

32-无实例无真相：基于LoadRunner实现企业级服务器端性能测试的实践（上）

2018-09-10 茹炳晟

软件测试52讲

[进入课程 >](#)



讲述：茹炳晟

时长 15:36 大小 7.15M



你好，我是茹炳晟。今天我和你分享的主题是：无实例无真相之基于 LoadRunner 实现企业级服务器端性能测试的实践（上）。

从今天开始的两篇文章，我将介绍如何基于 LoadRunner 实际开展企业级服务器端的性能测试。分享这个主题时，我会从最开始的性能需求获取开始讲起，带你完整地经历一个实际服务器端性能测试项目。通过这个过程，我可以帮助你快速建立服务器端性能测试的全局观，并了解各主要步骤的关键技术细节。

听到这里，你可能就有些困惑了。我在分享 [《工欲善其事必先利其器：后端性能测试工具原理与行业常用工具简介》](#) 这个主题时，曾经说到：LoadRunner 比较适合于传统软件企业

开展性能测试，而 JMeter 更适用于互联网企业的软件性能测试。那么，为什么我没有选择以 JMeter 为例来展开后端性能测试呢？

我选择 LoadRunner，是经过深思熟虑的，主要原因包括：

1. **JMeter 的官方文档对其使用方法介绍得很详细，而且其操作基本属于“傻瓜式”的。**

JMeter 使用的难点在于：如何支持海量并发，以及实现更好的 load 控制，解决这个问题你可以参考 LoadRunner 的实现方式，然后从你所在企业的实际业务场景出发，进行二次开发。

2. **互联网企业和传统软件企业的软件产品的后端性能测试，在原理以及基本方法上是基本一致的，区别较大的只是全链路压测。**所以，我以传统企业的软件产品为例展开分享，你因此学到的原理以及测试方法将同样适用于互联网软件产品的性能测试。

3. **关于互联网软件产品的全链路压测，由于需要实现海量并发以及流量隔离等操作，所以目前只有一些大型企业在做，比如饿了么、淘宝、ebay、美团等超级大的网站。**但是，如果你也想了解全链路压测的话，我也会准备一篇“加餐”文章，和你分享开展全链路压测的难点，以及应对方案。我会更新完性能测试这个系列以后，为你准备这篇“加餐”文章。

为了让你在服务器端性能测试时更充分地利用好 LoadRunner，所以在正式开始讲解这个测试案例前，我会先给你简单介绍一下 LoadRunner 的基本原理，以及主要的功能模块。这些功能模块不仅在这个案例中会用到，也会在实际工程项目被经常使用，所以如果你有什么不理解的地方，欢迎给我留言。

LoadRunner 的基本原理

你还记得我在 [《工欲善其事必先利其器：后端性能测试工具原理与行业常用工具简介》](#) 这个主题中，介绍过的后端性能测试工具的基本原理吗？

我们先一起来回忆一下吧：

后端性能测试工具首先通过**虚拟用户脚本生成器**生成基于协议的虚拟用户脚本，然后根据性能测试场景设计的要求，通过**压力控制器**控制协调各个**压力产生器**以并发的方式执行虚拟用户脚本，并且在测试执行过程中，通过**系统监控器**收集各种性能指标以及系统资源占用率，最后通过**测试结果分析器**展示测试结果数据。

LoadRunner 的基本原理，与上面的描述完全一致。在 LoadRunner 中，Virtual UserGenerator 对应的就是虚拟用户脚本生成器，Controller

对应的就是压力控制器和系统监控器，Load Generator 对应的就是压力产生器，Analysis 对应的就是测试结果分析器。

为了帮助你理解 LoadRunner 的工作原理和模块，先撇开这些名词不谈，设想一下如果没有专用的后端性能测试工具，我们如何开展后端性能测试。

其实，“人肉”开展后端性能测试也不算太难。这个过程大致是这样的：

首先，我们需要一批测试机器，每台测试机器雇佣一个测试人员；

然后，我们需要一个协调员拿着话筒发号施令，统一控制这些测试人员的步调，协调员会向所有测试人员喊话，比如“1 号到 100 号测试人员现在开始执行登录操作，100 号到 1000 号测试人员 5 分钟后开始执行搜索操作”，同时协调员还会要求每个测试人员记录操作花费的时间；

