可输入 65535 个字符。格式如下：

```
name=test&passwd=123456&itemid=2
```

对 Body 进行 decode : 如果定义的 Body 已编码, 例如 "%25" 即为 "%" 编码后的值, 那么压测前需要先对 Body 解码, 以避免 PTS 对已编码的请求再次编码引起错误。单击编辑框右上角的**对 Body 进行 decode** 并选择解码格式 UTF-8 或者 GBK。

JSON (application/json) 示例 :

```
{ "code":200,"data":{"items":[{"id":"123","name":"yaozhuang"}, {"id":"456","name":"fuzhuang"}]}, "requestId":"Cf2HJKEf-197B-JK47-79E9-FA82BH64KA40", "successResponse":false}
```

文本 TEXT (text/plain) 示例 :

```
<?xml
version="1.0"?> <resource> <id>123</id> <params> <name> <value>test</value> </name> </params> <
/resource>
```

出参和断言

出参是从应答 (Response) 中截取需要的内容, 作为断言, 或者在后续的请求中作为参数使用。出参支持的形式有 Body : JSON、Body : TEXT 和 Header : K/V。

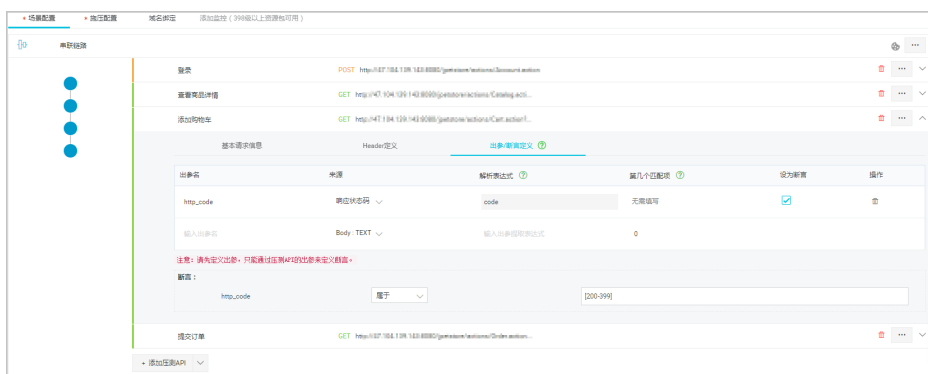
断言提供了业务维度成功与否的判断功能, 即将某个出参的临界值定义为异常判断标准, 类似检查点。

示例 1 : 将出参设置为断言

当请求的响应状态码为 200 至 399 时, 则代表业务正常。

基于请求的响应详情, 提取出响应状态码作为出参。填写出参名 (如 http_code) 和具体的解析表达式 (如 code) 。

勾选**设为断言**, 并在**断言**区域框中填写条件和数值, 如下图所示。



更多出参及断言的使用, 参见出参与断言。

示例 2：在后续请求中使用出参

理财业务中，需要根据用户的消费能力，推荐适当产品。则需要在前置接口中提取出消费能力等级作为出参，在当前的产品推荐接口中，使用该出参。

操作步骤如下：

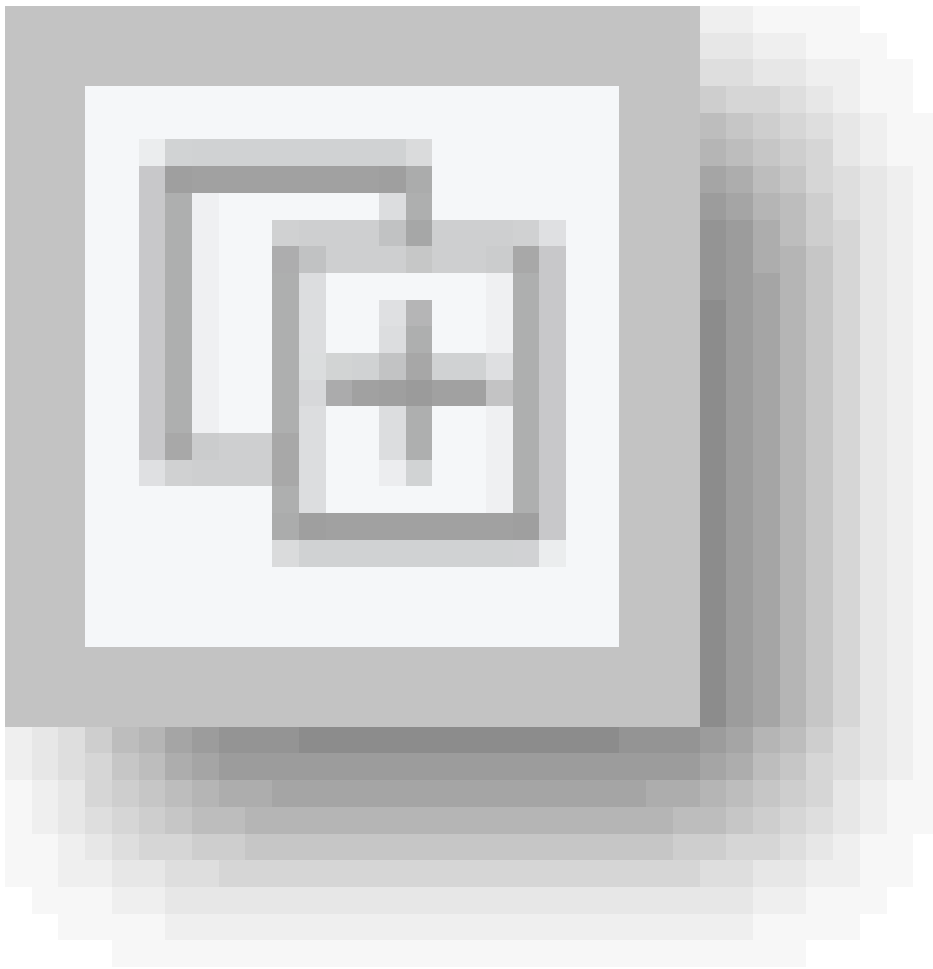
在**创建压测**页面，基于请求的响应详情，提取消费能力信息作为出参。填写出参名（如 output）和具体的解析表达式（如 data.items[0].value）。



单击**创建压测**页面右上角的**参数列表**，可查看已创建的出参参数。



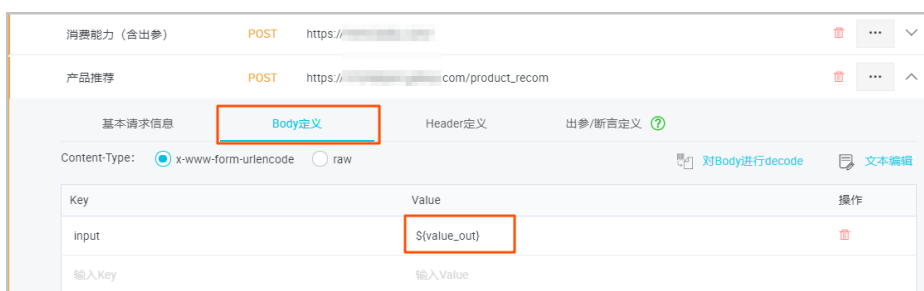
单击出参参数名称或



复制参数内容。

图标，系统将自动

假设需要在产品推荐 API 的 Body 中带入已提取的出参，在 **Body 定义** 编辑框内，粘贴参数内容。



此外，您也可以对 Body 内容进行编辑，如组合字符串、参数或函数等。

关于出参与断言的更多信息，参见出参与断言。

出参与断言

出参是从当前请求的 Response 中截取需要的内容，设置为断言，或在后续的请求中作为参数使用。断言用于判断业务是否成功，将某个出参的临界值定义为业务异常判断标准，类似检查点。

本文详细介绍如何提取出参、配置断言。

出参

在一个压测请求 API 中可以定义多个出参，示例如下：



出参的配置项说明：

配置项	说明
出参名	只能包含英文字母、数字和下划线
来源	标记 Response 的解析方式，支持选择以下： <ul style="list-style-type: none"> - Body : JSON (以 JSON 格式解析 Response Body) - Body : TEXT (以 TEXT 格式解析 Response Body) - Header : K/V (以键值对格式解析 Response Header) - Cookie : K/V (以键值对格式解析 Cookie) - 响应状态码 (提取 Response 中的状态码)
解析表达式	从 Response 截取需要的内容，对应到当前变量。
第几个匹配项	仅用于来源为 Body : TEXT 时。若上一步定义的解析表达式在 Response 中有多个匹配时，指定第几个字符串作为出参。 从 0 开始，-n 表示倒数第 n 个，取值区间为 [-99,99]。如果想要取随机匹配项，请填写 random 。
设为断言	具体用法，参见断言小节。

注意： 为确保能提取成功，Response Headers 中数据格式必须是标准的 MIME 类型。

关于解析表达式的编写，您可以在场景调试时，查看请求响应详情，测试出参解析表达式是否正确，并直接设置出参。请参考场景调试的“测试出参”小节。

Body : TEXT

TEXT 格式解析支持任意文本格式且可使用正则表达式提取。当一个正则表达式有多个匹配项时，可指定第几个匹配项，默认用 0 表示匹配第一个。

API 的 Response 示例如下：

```
<input name="id" value="347290348">
<input name="token" value="acdfo4dfopasdf44ddvwm">
...
<script>
var planId=4587;
var planId=5689;
var planId=8906;
</script>
```

出参的值	解析表达式	匹配数字
获取 name 为 id 的 value	<input name="id" value="([0-9]*)">	0
获取 name 为 token 的 value	name="token" value="([A-Za-z0-9]*)"	0
获取 第三个 planId 的值	var planId=([0-9]*);	2
随机 planId 的值	var planId=([0-9]*);	random

注意：

顺序匹配原则：按照顺序匹配出参信息。一旦匹配到之后即赋予该出参变量值，不会再匹配后续的信息，故需保证解析表达式的唯一性（除非多个中选第一个即可）；

被匹配的文本内容不能包含大括号 {}、小括号 () 等特殊字符。

使用正则表达式提取 JSON 类型的 Response 时，无需在 K/V 间隔的冒号后面加空格（JSON 格式化展现中加入空格是为了展现优化）。可查看 **Response 原文**信息，参考下图：



Body : JSON

JSON 格式解析支持两种类型的返回信息：application/json 和 text/json。

API 的 Response Body 示例如下：

```
{
  "info": "success",
  "message": "处理成功",
  "data": {
    "id": 13509, "code": 0,
    "items": [
      {"name": "name1", "value": "1234"},
      {"name": "name2", "value": "8448"},
      {"name": "name3", "value": "1298"},
      {"name": "name4", "value": "3049"},
      {"name": "name5", "value": "7648"}
    ]
  }
}
```

期望获取出参的位置	解析表达式
获取 info 的值	info
获取 data 的 id 的值	data.id
获取 items 第一个对象（支持相对位置）的 value 的值	data.items[0].value
获取 items 倒数第二个对象（支持相对位置）的 value 的值	data.items[-2].value
获取 items 整个数组	data.items[]
获取 items 数组中随机一个对象	data.items[random]

Cookie : K/V 和 Header : K/V

Cookie : K/V 和 Header : K/V 分别用于提取 Cookie 字段和 Header 字段，解析表达式中直接写需要提取的 Key 即可。例如，Cookie 信息为 “token=1234;path=/"，需要提取 token 的值，则解析表达式直接填写 “token” 即可。

响应状态码

响应状态码支持提取请求的状态码，一般情况下用于断言或者条件跳转（不同状态码返回不同页面信息）。

断言

断言提供了从业务维度判断成功与否的功能，将某个出参的临界值定义为业务异常判断标准，类似检查点。定义之后，在压测实时报表和压测报告中都会展现业务成功或失败的信息，为判断是否继续压测提供参考。

设置出参后，勾选**设为断言**，并在下方定义断言判断的 Value。

出参名	来源	解析表达式	第几个匹配项	设为断言	操作
test	响应状态码	code	无需填写	<input checked="" type="checkbox"/>	删除

输入出参名: Body: TEXT

输入出参提取表达式: 0

注意: 请先定义出参, 只能通过压测API的出参来定义断言。

断言:

test 属于 [200-299]

配置说明：

- 多个断言之间是“且”的关系。
- 当 Value 为字符串（非纯数字）时，支持“等于”、“不等于”、“包含”、“不包含”、“属于”、“不属于”。
- 当 Value 为数字时，支持“大于”、“大于等于”、“等于”、“小于等于”、“小于”、“属于”、“不属于”。
- “存在”、“不存在”所有类型数据均可使用，仅判断出参是否存在。

“属于”、“不属于”支持正负整数、0、字符串：

数值：请按照标准的数学表达式规范填写。示例如下：

- 表示闭区间：[-20,20]
- 表示集合：{[-200,200],-1,2}

说明：当数值集合中不包括负整数时，旧表达式仍然生效，但不推荐使用。例如，[200-300,1]，可表示包括 1、200 至 300 闭区间的整数的集合。

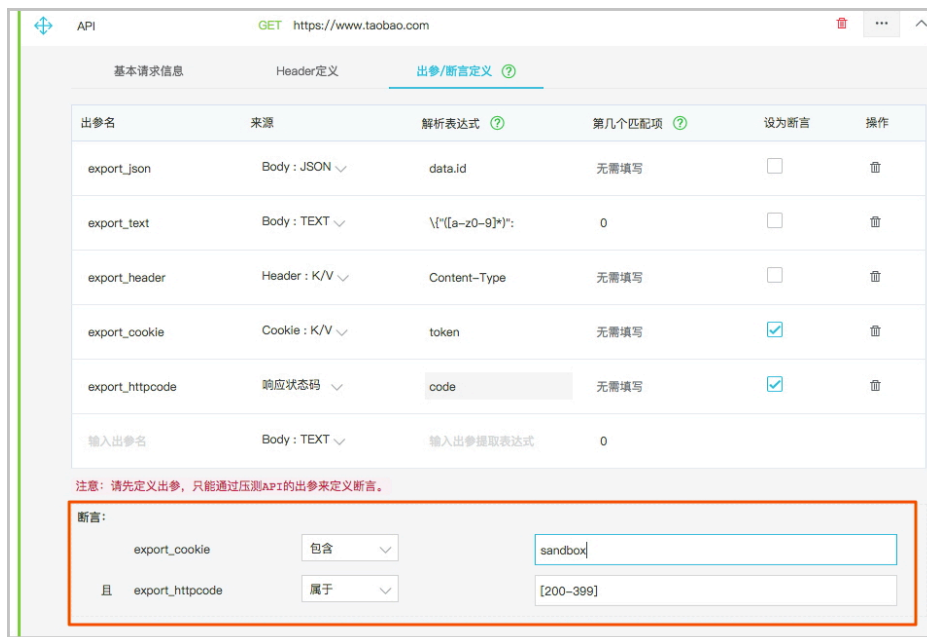
字符串示例：{ "abc" , " bcd" , " 200" , " -200" , " 已报名" , " 报名成功" }

示例：

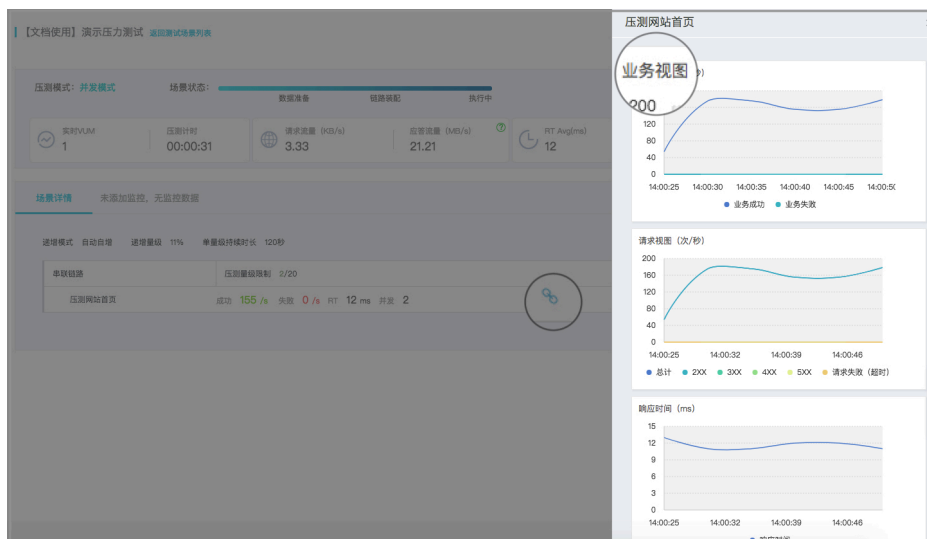
假设某业务中，需要判断请求用户来源环境是否同时满足以下两个条件：

1. export_cookie 符合预期，如 export_cookie 包含 sandbox。
2. 响应状态码在 200-399 范围内。

那么对应 API 的断言配置，如下图所示。



带断言的请求压测过程中，可以查看业务成功率：



参数化使用方法

在一些业务场景下，压测的请求中动态可变的部分是需要参数化的，比如以下常见场景：模拟不同用户登陆、当前接口依赖前置接口的部分返回信息、请求参数中需要带上当前的时间戳甚至还需要二次加签等。

本文将介绍参数类型、如何创建各类型参数和使用参数。

参数类型

压测请求配置时，能够为某个 API 的 URL、Header 或 Body 添加特定的参数，使 URL、Header 或 Body 动态变化。

PTS 支持配置以下参数来源及他们之间的组合。

适用范围	参数类型	说明
全局	文件参数	将存在关联的参数放在一个文件的不同列中，通过文件上传相应的参数值。
全局	自定义参数	将字符串、系统函数组合成一个自定义的变量，便于在全局使用，适用于全局需要经常调整某部分设置（如某个 Header 设置）。
全局	系统函数	包括加密函数、随机函数、四则运算和其他常用函数类型。可以对字符串、全局自定义变量及串联链路变量做函数再处理。例如，对前置接口返回的 UUID 输出做 MD5 加签。
全局	全局 Header	为压测全场景的所有 API 设置同样的 Header。具体用法参见 Header 定义。
串联链路	串联链路 Header	为串联链路下的所有 API 设置同样的 Header。具体用法参见 Header 定义。
-	数据导出参数	导出串联链路中的参数后，供其他串联链路使用，具体参见数据导出。
串联链路	自定义参数	通过数据指令定义的自定义参数，具体参见数据指令。
串联链路	出参	在创建串联链路时，将前置接口的部分返回信息作为参数（即出参）。

使用系统函数

下面以对 Body 进行参数化为例（URL 和 Header 类似）介绍如何使用系统函数。通过复制所需函数变量，粘贴至可编辑的文本区域。同时，可以按需进行组合、替换、嵌套或者追加需要的字符串等。

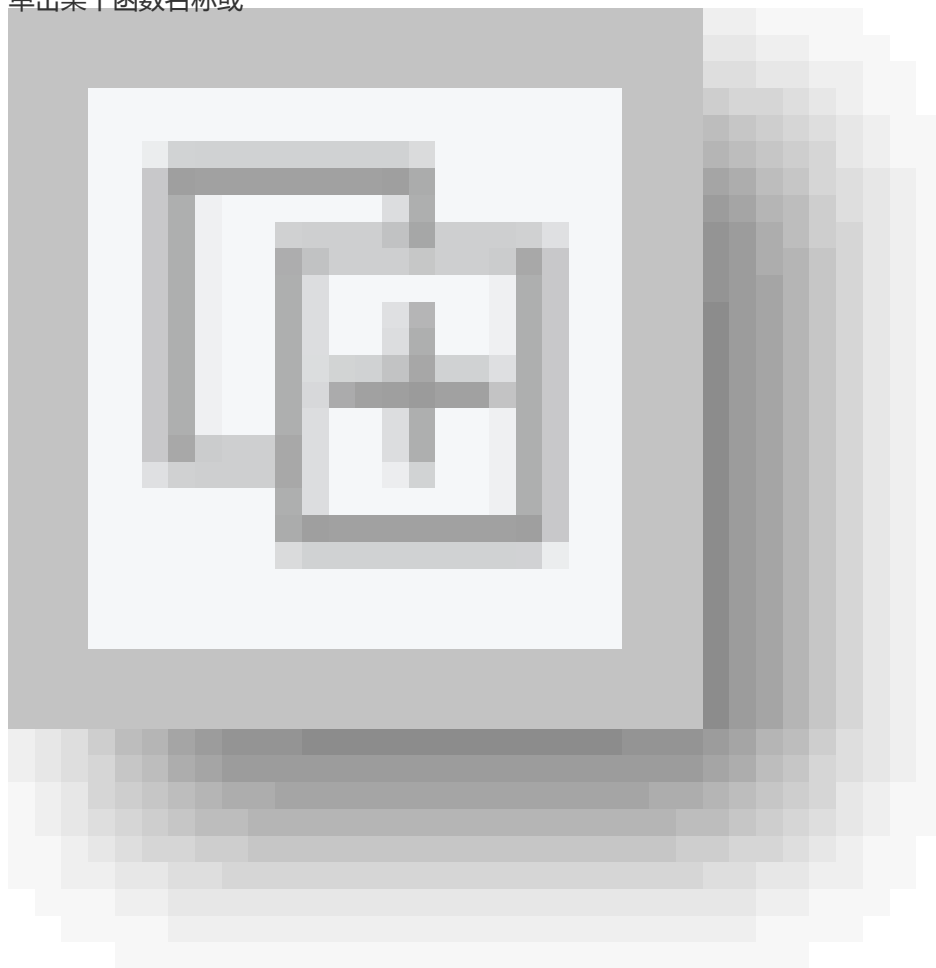
在 PTS 控制台**创建压测**页面，填写压测 API 名称和 URL。

选择**请求方法**（如 POST）和**请求参数类型**。

说明：在本示例中，将对 Body 参数化进行举例，请求方法选择 **POST** 或 **PUT**。在实际压测中，请根据您的具体情况进行选择。

单击**创建压测**页面右上角的**系统函数**，查看系统支持的函数列表。

单击某个函数名称或

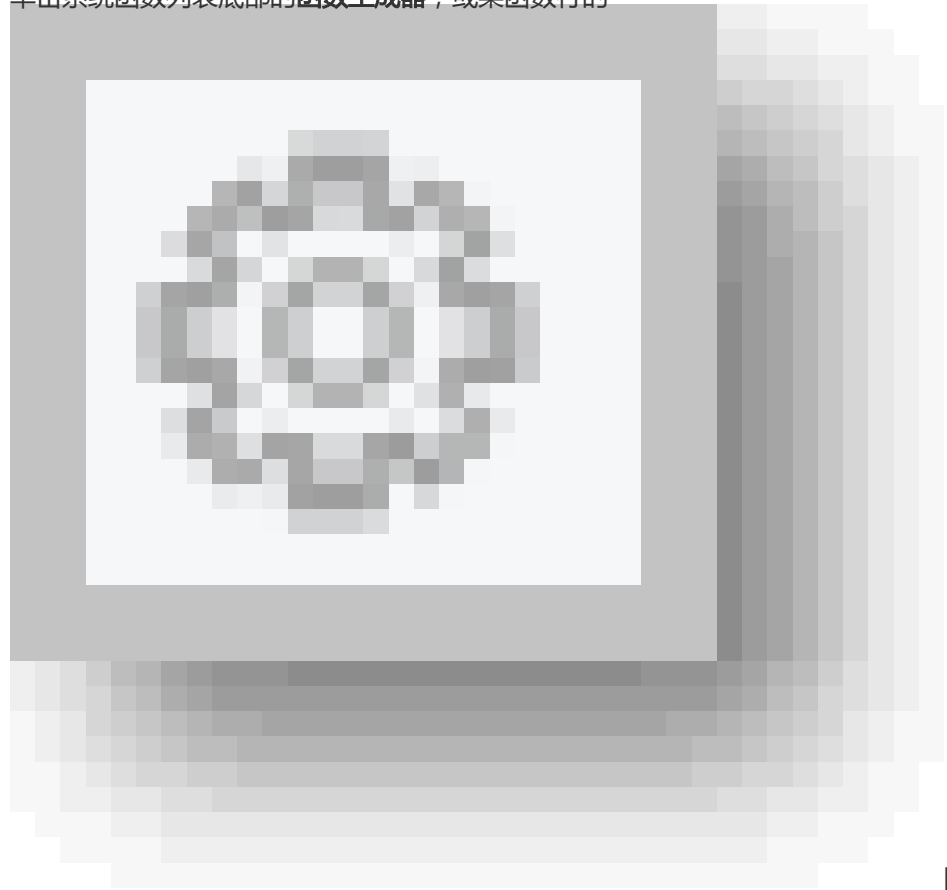


复制函数样例。

图标，系统将自动

（可选）函数生成器通过参数示例及说明，辅助您编写函数表达式。

单击系统函数列表底部的**函数生成器**，或某函数行的



图标

，打开**函数生成器**。



根据实际情况，编辑函数表达式，将示例值替换为实际压测请求使用的值。单击**调试**，可查看调试结果是否符合预期。

单击**复制并添加**，复制该函数表达式。

(可选) 勾选**添加为全局自定义参数**，并输入变量名。再单击**复制并添加**，复制函数的同时，此函数将添加到全局变量列表中，供整个压测场景调用。

在 API 的 Body 编辑框内，粘贴函数内容，并进行编辑，如组合字符串、参数或函数等。

关于四则运算函数的使用，参见四则运算的使用。

使用已创建的参数

下面以对 Body 进行参数化为例（URL 和 Header 类似）介绍如何使用您已经创建的参数，包括文件参数、自定义参数、出参参数。通过复制所需参数变量，粘贴至可编辑的文本区域，同时，可以按需进行组合、替换、嵌套或者追加需要的字符串等。

具体参数创建方法，参见：

- 创建文件参数

- 创建自定义参数
- 创建出参参数

使用参数的操作步骤如下：

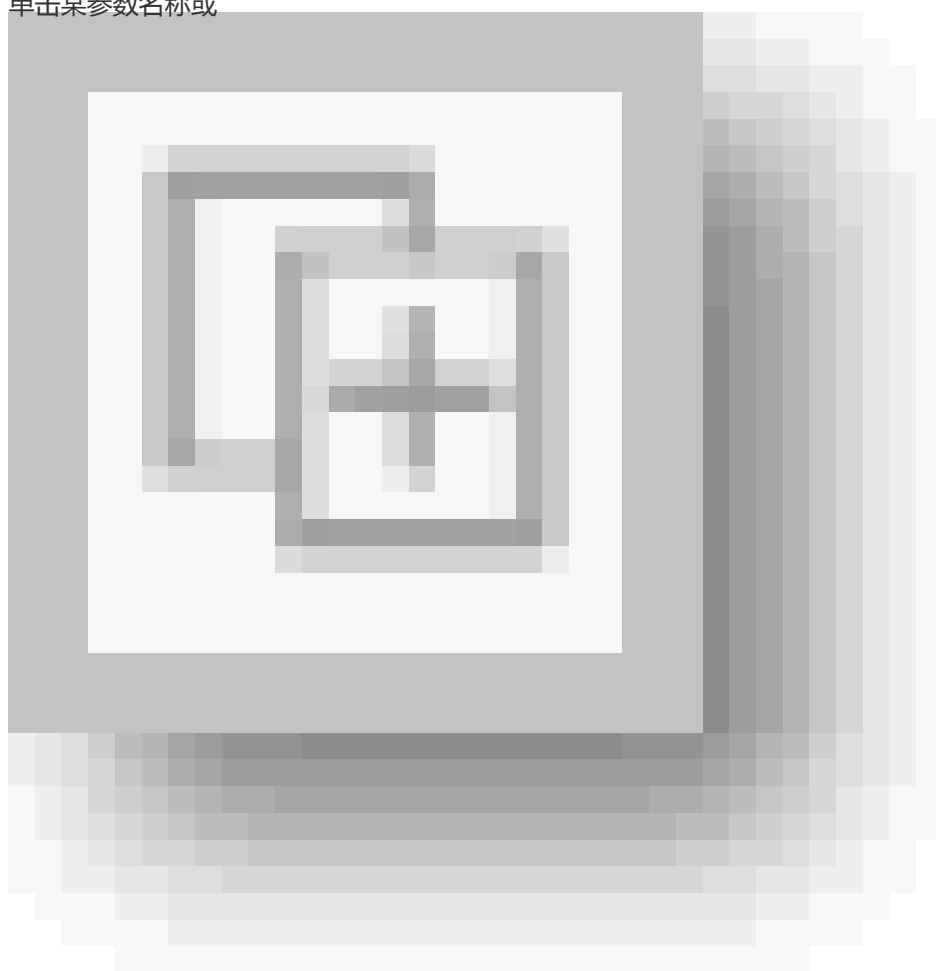
在 PTS 控制台**创建压测**页面，填写压测 API 名称和 URL。

选择**请求方法**（如 POST）和**请求参数类型**。

说明：在本示例中，将对 Body 参数化进行举例，请求方法选择 **POST** 或 **PUT**。在实际压测中，请根据您的具体情况进行选择。

单击**创建压测**页面右上角的**参数列表**，查看所有已创建的参数。

单击某参数名称或



复制参数内容。

图标，系统将自动

在 API 的 Body 编辑框内，粘贴参数内容，同时对 Body 内容进行编辑，如组合字符串、参数或函数等。



(可选) 当使用参数文件来组装压测请求参数时，勾选**数据轮询一次**，能够保证组装后的每条请求信息都不重复。



创建文件参数

在**创建压测**页面，单击右上角的**全局自定义变量**。

在**文件参数**定义部分，点击**+上传文件**来上传参数文件。

文件限制：

文件格式支持 .csv 及 .zip 格式，.zip 压缩包内的文件需是以 .csv 结尾的单个文件。

注意：请不要直接修改 **xlsx** 文件的后缀。推荐使用 EXCEL、Numbers 等软件导出，或使用 Apache Commons-csv 程序生成。

建议您在参数文件第一行不写列名，直接写参数数据即可。如果参数文件第一行是列名（不是参数数据），可以使用**首行作为参数名**功能，则导入的文件参数将不计算第一行。

可上传多个文件，单个文件大小限制在 60 MB，超过 60 MB 可以尝试压缩为 .zip 文件。

文件中单行不能超过 2 万个字符。

若需要多个数据关联组合为一组参数，请将这几个入参对应的几列数据放置在一个数据文件中。例如，下面示例中的两列分别表示 Username 和 Password：

```
john,123456
mike,234567
lilei,345678
```

