

性能测试 PTS

产品简介

产品简介

产品概述

性能测试 PTS (Performance Testing Service) 是具备强大的分布式压测能力的 SaaS 压测平台，可模拟海量用户的真实业务场景，全方位验证业务站点的性能、容量和稳定性。

PTS 目标是将性能压测本身的工作持续简化，使您可以将更多的精力回归到关注业务和性能问题本身。在 PTS 平台上，您可以用最低的人力和资源成本，构造出最接近真实业务场景的复杂交互式流量，快速衡量系统的业务性能状况，为性能问题定位、容量最佳配比、全链路压测的流量构造提供最好的帮助。进而提升用户体验，促进业务发展，最大程度实现企业的商业价值。

业务场景

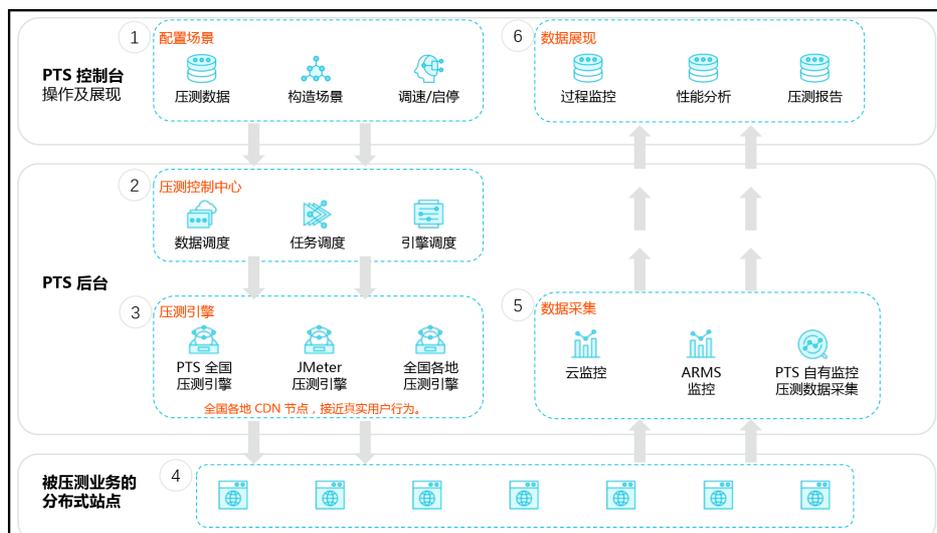
PTS 广泛应用于各种压力测试和性能测试场景，包括但不限于以下场景：



PTS 孵化于服务阿里巴巴全生态五年以上的单链路、全链路压测平台，是阿里巴巴内部最佳实践的输出。该平台对内除了支持日常的外部流量压测之外，同时支持了大大小小的促销活动，如天猫双 11、双 12 和年货节等。

压测流程

PTS 提供全面高效的压测流程：



压测流程说明：

在 PTS 控制台上，准备压测 API 数据，构造压测场景，定义压测模式、量级等；支持随时启停压测，压测过程中可调速。

压测启动后，PTS 后台的压测控制中心将自动调度压测数据、压测任务和压测引擎。

通过随机调度全国上百个城市和运营商的内容分发网络 CDN（Content Delivery Network）节点，发起压测流量。保证从虚拟用户并发量、压测流量的分散度等维度都接近真正的用户行为，压测结果更加全面和真实可信。

