宫斌凯 秦哲煜

**华东师范大学 计算机科学与软件工程学院**

性能测试计划

目录

[1 项目概要介绍 3](#_Toc501354155)

[1.1 项目简介 3](#_Toc501354156)

[1.2 项目成员 3](#_Toc501354157)

[1.3 测试范围 3](#_Toc501354158)

[1.3.1 范围内 3](#_Toc501354159)

[1.3.2 范围外 3](#_Toc501354160)

[1.4 前提假设 3](#_Toc501354161)

[1.5 测试目标 3](#_Toc501354162)

[2 性能测试策略 3](#_Toc501354163)

[2.1 性能测试模型 3](#_Toc501354164)

[2.2 性能测试场景 4](#_Toc501354165)

[2.3 重点测试策略 4](#_Toc501354166)

[2.3.1 重点测试原则 4](#_Toc501354167)

[2.3.2 重点测试交易 5](#_Toc501354168)

[3 测试案例设计 5](#_Toc501354169)

[3.1 生产压力分析 5](#_Toc501354170)

[3.2 场景通过标准 5](#_Toc501354171)

[3.3 测试场景设计 5](#_Toc501354172)

[3.3.1 独立场景 5](#_Toc501354173)

[3.3.2 混合场景 5](#_Toc501354174)

[3.3.3 峰值场景 6](#_Toc501354175)

[3.3.4 容量场景 6](#_Toc501354176)

[3.3.5 疲劳场景 7](#_Toc501354177)

[3.3.6 批处理场景 8](#_Toc501354178)

[3.3.7 极限测试 8](#_Toc501354179)

[3.3.8 异常场景 8](#_Toc501354180)

[4 测试实施安排 8](#_Toc501354181)

[4.1 测试进度 8](#_Toc501354182)

[4.2 测试流程 8](#_Toc501354183)

[4.3 测试报告需求 8](#_Toc501354184)

[4.4 性能缺陷管理 8](#_Toc501354185)

[5 性能测试标准 8](#_Toc501354186)

[5.1 启动标准 8](#_Toc501354187)

[5.2 中止标准 8](#_Toc501354188)

[5.3 通过标准 9](#_Toc501354189)

[6 测试环境规划 9](#_Toc501354190)

[6.1 部署环境 9](#_Toc501354191)

[6.2 执行环境 9](#_Toc501354192)

[7 测试风险分析 9](#_Toc501354193)

[8 角色与职责 9](#_Toc501354194)

表目录

[Table 1性能测试类型 5](#_Toc501358463)

[Table 2 场景通过标准 7](#_Toc501358464)

[Table 3 独立场景设计 7](#_Toc501358465)

[Table 4 混合场景设计 8](#_Toc501358466)

[Table 5 峰值场景设计1 8](#_Toc501358467)

[Table 6 峰值测试设计2 8](#_Toc501358468)

[Table 7 容量测试场景设计 9](#_Toc501358469)

[Table 8 疲劳测试场景设计 10](#_Toc501358470)

[Table 9 测试报告需求 10](#_Toc501358471)

# 项目概要介绍

## 项目简介

Timeline项目，包括注册、登录、发布文字和图片、获取消息功能。

## 项目成员

邹璇霖 宫斌凯 徐翼扬 秦哲煜

## 测试范围

### 范围内

（1）网站提供的主要功能，包括登陆注册，发送与阅览消息，以及它们各种情况下的组合。

（2）峰值性能测试

（3）疲劳测试

### 范围外

## 前提假设

假设网站正常运营，服务器正常运转，测试工具使用正常。

## 测试目标

1) 系统在单步操作响应时间方面，能满足用户当前及未来1到3年的发展需求；

2) 系统的批量作业运行稳定，处理能力能满足用户当前及未来1到3年的发展需求；

3) 在业务处理能力方面，能满足当前及未来1到3年的业务增长需求；

4) 发现并解决宕机，内存泄漏等严重问题，使系统具备良好的稳定性、健壮性；

5) 系统在资源使用方面比较合理，各项资源平均利用率在30%左右。

# 性能测试策略

与其它测试类型一样，性能测试周期分为：性能测试需求分析、性能测试设计、性能测试实现和性能测试报告四个存在迭代的阶段。其中，性能测试需求和性能测试设计由测试人员手工完成，而性能测试实现需借助性能测试工具达成，通过对性能测试报告的分析判定系统是否存在性能缺陷。