测试完成后，测试协调员会要求性能工程师分析测试过程中记录的数据。

这个过程，如图 1 所示。

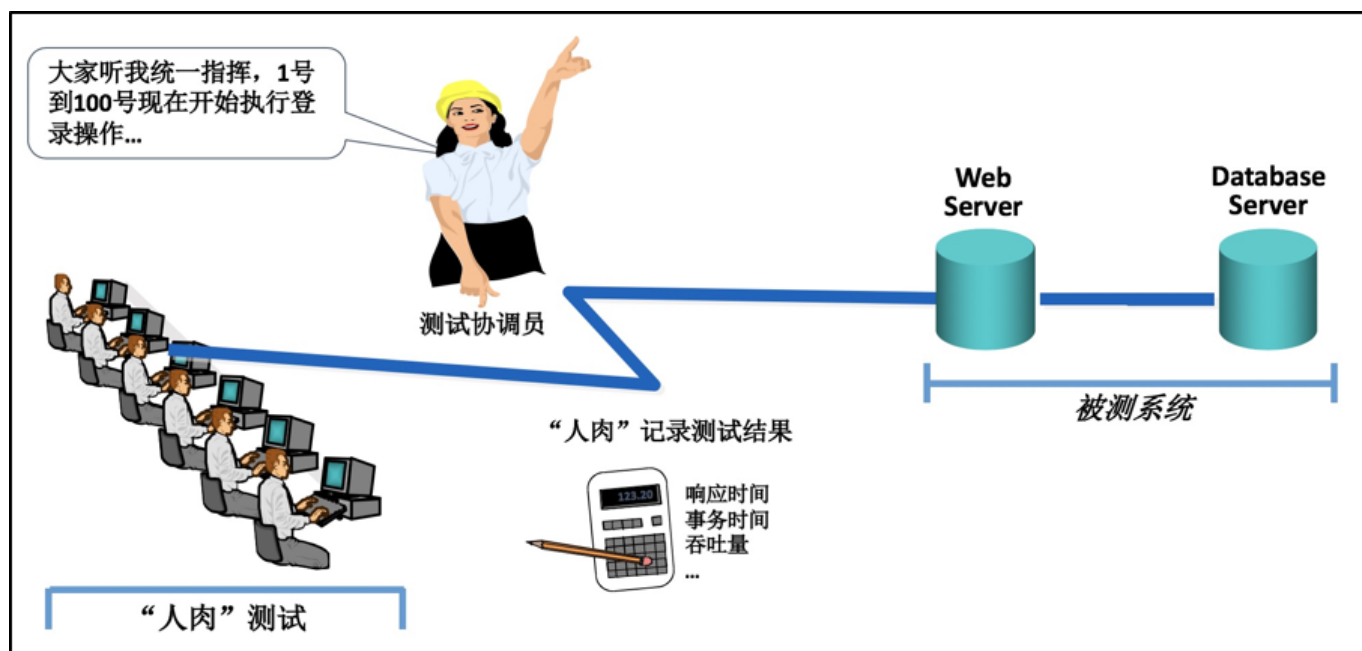


图 1 如果没有专用的后端性能测试工具，如何“人肉”开展后端性能测试

理解了这种“人肉”模式的后端性能测试后，我们再回过头来看 LoadRunner 的各个模块就豁然开朗了。

测试协调员以及完成数据记录的部分就是 Controller 模块；

大量的测试机器以及操作这些测试机器的人就是 Load Generator 模块；

操作这些测试机器的人的行为就是 Virtual User Generator 产生的虚拟用户脚本；

对测试数据的分析就是 Analysis 模块。

LoadRunner 的主要模块

通过对“人肉”模式和 LoadRunner 工具的类比，我们可以很清楚的看到，使用 LoadRunner 进行性能测试，主要需要 Virtual User Generator、Controller（这个模块包含了 Load Generator），以及 Analysis 这三大模块组合使用。接下来，我再和你详细聊聊这三大模块的作用，以及需要注意的问题。