注意：如果参数文件中有多列数据，请保持每一行中多个列的数据是完整的，否则会出现空字段，可能影响压测数据。

如果引用多个文件中的参数，根据行数少的文件组合后进行重复。具体说明，可参考并发模式和 RPS 模式下，是如何读取多文件参数的。

上传文件后，系统自动根据文件中的列生成多个变量。可自定义每个变量的**变量名**。

说明：如果参数文件第一行是列名（不是参数数据），可以使用**首行作为参数名**功能，则导入的文件参数将不计算第一行。



单击**确定**，完成创建。

创建自定义参数

在**创建压测**页面，单击右上角的**全局自定义变量**。

在**自定义参数**列表中，填写参数名和参数逻辑，单击**确定**。

说明：

- 自定义参数只能是字符串与系统函数的结合，不可使用文件参数和出参。
- 当自定义参数中使用了系统函数，且在串联链路中的两个 API 上使用，那么**同一个虚拟用**

户的一次串联链路（事务）操作里的两个 API 里的值是相同的。



创建出参参数（前置接口的返回截取）

在一些业务中，多个请求参数之间需要有前后传参关系。比如理财业务中根据用户的消费能力，推荐适当产品，这时就需要把消费能力等级从前置接口的返回中截取出来，并传到当前的产品推荐接口中。

前置 API 的出参提取有多种方式，本文以 Body : JSON 为例。其他出参提取方式，参见出参与断言。

API 的 Response 示例如下：

```
{
  "info": "success",
  "message": "处理成功",
  "data": {
    "id": 13509, "code": 0,
    "items": [
      {"name": "name1", "value": "1234"},
      {"name": "name4", "value": "7648"}
    ]
  }
}
```

期望获取出参的位置	解析表达式
获取 info 的值	info
获取 data 的 id 的值	data.id
获取 items 第一个对象的 value 的值	data.items[0].value

例如提取 items 第一个对象的 value 值，设置为出参变量为 “value_out”，则设置如下：



说明：

- PTS 仅提供数据生成能力，不对依据您的指令（包括数据生成方式）进行测试后产生的目标服务器或者业务异常承担任何责任。详情请见：阿里云产品服务协议(通用)。

登录（Cookie）的支持

在压测请求中，您可以构建用户登录的场景（即带 Cookie 的情况）。在一个串联链路（有先后顺序的事务）内或者全场景（N 个串联链路）内，可以共享登录返回的 Cookie。

串联链路内共享

若想在串联链路内传递（共享）Cookie 信息，则不需要任何额外设置。PTS 和浏览器类似，默认将获取到的 Cookie 在后续的请求中带上。

场景内共享

登录 Cookie 可以在同一个场景内进行共享。

同一场景内，只能指定将一个串联链路内的 Cookie 共享给其他串联链路使用。

示例

假设同一个压测场景内包含两个串联链路：

- 串联链路 1：登录->打开详情->添加购物车->下单
- 串联链路 2：查看收货地址->修改收货地址（PS：和串联链路1是同一批登录用户）

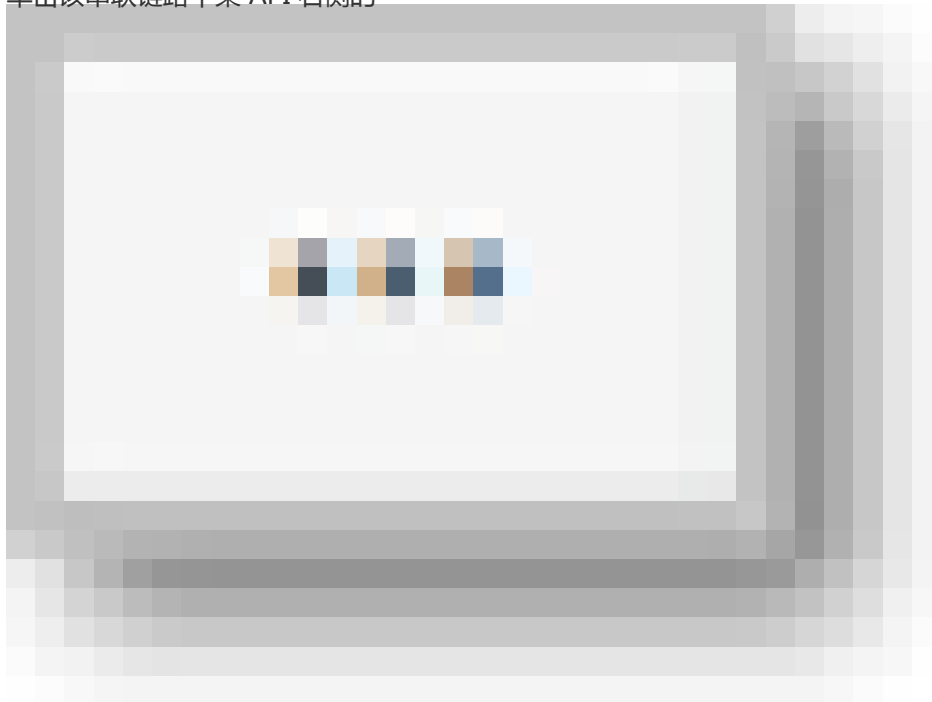
为了把串联链路 1 中的登录 Cookie 共享给串联链路 2，可以按以下步骤进行设置。

操作步骤：

在压测场景的**场景配置**页签下，新建**串联链路 1**，添加 API（如**登录**）模拟用户登录。

说明：您可以通过导入数据文件生成参数、通过 uuid 生成用户信息等多种方式设置 Cookie。

单击该串联链路中某 API 右侧的



图标，选择**添加数**

据导出。

选择数据为 **Cookie**，设置**导出量级**。该设置会为串联链路 2 提供相应的 Cookie 量级。

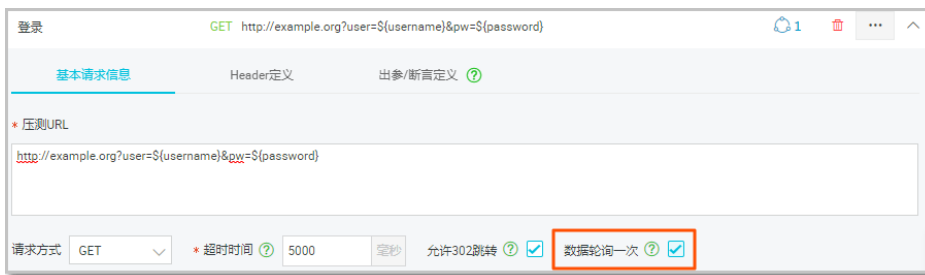


新建**串联链路 2** 并添加 API。串联链路 2 将自动带上串联链路 1 中设置的 Cookie 信息。

说明：该场景内创建的其他串联链路都会自动带上 Cookie 信息。

导出量级说明

若要保持共享Cookie的唯一性（即准备的 Cookie 信息不重复），设置的导出量级不可超过对应串联链路中登录 API 的文件行数（一般将登录信息放在数据文件中），且将登录 API 设置为**数据轮询一次**。否则输出结果可能会重复。



若未指定对应串联链路的 API 文件数据只轮询一次，则共享 Cookie 的数量可能会重复（当导出量级 > 文件行数时），系统默认不可超过 100 万。

若指定对应串联链路的 API 文件数据只轮询一次，则共享 Cookie 的数量 $\leq \text{Min}(\text{串联链路指定的唯一性文件中最小文件行数、填写的 Cookie 导出量级})$ 。

示例

串联链路 1 有 API1 和 API2，分别使用到文件的行数为 10 和 50；填写 Cookie 导出量级为 100，则：

均未指定两个 API 的文件数据只轮询一次，则最终 Cookie 量级为 100（会有重复数据）；

指定了 API2 (50行) 的文件数据只轮询一次，则最终 Cookie 量级为 50（因串联链路只执行 50 次结束）；

- 指定了 API1 (10 行) 或同时制定了两个 API 的文件数据只轮询一次，则永远不能准备到 100 量级，必须填写 ≤ 10 （因串联链路只执行 10 次结束）；

如何使用指令

参数化可以丰富压测请求，如果需要进一步模拟业务压测流量，可以使用指令功能。指令是一种可以改变、控制串联链路中压测行为和流程的功能组件。

指令包括思考时间、集合点、条件跳转、数据指令和数据导出。在场景配置页签，单击 API 右侧的



令。

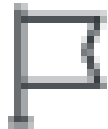
图标即可选择需要配置的指



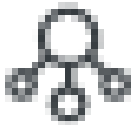
添加压测API



添加思考时间



添加集合点



添加条件跳转



添加数据指令



添加数据导出



复制API



粘貼API

思考时间

功能：模拟用户在前后两个节点间思考、反应花费的时间，支持多种模式。

使用限制：无

固定时间模式

经过的请求固定等待 N 秒之后继续，其中 N 为输入的**停顿时长**，单位为秒。

思考时间 固定停顿 10秒

定时器类型: * 固定时间 均匀分布 正态分布

停顿时长: * 10 秒

均匀分布

经过的请求固定等待 N 秒之后继续，其中 $N = [\text{延时基准}, \text{延时基准} + \text{可变跨度}]$ 区间内的任意随机值。例如延迟基准=3，可变的跨度=2，那么延时时间范围=[3,5]，请求会随机等待 3、4、5秒。

思考时间 停顿范围均匀分布在 3秒 与 5秒 之间

定时器类型: * 固定时间 均匀分布 正态分布

停顿时长: * 延迟基准 3 秒 可变的跨度 2 秒 延时时间范围=[延时基准, 延时基准+可变跨度]

正态分布

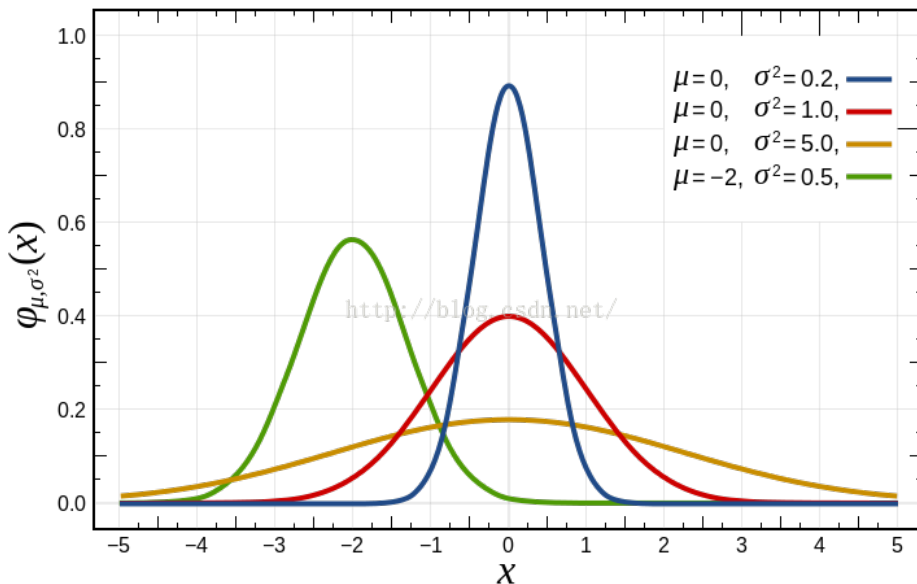
经过的请求固定等待 N 秒之后继续，其中 N 满足正态分布，即高斯分布。其中均值 (μ) 表示分布的中心值，方差表示距离均值的分布密度。方差越小密度越高、越集中；方差越大密度越低，越分散。

思考时间 停顿范围呈正态分布，均值为 10秒，波动方差为 2

定时器类型: * 固定时间 均匀分布 正态分布

停顿时长: * 均值 (u) 10 秒 方差 (q2) 2 秒

如下图均值 (μ) 和标准差的设定：



集合点

功能：使虚拟用户在集合点处等待，满足条件后一次性释放所有等待的用户，继续后续业务，例如整点秒杀场景。支持根据时间集合或者根据用户量集合两种工作模式。

使用限制：使用上无限制。需要注意，虽然集合点一次性释放所有用户，但释放的速率仍受后续节点的并发或者 RPS 限制。

根据时间集合

从压测启动开始计时，堆积用户请求，直到**等待时间**结束，一次性释放所有等待的用户。等待时间单位为秒，例如等待时间=600，那么集合点在 10 分钟之后释放。

集合点只能生效一次。

集合点 + 删除 ^

从压测启动持续 300秒 之后释放集合点，集合点仅能生效一次

定时器类型: * 根据时间集合 根据用户量集合

等待时间: * 秒

根据用户量集合

持续堆积用户等待，直到积压的用户数达到**用户量**时，一次性释放所有等待的用户，并且只生效一次。即实现上一个节点完成的请求数的累加。

集合点 + 删除 ^

当累计大于等于 10000个 用户之后释放集合点，集合点仅能生效一次

定时器类型: * 根据时间集合 根据用户量集合

用户量: * 人

条件跳转

功能：根据链路的出参判断改变串联链路请求的执行路径，支持**跳转**、**继续**和**结束**等。支持多个条件叠加（“且”的关系），按照顺序从前到后依次匹配。如果都没有匹配到，则执行**默认 Action**。您也可以在一个跳转条件的基础上继续增加条件跳转（条件跳转之间是“或”的关系）。

- 跳转：跳转到所在串联链路之前的其他链路；
- 继续：继续执行下一个节点，如果没有下一个节点则结束；
- 结束：中止当前用户执行流程。

示例：

假设串联链路为：用户登录>充值>加购>付费。在加购环节设置条件跳转，判断如果是实名认证用户但未充值，则跳回到充值环节；如果已充值，则继续；如果账户未充值且非实名认证，则流程结束。

使用限制：条件跳转不能位于串联链路的第一个。条件判断只对当前用户的当前请求有效，每个请求之间互不影响。



数据指令

功能：在串联链路中，可以添加一个或多个数据指令，对前置压测 API 的出参做二次处理，如拼接字符串或使用加密函数再处理，然后提供给该串联链路中的后续压测 API 使用。

同时，数据指令也支持对**数据导出**的参数做二次加工。

使用限制：

- 数据指令不支持二次加工全局参数，包括全局文件参数和自定义参数。
- 若对出参参数做二次加工，需要前置 API 已定义出参。关于如何定义出参，参见出参与断言。

示例：下文以前置 API 出参的二次加工为例，说明数据指令的使用方法。

某串联链路内包含请求：浏览产品>加购>付费。加购和付费请求（后续压测 API）都需要使用浏览产品（前置压测 API）的产品 ID，并截取产品 ID 的前五位，作为参数使用。

在 API “浏览产品” 出参/断言定义页签下，配置出参（如 ProductID）。



单击 API 右侧的
图标 > 添加数据指令，输入参数名。

编辑参数逻辑表达式。

该示例中，要实现“截取产品 ID 的前五位，作为参数使用”，您可以使用函数逻辑表达式：`${sys.substring("${ProductID}", 0, 4)}`。其中“`${ProductID}`”为上一步串联链路中定义出的参。



数据导出

功能：在压测场景中，可以导出某个串联链路中的数据（如 Cookie、出参或通过数据指令定义的参数），供其他串联链路使用，做到导出数据在全局共享。

示例：下文以导出 API 的出参为例，说明数据导出的使用方法。

串联链路 A 包含：打开登录页面->账号登录。

串联链路 B 包含：查看商品详情->加购->提交订单

串联链路 C 包含：查看收藏夹->加购->提交订单

其中 B 和 C 中使用到的登录信息来源于 A 登录后的 Body 里的 Token 信息，所以需要将 A 中 Token 做数据导出，供另外两个串联链路中 API 使用。



使用限制：

- 一个场景仅可以有一条数据导出指令；
- 一次只能导出一个数据，Cookie 或者某一个参数（出参或者数据指令定义的参数）；
- 导出数据量级不做限制，总大小不超过 100w 条；
- 导出 Cookie 后，其他串联链路会自动带上 Cookie 信息，无需手动设置；导出参数后，使用时需要按照参数使用方法进行设置（如\${token}）。

效率相关功能

PTS 提供了多种提升场景编排效率的功能，包括场景复制、串联链路的生效/失效、API 的复制。后续也会陆续开放更多的提效功能，敬请期待。

使用限制

目前这些功能只对购买了 898 及以上资源包的用户开放。

复制场景

在多个压测场景中部分内容一致或者类似的情况下，可以复制一个场景，并微调内容作为一个新的场景。

复制步骤：

1. 在 PTS 控制台左侧菜单栏选择**压测场景**。

找到需要复制的场景，单击右侧的**复制**选项。复制之后的场景名默认会带 `_copy` 的字样，显示在场景列表最上方。

场景ID	场景名	持续时间	状态	最后修改时间	操作
	new_scene_for_test_copy	11分钟	草稿	2018-07-10 15:01:14	编辑 删除 复制
	new_scene_for_test	11分钟	待启动	2018-07-10 14:53:36	操作 编辑 删除 复制

注意：

为了便于阅读，场景名限定为 30 个字，若复制后的场景名超过 30 个字（即加上 `_copy` 之后），便不会带上 `_copy` 字样。

串联链路的失效/生效

一个场景中可能包含了多个串联链路，有时仅需要对部分串联链路进行调试、压测。此时可以用串联链路的失效/生效功能控制调试和压测粒度，提升场景灵活使用的效率。

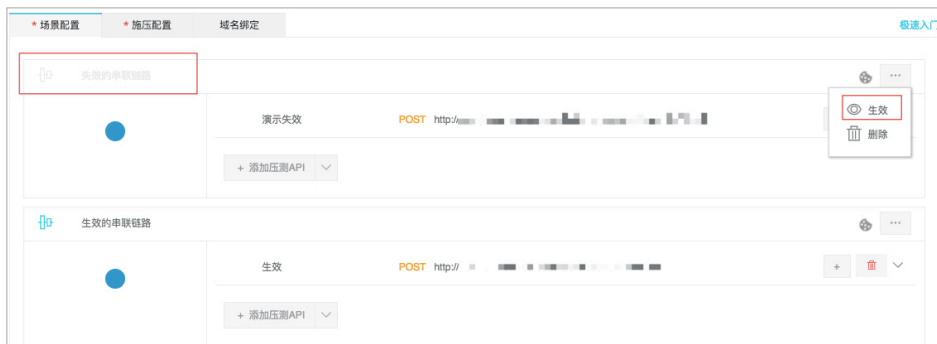
在原场景中，可将部分串联链路设置为失效，从而在调试、压测中略过该串联链路，但依旧保持已有的所有配置信息。

操作步骤：

在场景列表也，找到需要编辑的场景，单击右侧的**编辑**进入到场景编辑页面。

在串联链路右侧的...下拉列表中，选中**失效**或者**生效**即可。失效后的串联链路，名字和图标将变成灰

色。



复制 API

场景中的多个串联链路，有时会使用到相同的 API。例如，查看商品详情>加购>提交订单、查看购物车>提交订单、查看商品详情>提交订单这三个串联链路都有“提交订单”这个 API。此时可以直接复制 API，提升编辑效率和准确性。

操作步骤：

在**场景编辑**页面，单击**场景配置**页签。

单击需要复制的 API 右侧的加号“+”，单击**复制 API**。

在需要添加 API 的串联链路的场景编辑页面，单击前一个 API 右侧的加号“+”，在下拉列表中单击**粘贴 API**。这时复制的 API 会出现在该 API 下方。



您也可以在**+添加压测 API**的下拉列表中单击**粘贴 API**，将 API 粘贴到链路末尾。

绑定域名

域名绑定是指将域名与指定的 IP 地址关联。压测时，压测流量将直接访问绑定的 IP 地址，实现对目标设施的压测。

应用场景

以下情况，可能需要使用域名绑定功能：

公网测试环境中的压测

当压测业务请求需要通过公网访问，为了不影响线上业务，您可以使用测试环境的 IP 地址，与公网域名绑定，实现压测流量和线上流量的隔离。

PTS 的阿里云内网压测模式下的压测

PTS 阿里云内网压测模式，是指通过阿里云内网进行压测，目前支持经典网络 ECS 和 VPC 上挂内网 SLB 的方式，详见内网压测。为了便于压测时不用修改压测接口的域名，您可以在场景中为任意一个域名绑定阿里云内网 IP 地址，实现在阿里云内网压测的目的。

操作步骤

在创建压测或编辑场景界面，单击**域名绑定**页签。系统已自动识别出场景配置中所有的域名。

勾选**启动**，填写相应的 IP 地址。多 IP 地址间用英文逗号 (,) 分隔。



单击右上角**保存配置**，即可完成域名绑定。

绑定域名后，压测时将直接压测您已配置的 IP 地址。

压测场景的结构和数据分配

本文主要通过图例介绍常见的压测场景内的逻辑结构关系，并说明数据在压测中是如何分配的。

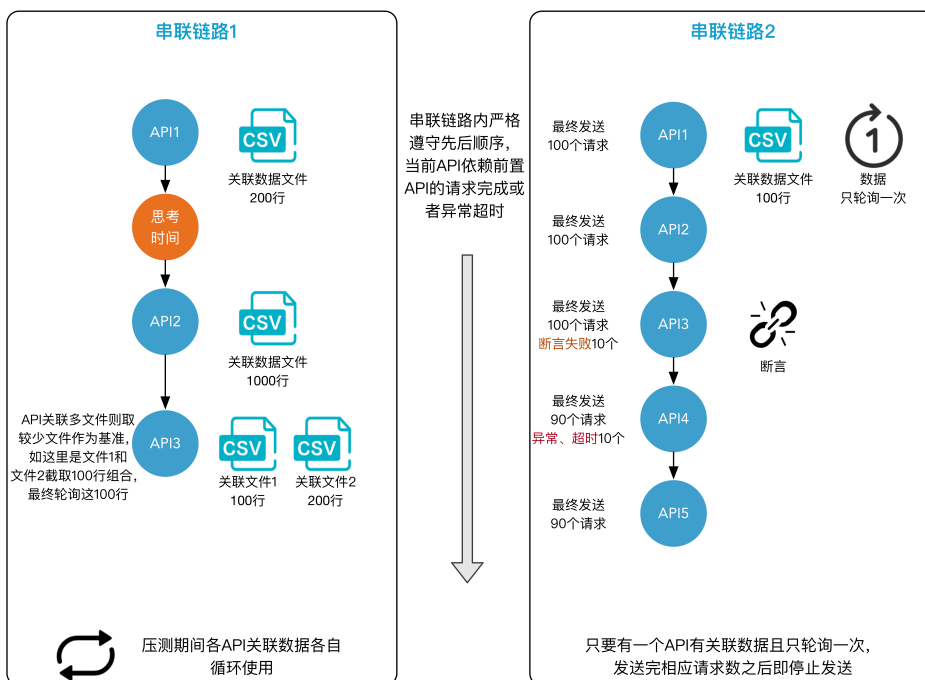
基本概念

- 压测 API：指由用户行为触发的一条端上请求，是压测中的必需元素。
- 串联链路：指一组压测 API 的有序集合（类似于事务），具有业务含义。
- 数据导出指令：用于导出某个串联链路中的数据（导出 Cookie 为典型应用），供其他串联链路使用，实现导出数据的全局共享。
- 关联数据文件：压测 API 使用了来自数据文件的参数，从而关联了相应的数据文件。如果使用了多个文件参数分别来自于不同数据文件则表示关联了多个数据文件。
- 断言：一般用于标记业务成功与否，从而验证压测请求的响应是否符合预期。
- 数据轮询一次：在请求中使用文件参数时，数据文件只轮询一次，以保证请求信息不重复。
- 出参：在创建串联链路时，将前置接口的部分返回信息作为参数。

常见的压测场景结构和数据分配规则

具体请参考下图：

压测场景中的串联链路之间是并行关系，各自独立。
只有使用数据导出指令的情况才会有有一定关联。



图例说明如下：

串联链路1 和 串联链路2 是并行关系。

串联链路类似于一个事务，串联链路内的所有压测 API，严格遵循先后顺序。当前 API 的请求发起依赖前置 API 的请求处理完成或者超时异常。

注意：后面的 API 最终发送的请求次数会少于前面 API 的，因为压测停止的时候总有一些请求还在串联链路的中间环节（API）上没有全部完成。

串联链路内，如果不同的压测 API 使用了不同行数的 CSV 文件里的参数，在实际压测时是各自轮询各自的数据文件的。

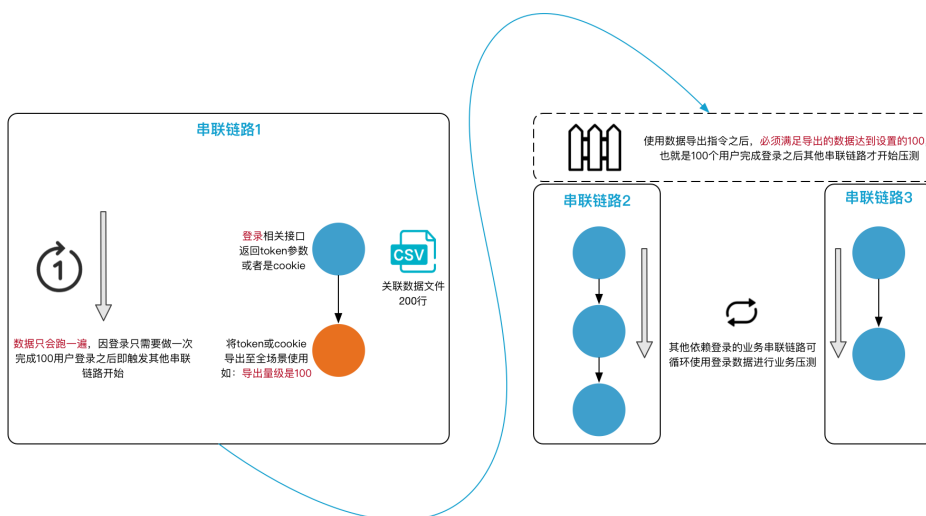
- 串联链路1 中 API1 使用了 200 行的数据文件，API2 使用了 1000 行的数据文件，API1 和 API2 压测时分别轮询各自的数据文件。
- 一般情况下，压测期间各压测 API 循环使用各自的关联文件数据。如果压测 API 勾选了**数据只轮询一次**，使用完由关联文件数据构造的请求数之后，将不再压测该 API，当前串联链路也将整体停止压测。

串联链路内，如果同一个压测 API 使用了不同数据文件里的参数，且各文件的行数不同，默认将以行数最少的文件作为基准文件，其他文件会被随机丢弃多余的行数。

设置了断言的情况下，如果当前请求断言失败，那么后置的 API 或者指令将不再执行，断言成功的没有影响。

依赖登录的压测场景结构

下图中串联链路1 的 API1 是登录业务相关接口，其典型配置如下：



图例说明如下：

- 数据导出指令一般应用于登录之后需要并行压测多个不同业务串联链路的情况，支持标准的 Cookie 导出或者是业务自定义的**出参**导出。

串联链路1 使用了数据导出指令，数据导出完成后其他串联链路才能开始压测，所以与其他串联链路不是并行的关系。

说明：只有使用了数据导出指令，才会出现串联链路之间不是全都并行的情况。

为保证用户登录信息不重复，需设置压测 API **数据只轮询一次**。串联链路1 中 API1 设置了数据只轮询一次。

一批用户登录完成后，将用户登录信息共享给场景内其他业务的串联链路使用，需设置数据导出的**准备量级**。达到该量级才会触发场景内剩余串联链路进行压测（它们彼此间还是并行的，数据分配逻辑和上面所述一致）。

