西安电子科技大学网络与信息安全学院

本科生毕业论文（设计）开题报告

（20 17 届）

学生姓名 唐天文

专 业 信息安全

学 号 13010310043

指导教师 张宁、唐青昊

2016年 12 月 18 日

（本表一式三份，学生、指导教师、学院各一份）

|  |
| --- |
| 一、论文名称及项目来源  名称: 基于覆盖导向原则的开源模拟器（QEMU）的半自动模糊测试框架  项目来源: 其它 |
| 二、研究目的和意义  目的：开发一套完整的QEMU模拟器fuzz测试框架。框架使用遗传变异产生数据，支持崩溃解析记录，支持样本数据记录，支持并发多任务测试，支持覆盖率检测，能够根据覆盖率优化变异脚本，从而提高覆盖率。  意义：QEMU模拟器是当今最知名的开源模拟器，企业可以应用QEMU来模拟多台主机和网络，搭建分布式系统和云平台。目前，多数云服务都是基于QEMU基础组件实现，而今年QEMU爆发出了多个严重的拒绝服务漏洞，数个可以造成虚拟机逃逸的漏洞，攻击者通过一台虚拟机就能达到控制所有云服务的目的。因此，及时发现QEMU的安全漏洞，帮助维护QEMU基础组件的安全可靠性，就能保障以QEMU为基础组件的服务安全可靠地运行。另外，挖掘QEMU安全漏洞的过程中，有很多重复零碎的工作，整合成一个有效的框架，不单可以应用在QEMU上，对于其他的模拟器的安全测试中也可以稍作修改后直接应用，大大提高安全研究人员的研究效率。 |
| 三、国内外研究现状和发展趋势  QEMU模拟器的安全研究现状：  目前国内针对QEMU进行安全研究的团队主要是360的Marvel Team，国外主要是个体研究人员在参与研究测试。[1]而针对QEMU的漏洞挖掘，主要是采用人工审计代码，然后编写特定的测试脚本结合动态调试的方式来发现安全漏洞。虽然相比于工具测试而言更加高效精准，但是费时费力，也会受制于研究人员自身的挖掘能力，忽视不少的安全问题，从而无法做到全面。  安全测试框架现状：  目前知名的各种fuzz框架主要有AFL，Sulley，Peach等。  AFL（American Fuzzy Lop）：AFL是一个安全测试框架，它采用了遗传算法和编译时检测的技术，具有适度的性能开销，高效的模糊策略等优点，可以无缝地处理的程序用例。目前，已经有人用AFL修改的版本对QEMU进行测试并发现了两个DOS漏洞。[2]  Sulley：Sulley是一个积极开发的模糊引擎和模糊测试框架，由多个可扩展组件组成。Sulley能够监视目标应用的状态，使用多种方法将其恢复到一个好的状态。Sulley还可以将所发现的错误进行检测、跟踪和分类，可以并行的进行模糊测试，能够自动的确定是哪个唯一的测试用例序列触发了错误。[3]  Peach：Peach是一款用Python写的开源的Smart Fuzz工具， 它支持两种文件Fuzz方法：基于生长(Generation Based)和基于变异(Mutation Based)。基于生长的Fuzz方法产生随机或启发性数据来填充给定的数据模型，从而生成畸形文件。而基于变异的Fuzz方法在一个给定的样本文件基础上进行修改从而产生畸形文件。[4]  以上举例的fuzz框架，AFL更适用于开源软件的模糊测试，后两者更多适用于文件和协议的fuzz测试。这些框架都是通过一个客户端来实现“样本生成”、“数据记录”、“崩溃解析”等功能，在目标测试程序崩溃时可以准确地得到栈回溯、测试参数等关键信息。但是，由于模拟器的特殊性，发送的数据需要从虚拟机系统内部发送，在模拟器外部监控，当模拟器崩溃之后，虚拟机系统就不再运行，而测试框架也已不复存在。由于缺乏外部监控端，在模拟器崩溃时，在虚拟系统重启后，变异数据可能部分丢失甚至完全无法保留，测试框架无法回溯崩溃时的堆栈回溯、传递的参数关系等关键信息。目前的主流测试框架都无法适用于模拟器的模糊测试，不能实现分离的数据生成和崩溃解析。  模拟器的模糊测试，需要开发一个分离式的监控、测试端，在虚拟系统内部调用测试样本，在外部检测覆盖并处理崩溃，主流的测试框架都无法做到。如果想应用主流的框架，就需要做个代理转发，把变异数据传递到虚拟机内部，然后由内部的代理来调用，因此，AFL测试框架在一番修改后，也是可以适用的。另外，QEMU不少模块包含了一个test功能，该功能基于qtest实现，只做到简单的功能测试，与模糊测试相去甚远，但是也可以以此为基础做一个模糊测试框架。  在2014年的时候，北京邮电大学的牛新立学者发表了一篇关于模拟器的模糊测试工具，测试了主流的3大模拟器软件，并挖掘出了数个安全漏洞，不过该工具并未发布到网上，也就无法了解更多细节。[5]  总结：模拟器的安全测试框架还处于少数人涉及的邻域，知名的fuzz测试框架也不再适用于模拟器的安全测试，需要专门定制。随着云服务的普及，云安全的重要性越来越受到企业的重视，QEMU模拟器作为当下最基础的云服务组件，它的安全性关乎着云服务的安全性，安全研究人员迫切需要一个测试框架帮助他们完成复杂困难的模拟器安全研究，一款模拟器的测试框架不仅可以在开源的模拟器上胜任工作，在获得了比较可靠的测试数据后，也能在不开源的平台上应用，未来的安全研究肯定不能局限于源代码的审计，半自动化以及全自动化必然是未来趋势。  参考文献:  [1] CVE Security Vulnerability Database. [OL][2016-12-22]  http://www.cvedetails.com/  [2] American fuzzy lop [OL] [2013-9-15] <http://lcamtuf.coredump.cx/afl/>  [3] Sulley project on Github [OL] [2015-10-18] https://github.com/OpenRCE/sulley  [4] 王清.2008.《0day安全：软件漏洞分析技术》(17.2章节).北京:电子工业出版社.358页  [5] [牛新立](http://s.wanfangdata.com.cn/Paper.aspx?q=%e4%bd%9c%e8%80%85%3a%22%e7%89%9b%e6%96%b0%e7%ab%8b%22). 基于模糊测试的虚拟化漏洞发掘工具的设计与实现 [D] 北京：北京邮电大学.2014 [6] [Jääskelä, Esa](http://jultika.oulu.fi/Search/Results?lookfor=%22J%C3%A4%C3%A4skel%C3%A4%2C+Esa%22&type=Author). Genetic algorithm in code coverage guided fuzz testing.[D]. Oulu : yliopisto, tieto- ja sähkötekniikan tiedekunta.2016 [7] IEEE. Vulnerability Analysis for X86 Executables Using Genetic Algorithm and Fuzzing.2008-11-13  [8] IEEE.Finding Software Vulnerabilities by Smart Fuzzing.2008-11-13  [9] 黄陇、于莉莉、李虎译.(美国)MichaelSutton、AdamGreene、PedramAmini著.《模糊测试：强制性安全漏洞发掘》.机械工业出版社 |
| 四、主要研究内容、要解决的问题及本文的初步方案  主要研究内容：模拟器模糊测试中，覆盖率的检测，样本的变异与优化，分离式监控测试端的开发。  要解决的问题：  问题1：如何有效地检测覆盖率？  问题2：如何通过覆盖率结果优化测试样本？  问题3：如何保留崩溃信息？  问题4：如何保证测试效率？  问题5：如何处理系统因为某些指令导致的重启及如何避免？  问题6：如何避免assert语句？  初步方案：结合gcov工具或打log完成覆盖率检测；结合编译器生成初始结构，检测路径分支；使用遗传算法来变异样本；通过tcp连接完成内外部数据交换（或者使用基于qtest实现）；调用gdb的api实现崩溃解析和堆栈回溯；利用fork来多任务测试。 |
| 五、工作的主要阶段、进度和完成时间  阶段一：完成基础框架  1.开发测试端 2016.12.26-2017.1.8  2.开发监控端的gdb接口 2017.1.9-2017.1.22  3.开发监控端的基础函数 2017.2.13-2017.2.26  阶段二：完成覆盖率检测和样本优化  1.完成覆盖率测试及学习遗传算法 2017.2.27-2017.3.12  2.编写遗传算法 2017.3.13-2017.3.26  3.配合覆盖率完成样本变异 2017.3.27-2017.4.9  阶段三：优化测试样本  1.解决assert问题 2017.4.10-2017.4.24  2.解决多任务问题 2017.4.24-2017.5.8  阶段四：整理资料，撰写报告  2017.5.8-2017.5.26 |
| 六、已进行的前期准备工作  1，搜集了语法分析相关开源代码和分析文章；  2，学习了模糊测试的一般方法和主流工具的变异方法； |

