上海交通大学致远学院官网

测试评估报告

版本 1.0

修订历史记录

|  |  |  |  |
| --- | --- | --- | --- |
| **日期** | **版本** | **说明** | **作者** |
| 28 /06/2020 | 1.0 | 性能测试报告 | 林舒怀、林褀龙、徐惠东、康艺潇 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

目录

1. 简介 4

1.1 目的 4

1.2 范围 4

1.3 范围 5

1.4 定义、首字母缩写词和缩略语 5

1.5 参考资料 6

1.6 概述 6

2. 测试过程与结果分析 7

2.1 关于培养方案查询的性能测试 7

2.1.1 测试过程与结果概要 7

2.1.2 分析结果与性能瓶颈 9

2.1.3 优化方案 11

2.2 关于搜索点击功能的性能测试 11

2.2.1 测试过程与结果概要 11

2.2.2 分析结果与性能瓶颈 14

2.2.3 优化方案 14

2.3 关于课程查询功能的性能测试 15

2.3.1 测试过程与结果概要 15

2.3.2 分析结果与性能瓶颈 16

2.3.3 优化方案 17

3. LoadRunner使用流程 17

4. 人员分工与贡献度 20

测试评估报告

# 简介

本测试报告是针对上海交通大学致远学院官网测试文档。致远官网是上海交通大学致远学院和致远荣誉计划的官方网站，本次测试为性能测试，针对并发性、吞吐量、响应时间三个评价指标设计了三种测试场景，并试图探索不同场景下，整体系统的性能瓶颈所在，最后尝试给出可行的优化方案。本次测试采用LoadRunner作为测试框架进行性能测试。

## 目的

致远官网 的这一“测试计划”文档有助于实现以下目标：

本文档针对上海交通大学致远学院官网（简称致远官网）的性能测试，记录相关测试内容，测试方法和测试环境。

* 本次性能测试共选取三个测试指标：并发性、吞吐量、响应时间。
* 针对上述三个不同的测试指标，我们分别设计了三个不同的测试场景：读取用户头像（并发性）、大规模上传文件（吞吐量）、高并发下的整站搜索（响应时间），并且对三种测试场景下的资源利用率进行分析。

## 范围

本次性能测试针对三个不同的测试指标（并发性、吞吐量、响应时间）设计了三个测试场景，并对三种场景进行资源利用率的分析，测试场景大致情况如下：

1. 查询培养方案（参数化测试）。我们模拟随机查询数学、物理学等专业方向的培养计划，由于涉及到 pdf 加载等信息所以对系统响应时间要求较高，且可能存在同时多人查询的情况，因此对系统并发性要求也较高。
2. 全文搜索并点击（参数化测试 + 关联测试）。我们在致远官网上进行全文搜索，并点击搜索出来的第一篇文章，因为全文搜索可能涉及到磁盘IO操作等，所以对系统的吞吐量有较高要求。
3. 查询课程（参数化测试）。我们根据所选条件查询课程，由于测试时信息检索会涉及到较为复杂的字符串匹配算法和数据库遍历操作，因此对系统响应时间和并发性有较高要求。

针对上述三种测试场景，我们也会通过对其资源利用率的使用情况，探索系统的性能瓶颈所在，并咋子分析后给出可行的优化方案。

## 范围

本次性能测试针对三个不同的测试指标（并发性、吞吐量、响应时间）设计了三个测试场景，并对三种场景进行资源利用率的分析，测试场景大致情况如下：

1. 针对并发性，我们选取的测试场景是大量用户对读取头像的过程，这个操作是加载二手交易信息和查用用户个人信息时的必要操作，因此对系统并发性的要求较高。
2. 针对吞吐量，我们选取的测试场景是大量用户上传二手交易信息与相应的文件（包括图片和视频），因为这个过程涉及到大型文件的传输，对系统吞吐量的要求较高。
3. 针对响应时间，我们选取的测试场景是大量用户对系统发起不同关键词的整站二手交易信息检索请求，因为整站检索功能是一个复杂的关键词匹配与查找功能，涉及到对数据库的访问和读取，对系统响应时间的要求较高。

针对上述三种测试场景，我们也会通过对其资源利用率的使用情况，探索系统的性能瓶颈所在，并咋子分析后给出可行的优化方案。

## 定义、首字母缩写词和缩略语

|  |  |
| --- | --- |
| 缩写词 | 含义 |
| 上海交通大学致远学院官网 | 本次测试所针对的接口类的名称 |
| 性能测试 | 针对系统整体性能的测试过程 |
| LoadRunner | 本次测试使用性能测试框架 |