准备量级需要小于等于登录接口的文件行数。如图所示，关联数据文件为 200 行，导出量级设置为 100。

注意：由于目前任务是拆分给施压agent单独执行的，故可能出现准备量级并没有达到设定值便有其他串联链路开始压测的情况。

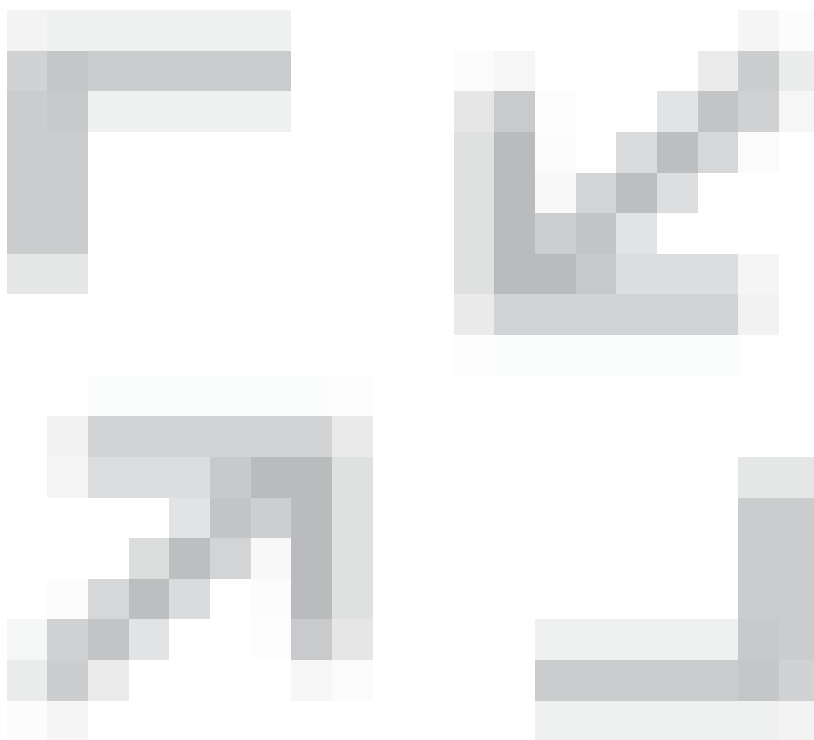
场景调试

编写好场景之后，可以先对场景进行调试。

在场景编辑页，单击右上角的**调试场景**按钮，开启调试窗口。

一般调试时间在 10 秒至 30 秒。

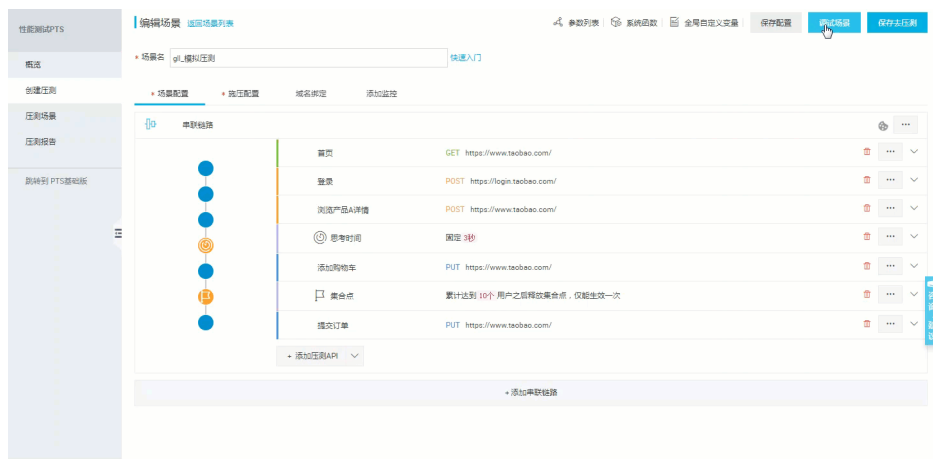
（可选）**最小化调试窗口**：单击右上角的



最小化窗口图标

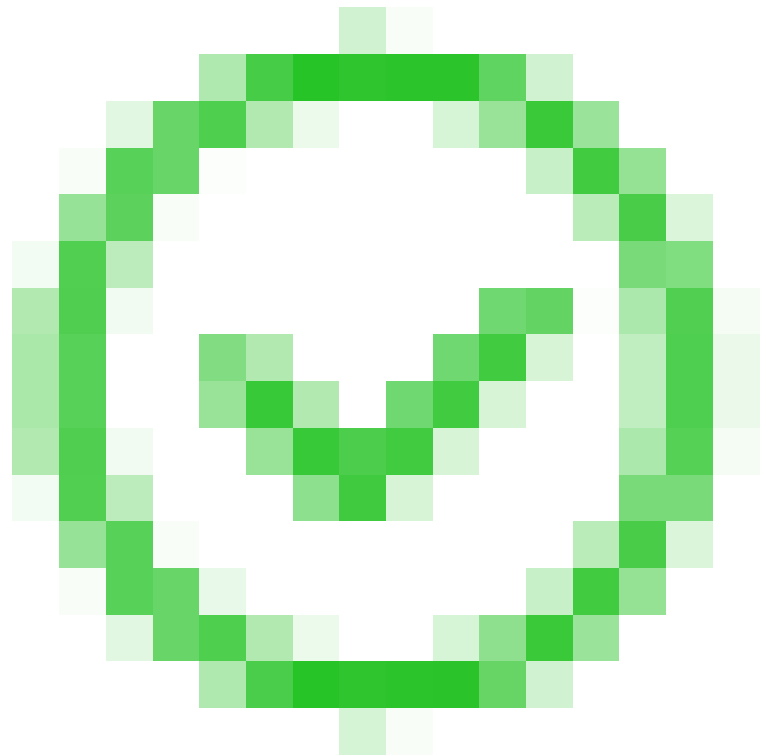
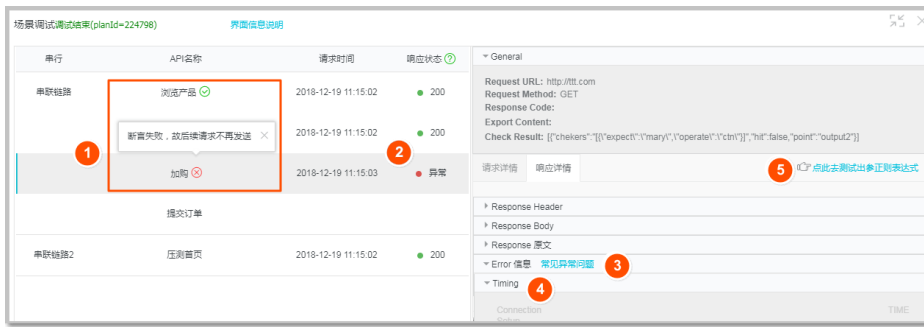
，在右下角以悬浮窗口随时关注调试状态。再次单击悬浮窗口，可切换回场景调试页面。

调试结束后，可以查看所有链路的请求日志。单击单条 API，可以查看这条 API 的调试详情，包括请求 URL、请求 Header、请求参数、断言情况（如果设置了断言）应答 Body、应答 Header 和应答的异常（如果存在）。

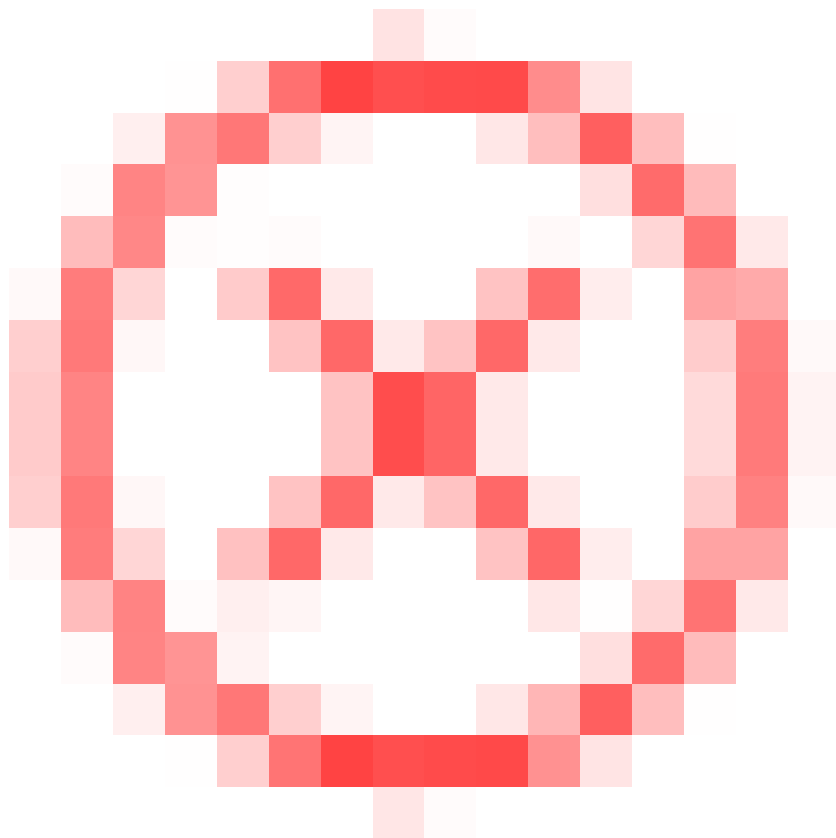


调试界面信息说明

如下图所示，API 名称和响应状态栏均有判断信息。以下是对图中带有数字标记字段的说明。



此处为断言判断，
表示调试的业务成功，



表示异常。选中该 API 后，即可在右侧查看响应详情。关于响应详情的含义，请参考“请求及响应详情解读”小节。

此处为响应状态，可能显示一个**状态码**（200、302、403、503、500 等）或**异常**字样。若显示**异常**，则表示未收到 Response，可能是因为请求超时，或因为请求的 URL（含 Body）有错误，例如函数无法识别、URL 不规范等。

此处为响应的错误信息，可显示异常响应详情，例如超时、参数错误、连接拒绝等。

此处为请求在核心的生命周期内的耗时详情，具体请参考查看采样日志中关于 Timing 信息的说明。

此处为出参表达式测试入口。具体操作，请参考“测试出参正则表达式”小节。

请求及响应详情解读

每条调试的 API 均可查看到请求与响应详情。针对上述异常现象，可以从以下几个方面排查：

若响应状态为**异常**，请在右侧**响应详情**标签页的 **Error 信息**区域查看异常信息。例如上图中的异常原因因为非法参数（系统函数使用不合法）；

若响应状态为除 **200** 以外的状态码，请查看具体返回信息，并结合服务器上的日志进行常规排查；

若断言出现异常，请在右侧**响应详情**标签页的**断言信息**区域查看断言信息。若当前值为空，则表示未获取到出参值。

常见 ERROR 信息解答

详情请移步常见问题。

测试出参正则表达式

当您需要基于请求的响应详情提取特定字符串时（一般是文本类型的响应体），可在调试窗口中测试正则表达式(由于application/json 和 text/json两种类型的JSON格式解析比较简单，暂不提供相关功能)是否正确，并直接设置出参。

1. 在调试窗口的右侧，单击[点此去测试出参正则表达式](#)。

在**测试出参正则**页签中，选择来源格式，填写正则表达式，指定为第几个匹配项。单击**测试表达式**。

通过下方的响应详情匹配结果，预判提取的内容是否符合预期。



单击**同步出参配置**，将此正则表达式同步到该 API 的出参列表中。

注意：场景调试结束后，如有同步的出参，需返回压测**场景配置**页，在对应的 API **出参/断言**定义页签下，设置出参名。

单击弹窗左侧的返回箭头，或左上角的[场景调试](#)链接，返回查看其它 API 调试结果。

施压配置

如何进行施压配置

每个 API 可以视为业务系统的一个节点，处理能力不同导致可承载的业务量也不一致。在**施压配置**中，您可以指定不同的压测模式、压测量级，并考虑压测数据是否允许重复等因素，对压测个性化设置，检验各业务节点的性能表现。

本文将介绍以下内容：

配置压测模式

- 是否智能压测
- 压力来源
- 压测模式
- 是否自动递增

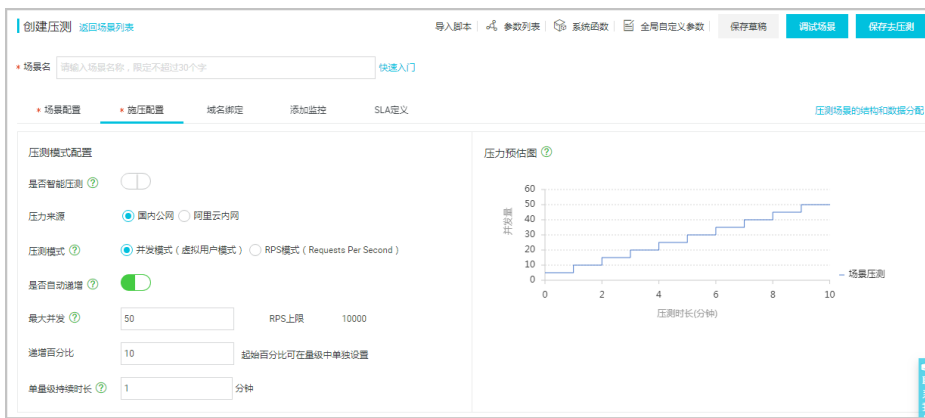
配置量级及数据

- 设置压测数值
- 来源 IP 扩展
- 指定压力来源

配置过程中，您可以通过查看压力预估图，预判测试压力分布。压力预估图的解读，参见压力预估图。

配置压测模式

压测模式配置区域如下图所示。



是否智能压测

智能压测可提供指标预估能力，判定实际值是否属于预估正常范围，帮助定位性能异常拐点。如何使用智能压测请参见创建智能压测。

压力来源

国内公网（默认模式）：发起压测的节点来自全国各地，根据压测的量级来分配对应的节点资源（IP）。

阿里云内网：使用阿里云经典网络或阿里云 VPC 网络来进行压测。若需使用阿里云内网压测，请参见阿里云内网压测。

选择**经典网络**时，压力来源来自该地域的阿里云经典网络施压机。

选择**VPC 网络**时，您需要继续选择地域、VPC-ID、安全组、虚拟交换机等信息，压力来源来自您选择的阿里云 VPC 网络施压机。

说明：

阿里云内网压测对持有 278 元及以上可用资源包的用户开放。

阿里云内网压测发起节点支持的地域以界面为准。如有其他地域的支持需求，您可以提交工单。

压测模式

压测模式有两种：并发模式（虚拟用户模式）、RPS 模式（Requests Per Second，每秒请求数，吞吐量模式）。

并发模式：

“并发”是指虚拟并发用户数，从业务角度，也可以理解为同时在线的用户数。

适用场景：如果需要从客户端的角度出发，摸底业务系统各节点能同时承载的在线用户数，可以使用该模式设置目标并发。

并发模式下，需要指定全场景的最大并发数，再设置各串联链路的并发权重。

串联链路内各 API 的响应速度不同（表现为响应时间不同），所以单位时间内 API 的并发数也会不同。API 响应速度越快，单位时间内累积在 API 上的并发用户数越少。

例如，共 100 个虚拟用户需要操作某个事务（即串联链路）。假设，该串联链路中共有 2 个 API，API 1 响应速度快而 API 2 响应速度慢。则更多的虚拟用户将等待在 API 2 上，API 2 需要更多的线程资源来处理更多的虚拟用户请求。

RPS 模式：

RPS (Requests Per Second) 是指每秒请求数。

适用场景：RPS 模式即“吞吐量模式”，通过设置**每秒发出的请求数**，从服务端的角度出发，直接衡量系统的吞吐能力，免去并发到 RPS 的繁琐转化，一步到位。

API 接口（如电商加购物车、下单等）主要用 TPS (Transaction Per Second, 每秒事务数) 来衡量**系统的吞吐能力**，选择该模式可以直接按照预期的 TPS 设置 RPS。如果希望检验“下单”接口是否能达到 500 TPS 的预期，那么设置 RPS 为 500，每秒发送 500 个请求，可检验系统的吞吐能力。

该模式下，请求无法及时响应时可能会导致较高的并发，异常情况请及时停止。

该模式仅支持非自动递增进行压测（即需在压测过程中手工调速）。

是否自动递增

自动递增（仅适用于并发模式）

按照固定比例进行压测量级的递增，并在每个量级维持固定压测时长，以便观察业务系统运行情况。

开启**是否自动递增**开关，并设置**最大并发**、**递增量级**（如 10%）和**单量级持续时长**（如 1分钟）。

说明：进行压测时，您可以随时对场景或串联链路进行手动调速，自动递增模式自动失效。

非自动递增

非自动递增即手动调速模式。不开启**是否自动递增**开关，设置**压测总时长**及**并发模式下的最大并发**。

配置量级及数据

设置好压测模式后，需要设置压测起始量级与最大量级，并支持使用来源 IP 扩展。

串联链路	最大并发权重	最大并发	起始百分比	起始并发
串联链路	3	15	10 %	1
串联链路2	7	35	10 %	3

压测数值

每个 API 可以视为业务系统的一个节点，处理能力不同导致可承载的业务量也不一致。并发模式与 RPS 模式施压的方式不同，故压测数值上的设置也会不同。

说明：无论选取何种压测模式，各场景最大值的总和不可超过该账户下对应资源包的最大 VU/RPS。

并发模式

在并发模式下，设置串联链路级别的**最大并发权重**和**起始百分比**。

最大并发权重：该串联链路占全场景的权重。如串联链路 1 置为 1，串联链路 2 置为 3，则表示各自在场景最大并发总占 25% 和 75%。

起始百分比：该串联链路自身的起始比例，默认 10%。

并发模式摸底的是系统承载并发用户数（从客户端业务考虑），而场景中的多个串联链路（事务）之间一般有一定的业务比例和实际吞吐关系（如下单和加购按照业务评估是 1 : 0.8），故只需在串联链路（事务）维度，依据比例关系，设置最大并发和起始并发。

RPS 模式

在 RPS 模式下，摸底的是单 API（接口）的服务端吞吐能力，故需要给每个 API 设置**最大 RPS**与**起始 RPS**。

批量设置

并发模式和 RPS 模式均可以使用批量设置。

勾选需要设置的所有串联链路或 API。

单击左下角**批量设置**，在弹框中输入数值。

3. 单击**确认**，完成设置。

来源 IP 扩展

仅在**国内公网**压测模式下使用有效。

来源 IP 是指发起压测流量的 IP 地址。来源 IP 扩展功能允许您根据并发量或 RPS，在一定范围内（该范围以控制台页面为准）增加施压 IP 地址数，减少因以下原因引起的压测异常。

说明：在以下情况，请优先按照建议处理方案执行；**来源 IP 扩展**只能帮助您减少问题发生的概率。

SLB 限制问题：首先请结合购买的 SLB 产品计费类型，关注规格限制和带宽限制，是否已经到达上限。另外，当 SLB 是服务最外层，且接口是 HTTPS 或者开启了 7 层会话保持功能，压测出现一些 503 而且后端并没有相关流量和日志，可能出现了 SLB 单 IP 地址限流问题。

高防/WAF 的问题：未关闭高防/WAF 可能触发流量拦截。建议您基于业务评估，压测时临时关闭高防/WAF。如果不想关闭，可以考虑增加PTS的UA放行规则来避免压测流量被拦截的情况。

使用 CDN 或全站加速：建议您接近或者超出已有业务峰值时，提前提交工单报备给这两个产品。

指定压力来源

在压力来源是公网的情况下，PTS 性能测试的压测流量是分布式的，从全国大小城市和运营商随机调度。通过流量定制功能，您可以指定压测流量发起的地域或运营商，并配置其占比。

说明：若您在压测时需要在一段时间内独占压测机器，可以使用独立资源池功能。详情请参见独立资源池。

具体操作步骤如下：

勾选**流量定制**。

单击**详细配置**。

在**地域定制**对话框中，配置以下项：

- **维度选择**：目前支持根据地区或运营商指定压力来源。其他功能将陆续开放。

IP数分配：可添加多条分配规则，指定具体的地域或运营商及其 IP 数占比。操作步骤如下

:

- a. 单击+**新地域/运营商分配**。
- b. 选择地域或运营商，并输入其 IP 数占比。

启动策略

- 优先匹配资源：只有匹配到指定资源时，才会启动压测。
- 优先启动压测：立即启动调度资源，指定资源不足时会随机分配资源，满足压测需求。

阿里云内网压测

PTS 的阿里云内网压测模式是指施压机与被压测服务器在相同地域，通过阿里云经典网络或阿里云 VPC 网络进行压测。具体操作流程如下：

限制条件

阿里云内网压测目前存在以下限制：

资源限制：

- 持有**278 及以上可用资源包**的用户可使用阿里云内网压测功能。

VPC网络压测目前面向于**1058及以上可用资源包**。

暂不支持压测量级超过 1 万 TPS的阿里云内网压测场景，所以并发模式中最大并发值不需要超过 1250，RPS模式下最大设置不需要超过1 万RPS。

IP 限制：

阿里云内网压测的场景内所有 API 必须最终都是阿里云内网地址，比如直接使用阿里云内网 IP 地址，或者给域名绑定了阿里云内网地址，否则无法正常压测。关于域名绑定操作，参见[域名绑定](#)。

阿里云内网压测 API 必须为经典网络。支持经典网络 ECS 或 SLB。如果需要压测 VPC 网络内的服务器，可以挂载在经典网络 SLB 上，具体请参看最后一部分的经典网络 SLB 内网压测。

同一场景最多支持一个地域。压测场景中所有压测 API 必须为一个地域内的内网服务，不能跨地域。

其他：

- 用户需要手动授权允许内网访问。使用前请了解并熟悉阿里云 ECS 安全组、负载均衡 SLB 产品的相关操作。

使用经典网络压测单机服务

当被压测的 IP 地址为经典网络中的单台 ECS 服务器时，需要先创建安全组，然后将 ECS 服务器加入到这个安全组中。才能使用阿里云经典网络进行压测。具体操作步骤如下：

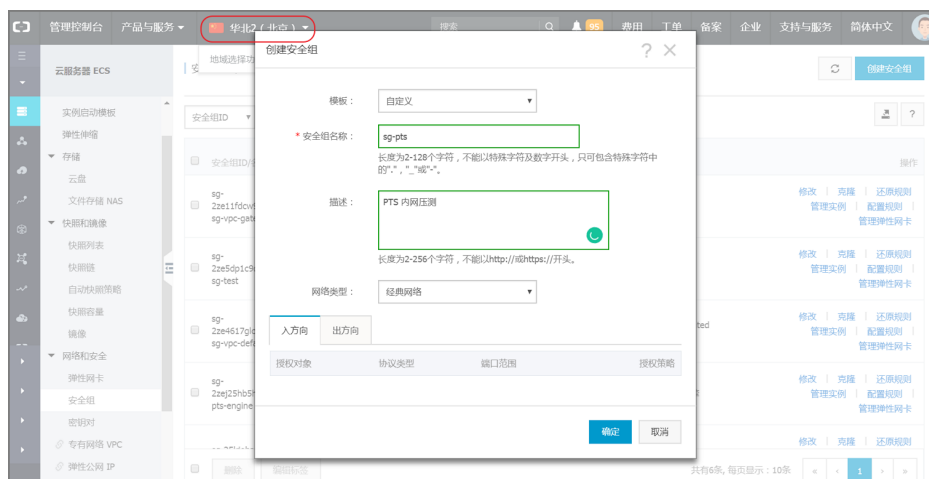
步骤一：创建安全组并添加 ECS

创建安全组

在 ECS 控制台，在左侧菜单栏选择**网络与安全**>**安全组**。

切换到阿里云内网压测**地域**，并单击**创建安全组**。

在**创建安全组**对话框，新建一个名为 **sg-pts** 的**经典网络**安全组。



找到新建的 **sg-pts** 安全组，点击 **配置规则**。



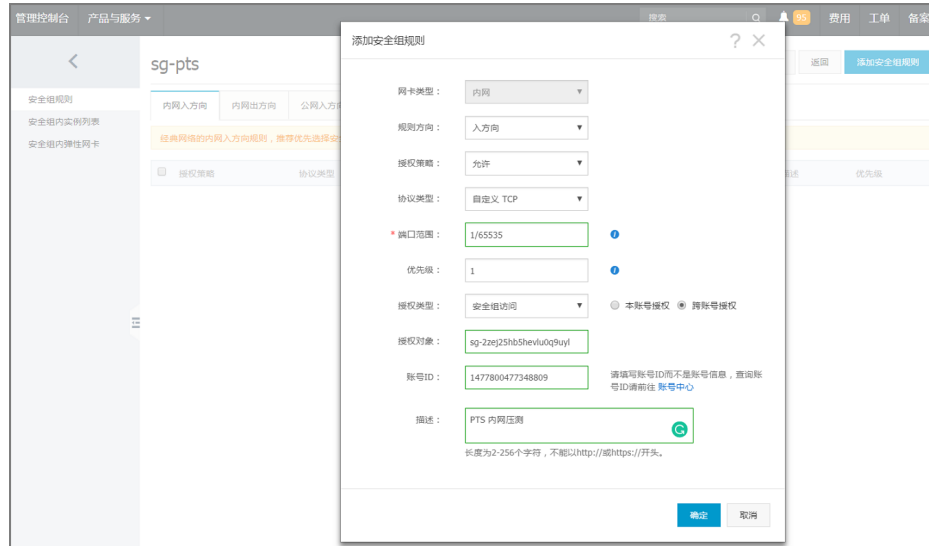
选择**内网入方向**页签，单击页面右上角的**添加安全组规则**。

- i. 端口范围：**1/65535**
- i. 授权类型：**安全组访问**，选择**跨账号授权**

ii. 授权对象：按地域配置

- 华北2（北京）：sg-2zej25hb5hevlu0q9uy1
- 华东1（杭州）：sg-bp1cusv1f4owdg4l78u7
- 华东2（上海）：sg-uf61ym290hfdtkswvkuz
- 华北1（青岛）：sg-m5e7xebvqkr1g6bj5w1
- 华南1（深圳）：sg-wz9fi3ce8nd9adqohurc

iii. 账号ID：1477800477348809



单击**确定**完成配置。

添加 ECS 服务器

在 ECS 控制台实例列表页，选择内网压测所在地域。

找到并勾选需要阿里云内网压测的经典网络 ECS 服务器，确保 ECS 内网地址为压测 API 地址，然后单击**更多**选项，依次选择**网络和安全组**>**加入安全组**。



选择已创建的名为 **sg-pts** 的安全组，单击**确定**。



创建成功后，在对应的 ECS 服务服务器右侧操作栏单击**更多**>**网络和安全组**>**安全组配置**，确认**安全组列表** 包含用于打通 PTS 内网压测的名为 **sg-pts** 的安全组。

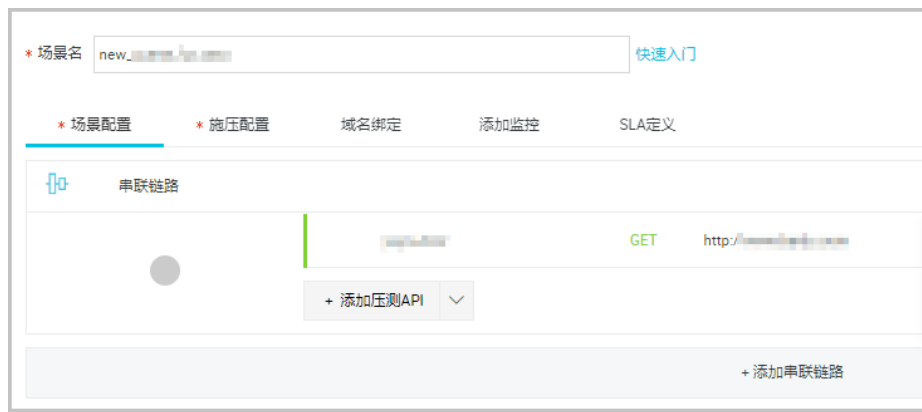


步骤二：启用阿里云经典网络压测

登录 PTS 控制台，在概览页面单击**快速压测**，或在左侧导航栏，单击**创建压测** > **快速压测**。

单击**场景配置**页签，然后完成场景配置，请参见创建压测场景。

注意：确保场景中所有 API 的压测 URL 为同一个地域经典网络下的 ECS 内网地址 或内网 SLB 地址。其 IP 地址通常为 10.* 或 100.*。或者为压测域名绑定符合上述要求的 IP 地址。关于域名绑定操作，请参见域名绑定。



单击**施压配置**页签，然后在**压测模式配置**区域将**压测来源**选择**阿里云内网**，**网络类型**为**经典网络**。然后在地域下拉列表中，选择施压机**的地域**。

所选地域必须与被压测的 ECS 服务器的地域相同。目前发起节点地域支持**华北2（北京）**，**华东1（杭州）**，**华东2（上海）**，**华北1（青岛）**，和**华南1（深圳）**。如果目前资源有限可能会出现资源不足导致的无法压测，请稍作等待。

根据需要配置其他参数，请参见**如何进行施压配置**。

单击页面右上角**保存去压测**，启动压测。

使用 VPC 网络压测单机服务

登录 PTS 控制台，在概览页面单击**快速压测**，或在左侧导航栏，单击**创建压测 > 快速压测**。

单击**场景配置**页签，然后完成场景配置，请参见**创建压测场景**。

注意：确保场景中所有 API 的压测 URL 为同一个地域 **专有网络** 下的 **ECS 内网地址** 或 **内网 SLB 地址**。其 IP 地址通常为 **10.*** 或 **100.***。或者为压测域名绑定符合上述要求的 IP 地址。关于域名绑定操作，请参见**域名绑定**。



单击**施压配置**页签，在**压测模式配置**区域将**压测来源**选择**阿里云内网**，**网络类型**为 VPC 网络，然后单击右侧的**VPC 授权**。



在**云资源访问授权**页面选中 **AliyunPTSAccessingENIRole** 后单击**同意授权**。



完成授权后，在**施压配置**页签中选择**施压机**的**地域**、**VPC-ID**、**安全组**和**虚拟交换机**等信息。然后根据需要配置其他参数，请参见**如何进行施压配置**。

注意：确保您的机器中空闲 IP 大于等于施压 IP 数。

VPC 信息栏下方会显示当前地域下的可用区，若您的交换机不在当前地域的可用区内，请单击**创建交换机**去创建交换机。

单击页面右上角**保存去压测**，启动压测。

使用阿里云内网压测多机服务

若需对多台 ECS 服务器进行阿里云内网压测时，可先根据实际创建负载均衡实例，然后按网络类型启用阿里云内网压测。具体操作步骤如下：

步骤一：创建负载均衡实例

登录负载均衡管理控制台，单击**创建负载均衡**，在**网络与实例类型**模块将**实例类型**选择**私网**，根据实际需要选择**网络类型**。

说明：若需使用经典网络压测，则**网络类型**选择**经典网络**；若需使用 VPC 网络压测，则**网络类型**选择**专有网络**。创建负载均衡详情请参见**创建负载均衡实例**。

The screenshot shows the configuration page for creating a private network load balancer instance. The 'Instance Type' (实例类型) is set to 'Private Network' (私网), which is highlighted with a red box. The 'Network Type' (网络类型) is set to 'Dedicated Network' (专有网络). The 'Billing Method' (计费方式) is set to 'By Usage' (按使用流量计费). The page includes various configuration options like Instance Specification (实例规格), IP Version (IP 版本), and Virtual Switch (虚拟交换机).

配置负载均衡监听端口时，参见**监听介绍**。请注意**不要启用访问控制（默认未启用）**。

注意：

创建负载均衡实例将产生额外的配置费用。

负载均衡规格不够时可能成为压测瓶颈。

步骤二：启用 PTS 阿里云内网压测

若需使用阿里云经典网络来压测多机服务，请参见启用阿里云经典网络压测。

若需使用阿里云 VPC 网络来压测多机服务，请参见启用阿里云 VPC 网络压测。

创建智能压测

若您的应用在特殊场景（如秒杀场景）需要探索可承受压力顶点时，可以使用智能压测功能。智能压测功能可以提供指标预估能力，判定实际值是否属于预估正常范围，定位性能异常拐点。本文将介绍如何创建智能压测的方法。

操作步骤

登录 PTS 控制台，在概览页面单击**快速压测**，或在左侧导航栏，单击**创建压测** > **快速压测**。

单击**场景配置**页签，然后完成场景配置，请参见创建压测场景。

单击**施压配置**页签，然后进行以下配置：

在**模式配置**模块将**是否智能压测**选择为**是**，然后根据实际配置**压力来源**、**压测模式**、**总并发量**和**递增模式**等信息。

The screenshot shows the 'Stress Mode Configuration' (压测模式配置) section of the PTS console. It includes the following settings:

- 是否智能压测** (Is intelligent stress testing): Enabled (toggle switch).
- 压力来源** (Stress source): Domestic Public Network (国内公网) selected, Alibaba Cloud Internal Network (阿里云内网) unselected.
- 压测模式** (Stress mode): Concurrent mode (Virtual User mode) (并发模式 (虚拟用户模式)) selected, RPS mode (Requests Per Second) (RPS模式 (Requests Per Second)) unselected.
- 最大并发** (Maximum concurrent): 50.
- RPS上限** (RPS limit): 2000.
- 递增模式** (Increment mode): Mode 1 (10% increment, single level lasts 1 minute) (模式1 (10%递增, 单量级持续1分钟)) selected. A note indicates the starting volume is fixed and cannot be modified (起始量级固定, 不可修改).

