软件系统测试

清华大学软件学院 刘强



教学提纲

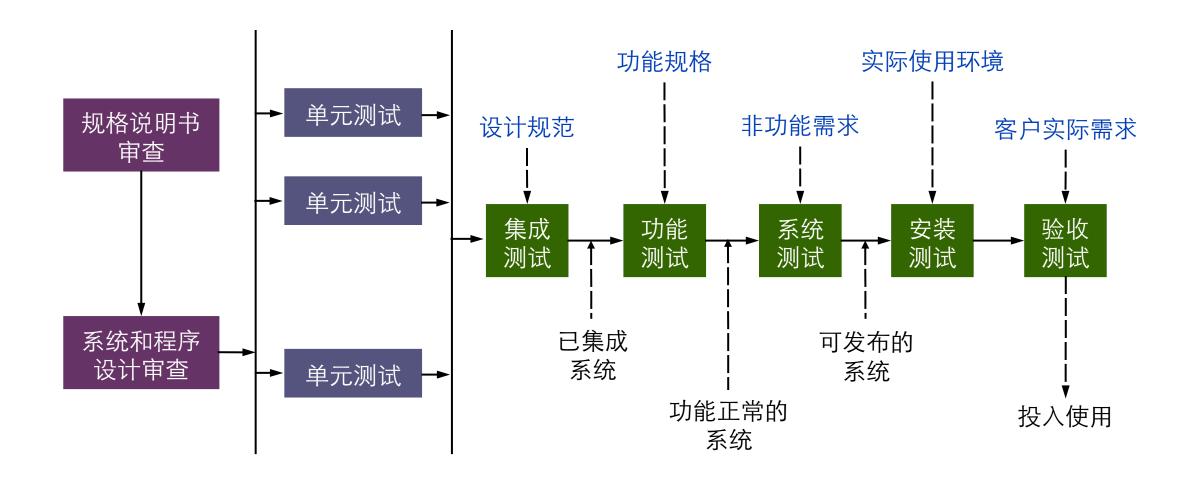


教学提纲

1

- 集成测试概念
- 基于功能分解的集成
- 持续集成
- 集成测试原则

回顾:软件测试阶段



集成测试



- 0000000000000000

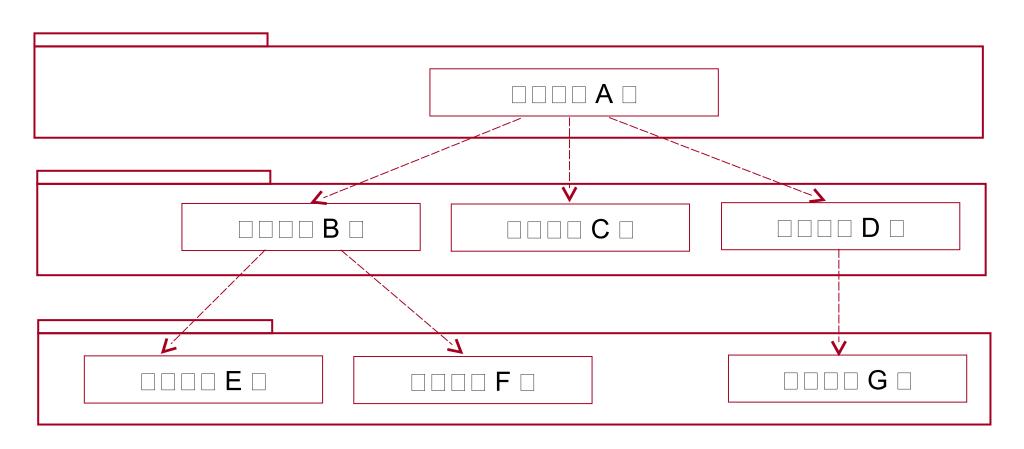


单元测试、集成测试与系统测试

测试类型	对象	目的	测试依据	测试方法	
单元测试	模块内部 的程序	消除局部模块 逻辑和功能上的 错误或缺陷	模块逻辑设计 模块外部说明	白盒测试	
集成测试	模块间的集成和 调用关系	找出和软件设计相关的 程序结构、模块调用关 系以及模块间接口方面 问题	程序结构	结合使用白盒测试 和黑盒测试,较多 采用黑盒测试	
系统测试	整个系统包括硬件、支持软件、 人员等	对整个系统进行 一系列整体的、有效性 测试	系统结构设计 目标说明书 需求说明书	黑盒测试	

集成测试策略

问题: 下面的三层结构系统如何进行集成测试?



集成测试策略

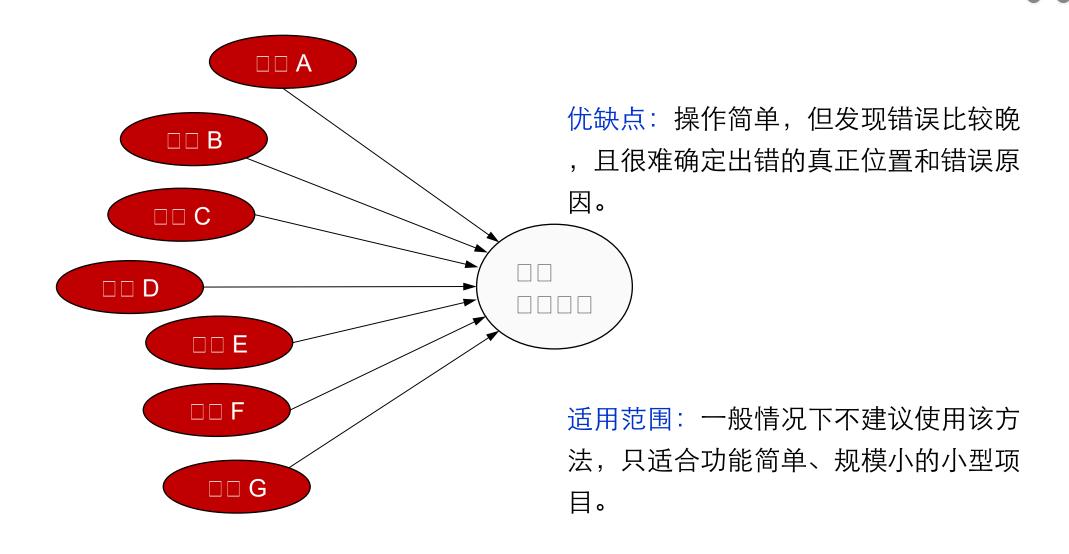
爆炸式集成

•将所有通过单元测试的模块一次性地集成到一起进行测试,也称为一次性组装或整体拼装。

递增式集成

将要测试的模块逐渐集成 到已经测试好的模块里面 ,边集成边测试,测完一 部分再连接一部分。

爆炸式集成

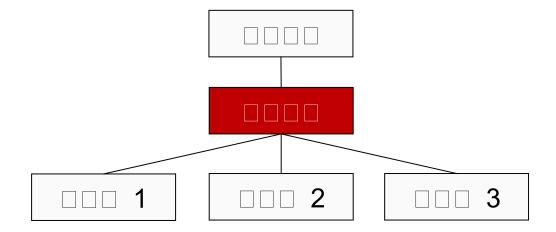


基于功能分解的集成

- 自顶向下集成: 从顶层开始采用同设计顺序一样的思路对被测系统进行测试, 一般集中于顶层的组件, 然后逐步测试处于底层的组件, 被上层单元调用的下层单元以桩出现。
- 自底向上集成:从最底层组件开始,按照分解树的结构,逐层向上集成,调用下层单元的上层单元以驱动出现。
- 三明治式集成: 这是一种混合增殖式测试策略,它将自顶向下和自底向上的集成方法有机地结合起来,从顶层和底层向中间集成,可以减少桩模块和驱动模块的开发。