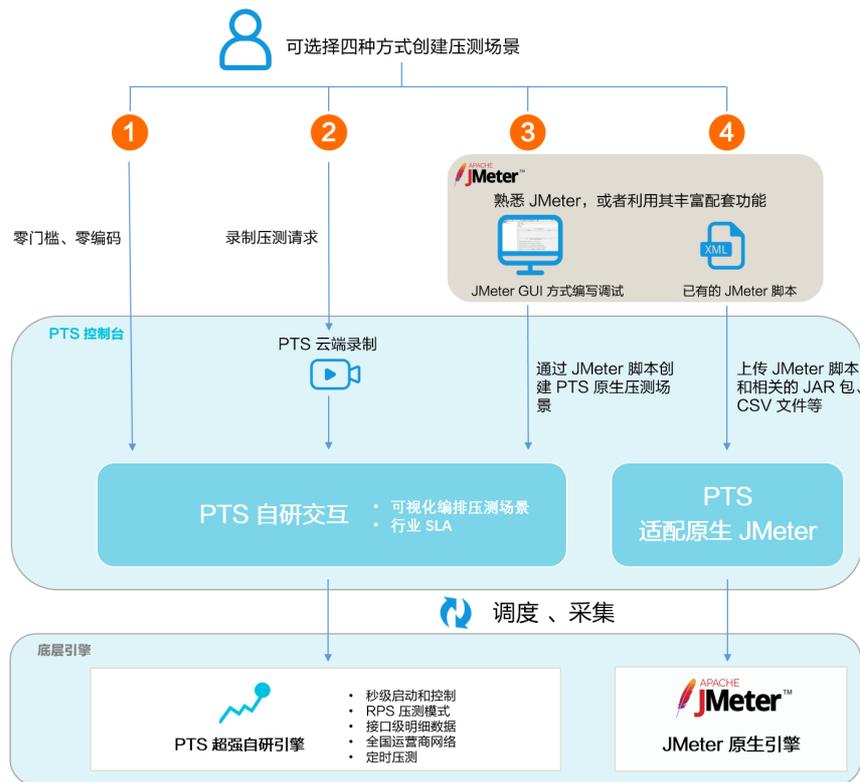
通过压测引擎向您指定的业务站点发起压测。

压测过程中，通过集成云监控、ARMS（应用实时监控服务）产品，结合 PTS 自有的监控指标，实时采集压测数据。

在 PTS 控制台，实时展现压测数据，进行过程监控；压测结束后，生成压测报告。基于整个压测场景的性能表现，定位性能问题、发现系统瓶颈。

压测创建方式

PTS 支持以下 4 种方式创建压测场景（或称压测用例），如图所示：



说明：

方式一：PTS 自研零编码可视化编排，使用自研强大引擎压测。

方式二：使用 PTS 自研云端录制器，零侵入录制业务请求并导入 1 中的自研交互中进行进一步设置。

方式三：将已有的 JMeter 脚本导入 1 中的 PTS 自研交互中，使用 PTS 自研引擎。

方式四：导入 JMeter 脚本并使用原生 JMeter 引擎进行压测，PTS 提供自定义的压力构造和监控数据汇聚等产品服务。

其中，方式一、二、三由于使用了 PTS 的自研引擎，具备 RPS (Requests per Second) 吞吐量压测模式、秒级启动、实时控制、定时压测和流量遍布全国运营商网络的差异化能力。

方式一是 PTS 最核心的一种压测场景创建方式，所有资源包均可使用。其他几种创建方式面向不同规格资源包开放，详见产品价格说明，并购买相应资源包。

学习路径图

您可以通过 PTS 产品学习路径图快速了解产品，由浅入深学习使用 PTS。

产品功能

性能测试 PTS (Performance Testing Service) 以互联网化的交互为基础，提供丰富的 API 编排功能。结合阿里巴巴强大的自研平台和引擎，支持按需设定压测模式、压测量级、压测时间，快速发起压测，监控压测过程并生成报告，确保流畅的压测体验，同时也兼容开源工具 JMeter。

压测场景构建

要发起一次性能压测，首先需要创建一个压测场景，进行业务 API 的编排。主要功能如下：

- 支持多个 API 并行或者有序串行。
- 支持 API 地址中添加参数，实现请求的动态变化。参数主要包括系统函数、文件参数、字符串、出参、其他自定义参数和它们彼此之间的组合。
- 支持 Cookie 传递，模拟用户登录场景。
- 提供丰富的指令功能，如集合点、思考时间等，扩展场景的仿真度。
- 支持压测前的场景调试，可进行复杂场景的数据流向校验。
- 提供云端录制器，便于移动端的请求抓取，可一键导入到压测场景中。

压测流量控制

压测流量的控制是性能压测中最重要的一环。主要功能如下：

- 支持两种压测模式：并发模式（虚拟用户并发）和 RPS 模式（Request per Second，即吞吐量模式）。RPS 压测模式为 PTS 独有，能够更精准地衡量服务端系统的能力。
- 随机调度遍布全国各地的压测引擎，一分钟内快速启动压测。
- 支持定时压测，可指定启动压测的日期、时间及循环周期等。结合服务等级定义 SLA (Service Level Agreement) 指标监控，触发告警或停止压测，实现智能压测。
- 支持两种调速模式：自动递增和手动调整。压测流量的调整秒级生效。
- 支持最高千万级的流量瞬时脉冲，多重机制确保压测流量及时停止。