## 参考资料

《Software Testing - A Craftsman’s Approach》

## 概述

本次测试报告包含针对测试用例设计的说明，测试结果和分析讨论，以及对被测试软件改进的建议

# 测试过程与结果分析

## 关于培养方案查询的性能测试

### 测试过程与结果概要

**测试网站**： <https://zhiyuan.sjtu.edu.cn/html/zhiyuan/service_list.php?bg=jxfw>



图1： 致远学院培养计划查询网站

**流程**

1. 选择专业 （数学、物理学、生命科学、计算机科学）
2. 选择年级 （2011级 - 2020级）
3. 点击查询

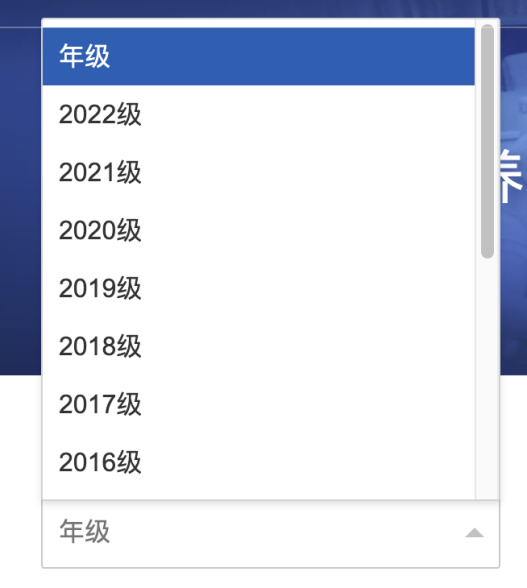
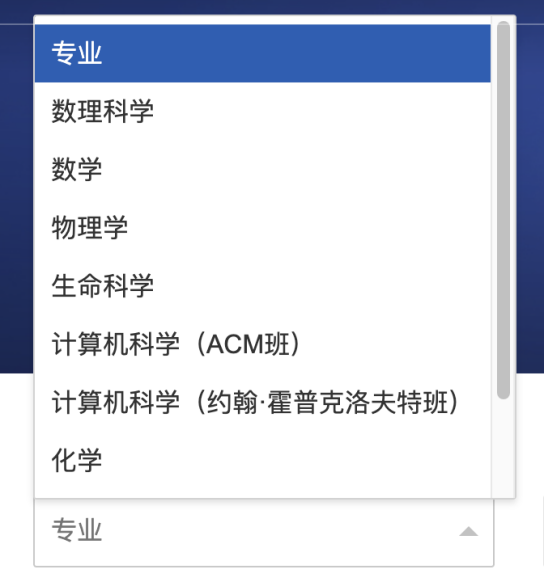