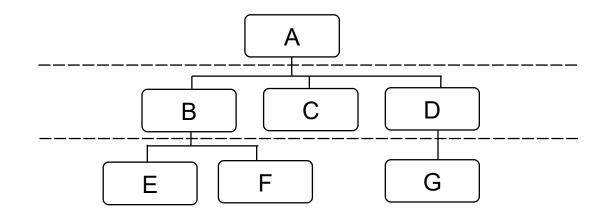
驱动模块与桩模块





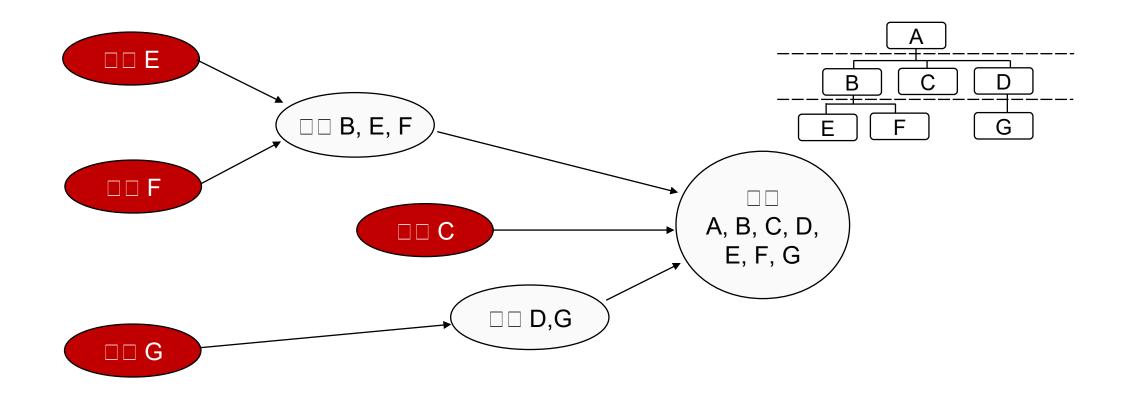


自顶向下集成



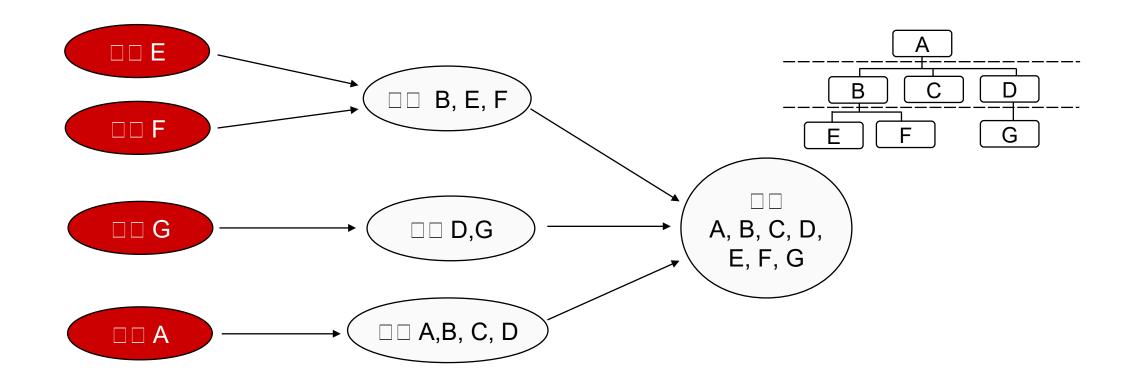


自底向上集成



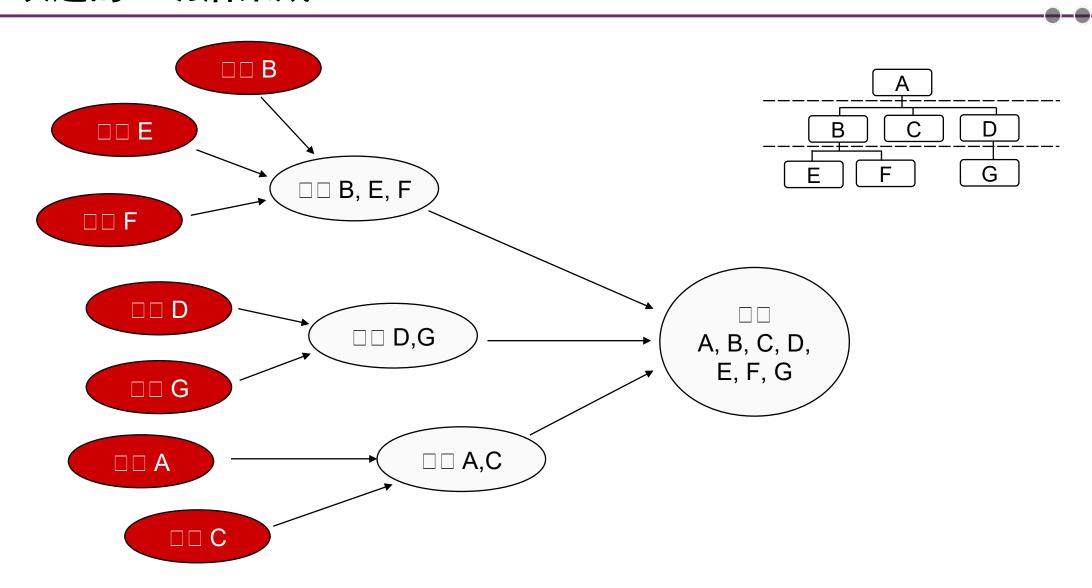
总结:该方法需要驱动模块,可以尽早地验证底层模块的行为,但不能及时发现高层模块设计上的错误;适合底层模块接口比较稳定、高层模块接口变更比较频繁的系统。