压测数据监控

PTS 提供压测过程监控，展现实时数据。主要功能如下：

- PTS 的监控指标包括每个 API 的并发、TPS (Transactions per Second)、响应时间、采样的日志等。同时从不同细分维度，统计了 API 请求的成功、失败情况和响应时间。
- 支持添加阿里云生态内的云监控产品，可提供包括 ECS、SLB 和 RDS 在内的各产品的性能指标监控。
- 支持添加阿里云生态内的性能管理类产品，ARMS，提供应用级别的监控，为压测提供定位的闭环能

力。

- 支持创建服务等级协议 SLA (Service Level Agreement) 规则并关联到不同的业务场景，对业务压测场景进行更智能的控制和更全面合理的评价，同时 PTS 也提供了大量 SLA 模板。

压测报告导出

在压测停止之后，系统会自动收集压测过程中的监控数据（包括云监控、ARMS 的数据），形成压测报告，供您查看和导出。

开源工具 (JMeter) 集成

除了强大的原生功能，同样支持开源压测工具 JMeter 引擎发起的压测。您只需要在本地完成 JMeter 脚本调试，即可在 PTS 上快速进行自定义并发的压测。

产品优势

性能测试 PTS (Performance Testing Service) 与一般的压测工具相比，具有以下优势：平台稳定可靠、功能强大、流量真实和配套完善。

平台稳定可靠

- 阿里巴巴中间件技术部高可用团队倾心打造，经过内部五年以上的全生态沉淀，平台及技术稳定性高。
- PTS 是基于支持阿里巴巴全生态多达五年的单链路/全链路压测平台的再加强版本。
- PTS 支持了多个行业，涉及电商、多媒体、金融保险、物流快递、广告营销、社交等等。

功能强大

- 全 SaaS 化形态，无需额外安装和部署。
- 0 安装的云端录制器，更适合移动端 APP 场景。
- 数据工厂功能，0 编码实现压测的 API/URL 的请求参数格式化。
- 复杂场景的全可视化编排，支持登陆态共享、参数传递、业务断言，同时可扩展的指令功能支持多形态的思考时间、流量蓄洪等。
- 独创的 RPS /并发多压测模式。
- 流量支持动态秒级调整，百万 QPS 亦可瞬时脉冲。
- 强大的报表功能，将压测客户端的实时数据做多维度细分展示和统计，同时自动生成报告供查阅和导出。

- 压测 API/场景均可调试，压测过程提供日志明细查询。

流量真实

- 流量来源于全国上百城市覆盖各运营商（可拓展至海外），真实模拟最终用户的流量来源，相应的报表、数据更接近用户真实体感。
- 施压能力无上限，最高支持千万 RPS 的压测流量。

配套完善

- 除了压测平台之外，可付费增值提供全链路压测解决方案输出，全方位保障站点平稳应对业务峰值。

应用场景

性能测试 PTS (Performance Testing Service) 具备强大的分布式压测能力，应用十分广泛。

适用于多业务场景

不论您处于哪个行业，在以下业务场景（但不限于），PTS 都是您最佳的性能测试工具。



- 新系统上线：通过 PTS，准确探知站点能力，防止系统一上线即被用户流量打垮；
- 技术升级验证：大的技术架构升级后进行性能评估，验证新技术场景的站点性能状态；
- 业务峰值稳定性：大促活动等峰值业务稳定性考验，保障峰值业务不受损；
- 站点容量规划：对站点进行精细化的容量规划，分布式系统机器资源分配；

- 性能瓶颈探测：探测系统中的性能瓶颈点，进行针对性优化。

适用行业广泛

PTS 应用行业广泛，涉及电商、多媒体、金融保险、物流快递、广告营销、社交等等。

PTS 服务阿里巴巴全生态多年，支持了天猫双 11、双 12、年货节等大促活动。植根于电商行业的 PTS，对电商的典型业务模型支持得更友好，压测来源更广泛，脉冲能力和流量掌控能力更强。

PTS 自商业版发布以来，吸引了来自多媒体、金融保险、政务等众多行业的用户，以其强大的压测场景编排能力和报表能力，帮助用户快速发现问题，进行针对性地调优，提升了系统承压能力。