压力来源：可选择**国内公网**和**阿里云内网**，详情请参见如何进行施压配置。

压测模式：智能压测中仅支持**并发模式**。

总并发量：全场景可达到的最大并发值。总并发量 $\geq (1/\text{模式百分比}) * 2 * N$ 。若选择模式 1，其中 1/10% 表示 10 次递增；为保证算法准确性，确保每个串联链路的最大并发数大于等于 2；N 表示 N 个串联链路。

递增模式：可选择**模式一**和**模式二**。

模式一：10% 递增，单量级持续 1 分钟。

模式二：5% 递增，单量级持续 1 分钟。

在**量级及数据配置**模块配置每个串联链路的最大并发权重。

注意：为保证算法准确性，确保每个串联链路的起始并发数大于等于 2。若串联链路的最大并发数小于 2 时，请返回上一步重新配置总并发量。

单击**场景 SLA** 页签，按实际需求配置场景 SLA。

创建 SLA 请参见**创建 SLA**。

导入 SLA 请参见**配置场景 SLA**。

注意：为保证您的应用安全，配置场景 SLA 时，必须配置停止压测的 SLA 规则。

单击页面右上角**保存去压测**，启动压测。

说明：压测过程中，不支持调速。出现紧急情况时可以停止压测。

后续操作

压测完成后，需要查看压测报告来分析本次压测结果。如何查看压测报告请参见**压测报告**。

压力预估图

通过**施压配置**页面的压力预估图，您可以预判性能测试中压力的分布情况。压力预估图会随着每个施压配置项的调整而变化。本文将说明如何解读压力预估图，并综合说明压测模式、Cookie 共享、压测量级等配置的不同，对性能测试压力的影响。

阅读本文前，您应该已了解如何进行施压配置的基本操作。

压力预估图说明（以并发模式、自动递增为例）

根据压测场景中是否使用数据导出 Cookie 功能，对两种情况分别说明：

第一种情况：未使用数据导出 Cookie 时，全部串联链路同时发起压测。



压力预估图说明：

压测量级（图标 ①）：

自动递增模式下，全场景并发量按照递增百分比逐级增加至最大并发量。

起始并发：

所有串联链路的起始并发量之和。例如，上图中两个串联链路的起始并发（图标 ②）都是 5，所以压力预估图中的起始并发是 10。

压测时长（分钟）：

表示整个压测持续的时间，计算方法如下：

$$\text{压测时长(分钟)} = \text{单量级持续时长} * [(100\% - \text{起始百分比}) / \text{递增百分比} + 1]$$

其中， $(100\% - \text{起始百分比}) / \text{递增量级百分比}$ 表示递增次数，在计算中需向上取整。

根据上图中的配置（图标 ③），单量级持续时长为 1 分钟，起始百分比为 10%，递增百分比为 20%，则压测时长为 $1 * [(100 - 10) / 20 + 1] = 6$ 分钟。其中，递增次数 $(100 - 10) / 20$ 向上取整得出 5。

说明：如存在多个串联链路且耗时不同的情况，总压测时长以耗时最长的串联链路为准。

第二种情况：如果使用了数据导出 Cookie 功能，被导出 Cookie 的串联链路需要先进行压测，完成数据量级（登录用户量级）的准备，其他链路再开始压测。

例如，将 **串联链路 1** 的 Cookie 共享至全场景，压力预估图将以数据准备完成为划分点，将压测分为两个阶段。



压力预估图说明：

压测量级：

串联链路 1 设置了导出 Cookie，需要先压测完成数据量级准备。压力预估图中会假定该串联链路递增至最大并发量（即 15）时，完成数据量级准备。

说明：实际压测中，数据准备完成的时刻由**串联链路 1** 的吞吐速率决定。

起始量级：**串联链路 1** 的起始并发（即 1）。

- 数据准备完成后的起始量级：其他串联链路的起始量级之和。例如，上图中数据准备完成后的起始量级是 3，即串联链路 2 的起始并发量。

压测时长：

表示整个压测持续的时间，计算方法如下：

压测时长(分钟)=数据量级准备时间（即共享 Cookie 的串联链路的压测时长）+ 其他串联链路压测时长

每个串联链路压测时长(分钟)=单量级持续时长 * [(100% - 起始百分比) / 递增百分比 + 1]

其中, (100% - 起始百分比) / 递增量级百分比表示递增次数, 在计算中需向上取整。

根据上图中的配置, 数据准备时长和压测时长均为 $2 * [(100-10) / 20 + 1] = 12$ 分钟。

压力预估图说明 (以 RPS 模式、非自动递增为例)

压测配置为非自动递增模式时, 整个压测过程中压测量级会在起始并发与最大并发的范围内波动, 故压力预估图使用区间示意线提供预览。

根据压测场景中是否使用数据导出 Cookie 功能, 对两种情况分别说明:

第一种情况: 未使用数据导出 Cookie 时, 全部串联链路同时发起压测。

压测量级, 是全场景压测量级之和。

第二种情况: 如果使用了数据导出 Cookie 功能, 被导出 Cookie 的串联链路需要先进行压测, 完成数据量级 (登录用户量级) 的准备, 其他链路再开始压测。

例如, 将 **串联链路 1** 的 Cookie 共享至全场景, 压力预估图将以数据准备完成为划分点, 将压测分为两个阶段。



压力预估图说明:

压测量级:

数据准备完成前的量级 = 串联链路 1 的总 RPS 范围内。

如图所示，串联链路 1 中所有 API 的起始 RPS 之和为 2，最大 RPS 之和为 20，所以数据准备完成前的压测量级预估是介于[2，20]之间。

数据准备完成后的量级 = 其他串联链路的总 RPS 范围内。

如图所示，其他串联链路只有串联链路 2，它的起始 RPS 是 1，最大 RPS 是 10，所以数据准备完成后的压测量级预估是介于[1，10]之间。

压测时长：

示例中，**串联链路 1** 的 Cookie 共享至全场景。非自动递增模式下，压力预估图中会假定**串联链路 1** 压测时长的 20% 时刻（如图中为 2 分钟时），完成数据量级准备。系统规定，数据量级准备时间最少为 1 分钟。

说明：实际压测中，数据准备完成的时刻由**串联链路 1** 的吞吐速率决定。

监控集成

云监控

性能测试过程中，往往需要实时观察各项指标是否正常，包括 PTS 本身已经具备的强大的客户端指标，以及应用服务器、数据库、中间件、网络入口等方面的指标。更重要的是，监控的过程是发现系统瓶颈的过程，监控数据是性能基线管理、容量规划甚至是高可用架构的重要基础。

目前 PTS 支持添加云监控，用于查看各项指标，更好地保证测试前提，记录相关数据，输出最终结果。如果您使用了阿里云基础服务（ECS、RDS、SLB），均可通过添加监控的方式，在压测及报告中便捷地查看相应的监控数据。

限制条件

- 云监控目前针对不同资源包用户，提供不同数量的可监控实例数，详见下表。若您当前资源包对应的上限不满足需求，可以先升级资源包。

监控实例 类型\资源	278	628	1058	8398	13998	58158 及 以上
---------------	-----	-----	------	------	-------	---------------

包大小						
ECS 实例数	4	8	8	16	16	32
RDS 实例数	1	2	2	4	4	8
SLB (IP : Port) 组	1	2	2	4	4	8

PTS 集成的 ECS 监控是操作系统级别的，监控项是通过插件采集的，采集频率为 15 秒一次。故需要在云监控控制台上安装插件，并确保插件在运行中的状态。

说明：关于插件安装的权限，主账号可以直接安装，子账号需要被授权 AdministratorAccess 权限。



云监控范围

目前支持的云监控实例类型包括 ECS，RDS，SLB 三种，所有当前使用 PTS 账号对应的主账号（若有）或该账号本身购买的实例，均可被添加。

PTS 的云监控集成不会计入云监控的查询监控数据API调用数量的计费指标，请放心使用。

添加监控

创建云监控分组，在监控分组内添加 ECS、RDS、SLB 实例，方便您将提供同样服务的多个实例做聚合统计。在查看具体监控信息时，仍然按照实例类型区分查看。每个压测场景中，各自建立监控分组，互不影响。

1. 在**创建压测**或者**编辑场景**页面，选择**添加监控**页签。
2. 在**云监控**页签下，单击**创建分组**，填写分组名，单击**确定**。



1. 选择新建的分组，单击**添加产品**。一次仅能添加一种产品。

在**添加资源**对话框中，选择产品类型（ECS、RDS 或 SLB），并勾选目标实例。单击对话框右下角**添加**。

当后台新增了云监控产品或实例时，单击**手工同步**，手动同步资源。



说明：

SLB 后端 RealServer 的 IP 信息需要配置完整，否则 SLB 不能被选中添加。

SLB 的一组 IP : PORT 视为一个实例。

确保 ECS 实例的 Agent 状态是 **正在进行中**，否则无法进行监控。如 Agent 处于非运行状态，需要在云监控控制台安装插件，或排查问题。

(可选) 需要为已有产品添加实例时，单击**添加实例**。

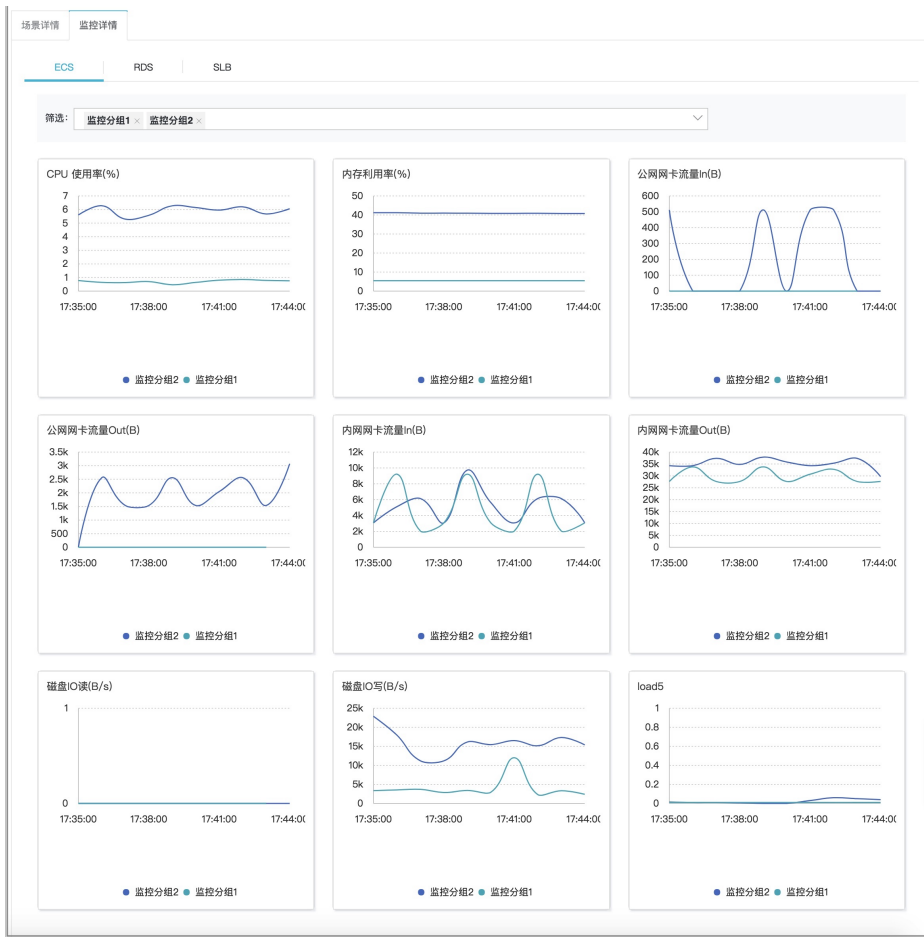


查看监控数据

压测过程中，可以实时查看监控数据，以便时刻观察线上负载情况，进行业务合理性判断及异常分析。

在压测过程中页面，单击**监控详情**，再单击不同实例页签查看具体数据。

以下以 ECS 的监控查看为例，做查看页操作介绍。



说明：

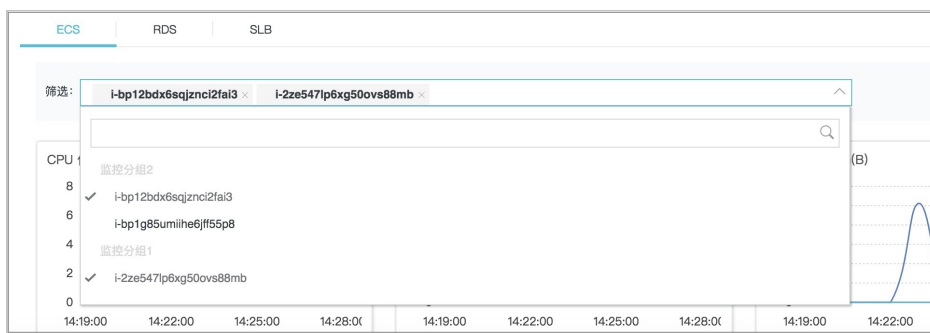
默认展现全部已添加的实例信息，点击筛选下拉框可调整查看对象；

同时查看的对象仅可以是分组或实例，分组与实例不可同时在趋势图中展现；

各实例的监控维度：

- ECS：CPU 使用率、内存利用率、公网网卡流入/流出流量、内网网卡流入/流出流量、磁盘 IO 读/写、load5（5 分钟平均负载）；
- RDS：连接数使用率、CPU 使用率、磁盘使用率；
- SLB：
 - 核心视图：端口新建/丢弃连接数、端口流入/流出带宽；
 - 其余视图：端口活跃/非活跃连接数，端口并发连接数，端口正常/异常ECS数。

因云监控的数据周期限制，ECS/SLB 的数据为 1 分钟一个采集点，RDS 数据为 5 分钟一个采集点。



监控报告

PTS 除了在压测过程中提供监控数据，也在报告中整合了压测数据，以便在压测结束后，复盘压测过程时对比业务数据与负载数据使用，帮助您快速地定位负载情况、系统是否异常等。

监控报告中的明细数据展现了在压测前配置的监控对象数据，与压测中的展现形式一致。报告中的数据趋势图区间为整个压测区间。

ARMS 监控

性能测试过程中，在监控硬件层面的指标（服务器、数据库、中间件、网络入口等）的同时，您可能也需要从业务角度监控应用的指标，从而更加有效地定位问题、发现系统瓶颈。因此，PTS 为您集成了 ARMS（Application Real-Time Monitoring Service，业务实时监控服务）的应用监控功能，ARMS 能够帮助您定位 JAVA 应用的出错接口、慢接口、重现调用参数和检测内存泄漏等。

前提条件

- 已开通 ARMS 服务。
- 已通过 ARMS 应用监控接入应用。参见 ARMS 创建应用监控。
- ARMS 监控集成，目前开放给 278 及以上资源包用户。

添加 ARMS 监控

在 PTS 控制台 [创建压测](#) 或者 [编辑场景](#) 页面，选择 [添加监控](#) > [ARMS 监控](#) 页签。

单击加号，添加 ARMS 监控应用。

选择 Region，当前 Region 已通过 ARMS 接入的应用会出现在**选择应用**下拉列表中。



在**选择应用**下拉列表中，勾选一个或多个应用。

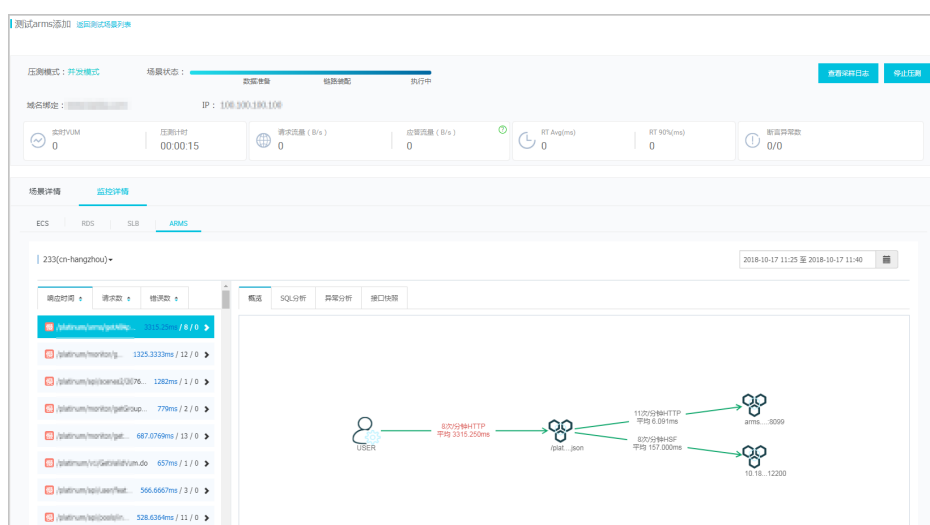
5. 单击**确定**，完成操作。

查看 ARMS 监控数据

在压测过程中页面和压测报告中都可以查看 ARMS 监控数据。

- 执行压测过程中，在压测页面，选择**监控详情** > **ARMS** 页签，实时查看应用的接口调用详情。

压测结束后，在压测报告中，选择**明细** > **ARMS** 页签，查看整个压测周期内的监控数据。



关于 ARMS 监控数据解读，参见 ARMS 应用接口调用监控。

场景 SLA

创建 SLA

服务等级定义 SLA (Service Level Agreement) 是判定压测是否异常的重要依据。压测过程中，通过监控核心服务状态的 SLA 指标数据，您可以更直观地了解压测业务的状态。

PTS 基于多年的业务沉淀，制定了一些行业通用 SLA 模板，您可以直接在压测中使用，具体可参见通用 SLA 模板。

除此之外，您也可以自定义符合业务需要的 SLA 规则，应用到特定的压测场景中。本文将介绍如何创建自定义的 SLA 规则。您可以设置量化指标、设置触发报警或停止压测的条件，便于实现智能化的压测。

名词解释

在使用该功能时，请先了解 PTS 对于 SLA 相关名词的说明。

SLA

服务等级定义 SLA (Service Level Agreement) 是判定压测是否异常的重要依据。压测过程中，通过监控核心服务状态的 SLA 指标数据，您可以更直观地了解压测业务或架构的状态。

SLA 指标

SLA 指标是指从哪些指标监控压测过程数据，主要包括业务质量相关指标、ECS、RDS 和 SLB 监控类指标。

PTS 将在后续迭代中逐步增加队列、SQL 连接等方面的 SLA 指标。

SLA 规则

SLA 规则是在 SLA 指标的基础上，添加了判断条件，以触发告警或停止压测。

SLA 模板

SLA 模板是 SLA 规则的集合，可包含一个或多个 SLA 规则。SLA 模板与行业类型绑定。

SLA 指标

SLA 指标是指从哪些指标监控压测过程数据，详细指标说明见下表。

SLA 指标类型	SLA 指标项	SLA 指标项说明	最小监控粒度
业务质量相关指标	RT (Response Time)	压测中 API 的响应时间。	1 秒
	RPS (Requests per Second)	压测中，每秒发出的压测请求数量。	
	成功率	压测中 API 的请求成功率。	
ECS 基础监控相关指标	CPU 利用率	即 CPU 使用率，指一段时间内 CPU 被占用的情况。	1 分钟
	内存利用率	一段时间内内存被占用的情况。	
	load5	5 分钟内系统的平均负荷。	
RDS 基础监控相关指标	CPU 利用率	即 CPU 使用率，指一段时间内 CPU 被占用的情况。	5 分钟
	连接利用率	连接数是指应用程序可以连接到 RDS 实例的数量。连接利用率即已经使用的连接数百分率。	
SLB 基础监控相关指标	丢弃连接数	端口丢弃的连接数，从 IP : Port 维度进行统计。	1 分钟
	异常后端 Server	例如，后端异常 ECS 实例个数	

自定义 SLA 操作步骤

除此通用 SLA 之外，您也可以自定义符合业务需要的 SLA 规则，应用到特定的压测场景中。自定义 SLA 模板步骤如下：

1. 登录 PTS 控制台，在左侧导航栏单击**SLA管理**。

在**SLA管理**页面，选择**自定义SLA**页签，单击右上角**自定义SLA**按钮。

说明：如果您要基于某个通用 SLA 模板自定义 SLA，可在**SLA管理**页**SLA模板**页签下，单击目标 SLA操作列的**复制到自定义**，克隆目标 SLA 至自定义 SLA 页面，再对其进行编辑。

在创建SLA页面，填写自定义的 SLA 名称、SLA 描述，并选择该 SLA 适用的业务类型。

单击**+添加SLA规则**，在**添加SLA规则**窗口中，设置具体信息。

- 规则名称
- 指标名称：选择具体指标，如业务指标 > RT。
- 阈值及报警级别：勾选并设定触发报警、停止压测的具体阈值。
 - 警告通知：满足条件时，通知报警接收人。
 - 停止压测：满足条件时，立即停止压测，并通知报警接收人。

设置每个条件的敏感度。

- **敏感，连续一次**：压测过程中，该指标只要一次符合报警或停止压测的条件，即

触发报警或停止压测。

- **可容忍，连续三次**：压测过程中，该指标**连续**出现 3 次，符合报警或停止压测的条件，即触发报警或停止压测。
- **不敏感，连续五次**：压测过程中，该指标**连续**出现 5 次，符合报警或停止压测的条件，即触发报警或停止压测。

说明：压测中每隔 5 秒统计一次 RT、RPS 和成功率的数据。

在添加SLA规则窗口中，单击**确定**。

在自定义SLA框中，再次单击**确定**，完成创建。

自定义SLA/新建SLA名称

SLA名称: * 8/20 ✔

SLA定义: 9/20

业务分类:

规则名称	指标类型	规则描述
RT规则示例	BUSINESS	RT > 1000 : 不敏感, 连续五次: 警告通知
RT规则示例	BUSINESS	RT > 5000 : 敏感连续一次:

+添加SLA规则

确定

在自定义SLA页签下的列表中，可以查看创建的 SLA。

后续步骤

使用 SLA

创建压测场景时，直接将通用 SLA 或自定义 SLA 应用到具体 API 上。具体操作步骤，参见配置场景 SLA。

编辑 SLA

在自定义SLA页签下的列表中，单击某个自定义 SLA 操作列的**编辑**，可修改名称、描述和规则等信息。

删除 SLA

在自定义SLA页签下的列表中，单击某个自定义 SLA 操作列的删除，可删除该 SLA。

更多信息

管理接收人

通用 SLA 模板

配置场景 SLA

服务等级协议 SLA (Service Level Agreement) 是判定压测是否异常的重要依据。您可以为压测请求设置 SLA，从而监控压测中服务状态、查看异常告警或及时终止压测。

本文将介绍如何在压测场景中定义 SLA。

前提条件

已创建告警接收人。参见管理接收人。

操作步骤

在 PTS 控制台的**创建压测**页面，单击 **SLA定义**页签。

选择一种合适的方式为该压测场景添加 SLA：

方式一：导入已有 SLA

如果您已经有符合该压测场景的通用 SLA 或自定义 SLA，您可以导入至压测场景中，并可以编辑导入的 SLA。具体步骤如下：

- a. 单击**导入已有 SLA**。
- b. 在**导入已有规则**下拉列表中，选择目标 SLA。一个压测场景仅能导入一个 SLA。
- c. (可选) 修改导入的 SLA 的具体信息，如修改 SLA 规则或添加 SLA 规则等。具体步骤，参见**创建 SLA中的自定义 SLA 操作步骤**。



方式二：创建 SLA

单击**创建 SLA**，在下面展开的编辑区，配置 SLA 及其具体规则。具体步骤，参见**创建 SLA 中的自定义 SLA 操作步骤**。



选择适用 SLA 规则的 API。

在每一个 SLA 规则的**适用对象**列，选择适用该规则的 API，全部 API 或部分 API。如果留为空，则默认是该场景中的全部 API。



注意：如果某 API 匹配了多个 SLA 规则，这些规则将同时生效。请您根据业务需要合理配置。例如，某 API 匹配了两个 SLA 规则：RT > 5000 ms 时触发告警和 RT > 3000 ms 时触发告警，则在实

实际压测中，两个规则同时生效，RT 大于 5000 ms 或者大于 3000 ms 时，均会触发告警。

在高级设置区域，选择告警消息的接收人和通知方式。

高级设置：	告警及操作	选择通知人	选择通知方式
告警通知		报警联系人A	<input checked="" type="checkbox"/> 钉钉 <input checked="" type="checkbox"/> 邮件 <input type="checkbox"/> 短信
停止压测		请选择	<input type="checkbox"/> 钉钉 <input type="checkbox"/> 邮件 <input type="checkbox"/> 短信

结果验证

启动压测后，如果请求的某个瞬时值达到设置的 SLA 告警或停止压测的条件，会出现如下提示。

The screenshot shows the PTS dashboard with the following data:

实时VUM	压测计时	并发(当前/上限)	RPS (当前/上限)	IP数 (当前/配置)	请求流量	应答流量	异常数 (请求/业务)
100	00:00:11	1/20	3/4000	1/1	108B	12.98KB	2/0

SLA Statistics:

API名称	API详情	SLA值	当前值
首页	http://...	RT > 500	11ms
访问	http://...	RT > 500	22ms
出错	http://...	RT > 500	3675ms

SLA Statistics Table (from image):

非自动递增模式	压测量级限制	成功	失败	RT	并发
首页	2/20	0.6/s	0/s	102 ms	0
访问	2xx	1/s	0/s	146 ms	0
出错	2xx	0/s	1/s	1618 ms	1

说明：

- 红色图标表示压测停止；黄色图标表示触发告警通知。
- 压测中，每隔 5 秒采集一次 API 的 RT、RPS 等数据。

压测中，单击 **SLA统计** 页签，可以看到具体的告警或停止压测的事件记录。

The screenshot shows the SLA Statistics page with the following data:

API名称	API详情	SLA值	当前值
首页	http://...	RT > 500	11ms
访问	http://...	RT > 500	22ms
出错	http://...	RT > 500	3675ms

Event Log (from image):

时间	事件
2019-03-29 15:19:51	出错 RT > 500 : 敏感连接一次 : 警告通知
2019-03-29 15:19:46	出错 RT > 500 : 敏感连接一次 : 警告通知

在压测报告中，您可以查看告警或停止压测的事件记录。

API名称	API详情	SLA值	平均值
首页	http://...	RT > 500	102ms
		成功率 < 99	100.00%
访问	http://...	RT > 500	146ms
		成功率 < 99	100.00%
出错	http://...	RT > 500	1618ms
		成功率 < 99	0.00%

最近事件 2
2019-03-29 15:16:36 出错 RT > 500: 敏感,连续一次: 警告通知
2019-03-29 15:16:36 出错 成功率 < 99: 敏感,连续一次: 停止压测

说明：压测报告的 SLA 详情中，每条 API 只展示整个压测周期内的平均值，不是瞬时值。

通用 SLA 模板

PTS 基于多年的业务沉淀，制定了符合行业特性的一些通用的 SLA 模板和指标，供您参考。

查看 SLA 模板：

登录 PTS 控制台，在左侧导航栏单击 **SLA 管理**。

在 **SLA管理** 页面的 **SLA模板** 页签下，单击 **查看**，可查看某个预定义 SLA 的具体信息。

在压测场景中使用 SLA 模板：

具体操作步骤，参见配置场景 SLA。

核心业务经验 SLA

与核心业务相关的 SLA，例如电商场景中，登录页面可能成为您的核心业务，需要为核心业务制定更严格的监控指标。

具体说明如下：

执行动作	满足条件	敏感度
警告通知	RT > 500ms	可容忍，连续三次
停止压测	成功率 < 99%	敏感，连续一次

金融业经验 SLA

金融业经验 SLA，具体说明如下：

执行动作	满足条件	敏感度
警告通知	RT > 1000ms	可容忍，连续三次
停止压测	成功率 < 95%	敏感，连续一次

保险业经验 SLA

保险业经验 SLA，具体说明如下：

执行动作	满足条件	敏感度
警告通知	RT > 3000ms	可容忍，连续三次
停止压测	成功率 < 99%	敏感，连续一次

传统制造业经验 SLA

传统制造业经验 SLA，具体说明如下：

执行动作	满足条件	敏感度
警告通知	RT > 1000ms	可容忍，连续三次
停止压测	成功率 < 99%	敏感，连续一次

电商业务经验 SLA

电商业务经验 SLA，具体说明如下：

执行动作	满足条件	敏感度
警告通知	RT > 500ms	可容忍，连续三次
停止压测	成功率 < 95%	敏感，连续一次

启动压测并调速

场景调试没有问题，则可以进入场景压测页面。您可以实时监控整个场景和每个 API 的压测指标、调整压测值或记录操作记录。

注意：同时启动的场景数受您购买的资源包规格限制，参见资源包说明。

启动压测

在创建压测页面，单击保存去压测按钮，进入场景压测。

- 立即执行：默认情况，立即开始压测。
- 定时压测：如需指定时间开始压测，勾选定时压测，并选择相关配置项。更多说明，参见定时压测。

单击确定。

如选择立即执行，稍等数秒，操作页会进入压测态（如下图所示）；如选择定时压测，在您指定的时间，将开始压测（如下图所示）。



页面分为以下几个区域：

场景实时状态区域：显示当前场景的压测状态，本次压测使用的 VUM 计量，场景整体发送和接受的报文大小。

调速区域：支持三个维度的调速：全场景维度、串联链路维度和单个链路维度。每次调速实时生效。

- **单链路（API）实时状态区域：**单击单个链路的**监控**图标，会显示当前链路的请求、响应时间和使用并发数等实时信息。

将鼠标悬浮在趋势图上，可以查看单个链路的历史状态信息。



自动递增模式下调速

自动递增模式下，点击**全场景调速**或者**串联链路调速**会将**自动递增切换到手工模式**，并按照输入压测值进行压

测。

- 若通过全场景调速切换到手工模式，则全场景（所有串联链路的全部 API）都切换到手工模式。
- 若通过串联链路调速切换到手工模式，则该串联链路切换到手工模式，其他未调速的串联链路仍然为自动递增模式。