三明治式集成

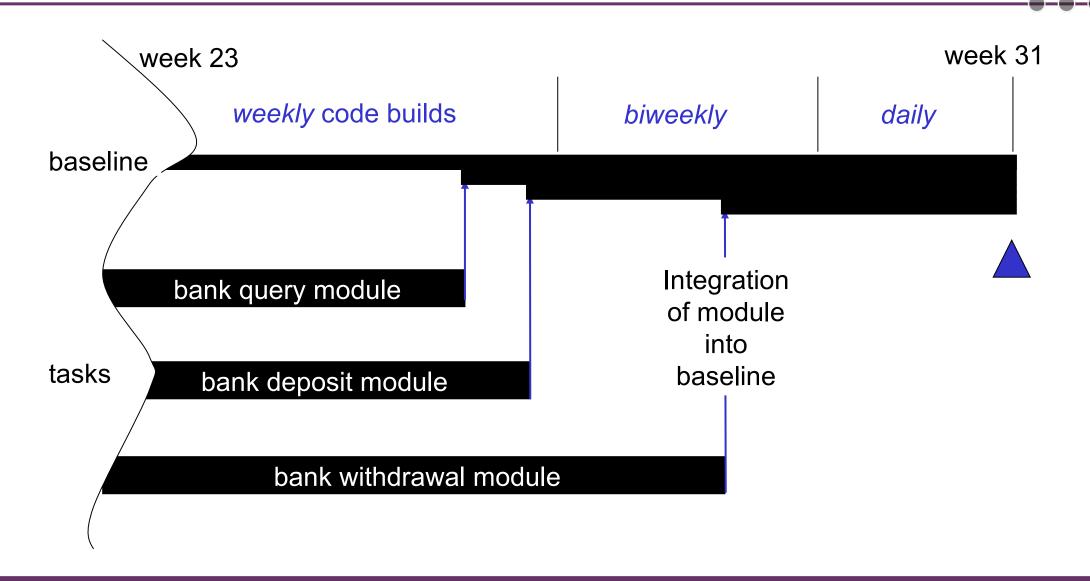


总结:该方法具有自顶向下和自底向上两种集成策略的优点,运用一定技巧可以减少桩模块和驱动模块的开发,在被集成之前中间层不能尽早得到充分的测试;适用于大多数系统。

改进的三明治集成

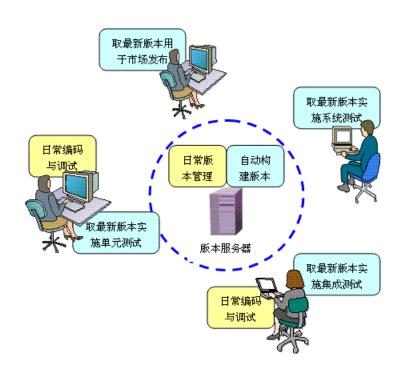


持续集成



持续集成

持续集成是一种软件开发实践,团队开发成员经常集成自己的工作,通常每个成员每天至少集成一次,即每天可能会发生多次集成。



- 所有开发人员需要在本地机器上进行本地构建, 然后再提交到版本控制库中,以免影响持续集成
- 开发人员每天至少向版本控制库中提交一次代码 ,至少从版本控制库中更新一次代码到本地机器
- 需要有专门的集成服务器来执行集成构建,并通过自动化的构建(包括编译、发布、自动化测试)来验证,从而尽快地发现集成错误

持续集成

手工构建

日常调试通常面对随 意的且尚在变化中的 测试脚本,经过整理 后可成为规范的测试 集,直接用于回归测 试。

日常调试

自动构建

通常冒烟测试与每日 构建配合应用,以每 天为单位开展自动测 试,运行一些基础用 例。

冒烟测试

自动构建

回归测试是阶段性开展的工作,它选用的版本应由构建服务器自动生成,即自动取最新源码并自动构建

回归测试

集成测试原则

尽早测试关键模块:关键模块实现了系统的主要功能,尽早测试有利于提高系统的 测试效率。

关键模块的特征:

- •满足某些软件需求
- •在程序的模块结构中位于较高的层次(高层控制模块)
- •较复杂、较易发生错误
- •有明确定义的性能要求



集成测试原则

为了做好集成测试,需要遵循以下原则:

- •所有公共接口都要被测试到
- •关键模块必须进行充分的测试
- •集成测试应当按一定的层次进行
- •集成测试的策略选择应当综合考虑质量、成本和进度之间的关系
- •集成测试应当尽早开始,并以总体设计为基础
- •在模块与接口的划分上,测试人员应当和开发人员进行充分的沟通
- •当接口发生修改时,涉及的相关接口必须进行再测试
- •测试执行结果应当如实记录

教学提纲

2

- Web □□□□□

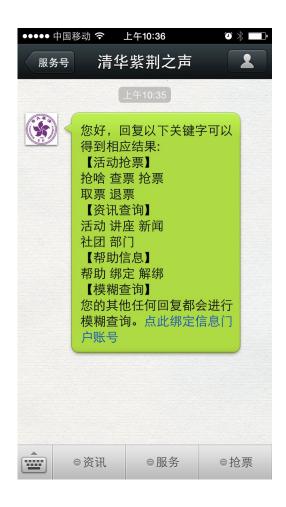
功能测试



功能测试

界面 数据 操作 逻辑 接口

- 程序安装、启动正常,有相应的提示框、错误提示等
- 每项功能符合实际要求
- 系统的界面清晰、美观
- 菜单、按钮操作正常、灵活,能处理一些异常操作
- 能接受正确的数据输入,对异常数据的输入有提示、容错处理等
- 数据的输出结果准确,格式清晰,可以保存和读取
- 功能逻辑清楚,符合使用者习惯
- 系统的各种状态按照业务流程而变化,并保持稳定
- 支持各种应用的环境
- 配合多种硬件周边设备
- 软件升级后,能继续支持旧版本的数据
- 与外部应用系统的接口有效



思考:

- •左图是首次关注后系统显示的信息,你有什么看法?
- •作为一个用户,你希望首次关注后看到什么信息?







上一篇 没有下一篇了

远去的人, 祝你们幸福。

-----题记

年少轻狂,有多少不眠的夜 晚, 幼稚地臆断着岁月的形 状,生命的远足。似乎,参 透了时光的流转, 命运的纠 缠,人世的纷纷扰扰,聚聚 散散。谁知道呢,华丽的辞

姜观? 清晰?







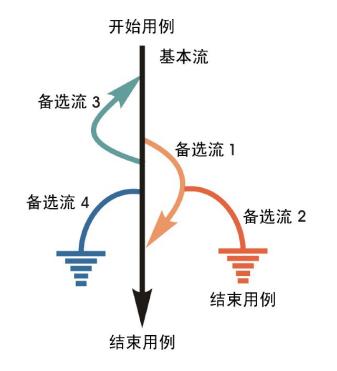
适合性测试

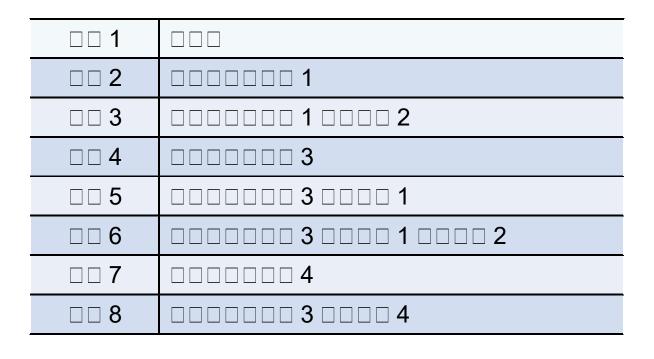
适合性测试是评估和确认软件产品是否能够帮助用户执行期望的任务或者实现了用户需要的功能。



适合性测试









- _____



















- 抢到多张票时,系统如何显示电子票?
- 抢票未开始时是否可以抢票?
- 抢票结束且活动未开始时是否可以抢票?
- 绑定不同账号进行抢票,互相有影响吗?
- 一个活动抢票时会影响其他活动状态吗?