图2： 选择范围

**用户模拟设置：**

总模拟用户数为20，增加用户阶段每10s增加一个用户，在用户退出阶段每15s退出一个用户。图示如下：

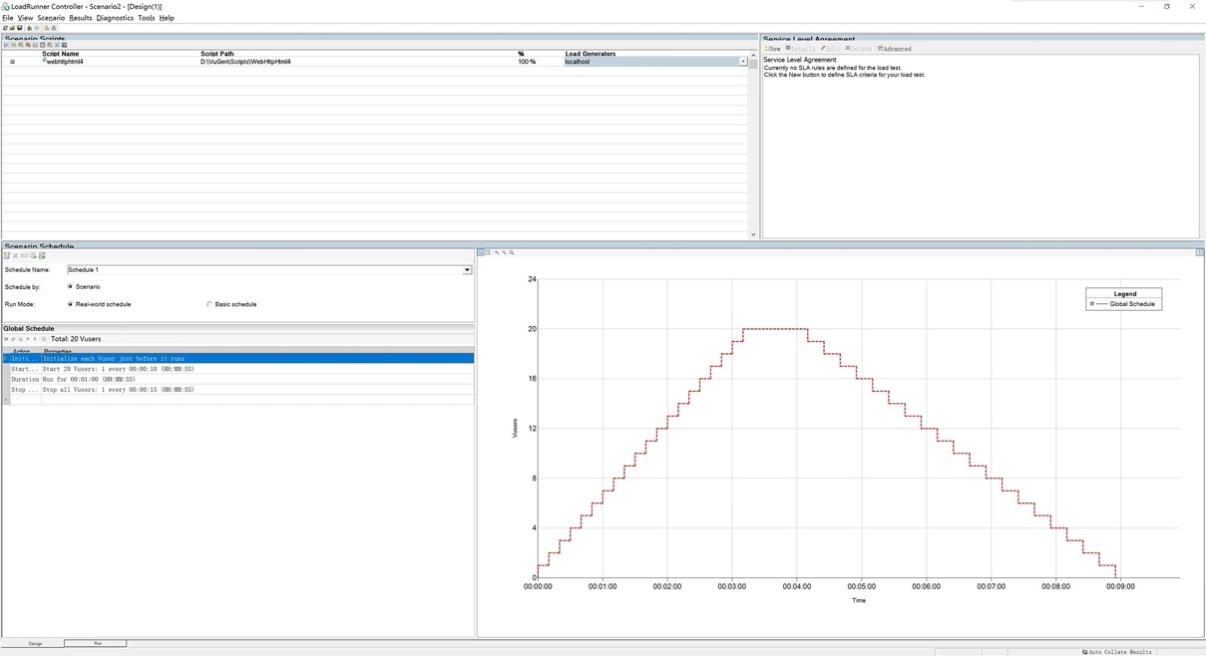
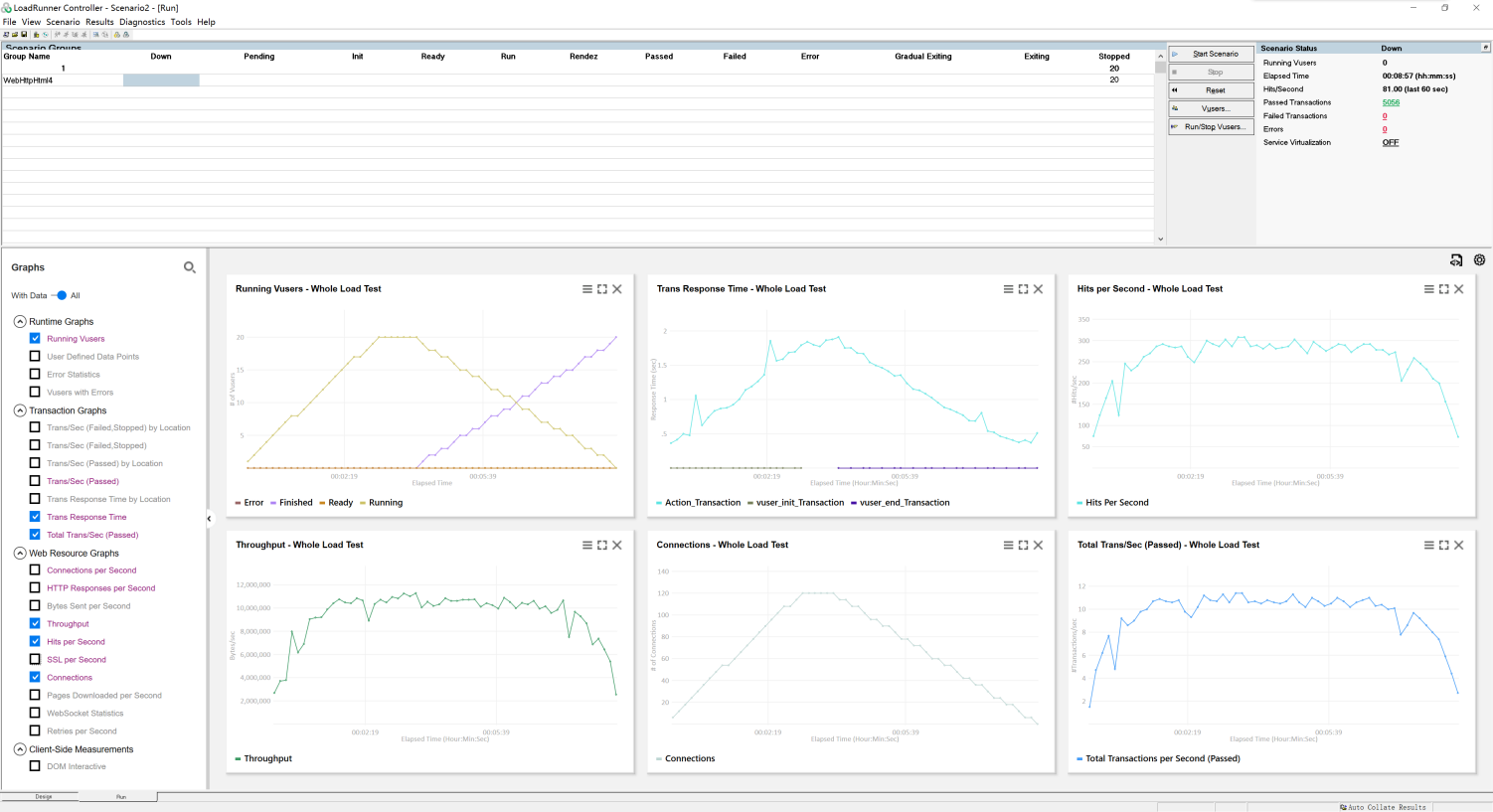
****