适用于多种网络环境

不论您的业务位于公有云、专有云、混合云、或者自建 IDC 中，只要能够通过公网访问，PTS 都能够通过遍布全国上百个城市和各运营商的 CDN 节点发起压测流量，最大程度地模拟真实业务场景。

适用于使用 HTTP/HTTPS/WebSocket 等协议的客户端

PTS 本身的 GUI 模式支持 HTTP/HTTPS 协议的压测，无论您的客户端是自研的 APP、移动端网页、PC 端网页、微信小程序还是 C/S 结构的软件，都可以使用 PTS 进行压测。PTS 同时集成了开源 JMeter，支持更多的协议和场景，比如您可以通过“JMeter + WebSocket 插件”的方式，对使用 WebSocket 协议的客户端进行压测（在 PTS 上传相应的插件 JAR 文件即可），其他协议以此类推。

名词解释

使用性能测试 PTS (Performance Testing Service) 之前，您需要了解 PTS 的相关名词。

3xx

表示客户端需要采取进一步的操作才能完成请求。通常，这些状态码用来重定向，后续的请求地址（重定向目标）在本次响应的 Location 域中指明。

4xx

表示客户端发生了错误，妨碍了服务器的处理。

5xx

表示服务器无法完成明显有效的请求。一般代表了服务器在处理请求的过程中有错误或者异常状态发生，也有可能是服务器意识到以当前的软硬件资源无法完成对请求的处理。

压测 API

指由用户行为触发的一条端上请求。压测 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。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。

并发用户数

同时发送压测请求的用户数量。一个用户在压测过程中可能是一个进程或者一个线程。

并发模式

虚拟用户模式，如果想要摸底业务系统能同时承载的在线用户数，可以通过该模式。

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。压测 API 之间只有在同一个串联链路中才能进行入参和出参关联（运行时数据传递）。两个不同的串联链路之间相互独立，通常不会存在参数的传递依赖（使用数据导出指令的情况除外）。

场景/压测场景

（压测）场景是若干个基于 HTTP/HTTPS 的 URL/API 的组合。URL/API 可能关联了数据文件表示不同用户。不同的 URL/API 表示不同的业务含义（比如登录、加入购物车），最终组合成一个接近用户各种真实行为同时具备一定用户量级的压测模型。

出参

从一个压测 API 的应答中截取需要的内容作为出参，供后续的压测 API 作为参数使用。

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。有时候响应码是 200 并不代表业务处理成功，有可能需要判断响应体内的内容。在 PTS 的串联链路中如果断言失败，当前请求就不会继续传递到下一个压测 API。另外，在压测实时报表和压测报告中都会相应展现业务成功或者失败的信息。

集合点

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。

TPS

Transaction per Second，系统每秒处理的事务数量。

条件跳转

根据出参，改变串联链路请求的执行路径，支持跳转、继续和结束等。

RPS 模式

吞吐量模式，指每秒固定发出设置的请求数量（RPS）。

SLA

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

是指从哪些指标监控压测过程数据。目前，SLA 指标主要包含业务质量相关指标 RT（Response Time）、

RPS (Requests per Second) 和成功率。PTS 将逐步增加性能 (云监控基础数据)、队列、SQL 连接等方面的 SLA 指标。

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

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。

思考时间

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。

数据导出

PTS 提供的指令之一。用于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，在全局共享导出数据。

数据指令

PTS 提供的指令之一。通过对前置压测 API 的出参、字符串或函数做二次处理，定义新的参数，提供给串联链路中的后续压测 API 使用。

VU

是虚拟并发数，表示压测能力。

VUM

是计量单位， $VUM = VU * Min$ 。

文件参数

将存在关联的参数放在一个文件的不同列中，通过上传该文件将其参数值导入 PTS。这些参数称为文件参数，可在压测 API 中使用。

响应时间 RT

从客户端发送一个请求开始，到客户端接收到服务端返回的响应所经历的时间。响应时间由请求发送时间、网络传输时间和服务器处理时间三部分组成。

75% 响应时间

在整个压测周期内 (压测启动到停止的时间内)，某个串联链路或者压测 API 的所有采样到的响应时间 (固定采样周期) 中 75% 的时间在这个值以内。

指令

是一种可以改变、控制串联链路中行为和流程的功能组件，可以更真实地模拟业务压测流量。

版本说明

本文列出了性能测试 PTS 的产品发布历史。您还可以在产品动态页，获取性能测试 PTS 的最新功能动态。

2019 年 9 月

功能

新增容量评估功能。容量评估可在压测过程中实时判定系统容量最佳压力值、极限压力值以及破坏压力值，帮助客户智能判定当前系统容量。同时，为后续智能压测埋下能力伏笔，提供限流降级阈值的建议值。【[相关文档](#)】