准确性测试

准确性测试是验证软件产品提供具有所需精度的正确或者相符的结果或者效果的能力。



• 预期的准确性:在运行期间执行任务的实际结果与预期结果的差别。

• 计算的准确性:由于计算本身的错误和数据精度方面的问题导致的错误计算结果。

准确性测试



商品名	单件积分	市场单价	网站单价	优惠	数量	删除
WENGER双肩电脑包	0	¥620.00	¥290.10	减¥91.10	1	删除
空间大师DPC003金属四层车	0	¥256.00	¥100.00	减至15.00	1	删除
欧姆龙电子血压计HEM-8102	269	¥438.00	¥269.00	无	1	删除
爱国者 aigo(4GB)U盘 8380	0	¥99.00	¥69.00	减至4.00	2	删除
好孩子木餐椅MY303A-E531绿色	339	¥369.00	¥339.00	无	1	删除
讲故事学折纸(全3册)	399	¥65,40	¥39.90	无	2	删除
		¥910.00 品积分; 140	6 商品金额	总计: ¥1101.80	结	算->

互操作性测试

互操作性测试是验证软件产品在多种指定的目标环境下(如硬件、软件、中间件、操作系统、浏览器等)是否可以正常工作。

- 输入互操作性:软件产品可以正常使用与之集成的其他软件产品或系统的输出(数据或协议格式等)。
- 输出互操作性:软件产品必须输出标准或协议规定的格式。
- 软件产品是否具有自动检测与该软件交互的系统的通信请求并作出相应的处理。





- 与学校信息门户连接(用户认证)
- 与学生清华系统连接(抢票结果)
- 不同品牌的手机
- 不同的操作系统
- 与扫码器连接(电子票、纸质票)
- 与打印机连接(打印电子票)

安全性测试

安全性测试是查找软件自身程序设计中存在的安全隐患,并检查软件产品对非法侵入的防范能力。

- 对应用程序或数据进行未授权的复制
- 未授权的访问控制
- 缓存区溢出
- 服务拒绝
- 在网络上窃听数据传输来获得敏感信息
- 破解保护敏感信息的加密代码
- 逻辑炸弹





XSS DECEMBER OF THE STATE OF TH



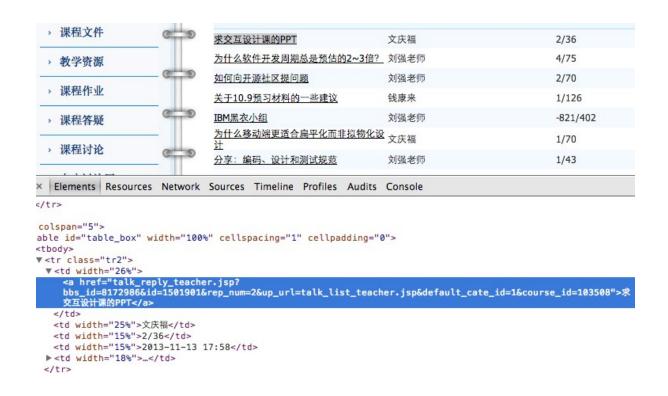
系统是否防止 755 注入

回复时间

2013-10-02 20:31



编辑



没有对删贴操作 进行权限验证!

http://learn.tsinghua.edu.cn/MultiLanguage/public/bbs/talk_manager_action.jsp? note_id=1501901&bbs_id=8172986&Commit=%E5%88%A0%E9%99%A4&course_id=103508&default_cate_id=1





安装测试应该包括哪些方面?

软件安装测试

- 安装过程中对于缺省目录及任意指定目录,是否都能正确安装;
- 若是选择安装,查看能否实现其相应的功能;
- 在所有能中途退出的位置退出安装程序后,验证程序并未安装成功;
- 软件安装后,对其它已经安装的软件是否有影响;
- 裸机安装后,各功能点是否可用;
- 在安装前,安装程序是否判断可用磁盘空间大小,如果不满足安装空间要求,安装程序能否继续 ;
- 安装时查看版权声明、版本信息、公司名称、 LOGO 等是否符合标准;
- 安装过程中界面显示与提示语言是否准确、友好;
- 重复安装时系统是否有提示、是否可以覆盖安装、是否可以升级安装、是否允许多版本共存;
- 是否有注册码或硬件加密狗,在没有它们(或错误)存在的情况下能否顺利安装。

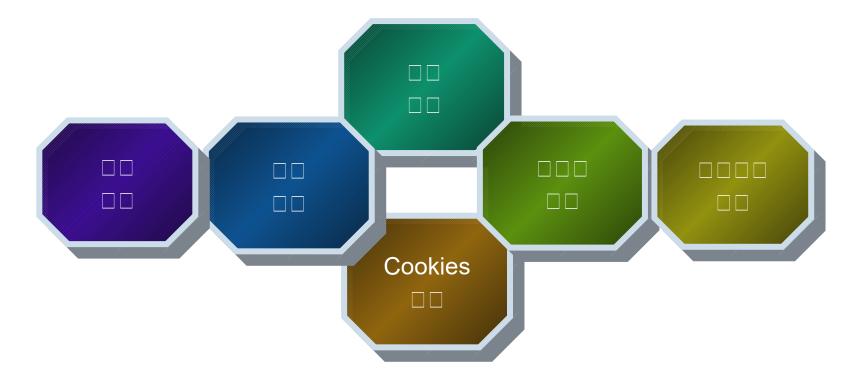
软件卸载测试

- 卸载后注册表中的注册信息及相关的程序安装目录是否能完全删除掉;
- 卸载过程中完全删除共享文件后,看其它程序能否正常运行;
- 卸载后,是否对其它已经安装的软件有影响;
- 系统卸载后用户建立文档是否保留;
- 软件卸载画面上的软件名称及版本信息是否正确;
- 在所有能中途退出卸载的位置是否能正确退出;
- 卸载过程中界面显示与提示语言是否准确、友好;
- 卸载后安装此系统能否打开原来保存的文件,并一切运行正常;
- 卸载程序如果要求重新启动机器,在重启之前是否提示用户保存现有的己运行程序的资料;
- 是否可以选择组件进行卸载;
- 卸载过程中,对意外情况的处理(掉电等)。
- 在卸载过程中,是否有终止或者结束按钮。