图3: 模拟用户增长和退出

**本样例针对的测试参数**：我们选取的测试场景是大量请求信息检索的场景, 对于查询功能用户比较关心查询所需要的时间，因此需要关注响应时间。同时存在很多人同时查询的可能性，所以也需要关注并发性。由于返回的数据为PDF文件（大概90KB），因此对并发下载的吞吐量也有要求。因此这个参数化用例主要关注**响应时间、****并发性和吞吐量**这三个参数。

**测试结果：**



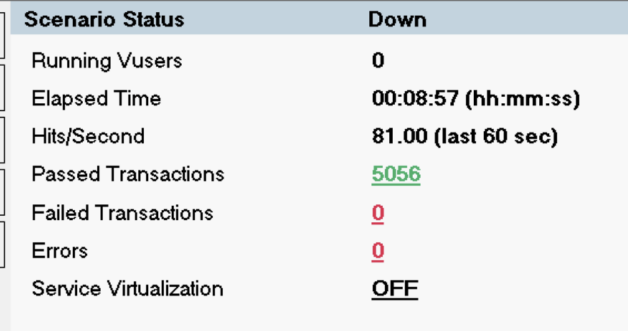


图4: 查询培养方案性能测试结果概要

### 分析结果与性能瓶颈

**响应时间：**

1. “事务平均响应时间”显示的是测试场景运行期间的每一秒内事务执行所用的平均时间，通过它可以分析测试场景运行期间应用系统的性能走向。常见标准为：优秀：<2s, 良好：2-5s, 及格：6-10s, 不及格：>10s
2. 可以看出响应时间随着用户数量的增加而变大, 当用户数比较小时小于0.5s, 在用户最多时接近2s。可以看到该网站在高并发的情况下响应迅速。
3. 在用户增大的过程中出现两个高峰，即相应时间突然变大的情况，考虑到是离散点的突然变高，可能是网速等不确定性影响。