## 性能测试模型

性能测试模型定义了性能测试涉及的各个待测方面，是性能测试设计的指南。



Figure 1 性能测试模型

性能测试类型：

Table 1性能测试类型

|  |  |  |  |
| --- | --- | --- | --- |
| **测试种类** | **测试方法** | **测试内容** | **场景类型** |
| 性能测试 | 以系统最大并发用户数的1-2倍作为上限对关键交易进行梯度加压测试，在压力时间内通过的交易量应接近峰值时段的交易量，甚至超过系统全天的交易量 | 核心业务，需求分析中发现的容易出现性能问题的业务 | 独立场景 |
| 压力测试 | 以系统预期最大并发用户数的1-2倍作为上限对常用/关键交易进行混合场景梯度加压测试，在压力时间内通过的交易量应接近峰值时段的交易量，甚至超过系统全天的交易量 | 存在一定关联关系的常用/核心业务 | 混合场景 |
| 峰值测试 | 按照系统峰值时预期最大并发用户数的2-3倍对峰值时段涉及的核心交易进行的混合场景梯度加压测试，在测试时间内通过的交易数量应接近或超过系统全天的交易量 | 生产上峰值场景中的核心业务 | 峰值场景 |
| 容量测试 | 通过阶段性不断增加系统并发用户的梯度加压方式，测试出系统对核心交易的最大并发处理能力以及系统的最大业务处理能力，测试时最大并发用户数为系统预期最大并发用户数的3-5倍 | 针对核心业务 | 容量场景 |
| 疲劳测试 | 以系统预期最大并发用户数的1-2倍作为上限对关键交易进行持续加压的稳定性与健壮性测试，在压力时间内通过的交易量应为全天交易量3倍以上 | 针对核心交易 | 疲劳场景 |

## 性能测试场景：

独立场景

混合场景

峰值场景

疲劳场景

## 重点测试策略

### 重点测试原则

为了全面评估系统性能，本次重点测试策略安排如下：

1. 对关键交易开展全面的独立场景测试；
2. 对关键交易开展全面的峰值、容量测试；
3. 多渠道发起以关键/常用交易为主的疲劳测试；
4. 在集成性能测试方面，对交易量大、容易出现性能瓶颈的模块，重点进行测试并安排专人在系统发布前一直跟进开发组的需求变更以便进行及时的回归性能测试。
5. 在渠道性能测试方面，对渠道场景按照业务量及重要性来划分优先级，按照优先级先后来执行，保证系统上线后关键业务的稳定性。

### 重点测试交易

注册、登录、查看消息、发布图片和文字。

# 测试案例设计

## 生产压力分析

由于服务器性能较差，应用的并发性可能较差。

假设最大并发用户数是50

## 场景通过标准

Table 2 场景通过标准

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **场景类型** | **单步操作响应时间** | **加压时间** | **操作账户数据** | **处理业务笔数** | **事务成功率** | **各后台利用率** | **并发/在线上限** | **在线思考时间设置** |
| 独立场景 | <1秒 | 30分钟 | >1万 | >5000 | 100% | <50% | 100 | 无 |
| 混合场景 | <1秒 | 1小时 | >1万 | >10万 | 100% | <80% | 200 | 无 |
| 峰值场景 | <2秒 | 1小时 | >1万 | >10万 | 100% | <80% | 200 | 无 |
| 容量场景 | <3秒 | 4小时 | >1万 | >50万 | 100% | <80% | 300 | 无 |
| 疲劳场景 | <1秒 | 12小时 | >1万 | >100万 | 100% | <50% | 100 | 无 |

## 测试场景设计

### 独立场景

对于关键场景，将分别对其采用渐进式加压的方式来进行独立场景测试。每个场景测试多组并发、并发数从1逐步增加到200。对于具体的场景，测试几组并发依据案例执行。案例需要按照业务量设计。对于使用频度较低的场景，可以采用独立/混合场景的方式进行测试，即在背景压力下以小并发的方式来进行长时间疲劳测试。