运行与关闭测试

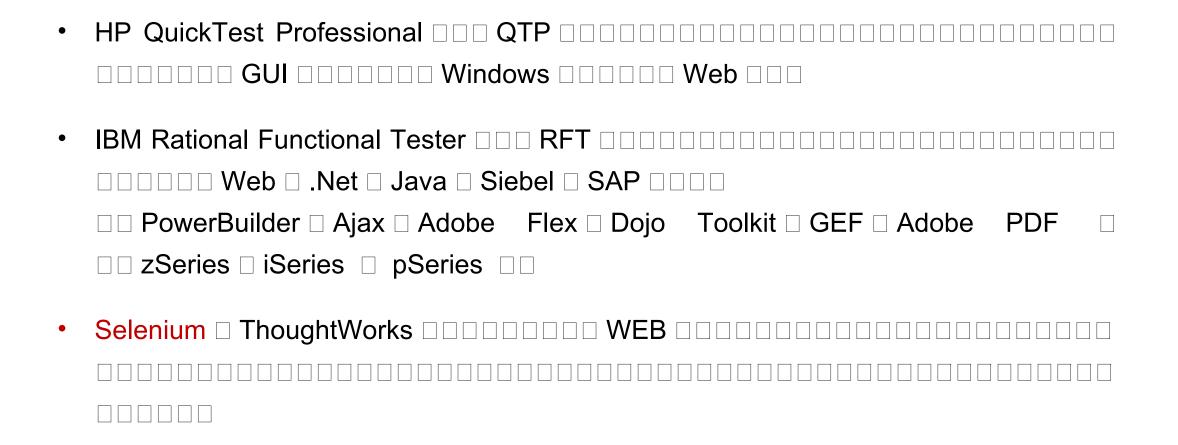
- 运行时是否与其它应用程序有冲突(内存冲突);
- 是否可以同时运行多个程序;
- 任务栏有无程序运行提示;
- 若有未保存的数据,关闭系统时是否有提示;
- 后台服务程序在点击关闭按钮时是否有确认提示;
- 运行时是否过份占用系统资源、退出时能否完成释放占用的系统资源。