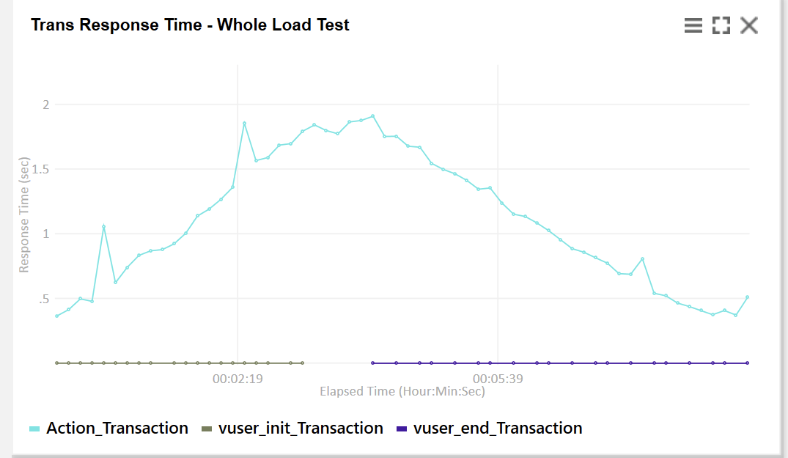


图5： Trans Response Time

**并发性：**

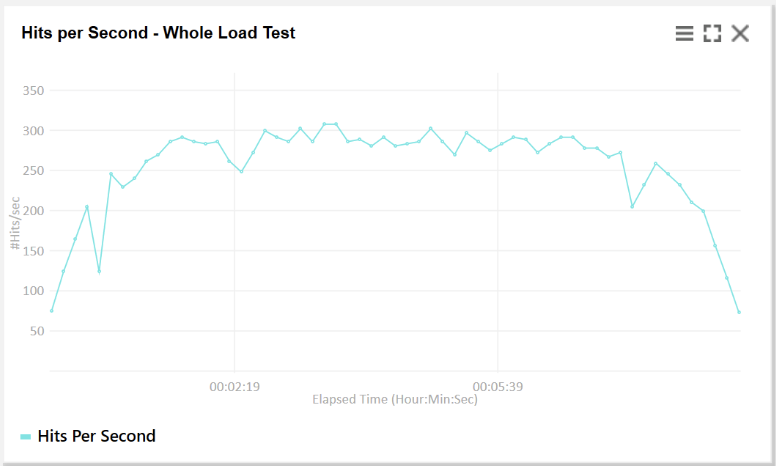


图6：Hits per Second

并发性主要关注 Hits per Second 参数 (每秒钟用户向WEB服务器提交的HTTP请求数。 这个指标是WEB应用特有的一个指标。WEB应用是"请求-响应" 模式, 用户发出一次申请, 服务器就要处理一次, 所以点击是WEB应用能够处理的交易的最小单位。如果把每次点击定义为一个交易, 点击率和TPS就是一个概念。容易看出, 点击率越大, 对服务器的压力越大， 这个参数阈值越大越好）。可以看到随着用户数的增加，Hits per Second 参数变大，在2分钟左右达到峰值（300Hits/Second）后周期性波动， 应该是受到了系统处理能力的限制。

**吞吐量**

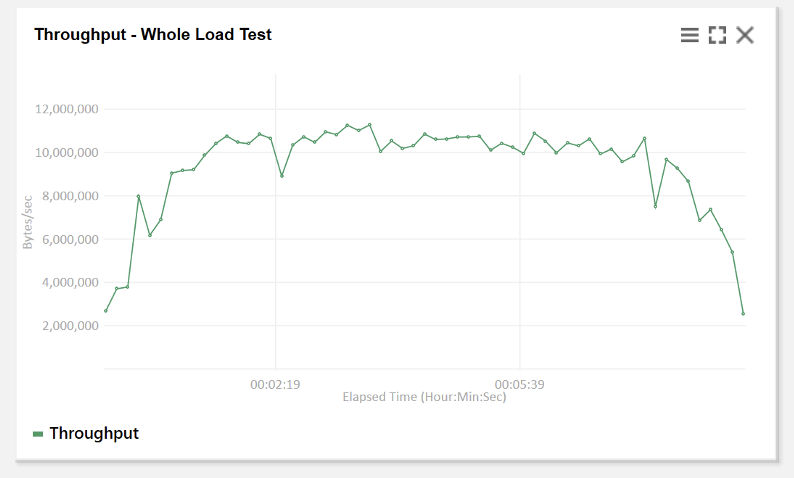
****

图7：Throughput

吞吐量指的是在一次性能测试过程中网络上传输的数据量的总和.吞吐量/传输时间,就是吞吐率。可以看出吞吐量逐渐增大，在2 min 左右达到瓶颈，大约在10，000，000左右。观察Throughput 和Hits per Second图像发现，两种图像的曲线都正常且基本一致，说明服务器能及时接受客户端的请求，并能返回结果。从图中可以看出，整体表现还是不错的。

### 优化方案

1. Hits per Second曲线在一分钟出现变化缓慢或者平坦，很可能是服务器响应时间增加，观察服务器资源使用情况，确定是否是服务器问题。

2. Throughput 曲线在1分钟出现变化缓慢并且平坦，由于现在的并发用户数为20，考虑到可能有大量用户的情况，建议增大Throughput的最大阈值。出现这种情况有两种可能：1.服务器硬件资源存在问题，需要拓展硬件或者优化应用。2. 若服务器硬件资源不存在问题，查看网络流量，可能网络带宽存在问题。

## 关于搜索点击功能的性能测试

### 测试过程与结果概要

交大官网可以根据关键词进行全文搜索，本测试针对这一功能进行参数化测试和关联测试，主要的参数化内容为全文搜索所用的关键词，而关联测试则是在进行全文搜索后点击匹配出的第一篇文章。

并发性的性能集中体现在一些高概率同时触发的应用场景下，而如果这一操作能够触发小批量的吞吐则更容易出现瓶颈，响应时间较高的服务请求则往往集中于要进行复杂运算或者大量磁盘 IO 的任务中。本测试场景涉及大量请求信息检索，当多名用户试图同时搜索同一关键词时会有较高吞吐，而全文搜索在后台更是依赖于数据库的查询等，因此本测试对响应时间、并发性还有吞吐量都有涉及。

