

软件性能测试初探*

陈志皓

(广东省计算中心 广东 广州 510033)

摘 要 随着当代软件开发技术的发展与成熟,越来越多复杂的软件系统应用于人们生活的各个领域,软件系统运行时的性能表现已经成为衡量软件产品质量的一个重要标准。研究了软件系统性能测试的整体的流程,并结合自动化测试工具 LoadRunner,对软件性能测试的相关信息进行分析。

关键词 软件测试;性能测试;LoadRunner

1 软件性能测试

根据测试的目的和内容的不同,性能测试主要包括以下方面:

(1)负载测试:确定在各种工作负载下系统的性能,目标是测试当负载逐渐增加时,系统各项性能指标的变化情况。

(2)强度测试:确定在系统资源特别低的条件下软件系统运行情况。

(3)容量测试:在用户可接受的响应范围内,确定系统可处理同时在线的最大用户数。

(4)压力测试:通过确定一个系统的瓶颈或者最大使用极限的测试。

(5)疲劳强度测试:以系统稳定运行情况下能够支持的最大并发用户数或者日常运行用户数,持续执行一段时间业务,通过综合分析交易执行指标和资源监控指标来确定系统处理最大工作强度性能的过程。

(6)大数据量测试:大数据量测试侧重点在于数据的量上,包括独立的数据量测试和综合数据量测试。独立的数据量测试针对某些系统存储、传输、统计、查询等业务进行大数据量测试,而综合数据量测试一般和压力性能测试、负载性能测试、疲劳性能测试相结合。

2 软件性能测试工具 LoadRunner

作为软件质量控制中的重要一环,性能测试已经越来越受到软件开发者和用户的重视,成为软件测试的重中之重。性能测试通常在系统测试阶段执行,常常与强度测试结合起来,一般需要使用测试工具。一个优秀的软件测试工具,不仅可以辅助测试工作,满足科学测试的基本要求,而且可以自动化测试过程,节约大量的时间、成本、人员和资源,提高软件产品的质量。目前市场上主要使用的测试工具有微软公司的 WAS(Web Application Stress Tool)、Compuware 公司的 QALoad、RadView 公司的 WebRunner、HP(Mercury)公司的 LoadRunner。下面以 LoadRunner 为例,介绍软件测试工具的工作流程:

LoadRunner 是一种预测系统行为和性能的负载测试工具。通过模拟上千万用户实施并发负载及实时性能检测来确认和查找问题,能够对整个企业架构进行测试。通过使用 LoadRunner,企业能够最大限度的缩短测试时间,优化性能和加速应用系统的发布周期。LoadRunner 能支持广泛的协议和技术,功能比较强大,可以为特殊环境提供特殊的解决方案。LoadRunner 由下面三部分组成:Virtual UserGenerator 用来录制脚本、编辑脚本;Controller 用来布置测试场景、执行测试场景;Analysis 用来对测试结果进行分析。

用 LoadRunner 进行负载测试的流程通常由五个阶段组成:计划、脚本创建、场景定义、场景执行、监视执行和结果分析。

(1)计划负载测试:定义性能测试要求,例如并发用户数、典型业务流程和所响应时间,根据软件项目相关需求,定义相关测试的细节,撰写性能测试报告。

(2)创建 Vuser 脚本:将最终用户活动捕获到自动脚本中。LoadRunner 的脚本是 C 语言代码,LoadRunner 有自己的一整套函数接口,可以供外部调用。脚本可分 INIT、ACTION、END 三部分,其中 INIT 部分可以理解为初始部分。ACTION 可以理解为事务部分,也是测试的主体,END 是退出结束。

当录制完一个基本的用户脚本后,在正式使用前我们还需要完善测试脚本,增强脚本的灵活性。一般情况下,我们通过以下几种方法来完善测试脚本:插入事务、插入结合点、插入注解、参数化输入。

(3)定义场景:使用 LoadRunner Controller 设置测试环境;录制好脚本之后,就可以把脚本加入到场景里面去了,这里首先介绍一下 LR 的场景类型,LR 有两种大的场景类型:

①Manual Scenario:该项要完全手动的设置场景,该项下面还可以设置为每一个脚本分配要运行的虚拟用户的百分比,可在 Controller 的 Scenario 菜单下设置。

②Goal-Oriented Scenario,如果你的测试计划是要达到某个性能指标,比如:每秒多少点击。每秒多少 transactions,能到

* 本文系广东省科技计划项目《广东省科技资源共享网络平台建设》(粤财教[2011]511 号)的阶段性研究成果。

达多少 VU, 某个 Transaction 在某个范围 VU(5D。-1000)内的反应时间等等, 那么就可以使用面向目标的场景。

(4)设置场景:

Design :设计测试场景的静态部分, 设置模拟用户生成器、模拟用户数量、模拟用户组等。

Run :设计测试的动态部分, 主要指添加性能计数器, 在脚本运行的过程中可以通过这些计数器反馈的数据。

建立了测试场景后, 我们可以对 Edit_Schedule 进行设置, 设置测试开始执行的时间, 对于手动设计的测试还可以设定它的持续时间, 以及何时起用或禁止调用模拟用户。

(5)运行场景: 通过 LoadRunner Controller 驱动、管理和监控负载测试。

设置完毕后, 点击“开始方案”运行场景。在运行过程中, 可以监视各个服务器的运行情况 (DataBase Server、WebServer 等)。监视场景通过添加性能计数器来实现, 下列数据需要特别关注:

①Transaction Summary(事务综述)

对事务进行综合分析是性能分析的第一步, 通过分析测试时间内用户事务的成功与失败情况, 可以直接判断出系统是否运行正常。

②Average Transaction Response Time(事务平均响应时间)

“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。

例: 随着测试时间的变化, 系统处理事务的速度开始逐渐变慢, 这说明应用系统随着投产时间的变化, 整体性能将会有下降的趋势。

③Transactions per Second(每秒通过事务数/TPS)

“每秒通过事务数/TPS”显示在场景运行的每一秒钟, 每个事务通过、失败以及停止的数量, 使考查系统性能的一个重要参数。通过它可以确定系统在任何给定时刻的时间事务负载。分析 TPS 主要是看曲线的性能走向。

将它与平均事务响应时间进行对比, 可以分析事务数目对执行时间的影响。

例: 当压力加大时, 点击率/TPS 曲线如果变化缓慢或者有平坦的趋势, 很有可能是服务器开始出现瓶颈。

④Total Transactions per Second(每秒通过事务总数)

“每秒通过事务总数”显示在场景运行时, 在每一秒内通过的事务总数、失败的事务总数以及停止的事务总数。

⑤Transaction Performance Summary(事务性能摘要)

“事务性能摘要”显示方案中所有事务的最小、最大和平均执行时间, 可以直接判断响应时间是否符合用户的要求。重点关注事务的平均和最大执行时间, 如果其范围不在用户可以接受的时间范围内, 需要进行原因分析。

⑥Transaction Response Time Under Load(事务响应时间与负载)

“事务响应时间与负载”是“正在运行的虚拟用户”图和“平

均响应事务时间”图的组合, 通过它可以看出在任一时间点事务响应时间与用户数目的关系, 从而掌握系统在用户并发方面的性能数据, 为扩展用户系统提供参考。此图可以查看虚拟用户负载对执行时间的总体影响, 对分析具有渐变负载的测试场景比较有用。

⑦Transaction Response Time(Percentile)[事务响应时间(百分比)]

“事务响应时间(百分比)”是根据测试结果进行分析而得到的综合分析图, 也就是工具通过一些统计分析方法间接得到的图表。通过它可以分析在给定事务响应时间范围内能执行的事务百分比。

⑧Transaction Response Time (Distribution)[事务响应时间(分布)]

“事务响应时间(分布)”显示在场景运行过程中, 事务执行所用时间的分布, 通过它可以了解测试过程中不同响应时间的事务数量。如果系统预先定义了相关事务可以接受的最小和最大事务响应时间, 则可以使用此图确定服务器性能是否在可以接受的范围内。

(6)分析结果: 使用 LoadRunner Analysis 创建图和报告并评估性能。

LR 的报表分析功能也异常强大, 有各种各样的报表, 甚至可以将单个报表组合, 也可以导出到 Excel 文件和 Html 文件。

3 总结

任何软件的测试结果都不是只与应用程序本身有关, 特别是性能测试, 还与其测试硬件环境、软件环境、测试方法及测试工具有关, 因此在进行测试之前, 一定要了解应用程序的使用及运行的约束条件。Loadrunner Controller 通过使用虚拟用户技术来达到并发的目的, 这个测试属于黑盒测试, 测试人员不需要对程序代码有很深刻的了解。通过模拟真实用户对系统的访问, 可以帮助系统分析员及时发现系统的瓶颈, 从而优化各部分的软硬件配置。

参考文献:

- [1]《GB/T16260.1-2006 软件工程产品质量第1部分 质量模型》。
- [2]《GB/T16260.2-2006 软件工程产品质量第2部分 外部度量》。
- [3]《GB/T25000.51-2010 软件工程软件产品质量要求与评价(SQuaRE) 商业现货(COTS)软件产品的质量要求和测评细则》。
- [4]《LoadRunner 实战》, 无忧测试电子杂志第四期。