Web

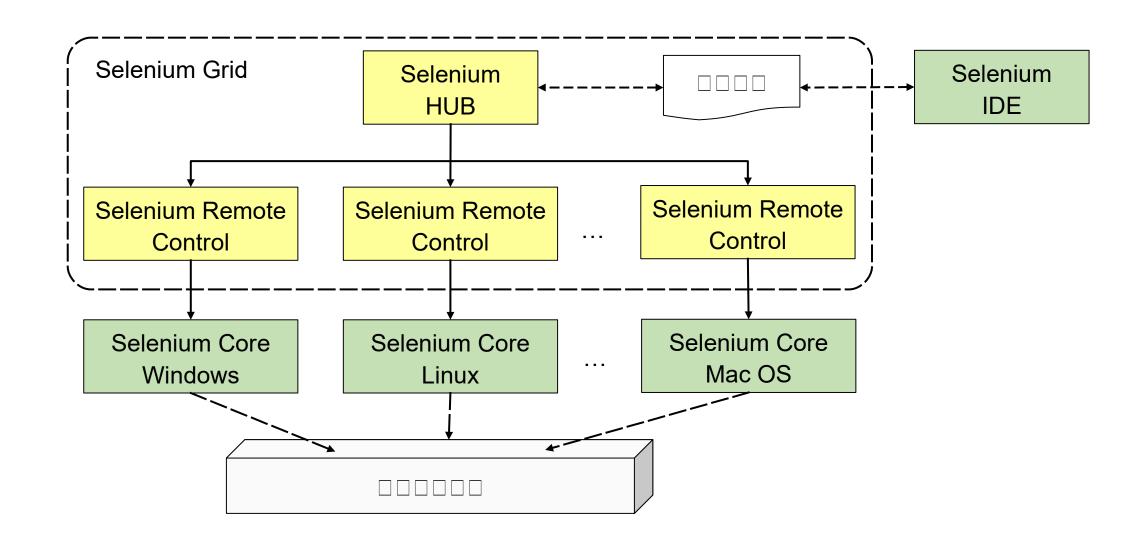




功能测试工具



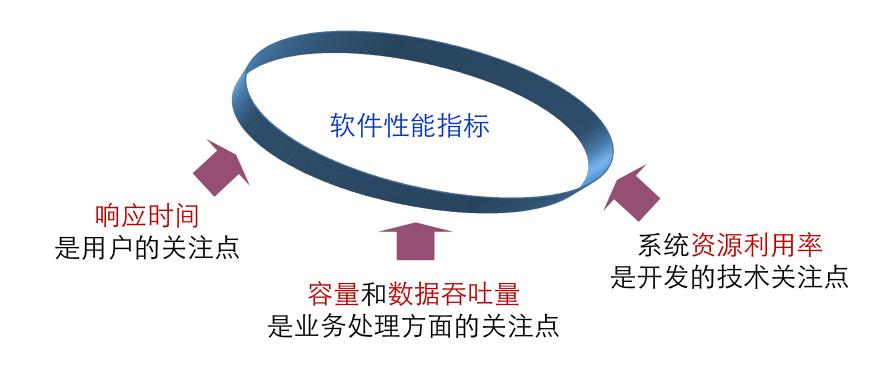
Selenium



教学提纲

软件性能

软件性能是软件的一种非功能特性,它关注的不是软件是否能够完成特定功能,而是在完成该功能时展示出来的及时性、资源占用、稳定性、安全性、兼容性、可扩展性、可靠性等。

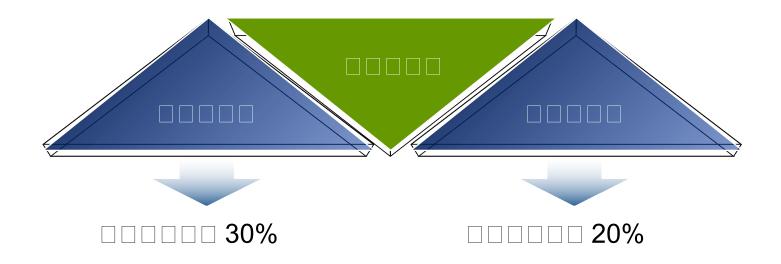


用户数

• 注册用户数: 所有在系统注册的用户数目

• 在线用户数: 所有正在访问系统的用户(不一定做操作)数目

• 并发用户数: 在某一给定时间内某特定时刻进行会话操作的用户数

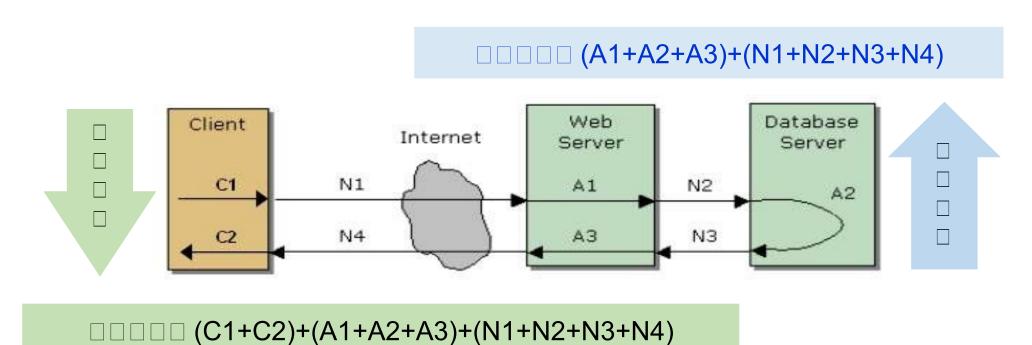


举例:新浪微博系统

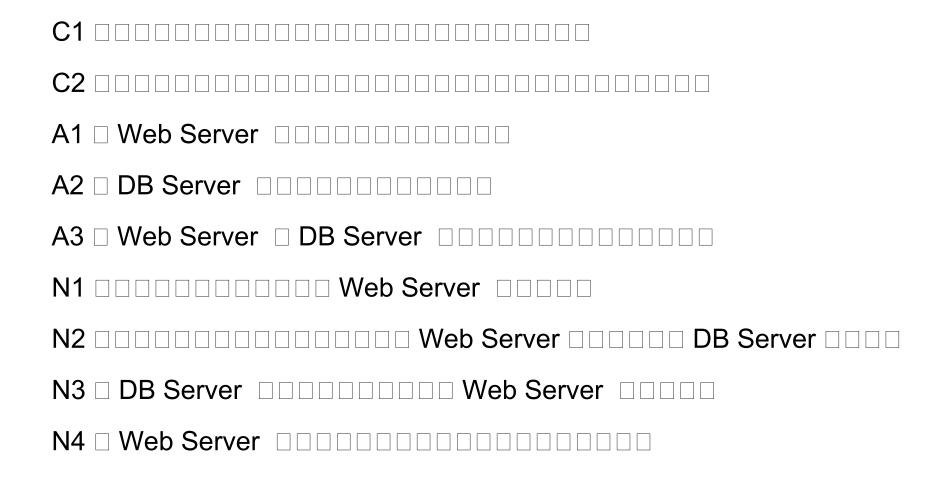
- •假设用户只有注册后才能使用,但注册用户并不是每时每刻都在使用该网站,因此具体一个时刻只有部分注册用户同时在线。
- •在线用户在浏览网站时会花很多时间阅读网站上的信息,因而具体一个时刻只有部分在线 用户同时向系统发出请求。
- •由于注册用户可能长时间不登陆网站,使用注册用户数作为性能指标会造成很大的误差。 在线用户数和并发用户数都可作为性能指标,但性能测试更多关心并发用户数。

响应时间

• 从客户端发出请求到获得响应的整个过程所经历的时间



52



吞吐量



资源利用率





性能测试是通过自动化测试工具或手段模拟多种正常、峰值以及异常负载条件对系统的各项性能指标进行的一种测试。

负载测试

通过逐渐增加系统负载,测试系统性能的变化,最终确定<mark>在满足性能</mark> 指标的情况下,系统能承受的最大负载量。

目标: 在特定的运行条件下验证系统的能力状况。

压力测试

通过逐步增加系统负载,测试系统性能的变化,最终确定在什么负载 条件下系统性能<mark>处于失效状态</mark>,从而获得系统能够提供的最大服务级 别的测试。

目标: 发现在什么条件下应用程序的性能会变得不可接受。

压力测试类型

- 稳定性压力测试: 高负载下持续运行 24 小时以上。
- 破坏性压力测试:通过不断加载的手段,快速造成系统的崩溃,让问题尽快地暴露出来。
- 渗入测试:通过长时间运行,使问题逐渐渗透出来,从而发现内存泄漏、垃圾收集或系统的其他问题,以检验系统的健壮性。
- 峰谷测试:采用高低突变加载方式进行,先加载到高水平的负载,然后急剧降低负载,稍微平息一段时间,再加载到高水平的负载,重复这样过程,容易发现问题的蛛丝马迹,最终找到问题的根源。

大数据量测试

- 独立数据量测试是针对某些系统存储、传输、统计、查询等业务进行的大数据量测试。
- 综合数据量测试是系统在具备一定数据量时,在负载压力测试下考察业务是否能够正常进行的测试。
- 目标:测试数据量较大时系统的性能状况。

疲劳强度测试





失效恢复测试

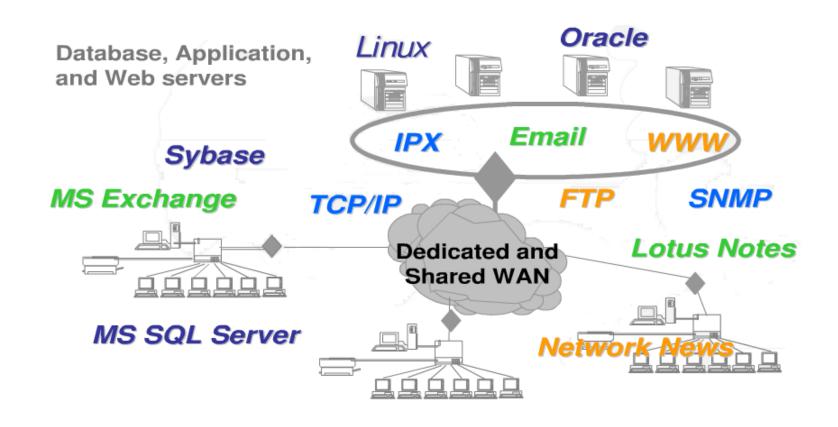
- 对于有冗余备份或负载均衡的系统来说,检验如果系统局部发生故障,系统灾备措施 是否可以正常启动,用户是否可以继续使用。
- 目标: 通过实施失效恢复测试,评估系统的健壮性和可恢复性。

基准测试



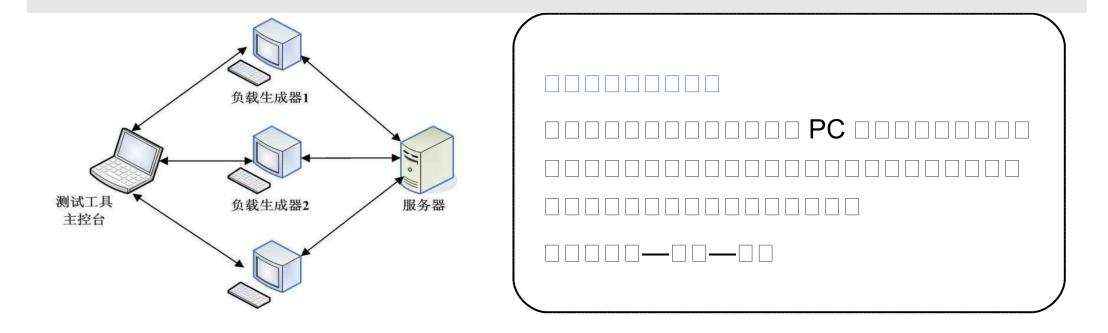


问题: 此何找出问题的根源?