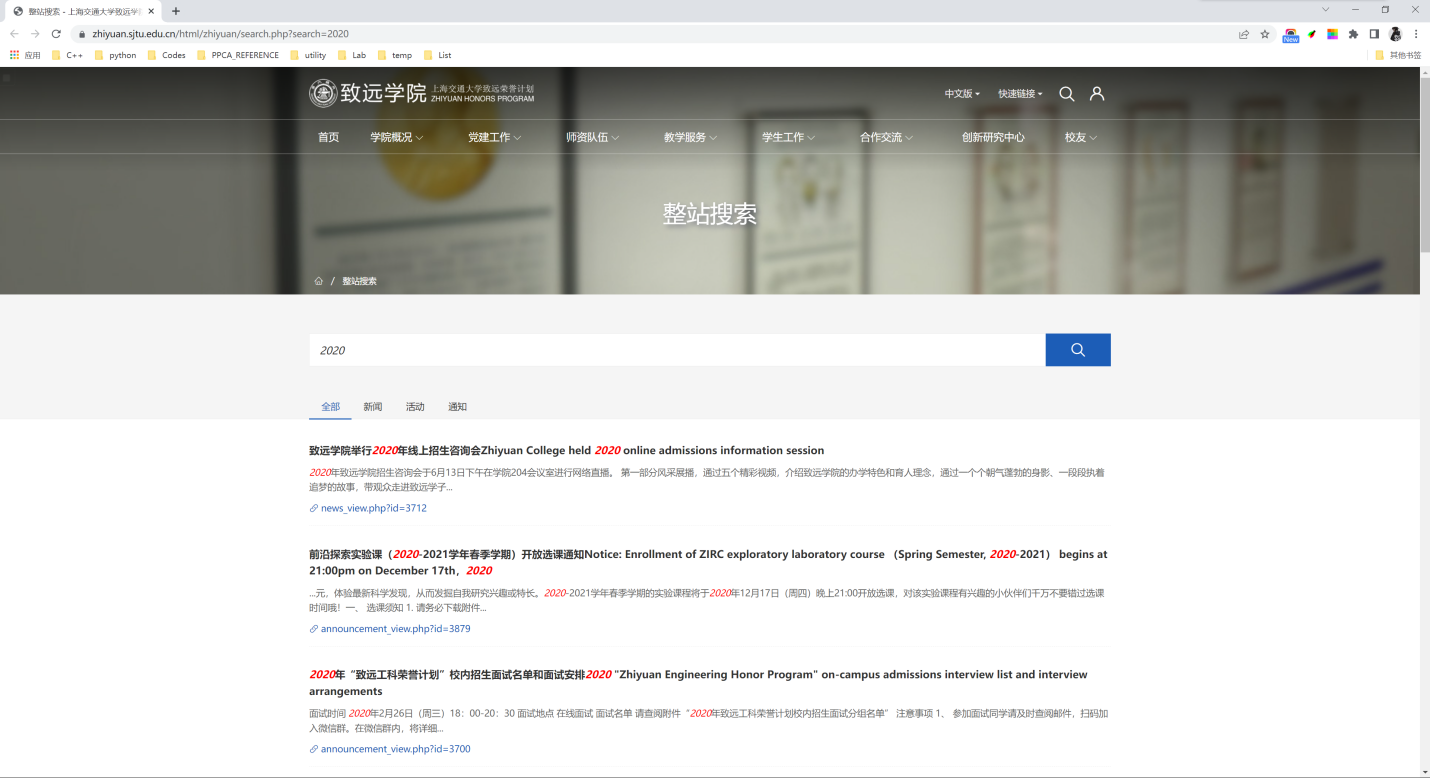


图8：致远官网的全文搜索界面

在具体测试用例上，我们设置模拟用户数目为20，并设置并发用户数量从零开始以每10秒增加1个并发用户的增速进行匀速爬坡，每个用户根据随机生成的条件循环发送课程查询请求，以此来观察响应时间和每秒请求数的变化。测试结果概要如下：

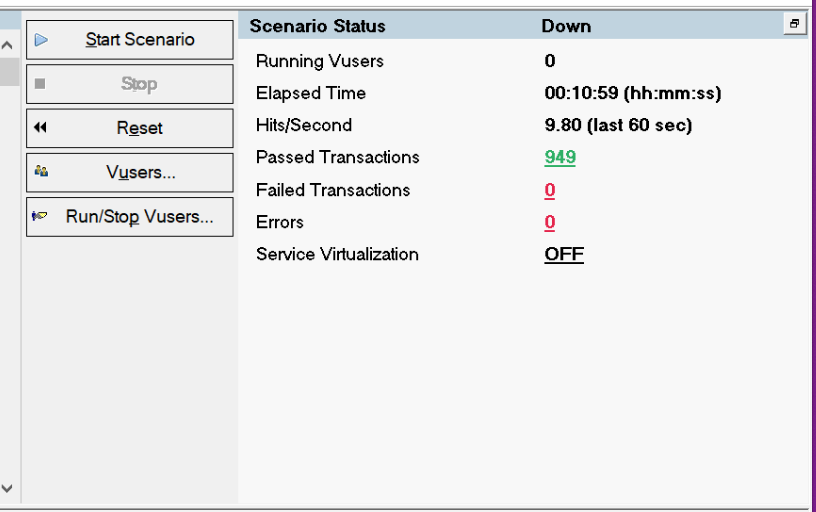


图9： 全文搜索并点击功能的性能测试结果概要

我们记录了正在运行的用户数量（Running Vusers）、每秒请求数（Hits per Second）、事务响应时间（Trans Response Time）随时间的变化，具体如下图所示：