手工模式下调速

手工模式下调速方式与上述一致。手工模式下若不进行调速，则按照当前压测量级持续施压。

操作记录

性能测试的过程中，您可以随时查看系统自动生成的该场景相关的操作记录，也可以随时记录操作笔记，保存在操作记录中。

操作步骤如下：

在压测中页面，单击场景详情右上角的**操作记录**。

压测的启停、手动调速等均会记录在操作记录中。



如果需要记录其他操作，可以输入操作笔记，单击**保存笔记**。

定时压测

场景创建并调试完成后，可以指定测试任务的执行日期、执行时间、循环周期和通知方式等，从而实现定时压测。

说明：JMeter 压测暂不支持定时功能。

您可以选择以下任一方式设置定时压测：

- 创建新的压测场景时，以定时任务的方式启动压测，参见新场景定时压测。
- 为已有压测场景添加定时任务，参见已有场景定时压测。

前提条件

已创建压测场景，并完成场景调试。具体操作参见创建压测场景和场景调试。

限制条件

- 定时压测功能及定时任务数量均受限于资源包规格。具体规格参见资源包说明。

- 一个压测场景仅支持添加一个定时压测任务。已配置定时任务的场景不可编辑。

新场景定时压测

在**创建压测**页面，启动压测时选择**定时压测**，即可添加该场景的定时任务。操作步骤如下：

在**创建压测**页面，单击右上角的**保存去压测**。

在**温馨提示**窗口中，**执行周期**选择**定时压测**，并填写相关配置项。

说明：为了保证顺利执行定时任务，您需要先完成压测场景调试，再进行定时压测。

温馨提示

您剩余资源包额度为 4699136 VUM，消耗VUM按照实际并发*时间计费，保底计费并发和保底计费VUM详见[价格说明](#)。

施压带来的大流量可能会对您按量付费的云产品产生费用，请确认是否要开始压测此场景？

启动策略

ⓘ 如果设置了定期执行，请确保服务可用性，避免定期执行施压造成业务异常。

执行周期 * 立即执行 定时压测

执行频率 * 仅执行一次 定期执行

执行通知方式 * 压测开始前通知 压测结束后通知

通知对象及通知方式

* 短信 邮件 钉钉

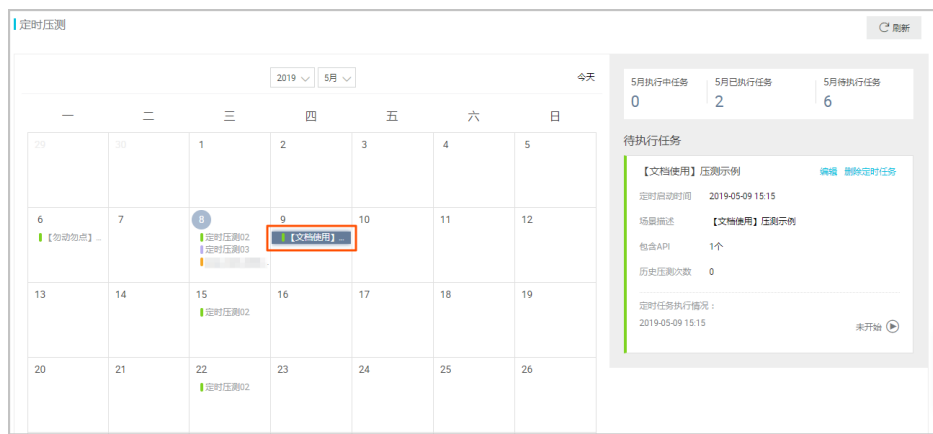
配置项说明如下：

配置项	说明
执行周期	选择 定时压测
执行频率	<p>选择仅执行一次，指定日期和时间，可具体到分钟，或选择定期执行：</p> <ul style="list-style-type: none"> - 单击日粒度，选择一个或多个日期执行压测； - 单击周粒度，选择在一周中的一天或几天执行压测； - 单击高级，编写 crontab 高级表达式自定义执行时间，参见 crontab 表达式。 <p>设置完成后，会自动显示最近五次压测的执行时间。</p> <p>说明：日粒度、周粒度和高级三个方式中，只能选择一种方式设置执</p>

	行时间。	
执行通知方式	可多选： - 压测开始前通知 - 压测结束后通知	
通知对象及通知方式	选择通知对象和方式： - 通知对象：选择在 PTS 中已创建的联系人作为通知对象。 - 通知方式：选择短信、邮件或钉钉作为通知方式。 说明： 关于如何创建联系人，参见管理接收人。	

单击**确定**，定时任务将自动添加到压测日历中。

从控制台左侧导航栏选择**定时压测**，在日历中，单击刚刚添加的定时压测任务，可预览待执行任务。



已有场景定时压测

在**定时压测**页面，可以将压测场景添加到指定日期，预约定时压测。操作步骤如下：

登录 PTS 控制台，在左侧导航栏选择**定时压测**。

在日历中，选择要执行压测的日期。

在右侧**待执行任务**场景区域，单击**+创建定时任务**。

在**场景名称**下拉列表中，选择压测场景。如果未调试压测场景，无法添加到定时任务中。

如需调试场景，可参见**场景调试**。

配置其他定时任务选项。

待执行任务

*** 场景名称** 🗑️

(18534) 【文档使用】 演示压力测试
▼

*** 执行频率**

仅执行一次

12:00 ✕

定期执行

日粒度
周粒度
高级

1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31	全部			

最近五次执行时间为
 2019-06-01 12:00
 2019-06-02 12:00
 2019-07-01 12:00
 2019-07-02 12:00
 2019-08-01 12:00

*** 执行通知方式**

压测开始前通知
 压测结束后通知

*** 通知对象及通知方式**

y****g ✕
▼

短信
 邮件
 钉钉

确定

放弃

配置项说明如下：

配置项	说明	
执行频率	选择 仅执行一次 ，指定日期和时间，可具体到分钟，或选择 定期执行 ： - 单击 日粒度 ，选择一个或多个日期执行压测； - 单击 周粒度 ，选择在一周中的一天或几天执行压测；	

	<ul style="list-style-type: none"> - 单击高级，编写 crontab 高级表达式自定义执行时间，参见 crontab 表达式。 设置完成后，会自动显示最近五次压测的执行时间。 说明：日粒度、周粒度和高级三个方式中，只能选择一种方式设置执行时间。 	
执行通知方式	<ul style="list-style-type: none"> 可多选： - 压测开始前通知 - 压测结束后通知 	
通知对象及通知方式	<ul style="list-style-type: none"> 选择通知对象和方式： - 通知对象：选择在 PTS 中已创建的联系人作为通知对象。 - 通知方式：目前仅支持钉钉通知。短信、邮件通知方式将即将上线。 说明：关于如何创建联系人，参见管理接收人。 	

单击**确定**，完成定时压测任务创建。定时任务将自动添加到压测日历中。

crontab 表达式

如果您需要从更细维度自定义定时任务的执行时间，可以使用 crontab 表达式。时间间隔的单位可以是分钟、小时、日(或周)、月、及以下的任意组合。

在创建定时任务时，选择**定期执行**，单击**高级**，设置 crontab 表达式：

定期执行

日粒度
周粒度
高级

cronab 表达式 (需填写所有空缺部分) [表达式示例](#)

59	分	23	时	1,2	日	*	月	?	周
----	---	----	---	-----	---	---	---	---	---

corntab 表达式展示 (不带秒) :59 23 1,2 * ?

最近五次执行时间为

2019-06-01 23:59

2019-06-02 23:59

2019-07-01 23:59

2019-07-02 23:59

2019-08-01 23:59

配置说明如下：

- 只能按照日或者周，指定执行压测时间。指定日或周之后，另外一个输入框内需填写？，表示不受限制。例如，上图所示，指定了每个月 1、2 号，周需填写？。
- 在 cronab 表达式内，请使用英文字符，如英文通配符（*、?）。
- 多个值之间用英文逗号（,）分隔。
- 值的区间用短划线（-）连接，例如：日输入框内填写 1-5，表示 1 号至 5 号。
- 填写周时，1 表示周日，7 表示周六；即每周从周日到周六的数字为 1-7；

cronab 表达式由 5 列数字组合而成，取值说明如下。

- 分（分钟）：0-59
- 时（小时）：0-23，0 表示子夜。
- 日（每个月的具体日期）：1-31
- 月：不可修改，目前仅支持每个月。
- 周（星期）：1-7，1 表示星期日，2 表示星期一，以此类推。

cronab 表达式示例：

每星期六、星期日的 23:00 执行：

crontab 表达式 (需填写所有空缺部分) [表达式示例](#)

0	分	23	时	?	日	*	月	7,1	周
---	---	----	---	---	---	---	---	-----	---

corntab 表达式展示 (不带秒) :0 23 ? * 7,1

最近五次执行时间为

2019-05-11 23:00

2019-05-12 23:00

2019-05-18 23:00

2019-05-19 23:00

2019-05-25 23:00

每月 5、15、25 日的 09:30 执行：

crontab 表达式 (需填写所有空缺部分) [表达式示例](#)

30	分	9	时	5,15,25	日	*	月	?	周
----	---	---	---	---------	---	---	---	---	---

corntab 表达式展示 (不带秒) :30 9 5,15,25 * ?

最近五次执行时间为

2019-05-15 09:30

2019-05-25 09:30

2019-06-05 09:30

2019-06-15 09:30

2019-06-25 09:30

每隔两天的零点执行：

crontab 表达式 (需填写所有空缺部分) 表达式示例

0	分	0	时	1/2	日	*	月	?	周
---	---	---	---	-----	---	---	---	---	---

corntab 表达式展示 (不带秒) :0 0 1/2 * ?

最近五次执行时间为

- 2019-05-09 00:00
- 2019-05-11 00:00
- 2019-05-13 00:00
- 2019-05-15 00:00
- 2019-05-17 00:00

管理定时任务

您可以通过定时压测日历，选择某天的定时任务，查看、编辑、删除该定时任务。

查看定时任务

进入 PTS 控制台概览页，通过压测日历，可快捷查看是否有定时压测任务。

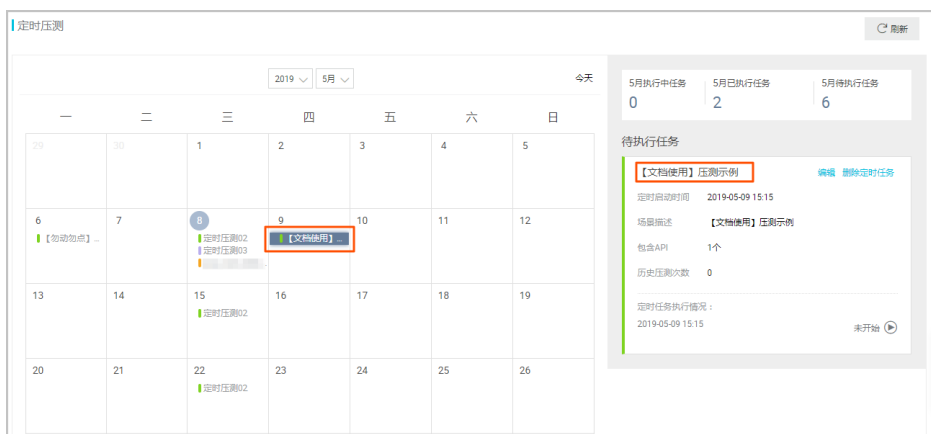


通过以下步骤，可以进一步查看定时压测任务的详情。

在 PTS 控制台左侧导航栏中，选择**定时压测**。

在定时压测日历中，单击某一个定时压测任务，可在右侧面板中查看详情。

单击场景名，可跳转至压测场景编辑页面。



编辑定时任务

在 PTS 控制台左侧导航栏中，选择**定时压测**。

在定时压测日历中，单击某一个定时压测任务。

在**待执行任务场景**区域，单击右上角**编辑**。

在展开的编辑区，修改该定时任务。

删除定时任务

删除定时任务将删掉该定时任务中的所有循环任务，无法单独删除某个循环中的子任务。但已经执行过的任务记录将保留到日历中，不会被清除。

您可通过两种方式删除定时任务：

方式一：

在 PTS 控制台左侧导航栏中，选择**定时压测**。

在定时压测日历中，单击某一个定时压测任务（日历视图中某一条详细的任务记录）。

在**待执行任务**区域，单击右上角**删除定时任务**，将删除该定时任务指定周期内所有未执行的循环任务（如果存在）。

方式二：

在 PTS 控制台左侧导航栏中，选择**压测场景**。

在**场景列表**中，找到需要删除的定时压测任务，在其**操作**列，单击**删除定时任务**。将删除该定时任务指定周期内的所有循环任务，但该压测场景仍会保留。

场景ID	场景名	持续时间	状态	最后修改时间	操作
47443	yuyu_mew	11分钟	待启动	2019-04-19 21:09:47	删除定时任务 查看 复制
45160	定时压测01	5分钟	待启动	2019-04-19 18:05:16	删除定时任务 查看 复制
46367	yuyu_test_sls	12分钟	待启动	2019-04-22 11:00:01	启动 编辑 删除 复制
27262	yuyu_test_jmeter_normal	10分钟	待启动	2019-04-22 10:45:44	启动 编辑 删除
44888	测试云监控-0328	10分钟	待启动	2019-04-22 10:36:48	启动 编辑 删除 复制

压测报告

压测报告

在压测停止之后，系统会自动收集压测过程中的数据（包括 API 采样日志），形成压测报告，供您查看和导出。

进入压测报告页面：

在 PTS 控制台左侧菜单栏，单击**压测报告**。

在**测试报告**列表页，可根据场景类型（PTS 或 JMeter）、时间、场景名称或报告 ID 等关键字进行搜索。

在对应的报告一行的**操作**栏，单击**查看报告**。

删除报告：

在**测试报告**列表页，可以删除或批量删除压测报告。

- **单个删除**：在对应的报告一行的**操作**栏，单击**删除**。
- **批量删除**：勾选要删除的报告，单击**批量删除**。

导出压测报告：

在压测报告页面，单击**报告导出**，根据需要选择是否包含**阿里云 PTS** 水印。

进入网页版报告页面，单击**下载报告**，可将压测报告（PDF）下载至本地。

查看采样日志：参见查看采样日志。

压测报告包括压测场景概览、串联链路信息和监控信息。

压测场景概览

压测报告第一部分显示整个压测场景的相关业务指标。

测试报告		返回报告列表				
场景名：		查看采样日志	报告导出			
执行时间：	2018-12-19 13:56:26 到 2018-12-19 14:07:26	运行时长：	11分钟			
施压模式：	并发模式	压力来源：	国内公网			
递增模式：	自动递增	递增量级：	10%			
		单量级时长：	1分0秒			
VUM 计费规则	并发 (峰值/上限)	RPS (峰值/上限)	来源IP (最小/配置)	平均流量	峰值流量	异常数 (请求/业务)
70447	11729/11500	10976/92000	23/23 分配详情	27.68MB	49.55MB	1266/409186

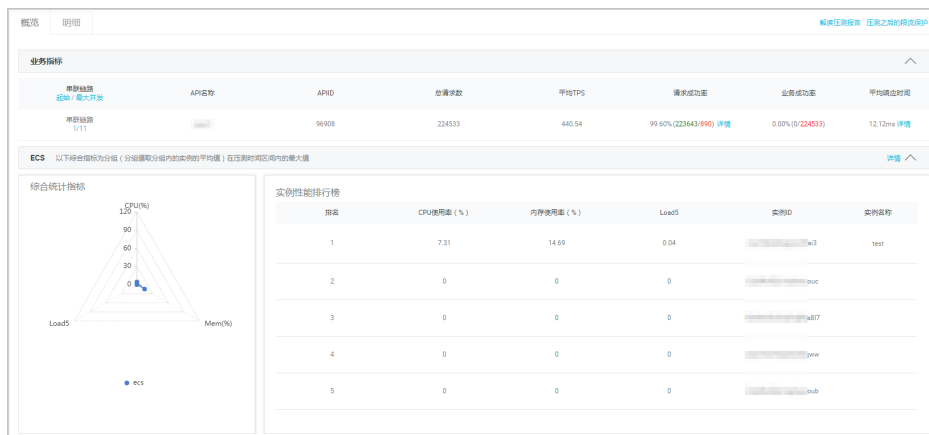
说明：若创建压测场景时，**递增模式**设置为**自动递增**，即使在压测过程中，进行了手动调速，此处显示的递增模式仍不变。

主要指标解读如下：

指标	说明
VUM	本次压测消耗的资源总数，单位是 VUM（每虚拟用户每分钟），参见计费规则。
并发	峰值：压测周期内场景的最高并发值 上限：根据该场景配置的压测量级，限制的最大并发数
RPS	Requests Per Second，即每秒发出的压测请求数量。 - 峰值：压测周期内，该场景的最高 RPS - 上限：根据该场景配置的压测量级，限制的最大 RPS
来源 IP	发起压测流量的 IP 地址个数。 - 最小：压测周期内，该场景的最小来源 IP 数； - 配置：场景施压配置中配置的来源 IP 个数） 单击 分配详情 ，可查看发起压测流量的 IP 地址所在的地理位置和所属运营商。该功能仅国内公网压测可见，并资源包规格需要是 628 元及以上。
平均流量	压测周期内的平均流量，PTS 采样统计的基于 7 层（HTTP）请求体或者响应体的最大值（与 4 层网络带宽完全不同），参见具体算法。
峰值流量	压测周期内的最大流量，算法同平均流量。
异常数	单击异常数，可快速查看其采样日志。 - 请求异常数：压测中，请求失败个数（包括 4XX、5XX、各种异常、超时等） - 业务异常数：压测中，设置了断言的 API 请求失败的个数

串联链路概览

压测报告的概览页，显示整个场景下所有的串联链路和 API 的压测情况；如添加了云监控产品，ECS、SLB 或者 RDS，您也可以看到这些产品在压测期间的性能表现。

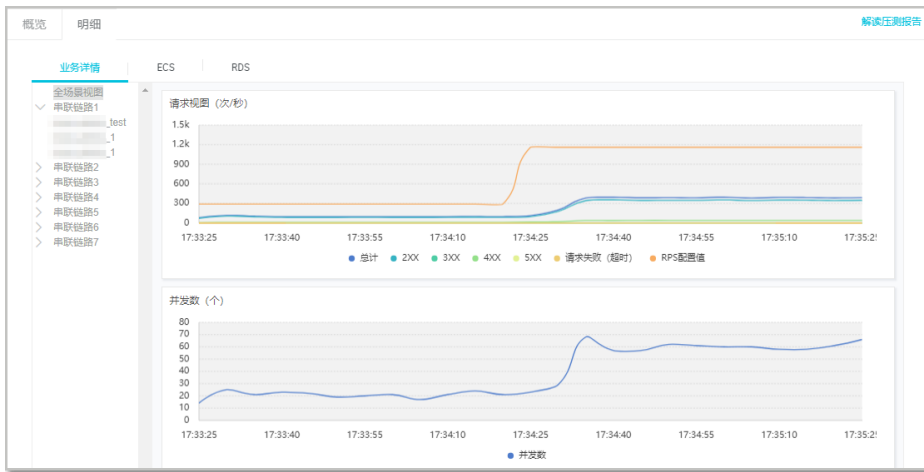


主要业务指标解读如下：

业务指标	说明
串联链路	并发模式下，串联链路的起始/最大并发量；RPS 模式下，API 的起始/最大RPS。
请求成功率	压测中此 API 的请求成功率。 - 单击请求成功/失败个数，可快捷查看对应日志； - 单击详情，查看 3XX、4XX、5XX 和其他异常导致的请求失败的个数。
平均响应时间	压测中此 API 的平均响应时间。单击详情，查看最大、最小及其各分位的响应时间。

串联链路业务详情

压测报告的**明细 > 业务详情**页签下，显示全场景视图和串联链路中单个 API 的业务详情，包括请求（TPS）视图、业务（TPS）视图、并发数、响应时间和指令详情等。对相关指标的解读，参见测试指标。



指令详情包括**思考时间**和**集合点**的指标。这两个指令会涉及到消耗 VUM，因为虚拟用户的思考和等待本身就是一个压测行为和资源消耗。

智能压测业务详情

在智能压测模式下，压测报告的**明细>业务详情**页签显示全场景视图和串联链路中单个 API 的业务详情，包括请求（TPS）视图、并发数和响应时间等。在单个 API 的业务详情，可以在各模块右上角选择**预测数据**，查看智能压测的预估阈值范围和压测实际值的时序曲线。



预估数据的时间间隔为 5s。

滑动鼠标至时序曲线中，可以查看当前的预估阈值范围和压测实际值。

智能压测时，将检测 `rt_avg`、`real_qps`、`qps_200`、`qps_fail`、`qps_success`、`qps_fail` 等 4-6 个指标的异常情况。当其中 1 个指标出现异常时，则对应的时序曲线被标红。

监控详情

若您添加了 ECS 监控、SLB 监控、RDS 监控和 ARMS 监控等，查看相关监控信息请参见**监控详情**。

SLA 统计

如添加了 SLA 指标监控，在压测报告的**明细页**，可查看压测过程中触发的 SLA 警告、停止压测的事件记录，以及各个 API 的具体表现。



相关文档

结合压测数据，您可以查看更多测试分析和调优方法指导。

查看采样日志

压测采样日志主要由压测 API 的请求详情、响应详情和请求核心生命周期的耗时信息三部分组成，您可以在压测执行时或者压测报告中，通过查看压测采样日志，快速定位问题。

采样日志按照 1% 频率进行收集，保存30天。

操作步骤

在压测进行页面或**测试报告**页面右上角，单击**查看采样日志**。

(可选) 在**查看日志**窗口，填写筛选条件(如下所示)，单击**查询**，筛选出相关日志。

- **压测 API**：只能选择单个压测 API。

异常状态：可在列表中选择多个，包括请求异常状态和业务状态(成功或失败)；**业务状态**仅对设置了断言的压测场景有效。

响应状态码：指定一个响应状态码。

关键字搜索：按照 **TransId** 进行模糊搜索；TransId（Transaction ID）用于标识当前日志属于哪个串联链路的第几个并发请求，格式为 **串联链路排序：并发排序**。例如，**0:1, 0** 表示第一个串联链路，**1** 表示该串联链路的第一个请求周期。同一个串联链路的同一个并发请求可能存在多个采样日志。

时间范围：指定日志发生的时间，最小单位是秒。

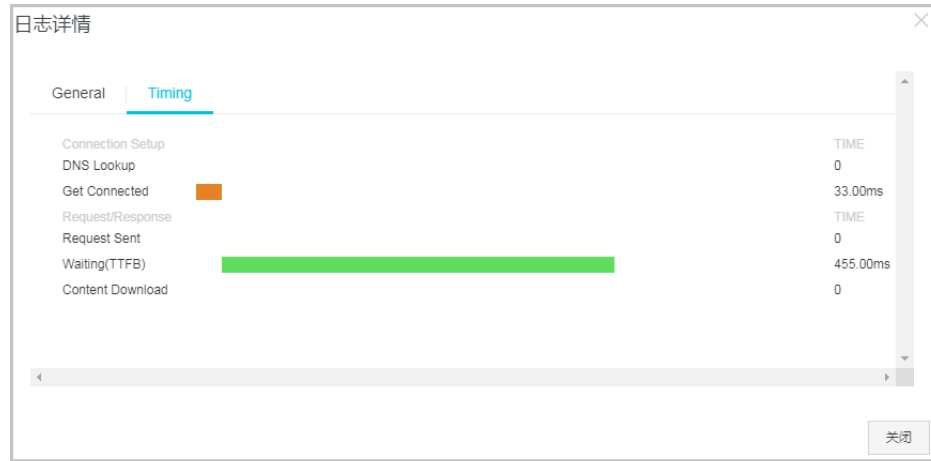
在下方列表中，找到目标日志，在其**操作**栏中，单击 **点击查看详情**。

在**日志详情**窗口，通过切换不同页签，General、Timing，查看具体信息。

- **General**：展示日志总览、Request Header、Request Body、Response Headers 和 Response Body。

Timing：展示该请求在核心生命周期的耗时情况，主要分为：

- DNS Lookup：执行 DNS 查询所用的时间。
- Get Connected：建立连接所用的时间。
- Request Sent：发出网络请求所用的时间。
- Waiting(TTFB)：等待初始响应所用的时间，也称为至第一字节的时间。此时间将捕捉到服务器往返的延迟时间，以及等待服务器传送响应所用的时间。
- Content Download：接收响应数据所用的时间。
- Reuse：复用连接所用的时间。



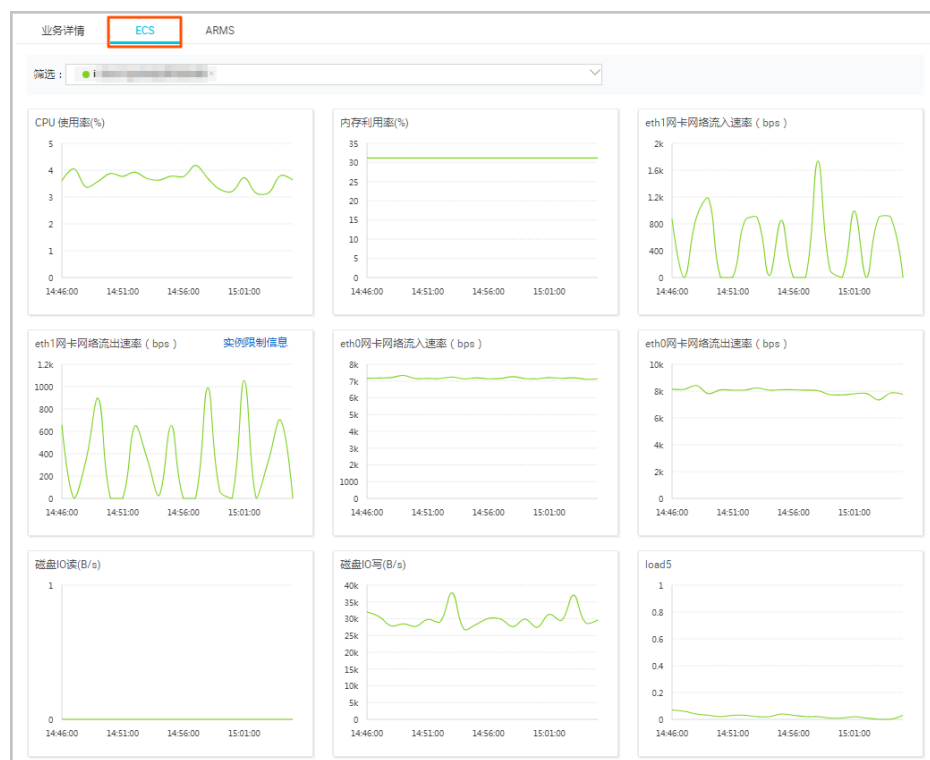
注意：采样日志中 Timing 仅统计请求在核心生命周期的耗时情况，方便您定位问题，未覆盖请求所有阶段的耗时。

监控详情

在压测报告中可以查看您添加的 ECS 监控、SLB 监控、RDS 监控和 ARMS 监控等信息。

ECS 监控详情

如添加了云监控 ECS，在压测报告的**明细页**，可查看整个压测时间内 ECS 的监控数据。具体解读，请参见云监控的监控项说明。



SLB 监控详情

如添加了云监控 SLB，在压测报告的**明细页**，可查看整个压测时间内 SLB 的监控数据。关于 SLB 监控详情解读，请参见负载均衡监控。



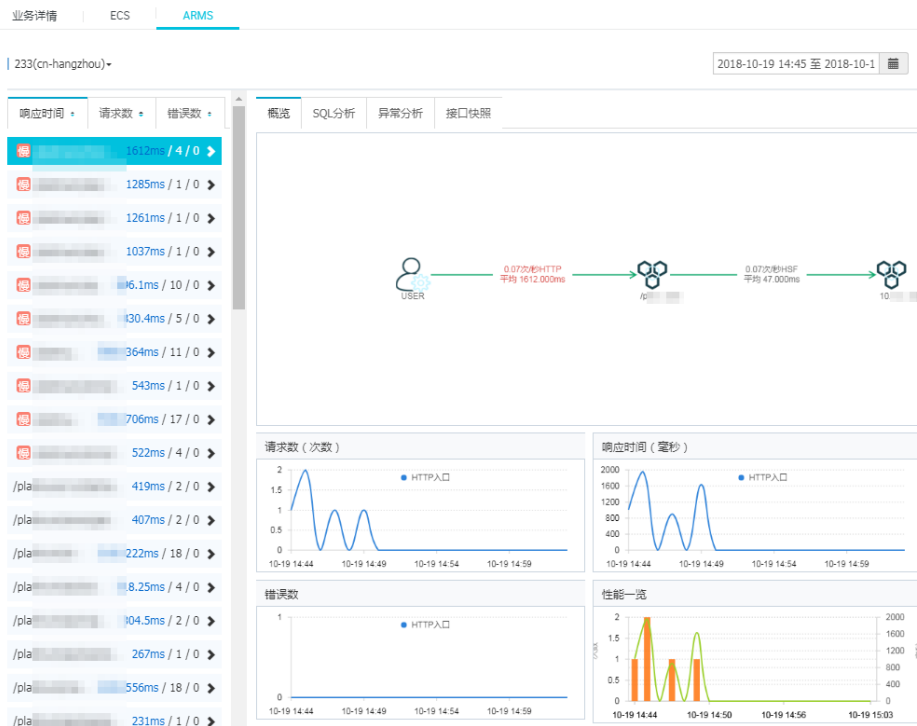
RDS 监控详情

如添加了云监控 RDS，在压测报告的**明细页**，可查看整个压测时间内 RDS 的监控数据。关于 RDS 监控详情解读，请参见云数据库RDS监控。



ARMS 监控详情

如添加了 ARMS 监控，在压测报告的**明细页**，可查看整个压测时间内 ARMS 基于应用接口的监控数据。关于 ARMS 监控详情解读，请参见 ARMS 应用接口调用监控。