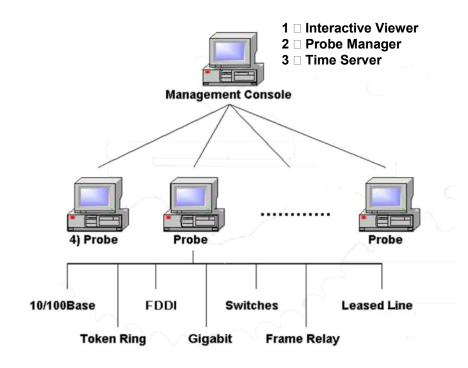


0

应用在客户端性能的测试:逐渐增加并发虚拟用户数负载,直到系统的瓶颈或者不能接收的性能点,通过综合分析交易执行指标、资源监控指标等来确定系统的并发性能



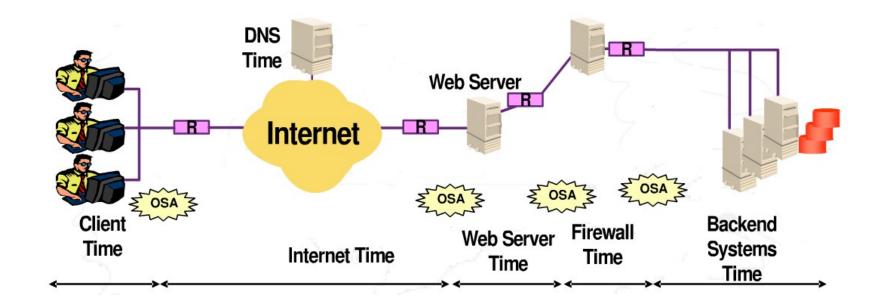




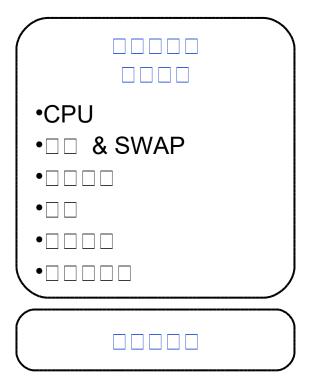


典型的故障诊断策略: 设置多点代理服务

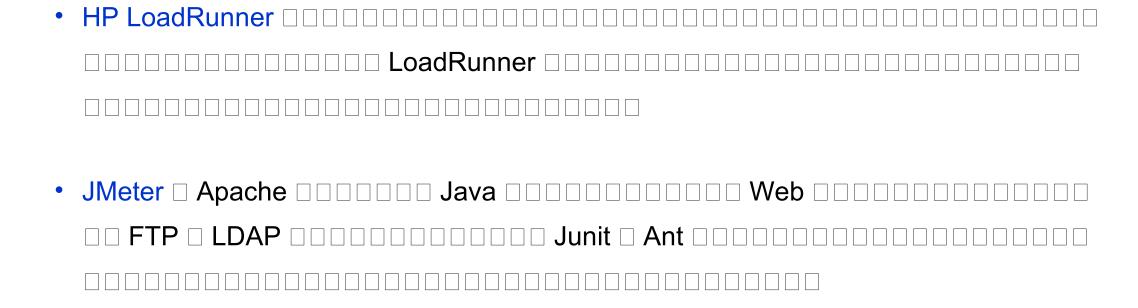
- •在应用逻辑路径上多点采集
- •在任何两个节点之间进行数据整合,测量分段响应时间,分析应用故障



应用在服务器性能的测试:测试关键点是资源占用情况、数据库性能和故障报警。对于应用在服务器上性能的测试,可以采用工具监控,也可以使用系统本身的监控命令。



性能测试工具



教学提纲



软件缺陷举例







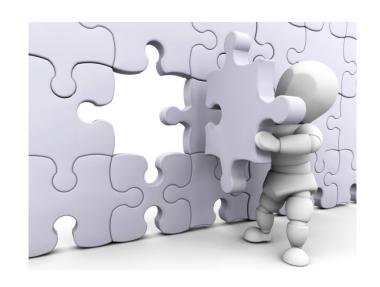




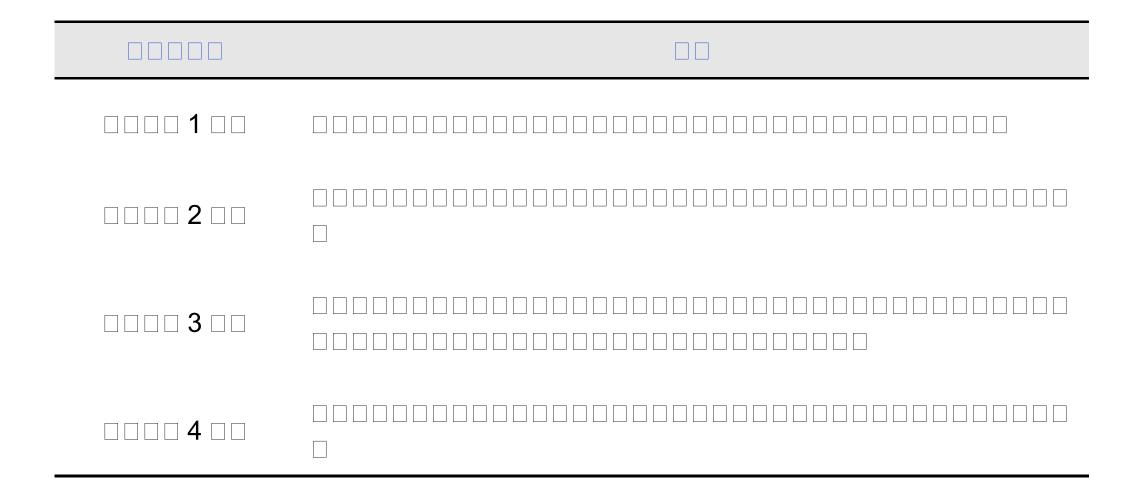
严重性与优先级

缺陷的<mark>严重性</mark>是指缺陷对软件产品使用的影响程度,缺陷的<mark>优先级</mark>是指缺陷应该被修复的紧急程度。

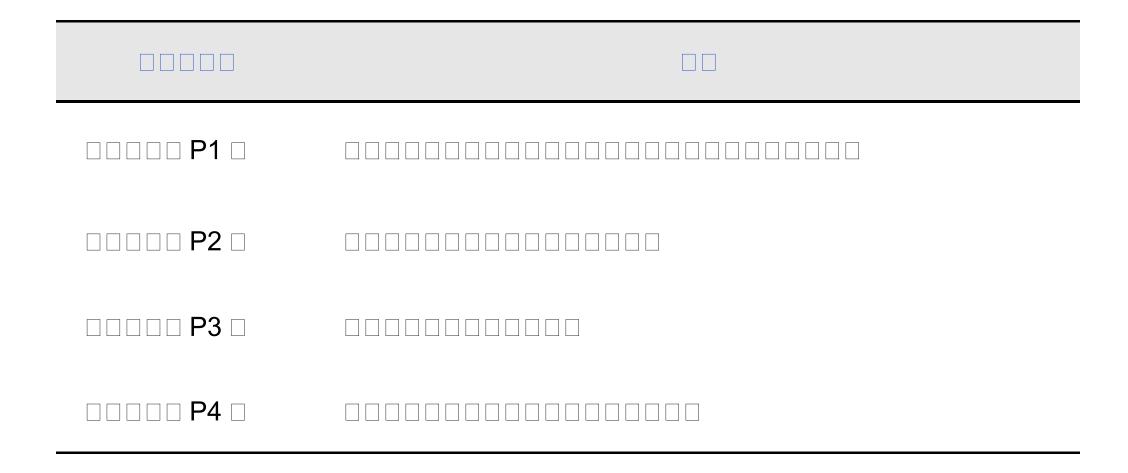
- 一般情况:缺陷越严重,越要优先得到修正, 缺陷严重等级和缺陷优先级相关性很强。
- 特殊情况:如有些缺陷比较严重,但由于技术 的限制或第三方产品的限制,暂时没法修正, 其优先级就会低。