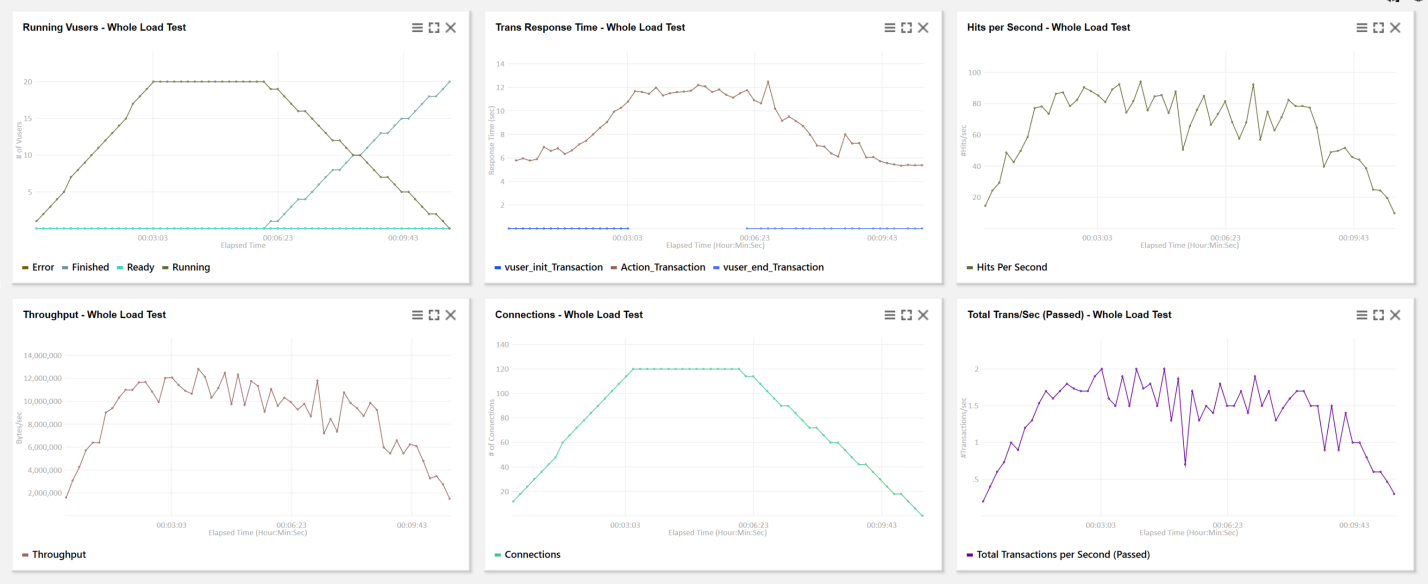


图10：全文搜索并点击功能的性能测试结果

### 分析结果与性能瓶颈

根据上述测试结果可以发现，随着模拟用户数量的上升，系统的每秒请求数、响应时间和吞吐量也在慢慢攀升，当用户数量稳定在20后，每秒请求数、响应时间和吞吐量则分别在80Req/s、12s、10Mb/s做周期性波动。当模拟用户数量逐渐减小后，每秒请求数和响应时间也逐渐减小。这可能是并发请求数量增加导致了系统在同时处理大量TCP的时候反应时间变长。另外如果并发请求数量超出了系统支持的TCP量，系统在取消TCP上需要一定的时延，因此导致测试过程中每秒请求数伴随着TCP用量被占满又释放而产生了周期性的波动。测试可知，该场景下网站不太能较好地应对高并发的全文搜索请求。

而且由于全文搜索需要进行大量查询，从测试结果可以看出吞吐量始终保持在10M/s左右较高水平值上下波动，所以可以推测出部分用于查询的索引文件可能存放于磁盘中。考虑以上因素，猜测性能瓶颈在CPU和磁盘 IO上。