云端录制器

云端录制器的使用

通过使用 PTS 录制器，您可以在 PC Web 浏览器或者移动端 APP 上模拟压测场景进行操作，从而轻松地记录下压测请求 URL（包括请求的 Header、Body 等信息）。

完成录制后可以将请求 API 导入到 PTS 压测场景中，快速发起一次压测，免除逐条录入压测请求的烦恼。您也可以对导入的请求 API 进行编辑，丰富压测场景。

录制器支持各种 PC 和手机端：Windows、MAC、iOS、Android。

使用限制

目前录制器功能在公测中，最高可用并发 1 万及以上的用户可以直接使用。若有需要请购买对应的 PTS 资源包。

录制的 HTTP 请求体的相关限制如下：

URL：最大6KB

BODY：最大60KB

HEADER：最大6KB

步骤一：进入录制器

登录 PTS 控制台。

在概览页中单击**录制器**，或在控制台左侧导航栏，单击**创建压测 > 云端录制**，即可进入录制页面。

说明：如需关闭录制器，可以在关闭录制器页面进行操作，或者停用代理。当前页面会被关闭，录制的请求不会被保留，请慎重操作。

步骤二：安装证书及设置代理

针对不同的设备端，代理设置和证书安装步骤有所不同，以下会逐一介绍操作流程。

代理设置			
证书下载地址:	http://record.aliyun.com/pts	点击下载	
代理服务器地址:	r.tb.cn	代理端口: <input type="text"/>	复制代理服务器和端口

说明：

当前录制器代理可持续运行 2 小时，2 小时之后录制器代理会关闭（下次再使用时是新的代理地址）。

首次使用时，请先安装证书再设置代理；证书安装一次后即可。

手机（移动设备）端

iOS 和 Android 操作系统的手机（移动设备）上均可运行，代理设置视操作系统略有区别。

iOS 安装证书及设置代理方法

Android 系统安装证书及设置代理方法

PC 端 Web 浏览器

PC 端浏览器版本较多，分别以 MAC/WIN 系统的 Chrome 和 Firefox 浏览器为例：

MAC 系统下 Chrome 及 Firefox 的证书安装与代理设置

WIN 系统下 Chrome 及 Firefox 的证书安装与代理设置

步骤三：开始录制

为便于分类录制需要的业务操作，可以使用建立步骤的方式，如登录、浏览、加购、提交订单等。

（可选）指定要录制的请求类型和域名。录制过程中，将自动抓取指定的请求。

新建步骤之后，单击**开始录制**，即可开始录制需要的请求。

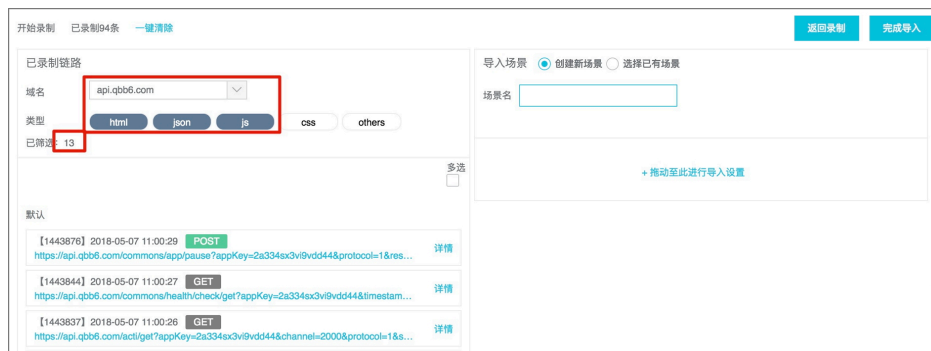


步骤四：导入场景

单击**暂停**后停止录制。

单击**导入**，将已录制的请求导入到场景中。

根据具体需求（域名以及 URL 类型）筛选出需要的请求进行导入：



导入到已有场景

选中**选择已有场景**，选择想要导入的场景，即可看到该场景中已有的信息，以灰色底色来表示该请求不可移除（已有场景中已保存的项）。



选择需要导入的请求（可多选）拖曳到右侧。若与已有请求没有顺序依赖关系，则拖曳到+ **新区域**即可。导入有两种方式：串行或者并行。

- **串行**：拖曳的请求会拖曳到一个串联链路中，有严格的先后顺序（顺序可再调整）
- **并行**：拖曳的 N 个请求之间均为并行模式。



导入到新场景

选中**创建新场景**，填写场景名称，拖曳需要的请求并设置导入方式即可。

完成导入

单击**完成导入**即完成本次导入，并创建或更新一个场景。导入后您可在场景编辑页面丰富压测场景。

Android 证书安装及代理设置

用 Android 手机录制压测请求前，需要安装证书并设置代理，操作方法如下（以华为手机为例）。

安装证书

登录 PTS 控制台。

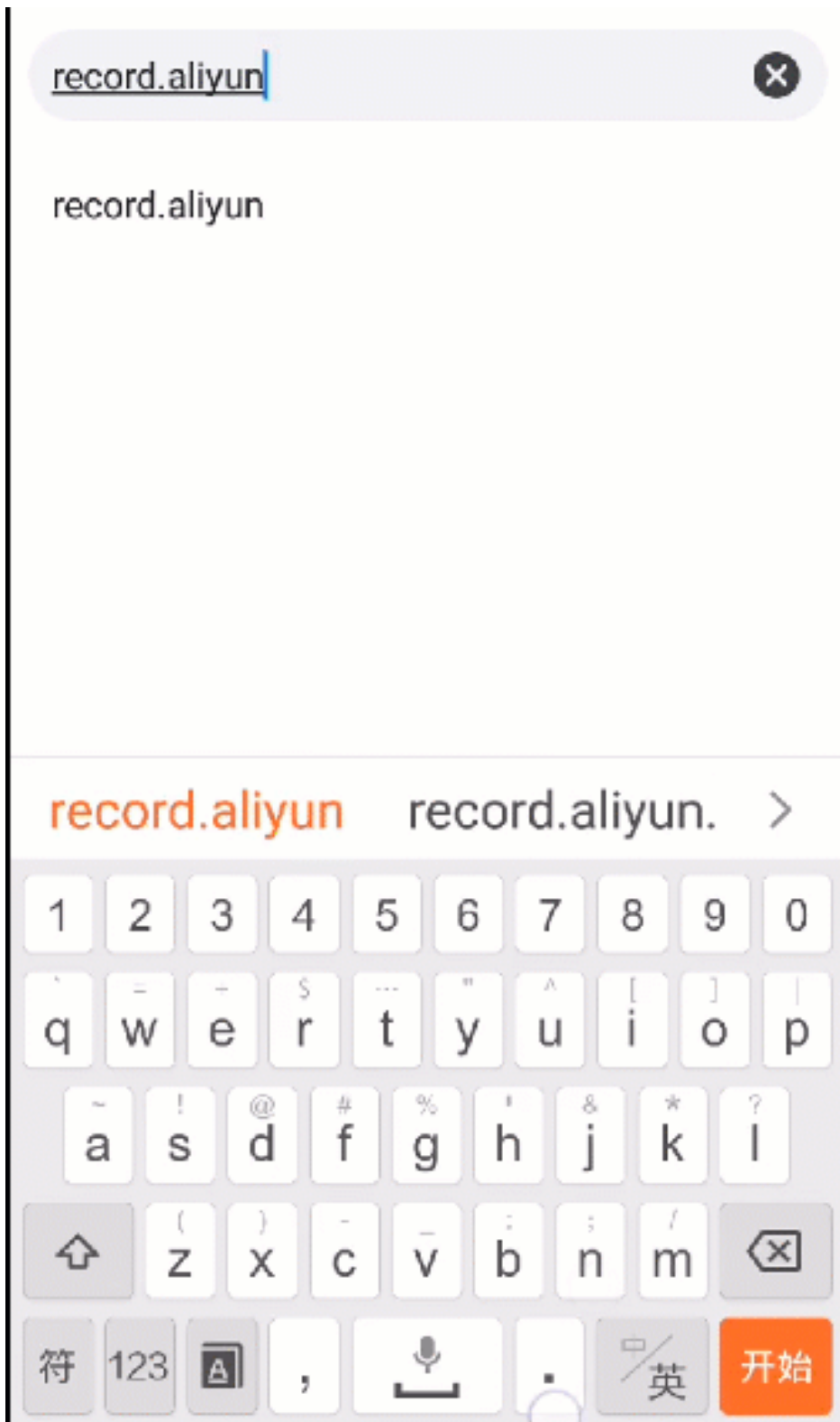
在概览页中单击**录制器**，或者在控制台左侧导航栏单击**创建压测 > 云端录制**。

在**录制场景**页面的**代理设置**区域框中，获取证书下载地址。

打开手机浏览器，输入证书下载地址，例如 <http://record.aliyun.com/pts>。

在手机浏览器界面上，单击 Android 图标，下载 PTS 录制器证书。

在手机弹出的证书安装器中，自定义证书名称，点击**确定**，进行描述文件的安装。



设置代理

在 PTS 控制台录制场景页面的代理设置区域，记录代理服务器地址和代理端口。

打开手机的**设置>WLAN**，长按所连 Wifi 并单击**修改网络**。

打开**显示高级设置**，配置代理为手工模式。此时需填写 PTS 分配的代理服务器地址及端口。



iOS 证书安装和代理设置

在 iOS 手机上录制压测请求前，需要安装证书并设置代理。

安装证书

登录 PTS 控制台。

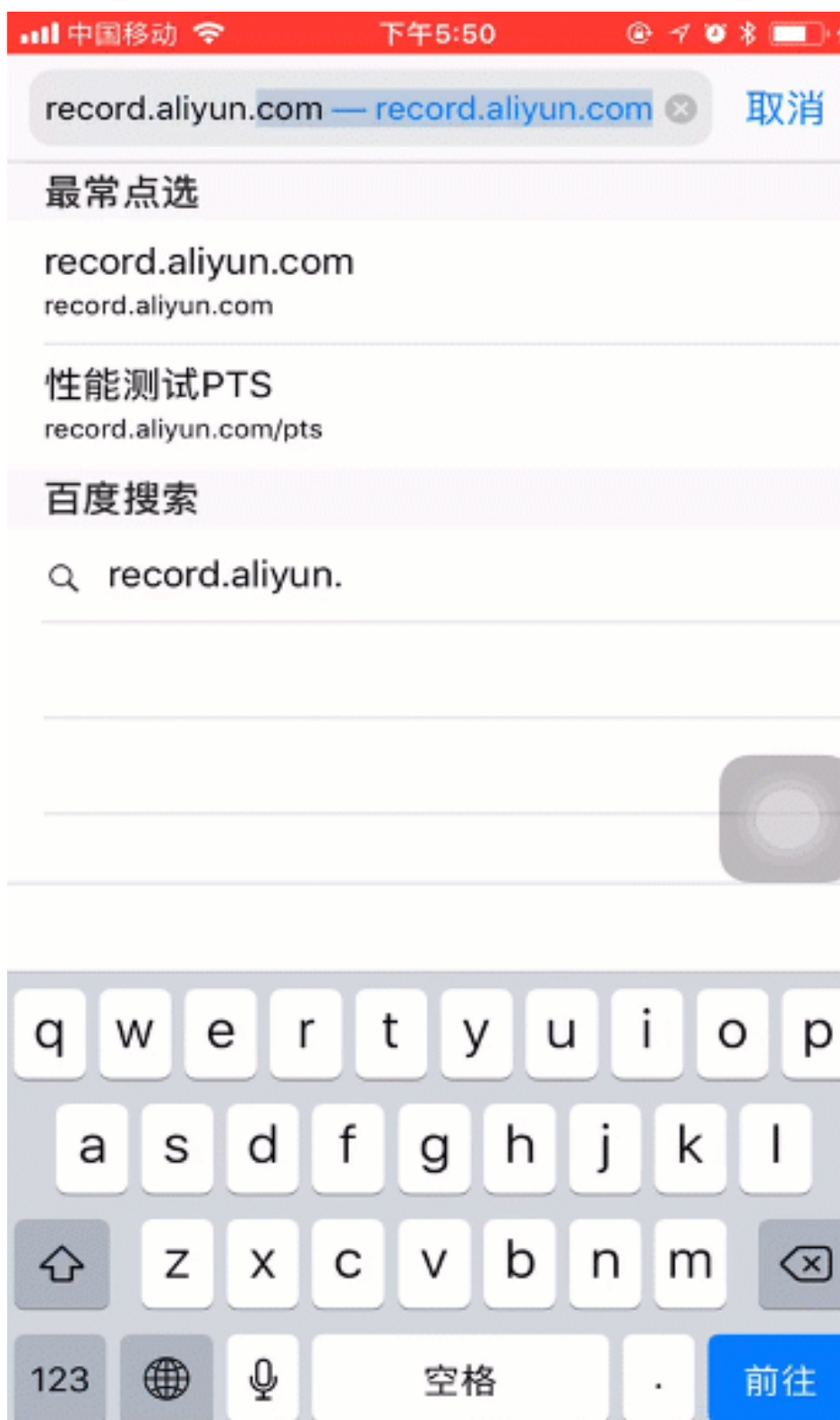
在概览页中单击**录制器**，或者在控制台左侧导航栏单击**创建压测** > **云端录制**。

在**录制场景**页面的**代理设置**区域框中，获取证书下载地址。

打开手机浏览器，输入证书下载地址，例如 <http://record.aliyun.com/pts>。

在手机浏览器界面上，单击 iPhone 图标，下载 PTS 录制器证书。

在手机浏览器弹框中，选择**允许**显示描述文件，单击**安装**。



iOS 设备还需要设置信任此证书，否则证书不能生效。在**设置>通用>关于本机>证书信任设置**中，打开**Ali-PTS**，**Record Proxy** 的信任。



设置代理

在 PTS 控制台录制场景页面的代理设置区域，记录代理服务器地址和代理端口。

打开手机的设置>无线局域网>所连 wifi 的更多设置（“i”标志）>配置代理。

配置代理选为手动。

在手机端配置代理界面，输入从 PTS 分配的代理服务器地址及端口。



Mac 下安装证书及配置代理

本文以 MAC 系统的 Chrome 及 Firefox 两款常用浏览器为例，介绍如何安装证书并设置代理。

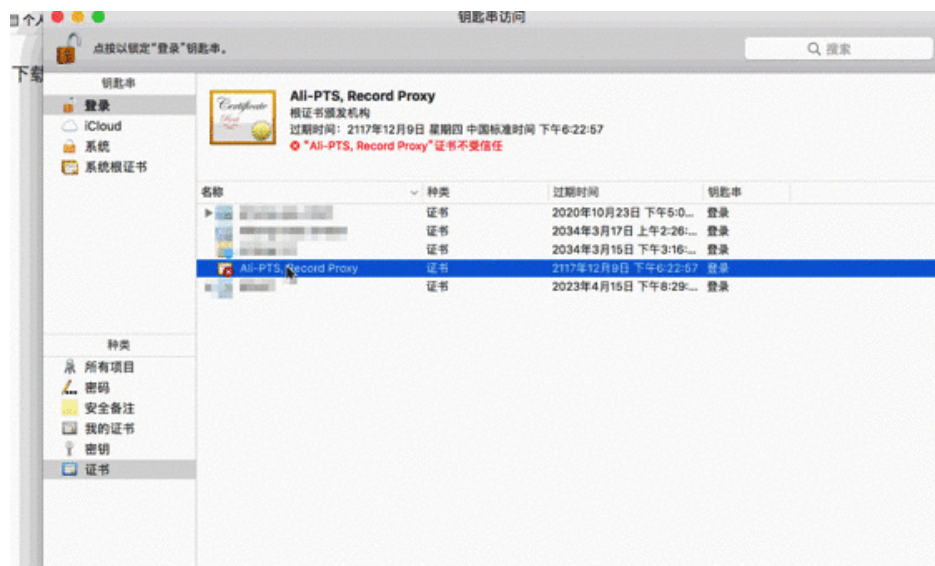
Chrome 浏览器

安装证书。

登录 PTS 控制台，在概览页中单击**录制器**，或者在控制台左侧导航栏单击**创建压测** > **云端录制**。

在**证书下载地址**后，单击**点击下载**，将自动下载适配您 PC 端的证书。

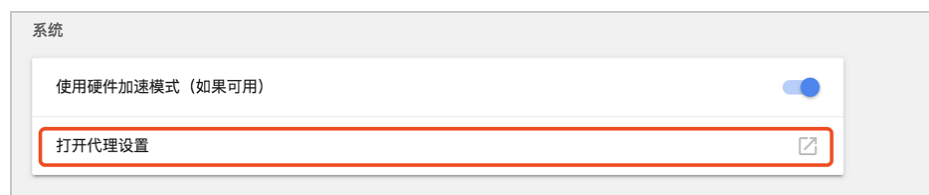
在 Chrome 浏览器中，导入已下载的证书，并添加信任，如下图所示。



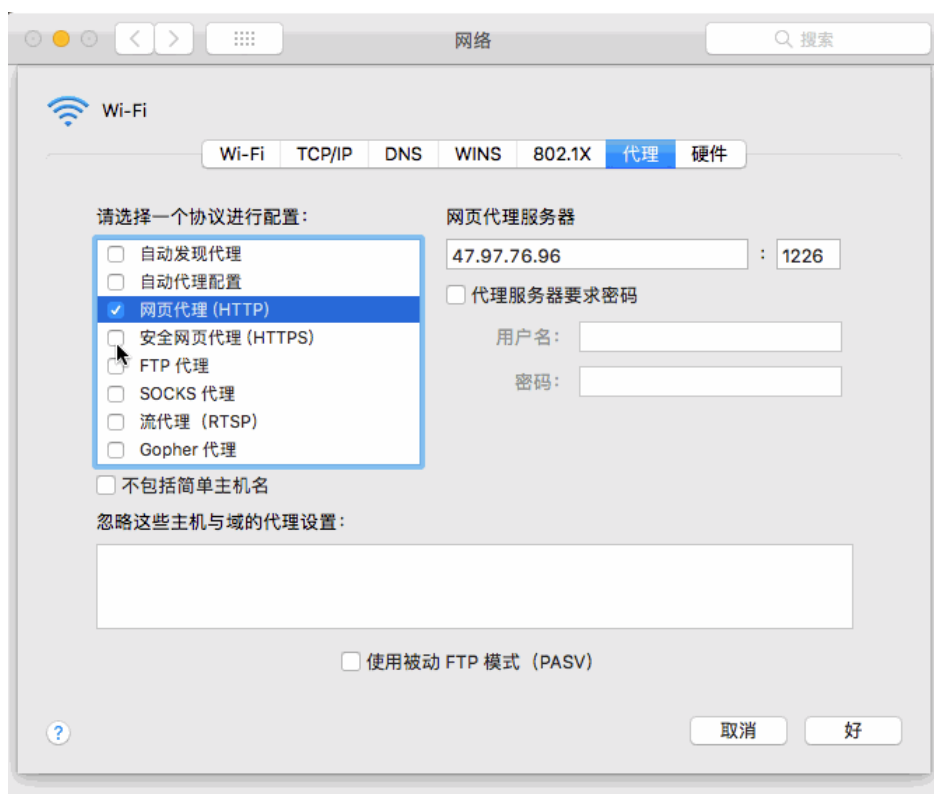
设置代理。

在 PTS 控制台**录制场景**页面，单击**代理服务器地址**后的**复制**，将复制代理服务器地址和端口。

在 Chrome 浏览器选项卡中，单击**设置** > **高级** > **系统** > **打开代理设置**。



填写代理服务器地址和端口，保存即可。



Firefox 浏览器

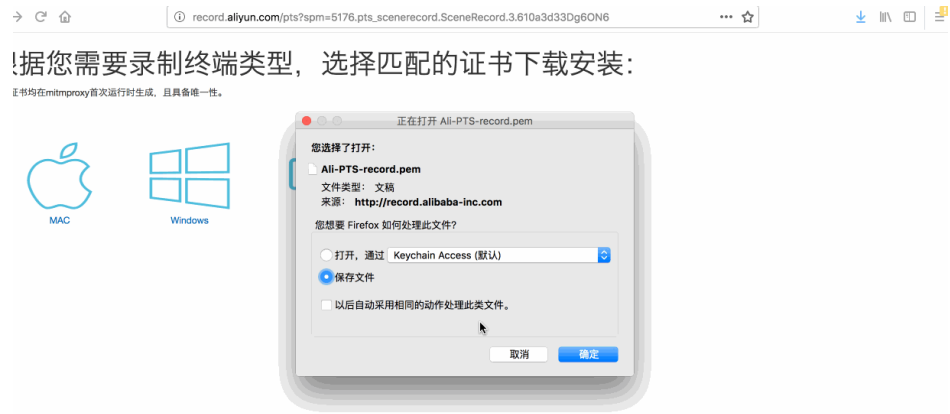
安装证书

登录 PTS 控制台，在概览页中单击**录制器**，或者在控制台左侧导航栏单击**创建压测 > 云端录制**。

在**证书下载地址**后，单击**点击下载**，将自动下载适配您 PC 端的证书。

在 Firefox 浏览器单击**首选项 > 隐私与安全 > 证书 > 查看证书**（页面底端），导入证书即可。

如下图所示：



!据您需要录制终端类型，选择匹配的证书下载安装：

证书均在mitmproxy首次运行时生成，且具备唯一性。

设置代理

在 PTS 控制台录制场景页面，单击代理服务器地址后的复制，将复制代理服务器地址和端口。

在 FireFox 浏览器中，选择首选项 > 常规（默认导航选项） > 网络代理（页面底端） > 设置。

选中手工代理配置，填写从 PTS 录制器界面获取的代理服务器地址及端口，保存即可。

如下图所示：



Win 下安装证书及配置代理

以 Win 系统的 Chrome 及 Firefox 两款常用浏览器为例，介绍如何安装证书并设置代理。

Chrome 浏览器

安装证书

登录 PTS 控制台，在概览页中单击**录制器**，或者在控制台左侧导航栏单击**创建压测** > **云端录制**。

在**证书下载地址**后，单击**点击下载**，将自动下载适配您 PC 端的证书。

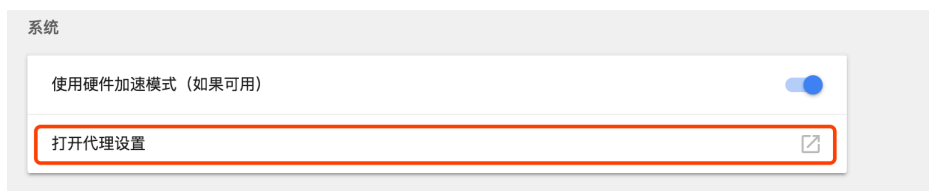
在浏览器中，选择**设置** > **高级** > **隐私设置和安全性** > **管理证书** > **受信任的根证书发布机构页签**，单击**导入**。按照提示选择上一步下载的 **Ali-PTS-record**，并逐步确认导入。



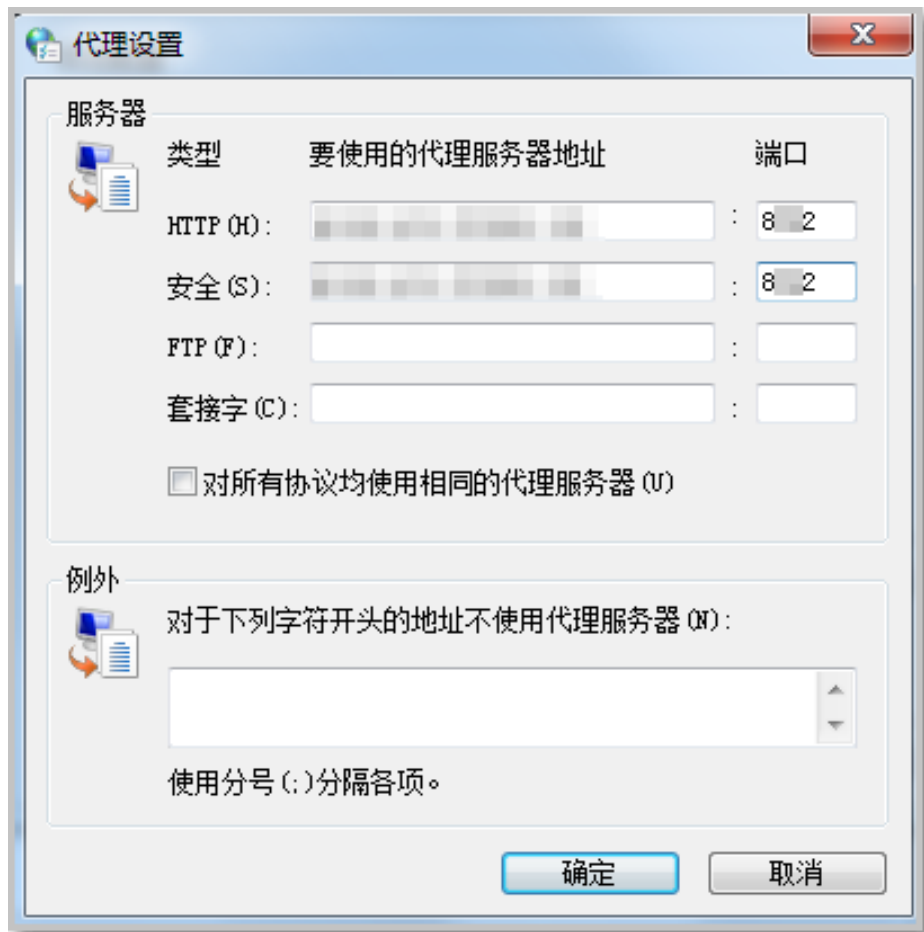
设置代理

在 PTS 控制台**录制场景**页面，单击**代理服务器地址**后的**复制**，将复制代理服务器地址和端口。

在 Chrome 浏览器中，选择**设置** > **高级** > **系统** > **打开代理设置**。



在 **Internet 属性** 对话框中，选择**局域网设置**>**代理服务器**>**高级**，填入在 PTS 控制台复制的代理 URL 地址。



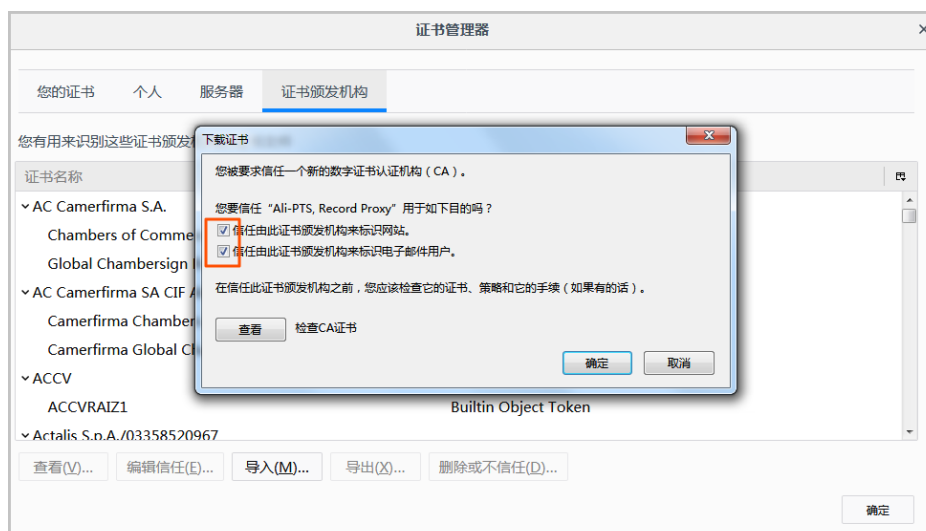
Firefox 浏览器

安装证书

登录 PTS 控制台，在概览页中单击**录制器**，或者在控制台左侧导航栏单击**创建压测** > **云端录制**。

在**证书下载地址**后，单击**点击下载**，将自动下载适配您 PC 端的证书。

在 Firefox 浏览器中，选择**选项 > 隐私与安全 > 证书 > 查看证书**（页面底端），在**证书颁发机构**页签下，单击**导入**，并勾选信任，将上一步下载的证书导入。



设置代理

在 PTS 控制台**录制场景**页面，单击**代理服务器地址**后的**复制**，将复制代理服务器地址和端口。

在 Firefox 浏览器中，选择**选项 > 常规**（默认导航选项）>**网络代理**（页面底端）>**设置**，勾选**手动代理配置**，在 PTS 控制台复制的代理 URL 地址。



JMeter

JMeter 原生压测

Apache JMeter 是 Apache 项目下的开源压测工具，有丰富的功能，支持参数化、断言等。PTS 支持 JMeter 原生引擎的压测，并在此基础上，为您提供便捷的横向扩展的能力，同时结合云监控集成功能持续补强 JMeter 在高并发和监控定位方面的能力和体验。

限制条件

- 最大并发用户数不能超过资源包的并发规格，参考产品价格。
- 目前该功能峰值并发用户数为 5 万；即使您购买的资源包最大并发用户数是 5 万及以上，也受此限制。
- 该功能和 PTS 原生功能共享并发用户数的额度。例如，您持有有一个 278（最大并发用户数 5000）的资源包，PTS 原生压测场景中使用了 1000 并发，那么同时进行的 JMeter 集成功能的压测只能使用 4000 最大并发来启动。
- 由于共享施压资源，如果碰到压测引擎不足的报错，请稍后重试。