第一，Virtual User Generator

Virtual User Generator，用于生成模拟用户行为的测试脚本，生成的手段主要是基于协议的录制，也就是由性能测试脚本开发人员在通过 GUI 执行业务操作的同时，录制客户端和服务端之间的通信协议，并最终转化为代码化的 LoadRunner 的虚拟用户脚本。

这样转化得到的虚拟脚本往往并不能被直接使用，还需要经历数据参数化（Parameterization）、关联建立（Correlation），以及运行时设置（Run Time Settings）等操作，然后才能用于性能测试场景中。

具体什么是数据参数化、什么是关联建立、运行时设置都有哪些可选项，我会在分享实例时再详细展开。

第二，LoadRunner Controller

Controller 相当于性能测试执行的控制管理中心，负责控制 Load Generator 产生测试负载，以执行预先设定好的性能测试场景；同时，它还负责收集各类监控数据。

在实际执行性能测试时，Controller 是和性能工程师打交道最多的模块，性能工程师会在 Controller 的 UI 界面上完成性能测试场景的设计、运行时的实时监控、测试负载的开始与

结束等操作。

第三，LoadRunner Analysis

Analysis 是 LoadRunner 中一个强大的分析插件。它不仅能图形化展示测试过程中收集的数据，还能很方便地对多个指标做关联分析，找出它们之间的因果关系。它最根本的目的就是，分析出系统可能的性能瓶颈点以及潜在的性能问题。

现在，你已经了解了 LoadRunner 的原理和各个模块了，接下来我们就开始实战吧。通过这个实战，我希望你可以掌握如何基于 LoadRunner 进行企业级的性能测试。

从宏观角度来讲，基于 LoadRunner 完成企业级性能测试，可以划分为五个阶段：

- 1. 性能需求收集以及负载计划制定；
- 2. 录制并增强虚拟用户脚本；
- 3. 创建并定义性能测试场景；
- 4. 执行性能测试场景；
- 5. 分析测试报告。

图 2 清晰地描述了这 5 个阶段的先后顺序，以及需要 LoadRunner 各模块发挥作用的部分。接下来，我和你详细聊聊每个阶段的具体工作，以及关键的技术细节。