### 优化方案

针对上述CPU的性能瓶颈，我们提出以下二种可能的优化方案：

1. 通过优化关键词匹配的算法，减少单次检索请求对CPU的占用。
2. 缓存一些常用的查询条件对应的课程查询结果，方便查询时快速响应。

针对上述磁盘IO的性能瓶颈，我们提出以下二种可能的优化方案：

1. 升级系统磁盘，例如将机械硬盘换成固态硬盘以提高磁盘的读写速度。
2. 使用更好的算法来进行索引文件的编写，旨在利用空间换去时间。

## 关于课程查询功能的性能测试

### 测试过程与结果概要

交大致远官网可以根据所选的条件查询课程。本测试针对这一功能进行参数化测试，主要的参数化内容为查询课程所用的条件。该测试场景涉及大量请求信息检索，会涉及到较为复杂的字符匹配算法和数据库遍历操作，偏重于对响应时间和并发性方面的测试。

社交网站的手机截图

描述已自动生成

图11： 致远官网的课程查询界面

在具体测试用例上，我们设置模拟用户数目为20，并设置并发用户数量从零开始以每10秒增加1个并发用户的增速进行匀速爬坡，每个用户根据随机生成的条件循环发送课程查询请求，以此来观察响应时间和每秒请求数的变化。测试结果概要如下：

图形用户界面, 应用程序, Excel, PowerPoint

描述已自动生成

图12： 课程查询功能的性能测试结果概要

我们记录了正在运行的用户数量（Running Vusers）、每秒请求数（Hits per Second）、事务响应时间（Trans Response Time）随时间的变化，具体如下图所示：

**图形用户界面, 应用程序, Excel, PowerPoint

描述已自动生成**

图14： 课程查询功能的性能测试结果

### 分析结果与性能瓶颈

根据上述测试结果可以发现，随着模拟用户数量的上升，系统的每秒请求数和响应时间也在慢慢攀升，当用户数量稳定在20后，每秒请求数和响应时间却开始围绕80Req/s和7s做周期性波动，并且存在一定的请求失败的情况。当模拟用户数量逐渐减小后，每秒请求数和响应时间也逐渐减小。这可能是并发请求数量增加导致了系统在同时处理大量TCP的时候反应时间变长。另外如果并发请求数量超出了系统支持的TCP量，系统在取消TCP上需要一定的时延，因此导致测试过程中每秒请求数伴随着TCP用量被占满又释放而产生了周期性的波动。测试可知，该场景下网站不太能较好地应对高并发的查询请求。

接下来我们尝试进行性能瓶颈的分析。首先，课程查询功能不涉及对文件的操作，因此基本不会对磁盘IO、内存占用和网络带宽产生很大的压力，从测试结果也可以看到吞吐量保持在6MB/s以下，相比前两个实验要小很多。

进一步思考我们的测试场景，因为没有涉及对文件读写的操作，因此大规模的运算需求大概率会成为系统的性能瓶颈，我们猜测性能瓶颈在CPU上。

### 优化方案

针对上述CPU的性能瓶颈，我们提出以下三种可能的优化方案：

1. 通过优化关键词匹配的算法，减少单次检索请求对CPU的占用。
2. 缓存一些常用的查询条件对应的课程查询结果，方便查询时快速响应。
3. 考虑对系统整体的硬件资源进行升级。

# Load Runner使用流程

该部分主要介绍列举性能测试各个流程步骤需要使用的LoadRunner的component及功能。

首先我们要知道LoadRunner三大组件：

（1）虚拟用户脚本生成器（Virtual User Generator）（图标有黄绳子的）

功能：录制脚本，编辑测试脚本会应用到

（2）压力调度控制台（Controller）（图标有绿绳子的）

功能：创建场景，运行这个场景、监控这个场景、然后手机一些测试的数据等。

（3）压力结果分析器（Analysis）（图标有蓝绳子的）

功能：把收集的测试数据会以图表的形式展示出来，生成测试报告

**测试过程**：

（1）首先打开Virtual User Generator新建文件和新增方案，协议选择Web-HTTP/HTML

（2）点击录制，配置集合点

（3）脚本录制结束，给录制脚本中插入事务或者检查点

图形用户界面, 应用程序

描述已自动生成

（4）参数化：右击需要参数化的值，选择“Replace with Parameter”

下方的Update value on有以下选项：

* Sequential：每运行一次选择一个用户
* Each itretion：每一次取新一行的变量
* Each ocurrence：每一次随机取值
* Onece：每一次都取相同的值

图形用户界面, 应用程序

描述已自动生成

（5）打开Analysis并载入之前录制的脚本

（6）设置模拟用户数量和启动方式

图表, 折线图

描述已自动生成

（7）运行测试脚本，生成测试报表

# 人员分工与贡献度

林祺龙：使用loadRunner进行参数化测试

徐惠东：撰写性能测试计划和性能测试报告

林舒怀：撰写性能测试报告

康艺潇：撰写性能测试报告

贡献度均为25%