JMeter 压测支持 VPC 内网压测。利用 JMeter 协议和方法的灵活性，支持 VPC 环境来源的压测，可以提供更多压测方法和能力。【[相关文档](#)】

支持 Postman 脚本导入的方式快速创建压测场景。【[相关文档](#)】

2019 年 8 月

功能

自定义参数支持对文件参数的二次定义，无需在线下做数据加密等操作后上传。

新增数据配置节点功能。数据配置节点中，可以指定任意文件的参数，并可设置读取方式、是否轮询一次、设置基准文件，免去之前线下组合和拼装的操作。【[相关文档](#)】

云端录制器支持展示 response 信息，提升出参定义的便捷性和可读性。【[相关文档](#)】

2019 年 7 月

功能

新增 VPC 内网压测。通过 VPC 内压测，使用 PTS 原生引擎或者 JMeter 引擎对 VPC 内的服务进行专项的压测时，无需考虑压力是否充足，免去搭建、二次开发和维护一套压测集群的烦恼，可以随时

发起高并发的压测。【[相关文档](#)】

支持智能压测模式。通过预测模式可预判压力变化下的业务系统表现，识别潜在瓶颈点。【[相关文档](#)】

SLA 规则中的 RDS 监控指标新增 IOPS 指标。【[相关文档](#)】

JMeter 集成压测支持不同的百分位数统计，包括 99%、90%、75%、50%、最大和最小，更便于观察整体的响应时间的分布和波动情况，排除干扰因素。【[相关文档](#)】

2019 年 6 月

功能

JMeter 压测支持指定施压机器数。场景配置时可以设置压测的机器数，将目标并发均匀分布到这些施压机器上，获得更多的施压资源，达到压测的目的。

SLA 规则中新增了 ECS、RDS 和 SLB 三款产品的个性化指标的异常条件设置（例如 SLB 的丢弃连接数的最大值），来及时提示压测中的异常。【[相关文档](#)】

2019 年 5 月

功能

支持按照全国地区和运营商维度深度定制压测流量。【[相关文档](#)】

JMeter 压测支持优雅停止的功能，减少强行中断带来的异常。

2019 年 4 月

功能

1. 通过创建服务等级协议 SLA (Service Level Agreement) 规则定义并关联到不同的业务场景，方便地对业务压测场景进行更智能的控制和更全面合理的评价，同时 PTS 也提供了大量 SLA 模板。
2. 支持定时压测。结合业务 SLA 等级定义，方便任务及时停止和异常提醒，也能让定时压测更安全。
3. JMeter 集成压测支持文件下载。

2019 年 2 月

功能

1. JMeter 增加设置循环次数进行压测的功能。
2. PTS 的 API 编辑时支持快速解码（UTF-8 和 GBK），方便用户快速进行解码恢复。

2019 年 1 月

功能

1. JMeter 集成压测支持 CSV 文件切割到不同施压机。
2. API 的出参调试支持 json、header、cookie 等更多类型。
3. JMeter 集成压测支持页面测试（含递归 HTML 子请求）请求详情透出，采样日志可查看子请求信息。
4. JMeter 集成压测的结果支持导出。
5. 支持全局或者串联链路维度设置一个公共的 Header。

2018 年 12 月

功能

1. API 的出参可以在场景调试中直接基于响应体进行调试校验，方便用户测试正则表达式的准确性。
2. 采样日志中可以查看请求的核心生命周期耗时情况，方便问题的定位和解决。

2018 年 11 月

功能

- PTS 原生压测功能支持已有的 JMeter 脚本转化，方便 JMeter 测试脚本直接使用 PTS 的强大能力。

体验优化

云监控交互升级2.0，更优体验。

JMeter压测增加 JMeter.log，更白盒的体验，更便于问题排查。

3. Body 区域支持文本和 K/V 模式的切换，同时支持 JSON 等代码高亮，定义请求更加清晰。

2018 年 10 月

功能

PTS 兼容 JMeter 压测方案

PTS 支持直接上传已有的 JMeter 测试脚本、CSV 或 JAR 文件，PTS 将利用原生 JMeter 引擎，结合 PTS 的平台能力进行即时和可伸缩的压测，同时压测结果数据会及时收集并展现到压测控制台和报告中，结合云监控等能力进一步提升压测效率和闭环价值。