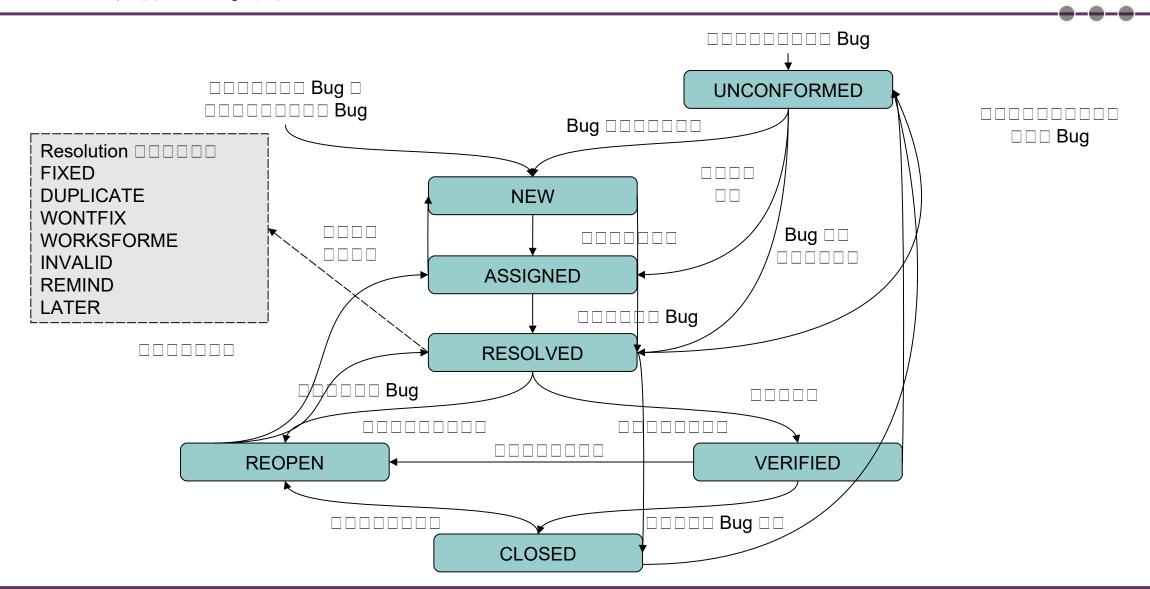
严重性与优先级



严重性与优先级



软件缺陷生命周期

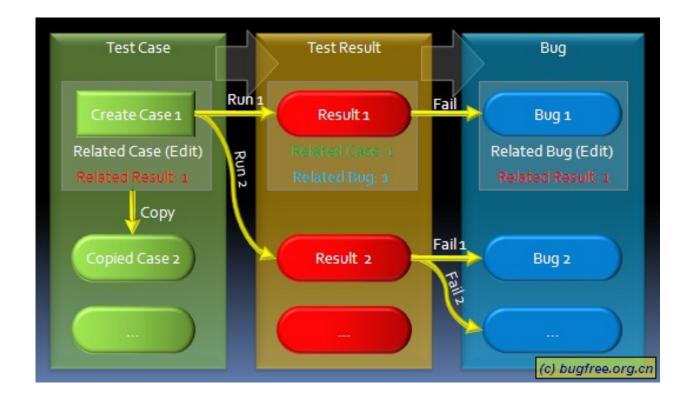


软件缺陷属性

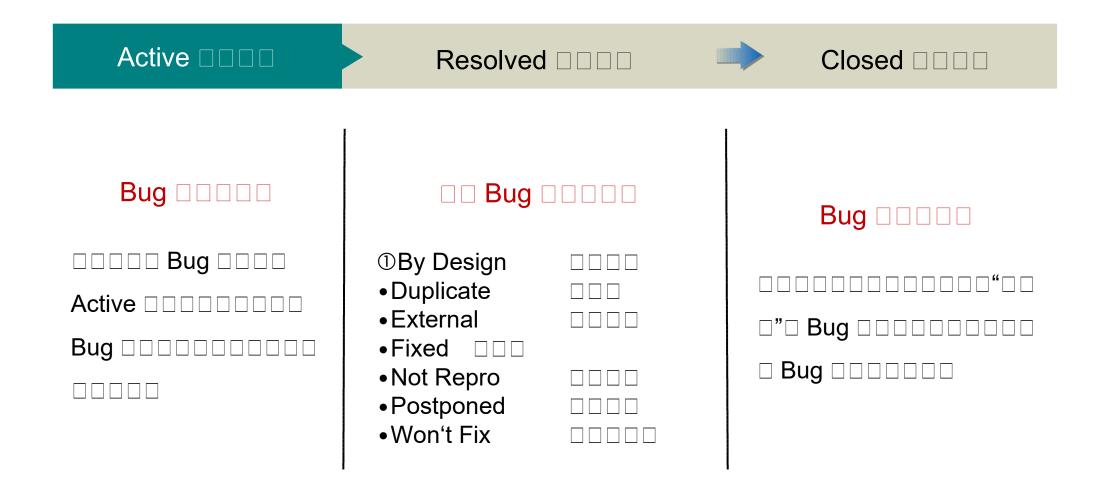
- 缺陷编号:编号是唯一的,可由缺陷跟踪系统自动生成。
- 缺陷标题: 简要说明缺陷的类型和内容。
- 缺陷所属: 指明缺陷所在模块或所属文档的名称。
- 缺陷严重程度: 缺陷对软件产品使用的影响程度。
- 缺陷优先级: 缺陷应该被修复的紧急程度。
- 缺陷处理角色:提交人、解决人、确认人
- 缺陷处理状态:
 - ① 确认结果:确认软件缺陷的修正工作是否有效
 - ② 解决结果:预计缺陷修改后能达到的结果
 - ③ 处理结果: 缺陷最后的实际处理结果
- 时间信息:提交日期、预计解决日期、确认日期、最终处理日期

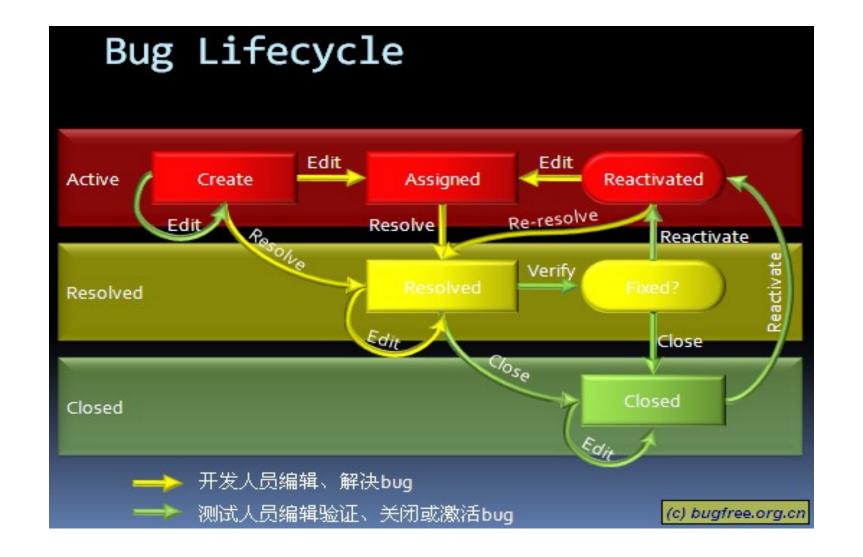
缺陷管理系统

BugFree DD PHP MySQL DD DD DD DD DD Linux Windows DD



Bug 🗆 🗆





□ □ Bug





□ □ Bug





. . .

谢谢大家!

THANKS