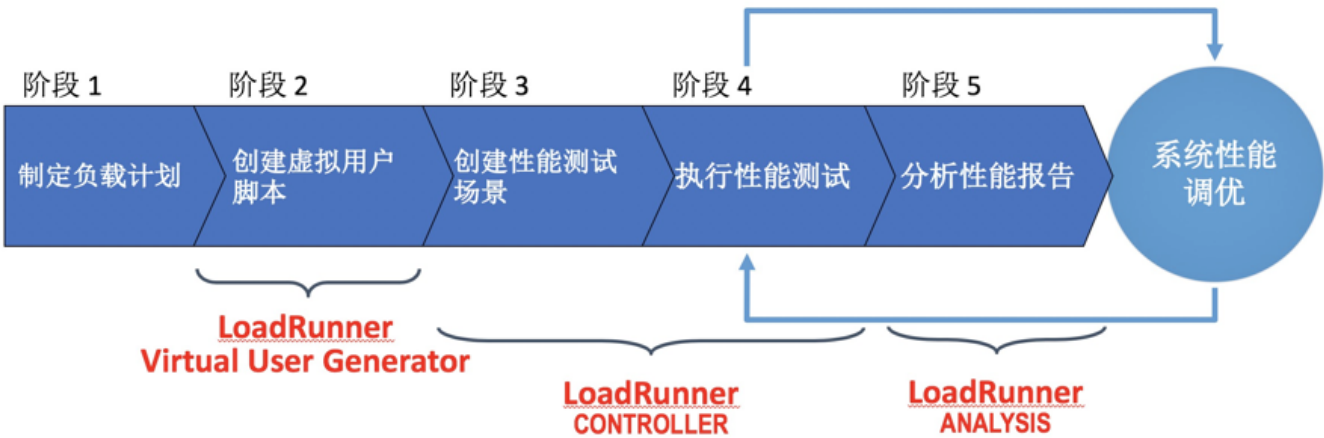


图 2 使用 LoadRunner 完成企业级后端性能测试的典型流程与步骤

阶段 1：性能需求收集以及负载计划制定

其实，无论是进行什么类型的测试，你的第一步工作都是要根据测试目的明确测试的具体需求。企业级的后端性能测试，当然也不例外。

一般情况下，企业级后端性能测试的具体需求，主要包含以下内容：

系统整体的并发用户数。比如，高峰时段会有 10 万用户同时在线；

并发用户业务操作的分布情况。比如，20% 的用户在做登录操作，30% 的用户在做订单操作，其他 50% 的用户在做搜索操作；

单一业务操作的用户行为模式。比如，两个操作之间的典型停留时间，完成同一业务的不同操作路径等；

并发用户高峰期的时间分布规律。比如，早上 8 点会有大量用户登录系统，晚上 6 点后用户逐渐退出；

达到最高峰负载的时间长度。比如，并发用户从 0 增长到 10 万花费的总时间；

...

完成这些点的测试，其实并不复杂。你只要按照这个已经明确的需求，开发后续的测试脚本、设计性能测试场景就可以了。

但是，如果你想要成长为更资深的性能测试工程师，或者已经是性能测试的设计者、资深的性能测试工程师了，那么你就需要全程参与到这些需求的获取和确定中。

其实，在我看来，获取这些测试需求时性能测试中最难的两个工作之一。另一个最难的工作是，测试结果分析与性能问题定位。而其他类似性能测试脚本开发、场景设计等工作看起来很有技术含量，但实际都是一些相对机械性的重复工作。

那为什么获取测试需求难做呢？因为绝大多数情况下没人会明确告诉你具体的性能需求。

对于软件的功能测试来说，如果需求不明确，你可以直接求助于产品经理。

而对性能测试需求来讲，产品经理通常无法准确告诉你用户的各个业务操作所占的百分比，也无法告诉你准确的用户行为模式。产品经理能做的，往往是给出定性描述，然后需要你去计算或者根据过往经验得到具体的定量需求。所以，我们经常听到产品经理对性能测试人员说：“你是性能专家，你来告诉我性能需求”。

那么，对于性能测试设计人员来说，到底如何获得这个明确的性能需求呢？说到这里，你应该明白了这是一个非常复杂的话题，因为测试目的不同，所用的方法也各不相同。所以，在这次分享中，我也只是可以给你准备一个实际的测试案例，和你分享获取具体测试需求的思考方式。

还记得我在第 29 篇文章[《聊聊性能测试的基本方法与应用领域》](#)中介绍的医院体检的例子吗？假设，产品经理对医院体检的性能要求是“每天支持完成 8000 个体检”，这个需求看似很具体，但是要转化成实际可操作的性能测试需求，还需要再细化。

首先，你要明确这里的“每天”是否指的是 24 小时。显然，这取决于产品本身的属性。比如，产品是为单一时区的用户提供服务，还是要面向全球所有时区的用户。那么，根据体检中心的属性，你很容易就可以确定“每天”一定是指 8 小时的工作时间。因为，体检中心一定是在一个确定的时区，并且不会 24 小时营业。

然后，你明确了这个 8 小时后，那么原始需求是不是可以转化为“每小时支持完成 1000 个体检”？

如果按照这个套路设计后续的性能测试的话，你会发现即使测试顺利完成，并且各项性能指标都达标了，但是一旦上线后，系统还是很有可能被压垮。因为实际情况是，验血往往需要空腹，所以上午往往是体检中心的高峰时段，体检者会在上午集中涌入体检中心。也就是说，这 8000 个体检并不是平均分布在 8 小时内完成的，而是有明显的高峰时段。