Table 3 独立场景设计

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **并发数** | **响应时间要求** | **TPS要求** | **加压时间** | **加压方式** |
| 1 | <1秒 | N/A | 3分钟 | 直接加压 |
| 10 | <1秒 | N/A | 10分钟 | 1个用户/5秒 |
| 20 | <1秒 | N/A | 10分钟 | 1个用户/5秒 |
| 50 | <1秒 | >100笔/秒 | 20分钟 | 1个用户/5秒 |
| 100 | <1秒 | >100笔/秒 | 20分钟 | 1个用户/5秒 |
| 200 | <1秒 | >100笔/秒 | 10分钟 | 1个用户/1秒 |

### 混合场景

对于相关联的关键交易，对其采用渐进式加压的方式来进行混合场景测试，以测试这些关联交易是否存在性能问题。每个场景测试多组并发，并发数从20逐步增加到200。对于具体的场景，测试几组并发依据案例来执行（案例主要依据业务量来设计）。对于使用频度较低的场景，可以采用独立/混合场景的方式进行测试，即在背景压力下以小并发的方式来进行长时间疲劳测试。

Table 4 混合场景设计

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **并发数** | **响应时间要求** | **TPS要求** | **加压时间** | **加压方式** |
| 20 | <1秒 | N/A | 10分钟 | 1个用户/5秒 |
| 50 | <1秒 | >100笔/秒 | 20分钟 | 1个用户/5秒 |
| 100 | <1秒 | >100笔/秒 | 20分钟 | 1个用户/5秒 |
| 200 | <1秒 | >100笔/秒 | 10分钟 | 1个用户/1秒 |

### 峰值场景

通过分析生成数据，多数系统日间各个峰值共有两个时段，上午9:00-12:00的3个小时；下午13:00-17:00点的4个小时，因此，峰值最长持续时间4-5小时。考虑到系统的扩展性及实际情况，峰值测试共测试三组：第一组测试正常峰值压力的处理能力；第二组测试2倍、4倍峰值压力下的系统处理能力，以满足未来的业务发展需求，测试时各组中的交易按照业务比例同时加载。

第一组：先用生成峰值时压力的70%加压1小时作为基准压力，之后采用当前生产上的实际峰值压力并发进行加压。

Table 5 峰值场景设计1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **并发数** | **响应时间要求** | **TPS要求** | **加压时间** | **加压方式** |
| 生产压力的70% | <1秒 | >50笔/秒 | 1小时 | 1个用户/5秒 |
| 生产压力 | <1秒 | >50笔/秒 | 4小时 | 1个用户/5秒 |

第二、三组：分别采用生产压力的2到4倍来进行加压，第二组达标后再进行后面的测试。

Table 6 峰值测试设计2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **并发数** | **响应时间要求** | **TPS要求** | **加压时间** | **加压方式** |
| 生产压力的2倍 | <1秒 | >50笔/秒 | 5小时 | 1个用户/5秒 |
| 生产压力的4倍 | <1秒 | >50笔/秒 | 5小时 | 1个用户/5秒 |

对于性能较好的系统，可以直接加压到最大并发数，持续加压5小时来测试系统的峰值处理能力。

### 容量场景

容量测试从生产压力作为基准压力来加压，逐步加压到生产压力的5-10倍，在这个过程中得出如下结果：

1. 最佳处理能力/最佳并发数：最佳处理能力是指不断增大压力的过程中，在业务处理能力与交易响应时间等性能指标满足用户性能需求的前提下，系统所能达到的最大处理能力，系统初始达到最佳处理能力时对应的并发数可以称为最佳并发数。可以看出，在满足用户性能需求这一原则下，最佳处理能力往往不是最大处理能力，因此，对于在很小压力下，业务处理能力或响应时间已经不满足用户性能需求的系统，不讨论最佳处理能力与最佳并发数两个性能指标。
2. 最大处理能力/最大并发数：最大处理能力是指系统系统所能支撑的最大业务吞吐量，可以用“万笔/小时“或”笔/秒“等来度量。对于性能较好的系统，最大并发数是指响应时间和业务吞吐量等性能指标满足用户性能需求的前提下，系统所能支持的最大并发数。可以看出在满足用户性能需求这一原则下，系统在加压过程中可能先达到最大处理能力，如果这是各项指标仍然满足性能需求，则需要继续加压才能测出最大并发数，即将后继加压过程中业务处理能力或交易响应时间等性能指标开始不满足用户需求前的并发数作为最大并发数。可以看出，对于性能较好的系统最大处理能力和最大并发数往往不存在对应关系。对于性能较差的系统，最大并发数习惯上定义为业务吞吐量达到最大值时所对应的并发数，还有一种做法是将加压过程中响应时间开始不满足用户需求前的并发数定义为最大并发数，实际上，对于响应时间与业务吞吐量等关键性能指标不能满足用户性能需求的系统，讨论最大处理能力和最大并发数没有实际意义，这一点与性能较差时不讨论最佳处理能力与最佳并发数这两个性能指标的出发点是一致的。对于此类性能较差的系统，首先需要做的是对系统进行全面优化，然后再进一步测试系统的最大处理能力/最大并发数。
3. 系统容量上限：系统交易成功率<99%时的并发数
4. 系统上限：系统变得极其缓慢、交易大量失败甚至停止服务时的并发数。对于各个渠道，定义为交易成功率<90%时的并发数。