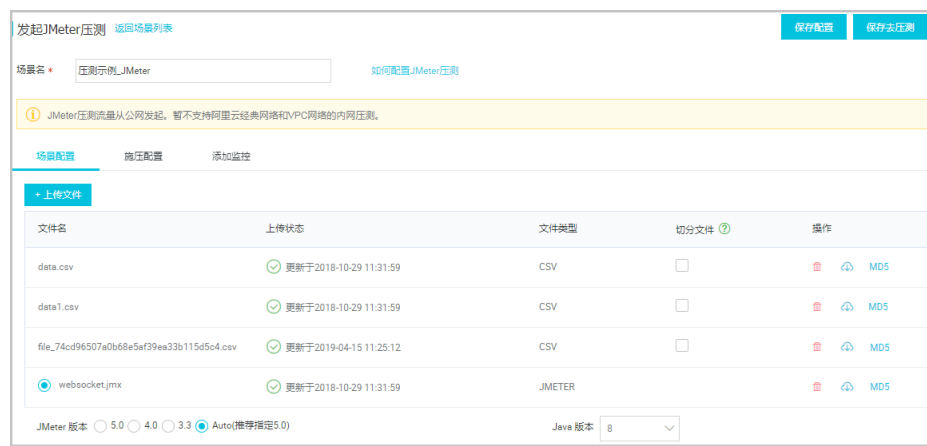
步骤一：导入 JMeter 脚本

登录 PTS 控制台，在概览页单击 **JMeter 原生压测**，或者在左侧导航栏选择 **创建压测 > JMeter 压测**。

在**发起 JMeter 压测**页面，输入场景名称。

在 **场景配置**页签下，单击**+上传文件**，添加 JMeter 脚本、JAR 文件（即放置在 JMeter 的安装目录 `.../lib/ext` 下的相关 JAR 文件）和其他数据文件。

注意：上传的同名文件将覆盖之前上传的文件。如需对比文件，可在该文件的**操作列**，获取 MD5 值，与您本地文件的 MD5 值进行对比，确认文件是否变更。



使用限制：

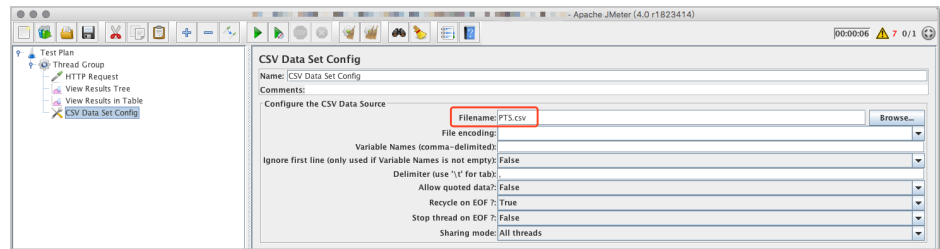
- **JMX 脚本文件：**文件大小不超过 2MB。可上传多个 JMX 脚本文件，但是压测仅能使用一个 JMX 脚本。发起压测前勾选一个使用。
- **JAR 文件：**文件大小不超过 10MB。上传之前，请在本地 JMeter 环境调试 JAR 文件，确保调试通过。

CSV 和 ZIP 数据文件：文件大小不超过 60 MB。ZIP 内打包的文件必须是以 .csv 结尾的单个文件。

注意：

- a. 准备 CSV 文件时，请不要直接修改 .xlsx 文件的后缀。推荐使用 EXCEL、Numbers 等软件导出，程序生成可以使用 Apache commons-csv。

如果 JMX 脚本中关联了数据文件，务必将 **CSV Data Set Config** 中的 **Filename** 改为文件名本身，**不要带路径**（如下图所示）。否则会读取不到数据文件。同理，如果使用了 `__CSVRe` 函数或者在 JAR 中有指定文件，均需使用文件名。



如上传了多个 JMX 文件，勾选一个 JMX 文件作为发起该压测场景使用的 JMeter 脚本。

（可选）为 CSV 文件勾选**切分文件**，可以保证该文件的数据在每个施压机不重复。不勾选该项，每个施压机上使用同一份数据。详见 JMeter 使用 CSV 参数文件。

选择 JMeter 版本 和 Java 版本。

- JMeter 版本：支持 Apache JMeter 5.0、4.0 和 3.3。
- Java 版本：支持 Java 8。

步骤二：施压配置

单击**施压配置**页签，配置以下项：

配置项	说明
指定 IP 数	指定施压引擎的数量，此项为可选项。引擎数增加会带来计费增加，计费并发 = 引擎数 * 500，且最大引擎数不超过 100。
并发数	指同时发送压测请求的用户数量。 并发数 受您购买的资源包限制。
指定 IP 数	选择该选项可扩展施压的引擎数。 注意： 1. 增加引擎数会带来计费增加，请关注 施压配置 页面右侧的预估消耗 VUM 数。 2. 可设置的最大引擎数受资源包限制，具体可参见 指定IP数 的帮助提示。
压测时长	指压测总时长，包括预热时间。
预热时间	指压测流量到达已配置的 并发数 的时间。在这个时

	间内，并发数递增，直到最大并发数。
循环次数	指每个并发的循环执行次数。 注意：循环次数达到设置值，或压测时长达到设置值，压测都会停止。

步骤三（可选）：添加监控

如果您有云监控或 ARMS 监控实例，通过添加监控的方式，您可以在压测及报告中便捷地查看相应的监控数据。

。

具体操作及监控数据的查看步骤，参见以下文档：

- 云监控
- ARMS 监控

步骤四：启动压测

完成基本配置之后，在**发起 JMeter 压测**页面右上角，单击**保存去压测**，发起压测。

稍等数秒，操作页会进入压测态。页面概览如下。



压测过程中，您可以监控压测数据，进行如下操作：

在**详情页**签下，单击单个链路的右侧图标



，查看其 TPS、成功率、响应时间、流量等实时数据。

选择 **JMeter 日志** 页签，可根据施压机、日志级别、时间范围或施压线程筛选日志，并查看日志详情。

单击页面右上角的**查看采样日志**，可根据压测 Sampler、响应状态、施压机等筛选日志，查看日志详情。

场景说明

场景概要的指标说明如下。

指标	说明
VUM	PTS 的计费单位
压测计时	压测执行的时长，包括预热阶段的时间
场景并发	当前施压的并发值，如果在预热阶段则未达到配置并发，预热结束后即为配置的并发值
场景 TPS (s)	所有 agent 统计周期内的平均 TPS 值再做一次平均所得
RT Avg (ms)	所有 agent 统计周期内的平均 RT 值再做一次平均

	所得（区分成功与失败）
成功率（%）	所有 agent 统计周期内请求合并计算的整体成功（无断言时 2XX、3XX 的比例；有断言时为断言成功的比例）情况

说明：监控数据的来源基于 Backend Listener 进行了简单的聚合计算。施压 Agent 的统计采样周期是 15 秒，数据汇总计算的周期也是 15 秒，故可能会有数据延迟的情况。

JMeter 日志详情

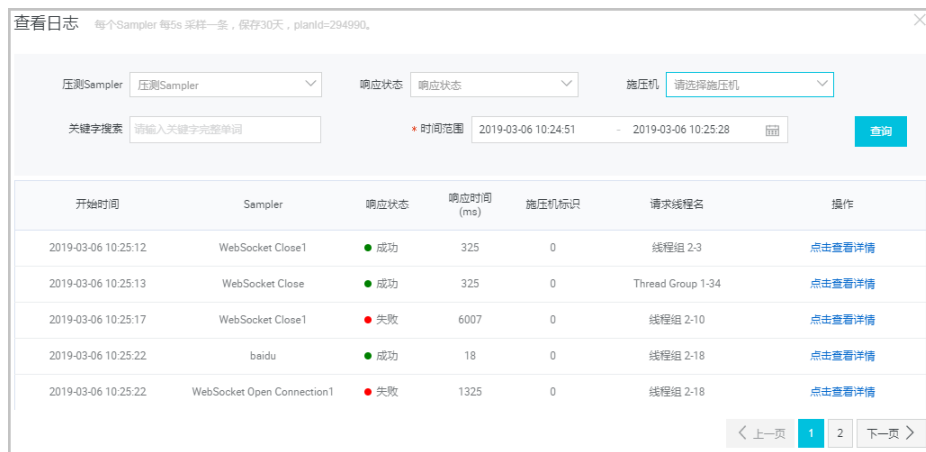
JMeter 日志详情包括日志时间、日志级别、线程名、logger 和具体日志信息。



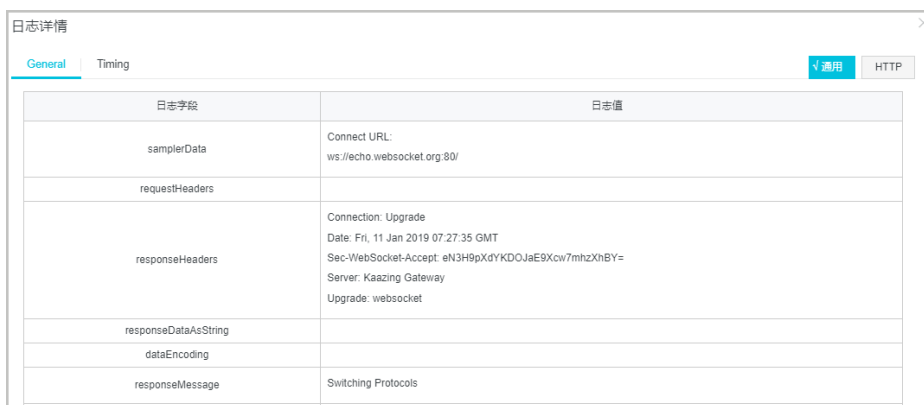
采样日志详情

单击报告页面右上角的**查看采样日志**，可筛选并查看日志详情。

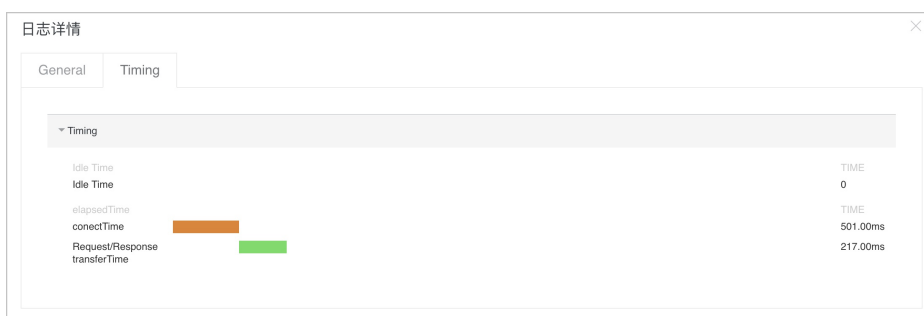
根据压测 Sampler、响应状态等筛选日志。在对应的日志操作栏，**点击查看详情**。



在日志详情对话框中，在 **General** 页签下查看日志字段和字段值。在对话框右上角切换显示格式：**通用**或 **HTTP 协议**样式。

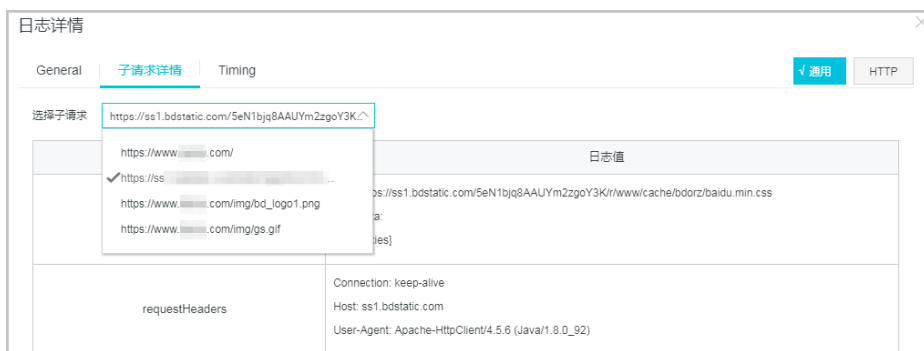


在 **Timing** 页签下查看耗时信息。

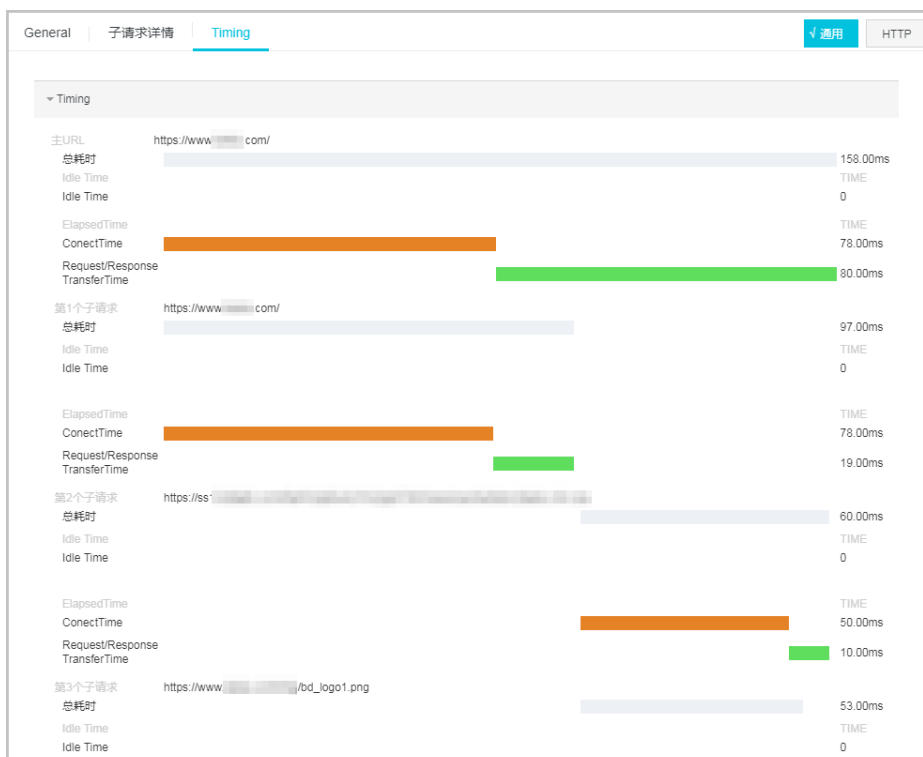


如果 JMeter 脚本中配置了从 HTML 文件嵌入资源 (Embedded Resources from HTML Files)，日志详情中会出现子请求详情页签。

通过选择特定的子请求，可筛选出对应的请求日志。



Timing 页签下也将显示总请求耗时和每一个子请求的耗时情况。



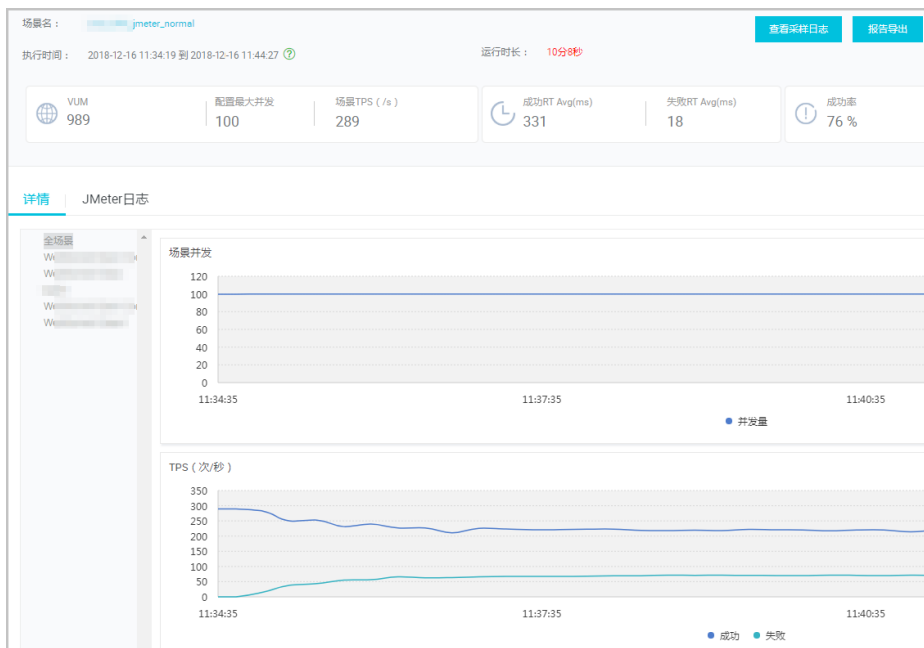
步骤五：查看压测报告

在压测停止之后，系统会自动收集压测过程中的数据，形成压测报告，展现 JMeter 压测场景和具体指令的压测动态信息。

在 PTS 控制台左侧菜单栏，单击**压测报告**，进入压测报告列表页。

在对应的报告一行的**操作**栏，单击**查看报告**。

说明：趋势图中，每个点数据的统计采样周期是 15 秒，最终数据报告上可能初期会有数据延迟的情况。



场景概要的指标说明如下。

指标	说明
VUM	PTS 的计费单位
配置最大并发	JMeter 脚本压测编辑时配置的并发值
场景 TPS (s)	压测周期内，当前场景的平均 TPS 值。TPS = 场景压测期间的全部请求数 / 压测时长。
RT Avg (ms)	压测周期内，所有请求的平均响应时间（区分成功与失败）
成功率 (%)	所有请求的整体请求成功（无断言时 2XX、3XX 的比例；有断言时为断言成功的比例）情况

（可选）在**场景详情**动态视图区域，您可以查看全场景或单个请求，在整个压测周期内的具体指标表现。

- （可选）在 **JMeter 日志** 页签下，可根据施压机、日志级别、时间范围或施压线程筛选日志，并查看日志详情。参考 **JMeter 日志** 详情。
- （可选）在页面右上角单击**查看采样日志**，可筛选日志并查看请求日志中展现的字段、字段值和耗时详情。参考**采样日志**详情。

步骤六（可选）：导出压测报告

如您需要，可以将当前压测报告导出至本地。

- 在**测试报告**页，单击右上角的**报告导出**，选择是否包含**阿里云 PTS 水印**版本。
- 进入网页版报告页面，单击**下载报告**，可将压测报告（PDF）下载至本地。

相关文档

- 关于更多测试指标的解读，参见测试指标。
- 关于测试分析和调优，参见测试分析及调优。

JMeter 导入 API 压测

除了利用 JMeter 脚本直接进行 JMeter 原生压测以外，您也可以选择将 JMeter 脚本转换为 PTS 原生的能力进行压测。PTS 具备更强的 RPS 压测模式，支持实时调速、细致到单接口的监控等。

支持导入的信息

对 HTTP 或 HTTPS 协议的压测，PTS 支持将 JMeter 脚本中有关请求的基础信息解析为压测场景中的 API。JMeter 脚本 与 PTS API 压测的字段对应关系如下：

Jmeter 脚本字段		PTS-API 压测字段
测试计划 Test Plan		压测场景
线程组 Thread Group		串联链路
采样器 Sampler (仅支持 HTTP Request)	Name	压测 API 名称
	Server Name or IP Path Port Number	压测 URL
	Method	请求方法
	Parameters/Body Data	Body 定义下的内容
HTTP Header Manager	Headers	Header 定义下的 Key 和 Value。 - 全局或者区域生效的 Header 会填充到对应区域的 API 中。
HTTP Cookie Manager	User-Defined Cookies	Header 定义下的 Cookie。 - 全局或者区域生效的 Cookie 会填充到对应区域的 API 中。

说明：

- HTTP Header Manager、HTTP Cookie Manager和HTTP Request Default导入时会取所有的并集到每个API中；
- Cookie相关的设置，domain和path不支持导入；
- 导入时Content-Type默认都为x-www-form-urlencoded，如果需要切换为raw请先复制，否则会清空

BODY ;

- 对于不支持导入的信息，您可以使用 PTS 提供的功能实现，例如断言、指令（思考时间、集合点等）、参数化等。

操作步骤

登录 PTS 控制台，在概览页单击**快速压测**，或者在左侧导航栏单击**创建压测** > **快速压测**。

说明：不支持在已有场景中导入 JMeter 脚本。

单击**创建压测**页面右上角的**导入脚本**。



单击上传图标，上传 JMeter 脚本。

说明：一个压测场景仅能导入一个 .jmx 脚本。在未保存场景配置前，可重新上传脚本覆盖之前的导入内容。

单击**确定**，可预览导入的脚本。可导入内容将以蓝色高亮条显示，如下图所示。



单击预览窗口右下角的 **导入 API 压测**。

导入成功后，在**创建压测**页面，您可以看到脚本内容已转化为对应的 API 压测信息。



后续操作

脚本导入成功后，您可以使用 PTS 提供的高级功能编排压测场景并发起压测，可参考以下文档：

- 出参与断言
- 如何使用指令（思考时间、集合点、条件跳转等）
- 参数化的使用方法
- 如何进行施压配置

相关文档

JMeter 原生压测

JMeter 使用 CSV 参数文件

在 PTS 原生压测中，如果您需要构造动态变化的业务请求，可以通过参数化实现。在使用 JMeter 压测时，如果有同样的动态压测需求，建议您在 JMeter 脚本中添加 Config Element：**CSV Data Set Config**，从 CSV 文件中读取参数。

同时，PTS 支持切分 CSV 文件，将 CSV 文件中的数据分配到每个压测引擎。

背景信息

CSV Data Set Config 是 JMeter 官方提供的配置元件，用于读取 CSV 文件中的数据并将它们拆分为变量。适用于处理大量变量的场景。

如果上传的 CSV 文件在 JMeter 脚本中没有对应的 **CSV Data Set Config** 配置，则切分 CSV 文件时默认作为不包含表头（即变量名）的 CSV 文件切分。为方便操作，并避免数据被错误切分，建议您始终通过配置 **CSV Data Set Config** 使用 CSV 参数文件。

配置 CSV 参数文件

为构建动态变化的压测请求，需要在 JMeter 脚本中配置 CSV 文件。

例如，为测试同时有 100 人登录某网站时，该网站登录接口的性能表现如何，则需要构造 100 条不同的用户登录数据。操作步骤如下：

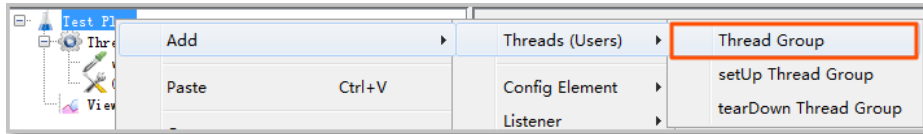
在本地新建一份 CSV 文件，包含 100 条用户数据。该 CSV 文件示例中第一行是表头，表示变量名，所以文件共 101 行。基本格式示例如下：

```
id,name  
1,ali
```

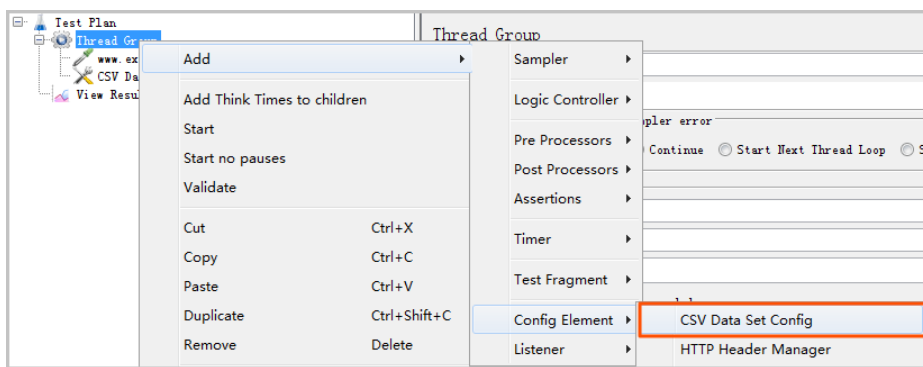
2,pts
3,jmeter

说明：如果 CSV 文件不含表头，CSV Data Set Config 中需要配置变量名。

启动 JMeter GUI，右键单击测试计划，选择 **Add > Threads (Users) > Thread Group**，添加线程组。

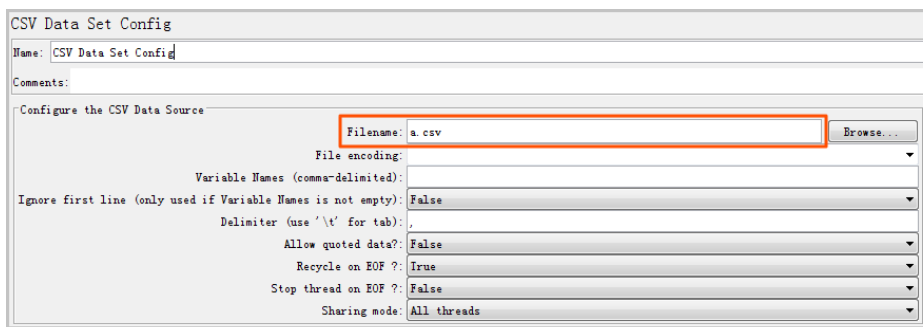


右键单击新添加的线程组，选择 **Add > Config Element > CSV Data Set Config**。



在 **CSV Data Set Config** 配置区域，填写 CSV 文件名，其他配置项保持默认。

JMeter 会自动从 CSV 文件的表头中读取参数名作为变量名。

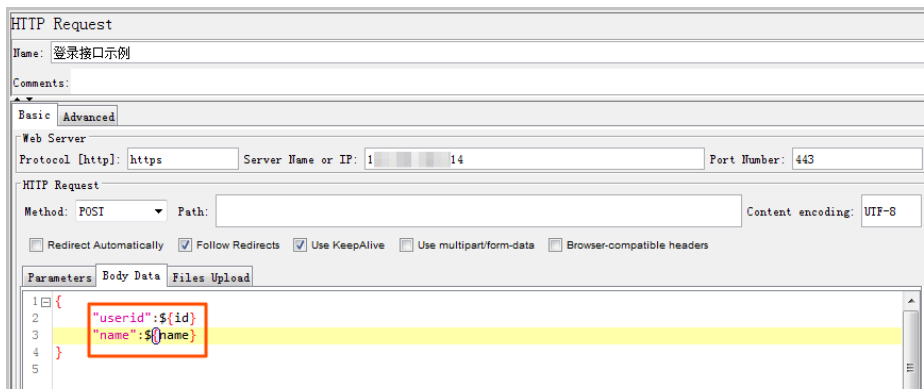


说明：如果 CSV 文件不含表头，需要在 CSV Data Set Config 的 **Variable Names (comma-delimited)** 中，输入 *id,name*。

右键单击线程组，选择 **Add > Sampler > HTTP Request**。根据具体业务接口，填写配置项，示例如下：

- **Name**：登录接口示例

- **Web Server**：填写具体的业务接口协议、服务名或 IP 地址和端口号。
- **HTTP Request**：选择请求方法，填写 Path 和 Content encoding 格式。
- **Body Data** 页签：使用 \${id}、\${name}，表示读取 CSV 文件中的参数。



保存该 JMeter 测试计划脚本。

将测试脚本和 CSV 文件上传至 PTS。具体步骤请参见 JMeter 原生压测。

勾选 CSV 文件后的切分文件复选框，可将 CSV 文件中的数据切分到每个压测引擎上。具体说明参见下一章节 CSV 文件切分。

如下图所示：



CSV 文件切分

PTS 自动根据场景并发数分配 JMeter 施压引擎。分配多个引擎时，PTS 支持切分 CSV 文件，将 CSV 文件中的数据均匀分配到每个压测引擎。

如果数据条数不能完全均分，则部分引擎可能比其他引擎多一条数据。

包含文件头的 CSV 文件切分

当原 CSV 带文件头时，切分后的每个文件都会包含文件头，但保证数据不重复。

如上述示例文件切分到 2 个引擎时：

第 1 个文件内容为：

```
id,name  
1,ali  
2,pts
```

第 2 个文件内容为：

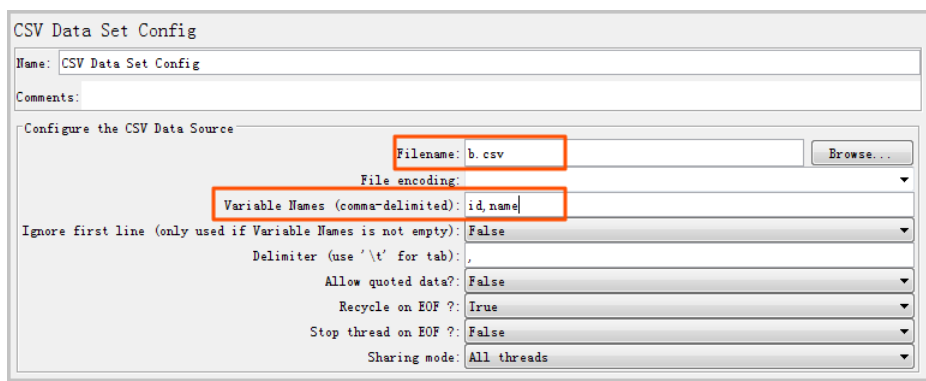
```
id,name  
3,jmeter
```

不含文件头的 CSV 文件切分

如果 CSV 文件不包含文件头（第一行即为 CSV 数据），例如，文件 “b.csv” 内容如下：

```
1,ali  
2,pts  
3,jmeter
```

则需在 JMeter 中配置变量名，如下图所示：



此时如果选择切分 CSV 文件，切分后的文件也不包含文件头。如切分到 2 个引擎时：

第 1 个文件内容为：

```
1,ali  
2,pts
```

第 2 个文件内容为：

3.jmeter

更多信息

更多 CSV 文件配置和使用说明，参见 JMeter 官方文档。

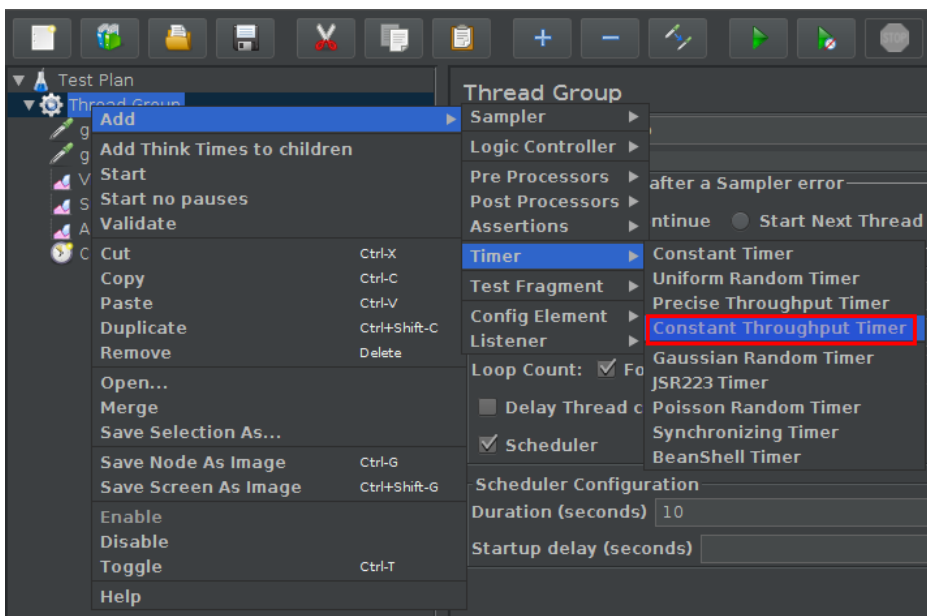
JMeter 配置 RPS 限制

RPS 即每秒请求数（Request Per Second），通常用来描述施压引擎实际发出的压力大小。更适合 API 接口测试的场景。下面介绍一下基于 JMeter 实现 RPS 控制的方法。