最后，你可以采用 80/20 原则对高峰时段的用户负载进行建模，比如 80% 的体检（6400 个）是发生在上午 20% 的时间（96 分钟）里。当然，为了使模型更接近真实的情况，你还应该分析历史数据，然后对该模型做进一步的修正，这也是目前被普遍采用的方法。

另外，在得到了负载模型后，性能测试设计人员往往还会在此基础上加入一定的负载冗余，比如在峰值的基础上再额外放大 20%，以增强系统上线后稳定运行的信心。

通过上面这个分析过程，你可以认识到，性能测试需求的定义与计划非常复杂，牵涉到项目的方方面面，不可能通过阅读一两篇文章就快速掌握这一技能，需要不断地沉淀在实战中获得的经验。

制订了性能测试计划后，接下来你就需要根据性能计划中涉及的用户业务操作来开发性能测试的脚本了。比如，前面提到“20% 的用户在做登录操作，30% 的用户在做订单操作，剩

下 50% 的用户在做搜索操作”，接下来你需要分别开发“用户登录”“下订单”和“搜索”这三个虚拟用户脚本。

在 LoadRunner 中，开发虚拟用户脚本的工作主要是基于录制后再修改的方式完成的。其中，录制由 Virtual UserGenerator 基于协议完成，录制后的修改主要是实现参数化、建立关联、建立事务、加入必要的检查点以及加入思考时间。

所以，我会在下次分享时，和你详细讨论这四部分工作的作用，以及具体如何完成。同时，我还会和你分享，企业级服务器端性能测试的后四个阶段如何实现。

总结

今天我和你讨论的主题是，如何基于 LoadRunner 实现企业级服务器端性能测试。

首先，我用一个“人肉”测试的流程类比，为你介绍了 LoadRunner 这个工具的基本原理，并分析了 Virtual User Generator、Controller（内含 Load Generator 模块），以及 Analysis 这 3 个模块的功能。

然后，从宏观角度，我把整个性能测试过程划分成了五个阶段：性能需求收集以及负载计划制定、录制并增强虚拟用户脚本、创建并定义性能测试场景、执行性能测试场景，以及分析测试报告。

在我看来，这五个阶段最难的两部分工作分别是：明确具体的性能测试需求，以及测试结果分析与性能问题定位。因为这两部分工作，要大量依赖于测试工程师的能力以及经验积累。所以，就像一名优秀的医生一样，优秀的测试工程师，需要在实际的工程项目中不断积累和总结经验。

最后，我以前面文章提到过的体检中心为例，和你详细讨论了如何收集性能需求，以及制定负载计划的内容。我也解释了为什么性能测试的需求不能直接从产品经理那里获得：因为产品经理定义的性能需求层次比较高、比较抽象，要落实到实际可执行的性能测试需求往往需要分析和细化。这也是为什么获取具体的性能需求比较难的一个原因。

思考题

你在实际工作中，是如何获取并细化性能测试需求的呢？

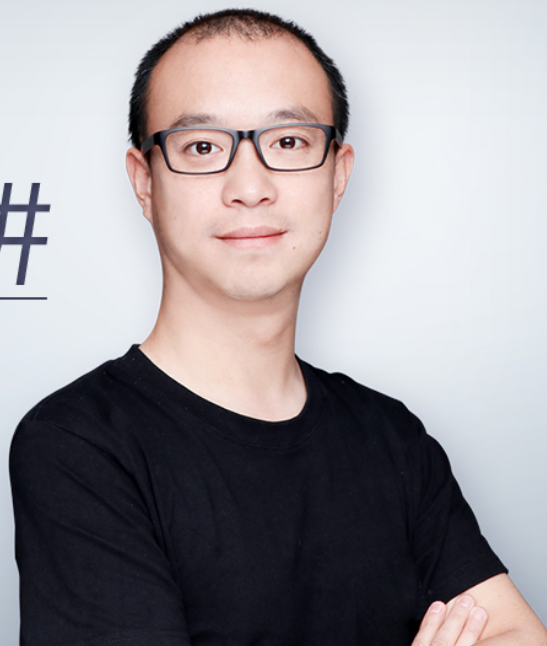
感谢你的收听，欢迎你给我留言。



软件测试52讲

从小工到专家的实战心法

茹炳晟 eBay中国研发中心
测试基础架构技术主管



新版升级：点击「 请朋友读」，10位好友免费读，邀请订阅更有**现金**奖励。

© 版权归极客邦科技所有，未经许可不得传播售卖。页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 31 | 工欲善其事必先利其器：前端性能测试工具原理与行业常用工具简介