PTS-APM 集成功能发布

PTS 的监控集成新增 APM 类型（阿里云 ARMS 监控），方便 java 类应用压测问题的迅速定位。

2018 年 9 月

功能

PTS 新增数据指令和数据导出指令

数据指令和数据导出两个指令可以方便的将串联链路中的出参进行二次加工或者导出到全场景共享

PTS 压测报告导出功能

压测完成的报告支持 PDF 格式导出。

2018 年 8 月

功能

外网压测支持 IP 的弹性扩展

外网压测时在默认 IP 分配策略上支持继续增加 IP，方便进一步提升压测来源的真实度。

录制器支持提前筛选

通过提前筛选报文/请求类型快速精准录制。

- 云端录制支持录制前的域名、请求类型等的维度的提前设定；

- 页面展示进行了优化；
- 录制器的状态透出；
- 增加了一键清除的功能。

发布云监控 (ECS/RDS/SLB) 集成

压测场景页中支持云监控 (ECS/RDS/SLB) 的集成，支持自定义分组展示，提供覆盖客户端及服务端的监控一站式能力，发现和定位问题更加便捷高效。

压测报告中增加施压配置信息

PTS 铂金版的压测报告中增加了施压配置中的重要信息，比如：压测的模式、相应的起始/最大压测的量级、内外网的环境等等。便于事后更好的复盘压测过程和分析问题。

体验优化

场景压测中页面展现全新升级

压测场景执行中页面交互全新升级，更清爽简易的视觉，信息更聚焦。

施压配置页全新交互发布

全新的施压配置页面，提供压力预览，并发模式和吞吐量模式区分度提升，增强不同设置之间的关联性，确保强功能的前提下更简易的使用。

2018 年 7 月

功能

- 支持场景内 API 复制和开启/关闭串联链路功能

2018 年 6 月

功能

出参提取支持正则表达式匹配多个时的指定获取；URL自动encode；POST的body支持回车；监控和报告的采样数据精准度提升。

云端录制发布

PTS 铂金版推出云端录制功能，手机端压测请求也能方便抓取并一键导入。

出参和断言支持更多类型的 json 响应格式

2018 年 5 月

功能

- 场景编排 2.0 发布，全新交互和功能体验。用户可基于场景编辑页完成压测编排所有相关设置。

2018 年 3 月

功能

- 数据工厂发布，免去压测 API 层面的编码成本。压测API/URL相关的请求不再需要逐行构建。

2018 年 1 月

功能

发布调试和日志功能。

在场景中增加了调试功能，将场景维度整体跑一条数据/一次请求，调试结果一目了然；场景压测和事后报告中都增加了日志查询功能（日志采样收集），方便查看明细和定位问题

发布可视化编排复杂业务场景功能。

PTS 铂金版推出全新可视化编排场景功能，快捷实现复杂全链路业务压测。

2017 年 7 月

全链路压测云上版本 PTS 铂金版发布。

功能

无限接近真实的流量：业务场景中无论是高并发要求还是发起端的分散度，覆盖三四线城市主要运营商的节点广度都能做到真正模拟用户行为，客户端到服务端间复杂的网络瓶颈也能暴露无遗，压测结果更加全面和真实可信。

超高并发能力：通过底层技术优化以及阿里的全国布点，轻松支持千万级虚拟用户并发，低成本提供您所需的并发场景。

操作零门槛：不需要专门的性能测试团队或者测试背景的积累，完全面向开发的交互设计，开发自测试，投入产出比高。

多维度施压：支持并发和 RPS（旧称 TPS）双维度。

压力动态调整：支持压测能力动态修改。

2016 年 6 月

性能测试新的计费模型诞生，让用户更多选择，足够省钱，放心压测。

2015 年 1 月

性能测试正式商业化，开始有偿为阿里云的用户提供性能测试服务。

2014 年 8 月

性能测试接入阿里云，进行为期 5 个月的公测，服务阿里云企业客户、开发者。

2014 年 4 月

性能测试服务发布，正式对外提供性能测试服务，主要服务于聚石塔 ISV。