添加 RPS 限制节点

为了复用已有的 JMeter 脚本，PTS 支持使用原生 JMeter 引擎进行压测。JMeter 引擎默认使用并发模式压测，施压 RPS 主要由并发数和服务器响应时间决定。并发数过低时可能达不到预期的 RPS，并发数过高时可能压力过大压垮服务器。为了避免 RPS 压力过大压垮服务器，JMeter 提供了 Constant Throughput Timer 组件，可限制最大施压 RPS 压力大小。

JMeter 脚本线程组下添加“Constant Throughput Timer”节点，即可对该线程组发出的 RPS 压力大小进行限制。操作如下：



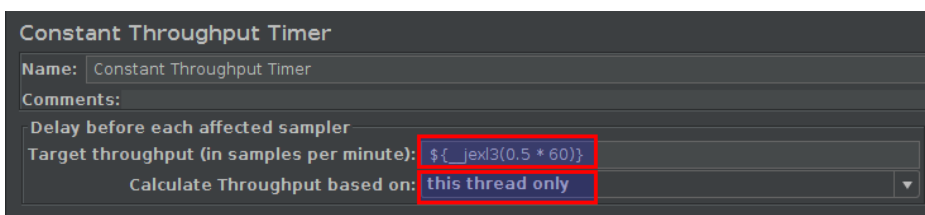
配置 RPS 限制

配置 “Constant Throughput Timer” 节点时有两点需要注意：

“Calculate Throughput based on” 必须设置为 “this thread only”，即按单个线程限制。PTS 自动根据场景并发数分配多个 JMeter 引擎，其他选项只能对单个引擎进行限制，无法有效的进行全局 RPS 限制。所以可以通过设置单个线程的限制和并发数达到整体RPS限制的效果。

“Constant Throughput Timer” 的 “Target throughput” 以分钟为单位配置，RPS 值换算成分钟时，需要乘以 60（1分钟=60秒），可使用 `__jexl3` 进行计算。如期望 **单线程 RPS 为 0.5**，则 “Target throughput” 可设置为 `$_jexl3(0.5 * 60)`。如果要指定全局 RPS 限制，可用全局 RPS 限制除以并发数，得到单个线程的 RPS 限制。

配置示例如下：



综上，可以看出JMeter的RPS控制比较复杂，不但依赖于设置合理的并发，而且无法实时调整并发，RPS不一定能完全实现需要的效果。对于RPS的压测模式，建议直接使用PTS原生场景进行压测，支持直接的RPS设置和秒级调整，完全不用考虑并发的概念。

JMeter 多个线程组的使用说明

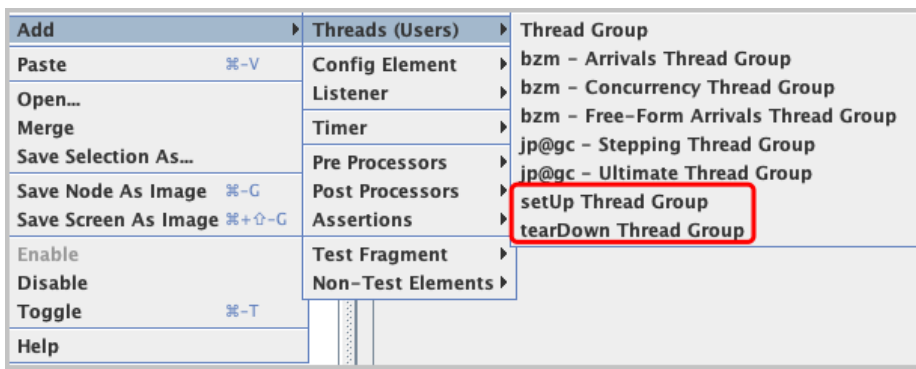
当 JMeter 测试计划中存在多个线程组，您需要了解如何结合 JMeter 和 PTS 配置参数，使多个线程组并行或串行压测。

前提条件

已掌握 JMeter 原生压测的基本操作。具体参见 JMeter 原生压测。

背景信息

JMeter 线程组包括 setUp 线程组、tearDown 线程组和主线程组。在 PTS 中的施压配置不会影响 setUp 和 tearDown 线程组，只影响主线程组。



在 PTS 中配置的并发数、预热时间、循环次数会覆盖主线程组在 JMeter 脚本的配置。



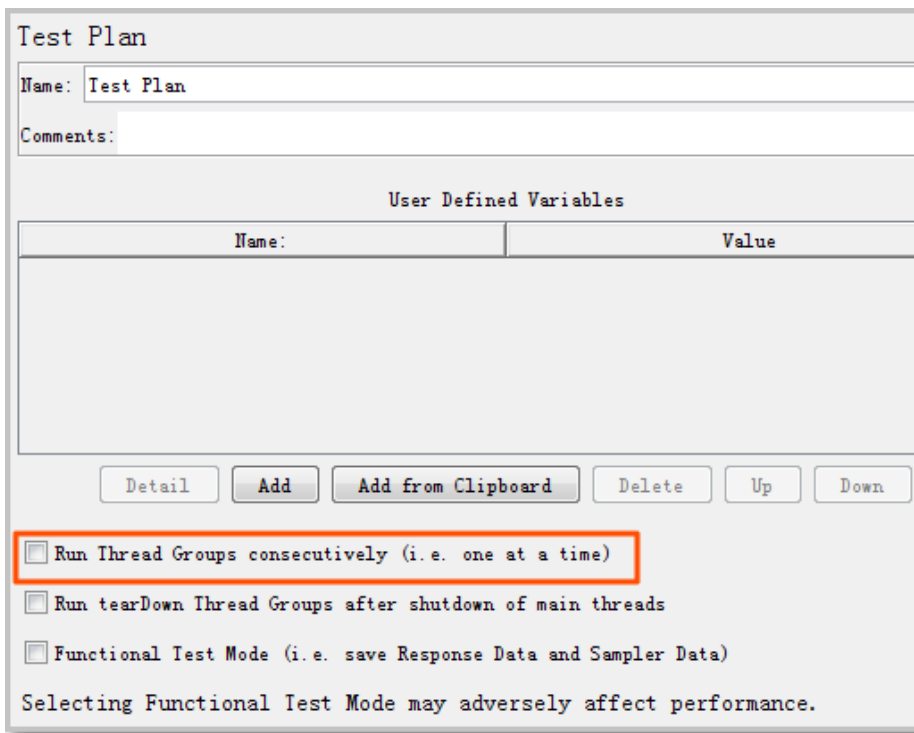
设置多线程组并行或串行

如果 JMeter 脚本中包含多个主线程组 (Thread Group) ，您可以参照以下说明设置多线程组的并行和串行关系。

PTS 施压配置中的**循环次数**会覆盖所有的线程组，而**压测时长**一旦到达，整个压测都会停止。

多线程组并行

多个线程组并行时，在本地 JMeter 脚本中，请**不要**勾选测试计划 (Test Plan) 属性中的**独立运行每个线程组** (例如在一个组运行结束后启动下一个) ，即 *Run Thread Groups consecutively(i.e one at a time)*。

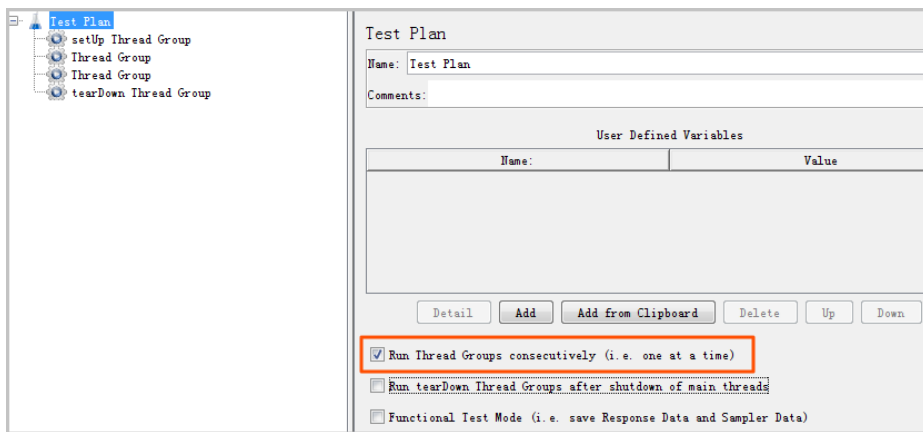


多个线程组并行时，执行的时间由循环次数和压测时长中先完成的决定。例如，压测时长设置为 10 分钟，而循环次数设置为 5 次，假设循环 5 次只需要 5 分钟，那么压测在 5 分钟时就停止了。

多线程组串行

串行的情况会稍微复杂一点，操作如下：

在本地调试 JMeter 脚本时，勾选测试计划（Test Plan）属性中的**独立运行每个线程组**（例如在一个组运行结束后启动下一个），即 *Run Thread Groups consecutively(i.e one at a time)*。



在 PTS 上设置**循环次数**。该循环次数会作用于每个线程组。

例如，测试计划中有 A、B、C 三个线程组，在 PTS 施压配置中，设置**循环次数**为 5，那么执行顺序

是先基于 A 的并发数循环 5 次，之后基于 B 的并发数循环 5 次，最后是 C 的并发数循环 5 次。

在 PTS 上设置的**压测时长**需要足够长，避免在压测中多线程串行的过程中被中断。

预估的压测时长 = 业务请求的 RT * 总请求数。您可以在预估的压测时长基础上适当延长。

更多信息

JMeter 原生压测

JMeter 压测报告

使用 JMeter 压测模式完成压测后，系统会自动收集压测过程中的数据，形成压测报告。本文将帮助您了解 JMeter 压测模式下的压测报告各部分详情。

功能入口

在 PTS 控制台左侧导航栏，单击**压测报告**。

在**测试报告列表**页面，可根据场景类型（PTS 或 JMeter）、时间、场景名称或报告 ID 等关键字进行搜索。

说明：JMeter 压测模式下的压测报告会有 JMeter 标签。



报告ID	场景名	峰值并发	消耗VUM	压测时长	压测开始时间	操作
529866	JMeter_1000并发	100	1000	10分1秒	2019-07-25 20:05:09	查看报告 删除
529849	PTS_1000并发	20	118	12分钟	2019-07-25 19:51:54	查看报告 删除
529769	JMeter_1000并发	100	1000	10分3秒	2019-07-25 18:57:58	查看报告 删除
529709	PTS_1000并发	1	100	40分钟	2019-07-25 18:19:03	查看报告 删除
529654	JMeter_400并发	400	4372	22分钟	2019-07-25 17:59:07	查看报告 删除
529700	PTS_50并发	50	100	39秒	2019-07-25 18:16:15	查看报告 删除
529646	PTS_1000并发	4	100	9分34秒	2019-07-25 17:55:19	查看报告 删除
529444	PTS_1000并发	40	360	12分钟	2019-07-25 16:37:01	查看报告 删除
529342	PTS_1000并发	10	100	2分钟	2019-07-25 16:08:00	查看报告 删除
528898	PTS_50并发	50	100	2分3秒	2019-07-25 11:39:52	查看报告 删除

在目标压测报告的操作列，单击**查看报告**。

单击**测试报告**页面右上角的**查看采样日志**，可以查看采样日志的日志字段和字段值，以及各请求的耗时信息。详情请参见采样日志。

单击**测试报告**页面右上角的**报告导出**，再根据需要选择**是否包含阿里云 PTS 水印**，导出 JMeter 压测模式的压测报告。

压测场景预览

压测报告第一部分显示整个压测场景的相关业务指标。



指标说明：

指标	说明
VUM	本次压测消耗的资源总数，单位是 VUM（每虚拟用户每分钟），参见计费规则。
配置最大并发	当前施压的并发值，如果在预热阶段则未达到配置并发，预热结束后即为配置的并发值。
场景 RPS (s)	所有 Agent 统计周期内的平均 TPS 值再做一次平均所得。
成功 RT Avg (ms)	所有成功请求的平均 RT。
失败 RT Avg (ms)	所有失败请求的平均 RT。
成功率	所有 Agent 统计周期内请求合并计算的整体成功情况（无断言时 2XX、3XX 的比例；有断言时为断言成功的比例）。
总请求数	整个场景在压测过程中发出去的总请求数。

业务指标

压测报告的**概览**页，显示整个场景下所有的全场景和所有 Sampler 的压测情况。

业务指标	总请求数	平均TPS	成功率	平均响应时间
全场景	2160	23.07	0% (346/0)	1735.49ms 详情
home	1096	11.60	0% (174/0)	2029.88ms 详情
download page	1064	11.47	0% (172/0)	1432.24ms 详情

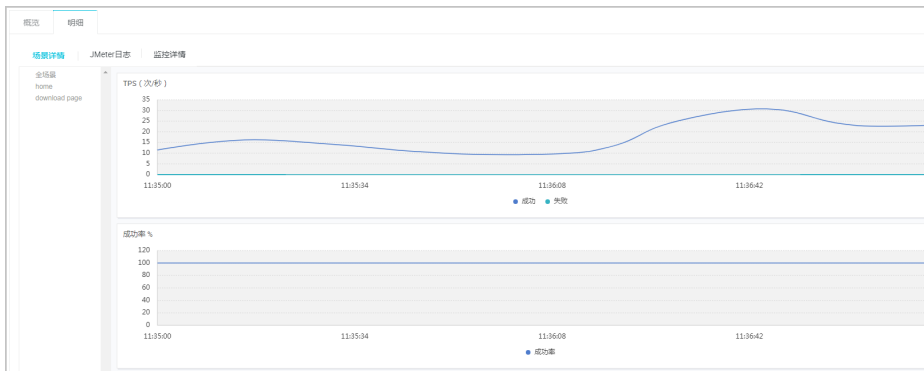
业务指标说明：

业务指标	说明
Samler 名称	包括全场景和所有 Sampler 的名称。
总请求数	整个场景在压测过程中发出去的总请求数。
平均 TPS	压测周期内，当前场景的平均 TPS 值。TPS = 场景压测期间的全部请求数/压测时长。
请求成功率	压测中此 API 的请求成功率。 - 单击请求成功/失败个数，可快捷查看对应日志； - 单击 详情 ，查看 3XX、4XX、5XX 和其他异常导致的请求失败的个数。
平均响应时间	压测中此 API 的平均响应时间。单击 详情 ，查看最大、最小及其各分位的响应时间。

说明：监控数据的来源基于 Backend Listener 进行了简单的聚合计算。施压 Agent 的统计采样周期是 15 秒，数据汇总计算的周期也是 15 秒，故可能会有数据延迟的情况。

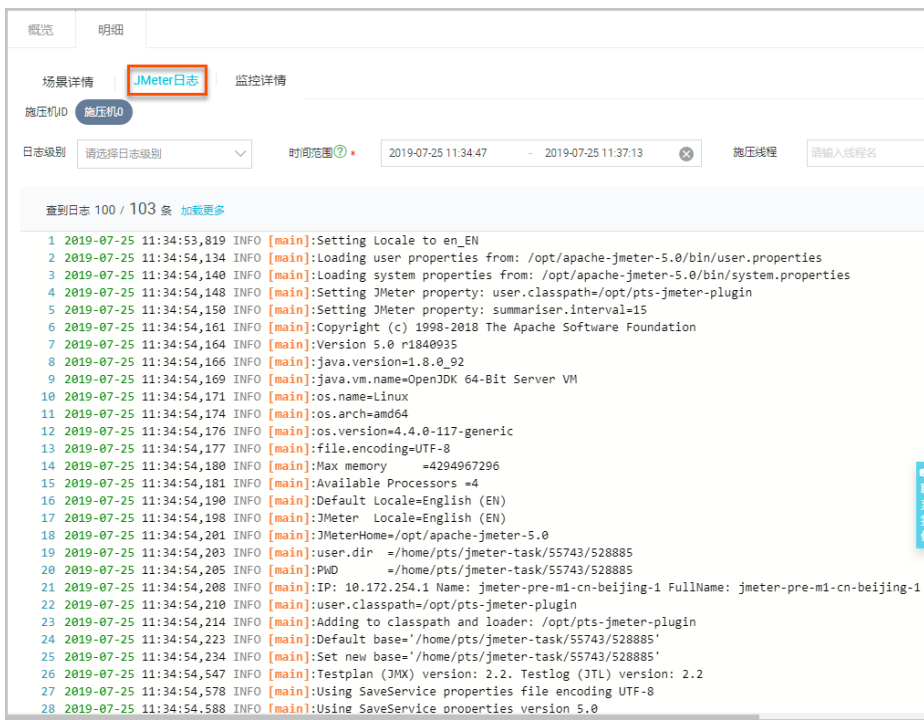
场景详情

压测报告的**明细** > **业务详情**页签下，显示全场景视图和各 Sampler 的名称中单个 API 的业务详情，包括请求（TPS）视图、成功率、响应时间、流量详情等。更多详细指标说明，参见**测试指标**。



JMeter 日志

JMeter 日志包括日志时间、日志级别、线程名、logger 和具体日志信息。



监控详情

若您添加了 ECS 监控、SLB 监控、RDS 监控和 ARMS 监控等，查看相关监控信息请参见监控详情。

采样日志

在压测报告页面右上角单击查看采样日志，根据压测 Sampler、响应状态等筛选日志，然后在对应的日志操作列单击查看详情，可以查看采样日志。在压测执行时或者压测报告中，通过查看压测采样日志，可以快速定位问题。

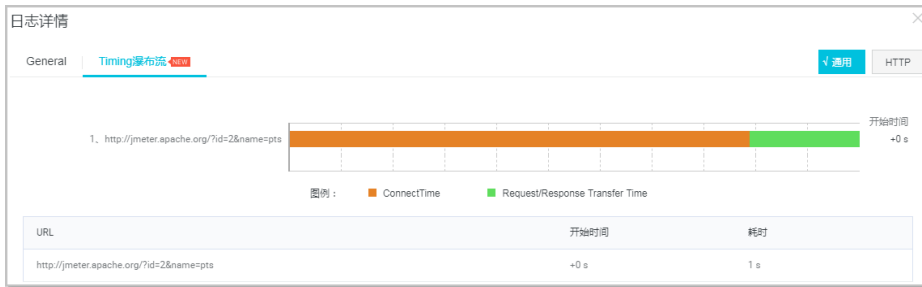
General

General 页签中展示了采样日志的日志字段和字段值。

日志字段	日志值
samplerData	Connect URL: ws://echo.websocket.org:80/
requestHeaders	
responseHeaders	Connection: Upgrade Date: Fri, 11 Jan 2019 07:27:35 GMT Sec-WebSocket-Accept: eN3H9pXdYKDOJaE9Xcw7mhzXhBY= Server: Kaazing Gateway Upgrade: websocket
responseDataAsString	
dataEncoding	
responseMessage	Switching Protocols

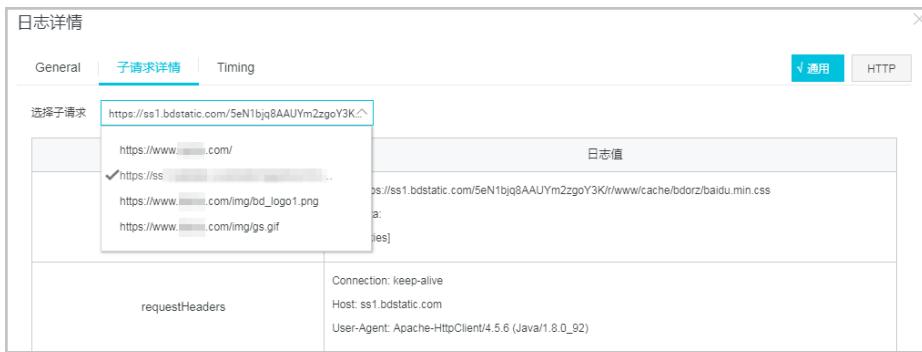
Timing 瀑布流

Timing 瀑布流页签中展示了各接口的耗时信息。

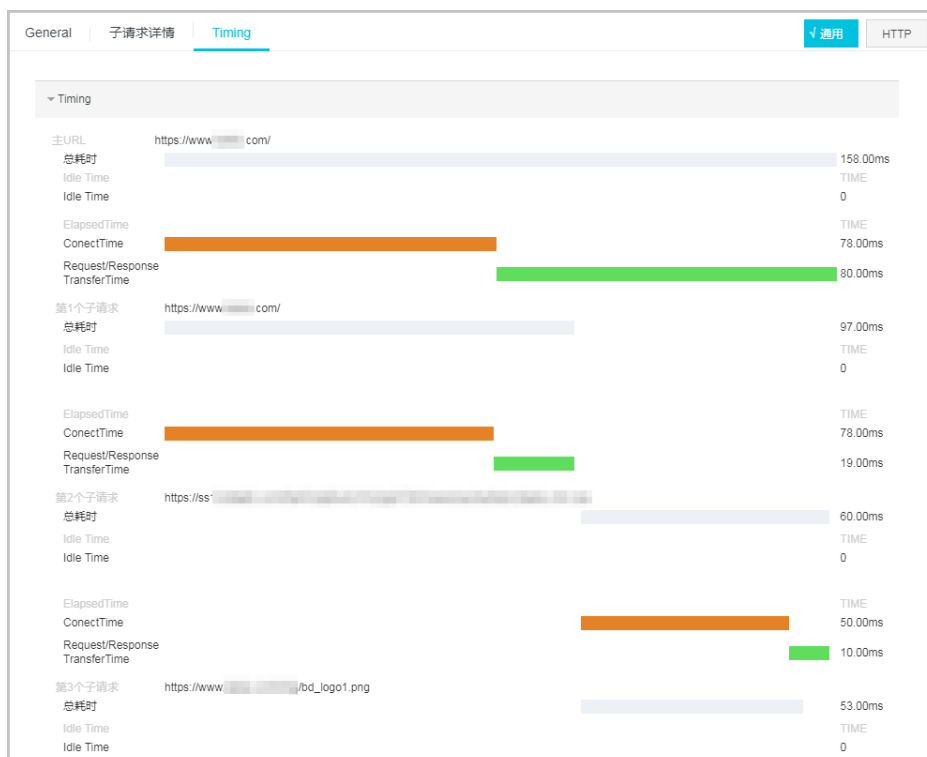


子请求详情

若 JMeter 脚本中配置了从 HTML 文件嵌入资源 (Embedded Resources from HTML Files)，日志详情页面会出现子请求详情页签。可通过选择特定的子请求，可筛选出对应的请求日志。



此时 Timing 页签下也将显示总请求耗时和每一个子请求的耗时情况。



接收人管理

管理接收人

在使用 SLA、定时压测等功能时，需要发送通知给特定人员。您需要通过**接收人管理**页面增加、修改、删除接收人信息。

新增接收人

新增接收人的步骤如下：

登录 PTS 控制台，在左侧导航栏选择**接收人管理**。

在**接收人管理**页面，单击右上角**新建联系人**。

在**设置报警联系人**窗口，设置以下信息：

用户名：接收人的名字

邮箱：接收人的邮箱地址

- 旺旺：接收人的旺旺名

钉钉机器人：钉钉机器人的接口地址

设置钉钉机器人后，可以通过在钉钉群中添加该钉钉机器人，允许钉钉群接收 SLA 报警、定时压测等通知。关于如何为钉钉群添加钉钉机器人，参见**设置钉钉机器人**。

编辑接收人

进入**接收人管理**页面，在目标接收人的**操作**列，单击**编辑**，可修改其具体信息。

删除接收人

进入**接收人管理**页面，在目标接收人的**操作**列，单击**删除**，可删除该接收人。

设置钉钉机器人

在钉钉群中添加钉钉机器人，同时在 PTS 设置钉钉机器人的通知方式，您可以在钉钉群接收压测报警、压测通知等。

操作步骤

获取钉钉机器人地址。

- i. 在 PC 版钉钉上打开您想要添加报警机器人的钉钉群，并单击右上角的群设置图标。

在群设置弹框中选择**群机器人**。



在群机器人页面单击**添加机器人**区域的 + 按钮，然后选择添加**自定义机器人**。



在添加机器人对话框中编辑机器人头像和名称，然后单击**完成**。



在**添加机器人**对话框中复制生成的机器人地址。



在 PTS 控制台上添加钉钉机器人为联系人。参见管理接收人。

设置报警接收人。

- 为场景 SLA 添加报警接收人，参见配置场景 SLA，选择钉钉机器人作为通知方式。

操作至此，您已成功设置一个钉钉机器人报警。当报警触发时，您将在设置接收报警的钉钉群中收到报警通知。

独立资源池

当您需要在这段时间内独占压测机器时，可以预约 PTS 提供的独立资源池，PTS 将按照您对压力来源地域（支持港澳台等地域）、运营商、生效时间等的要求预留施压机器，满足您的压测需求。本文将介绍如何使用独立资源池功能。

使用限制

5 万及以上的最大并发用户数的资源包规格才可以使用独立资源池功能。详见资源包规格。

操作步骤

您需要提前 7 天在 PTS 控制台申请独立资源池资源，申请成功后，系统会根据申请条件自动为您生成一系列 IP，您的压力来源将来自这些 IP 所属机器。然后再使用普通压测方法进行压测即可。具体操作步骤如下：

申请独立资源池

登录 PTS 控制台，在左侧导航栏中单击**独立资源池**。

在**申请列表**页面右上角单击**新建申请**。

在**新建申请**页面设置**开始使用时间**、**期望释放时间**、**主压测模式**、**峰值量级**、**分配类型**、**详细说明**等信息，并选择**是否透出 IP**。



The screenshot shows the 'New Application' form in the PTS console. The form is titled '申请列表 / 新建申请' and has a subtitle '目前独立资源池仅支持PTS原生、国内公网压测需求'. The form contains the following fields and options:

- 开始使用时间**: Select date
- 期望释放时间**: Select date
- 主压测模式**: 并发模式 (虚拟用户模式) RPS模式 (吞吐量模式)
- 峰值量级**: 填写当前账号下，全部场景运行时的总并发/RPS 预估总IP数为: 预估总VUM为: 0
- 分配类型**: 选择所需资源的分配类型，区域/省份/运营商/随机/自定义
- 详细说明**: 请填写详细的资源分配信息，比如按照区域类型，请输入每个区域需要的机器数量
- 是否透出IP**: 需要 不需要

At the bottom of the form, there are two buttons: '保存' (Save) and '保存并提交' (Save and Submit).

开始使用时间：资源池生效的时间。一般申请资源池后 7 天才能使用。

期望释放时间：释放资源池并停止计费的时间。若您在期望释放时间前已无需使用独立资源池功能，则可以在**申请列表**中释放资源池。

主压测模式：压测模式有两种：并发模式（虚拟用户模式）、RPS 模式（Requests Per Second，每秒请求数，吞吐量模式）。

并发模式：“并发”是指虚拟并发用户数，从业务角度，也可以理解为同时在线的用户数。

适用场景：如果需要从客户端的角度出发，摸底业务系统各节点能同时承载的在线用户数，可以使用该模式设置目标并发。

并发模式下，需要指定全场景的最大并发数，再设置各串

联链路的并发权重。

串联链路内各 API 的响应速度不同（表现为响应时间不同），所以单位时间内 API 的并发数也会不同。API 响应速度越快，单位时间内累积在 API 上的并发用户数越少。

例如，共 100 个虚拟用户需要操作某个事务（即串联链路）。假设，该串联链路中共有 2 个 API，API 1 响应速度快而 API 2 响应速度慢。则更多的虚拟用户将等待在 API 2 上，API 2 需要更多的线程资源来处理更多的虚拟用户请求。

RPS 模式：RPS (Requests Per Second) 是指每秒请求数。

适用场景：RPS 模式即“吞吐量模式”，通过设置每秒发出的请求数，从服务端的角度出发，直接衡量系统的吞吐能力，免去并发到 RPS 的繁琐转化，一步到位。

API 接口（如电商加购物车、下单等）主要用 TPS (Transaction Per Second, 每秒事务数) 来衡量系统的吞吐能力，选择该模式可以直接按照预期的 TPS 设置 RPS。如果希望检验“下单”接口是否能达到 500 TPS 的预期，那么设置 RPS 为 500，每秒发送 500 个请求，可检验系统的吞吐能力。

该模式下，请求无法及时响应时可能会导致较高的并发，异常情况请及时停止。

该模式仅支持非自动递增进行压测（即需在压测过程中手工调速）。

峰值量级：需同时运行的场景的总并发数之和或 RPS 数之和。例如某账户下共有 10 个场景，其中有 2 个场景会同时运行，则峰值量级为这两个场景的总并发数之和或 RPS 数之和。

说明：设置完峰值量级后，可以在右侧的 Tip 中看到该峰值量级下的预估总 IP 数和预估总 VUM。

分配类型：可选择区域、运营商、省份、随机和自定义。

详细说明：详细的资源分配信息。例如在区域类型下，每个区域需要的机器数量。

单击**保存并提交**。

若您暂时不需要提交，请单击**保存**，后续在**申请列表**中提交该申请。

发起压测，具体请参见**如何在一个月内发起压测**。

注意：独立资源池功能暂不支持 JMeter 原生压测和阿里云内网压测。

后续操作

在**申请列表**页，可以看到所有申请记录。您可以查看当前申请的状态、分配的 IP 等信息，还可以释放资源池。具体操作如下：

1. 在目标记录的**操作列**的**查看**。
2. 在**查看申请单**页面您按需查看单击 **IP 列表**，查看分配的 IP，您可以按需添加白名单。若您当前已不需要使用独立资源池功能，单击**释放资源池**释放资源并停止计费。