下一篇 33 | 无实例无真相：基于LoadRunner实现企业级服务器端性能测试的实践（下）

精选留言 (9)

写留言



李奋斗

2018-09-11

4

图里的测试协调员，DJ打碟的既视感！

展开

作者回复：好像是哦



Cynthia◆...

2



2018-09-12

性能测试一直做的不多，所以最近看的也晕晕的，感觉还是要多实践呀！

不过呢，读了文章之后感觉更有方向了，比如为什么用这个工具，工具中的哪个模块是有怎样的用处，尤其是人肉测试那段类比，作者真的是好会写文章！

展开 ∨



伪专家

2018-09-10

👍 2

在测试道路上践行渐远.....目前的情况及行业趋势看，测试去转开发了...原因有以下方面：

1:通过拉钩和Boss测试的工资越来越少，没有优势

2:大企业Api, Ui, 全连路压测, ui兼容性, 稳定性, 健壮性都已成性, 线上copy, 防用户测试都已成型...招聘职位中什么都要有,但是实际业务及公司组织架构的影响, 测试的优势更受局限.....

展开 ∨

作者回复: 是机遇也是挑战，这就会形成两级分化，原本的功能测试会比较尴尬，但是具有宽广知识面的技术型测试会供不应求



涅槃Ls

2018-09-18

👍 1

打卡32、准备虚拟机 安装LR 学习LR

展开 ∨



口水窝

2019-04-26

👍

以前做性能测试难点就是不知道为什么要并发用户数分为多个行为去操作，顺着茹老师的课程，知其然，知道怎么去分解，剖析，点赞，坚持打卡！



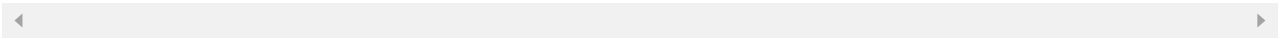
人心向善

2018-11-27

👍

loadrunner中有个说法是说并发和在线的关系大约是1:10，也就是说1个用户并发相当于10个用户同时在线，但是后来网上查了查确实是有这个并发关系，但比例不是1:10，而是在20%范围内浮动，我想向老师问下，这个是否是有依据的？谢谢！

作者回复: 没有依据, 如果你深入理解了并发用户数和用户行为模式的含义的话, 可以直接得出这个结论



小老鼠

2018-10-30



我们的性能是否通过的检查点为:

- 1, 系统中没有内存泄漏、磁盘、CPU、网络故障
- 2, 本版本获得的并发用户、容量的拐点 \geq 上次版本对应拐点的95%以内 (含95%)



jinjin

2018-09-30



最近第一次上手性能测试, 就遇到了需求不明确的情况。和老师举的体检中心的例子比较相似, 我得到的需求是看一下我们的系统index完10万篇 (每篇2M) 的文档的性能。对于并发用户数量到底该选取多少, ramp up时间设置多少合适以及需不需要加think time都很茫然。现在只能摸着石头过河, 不知道老师有什么好的建议吗?

展开 ∨



annie

2018-09-16



我们的产品类似WAF, 可以应用在不同行业, 但每个行业的应用场景不一样, 在大版本发布前测试了TPS, 吞吐量, 新建连接, 并发连接四个指标。后续版本测试只关注了TPS, 将TPS与之前版本做对比, 只要波动范围不低于10%就通过
但如何将测试结果的TPS指导不同行业客户部署上线, 还没有找到好的方法, 老师有没有思路可以指导一下呢? 谢谢

展开 ∨