Table 容量场景设计

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **并发数** | **响应时间要求** | **TPS要求** | **加压时间** | **加压方式** |
| 生产压力的70% | <1秒 | >50笔/秒 | 1小时 | 1个用户/5秒 |
| 生产压力 | <1秒 | >50笔/秒 | 1小时 | 1个用户/5秒 |
| 生产压力的150% | <1秒 | >50笔/秒 | 1小时 | 1个用户/5秒 |
| 生产压力的200% | <1秒 | >50笔/秒 | 1小时 | 1个用户/5秒 |

### 疲劳场景

联机与批处理每天先后轮流执行，连续测试一周。9:00-21:00执行联机交易，21:00-次日9:00执行EOD。测试过程可以根据实际情况进行调换。

联机交易场景采用正常压力持续加压，加压分为3个时段进行。这个过程的目的是为模拟全行一天24小时的业务。

Table 8 疲劳测试场景设计

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **并发数** | **响应时间要求** | **TPS要求** | **加压时间** | **加压方式** |
| 生产压力\*70% | <1秒 | >50笔/秒 | 1小时 | 1个用户/10秒 |
| 生产压力 | <1秒 | >50笔/秒 | 9小时 | 1个用户/10秒 |
| 生产压力\*70% | <1秒 | >50笔/秒 | 1小时 | 1个用户/10秒 |

# 测试实施安排

## 测试进度

2019.1.4 进行了独立场景和混合场景的测试

2019.1.5 进行了峰值测试

2019.1.6 进行了容量测试

2019.1.7 进行了疲劳测试

## 测试流程

（1）依次对独立场景、混合场景测试

（2）依次对峰值场景、容量测试、疲劳场景测试

## 测试报告需求

Table 9 测试报告需求

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **报告类型** | **响应时间要求** | **报告者** | **接受者** | **报告内容** |
| 邮件 | 每日 | 各渠道测试组负责人 | 测试中心，项目经理 | 每日结果总结报告 |
| 邮件 | 每日 | 各渠道测试组负责人 | 测试中心，项目经理 | 每日测试结果记录 |
| 邮件 | 每日 | 性能测试负责人 | 测试中心，项目经理 | 周报 |
| 文档 | 各测试阶段结束 | 各渠道测试组负责人 | 测试中心，项目经理 | 性能测试报告 |

## 性能缺陷管理

测试过程采用Quality Center进行缺陷管理

# 性能测试标准

## 启动标准

1. 测试环境满足计划需求
2. 基准参数配置完成校验
3. 关键交易通过冒烟测试

## 中止标准

1. 测试环境或关键系统不可用
2. 测试环境距生产标准差距太大
3. 缺陷周转周期不符合规定的时间
4. 出现宕机、不响应等严重的性能问题
5. 系统的交易成功率低于95%

## 通过标准

系统上线至少满足下面标准：

1. 系统无宕机、不响应类的严重性能问题
2. 系统响应时间80%达到系统的期望值
3. 系统的业务吞吐量达到预期目标，即当前生产需求的3倍以上
4. 柜台类系统要求通过7\*12小时以上的疲劳强度测试
5. 电子渠道类系统要求通过7\*24小时以上的疲劳强度测试

# 测试环境规划

## 部署环境

阿里云的性能测试工具

## 执行环境

阿里云的性能测试工具

# 测试风险分析

# 角色与职责

秦哲煜： 独立场景和混合场景

宫斌凯： 峰值场景、容量场景和疲劳场景