**毕业设计-----《开题报告》达成评价表**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **毕业要求指标点** | **比例** | **优** | **良** | **中** | **差** | **得 分** |
| 理解工程活动中获取相关信息的必要性与基本方法 | 15% | 12-15 | 9-11 | 6-8 | 0-5 |  |
| 能够将专业讲授课程中的创新性思想及意识在工程设计过程中体现 | 10% | 9-10 | 6-8 | 3-5 | 0-2 |  |
| 在解决工程专业复杂工程问题过程中，能基于网络空间安全专业相关背景知识进行合理分析，评价网络空间安全专业工程实践和复杂工程问题解决方案对社会、健康、安全、法律以及文化的影响 | 20% | 16-20 | 11-15 | 6-10 | 0-5 |  |
| 理解因方案实施可能会产生的后果及应承担的责任 | 15% | 12-15 | 9-11 | 6-8 | 0-5 |  |
| 能够理解和评价针对复杂工程问题的工程实践对环境、社会可持续发展的影响；理解用技术手段降低其负面影响的作用与其局限性 | 15% | 12-15 | 9-11 | 6-8 | 0-5 |  |
| 对网络空间安全领域相关的国际状况有基本了解 | 25% | 25-30 | 18-24 | 9-17 | 0-8 |  |
| **总 分** |  |  |  |  |  |  |
| 注：90-100分为优秀，80-90分为良好，70-80分为中等，60-70分为及格，59及以下为差。 | | | | | | |
| 指导教师意见  签名  年 月 日 | | | | | | |
| 学院审核意见  签名  年 月 日 | | | | | | |