

关于 软件稳定性测试 的思考与实践

邱鹏 Ricky Qiu

Staff QA Engineer, Trend Micro CDC

About Ricky

2005年加入趋势科技中国研发中心，现为企业级安全产品高级测试工程师。有丰富的软件性能测试和系统稳定性测试工作经验，并负责部门相关测试工具的评估和选型。毕业于华中科技大学，**CS Master**。

Contact me:

Blog: <http://blog.csdn.net/superqa>

MSN: qiupeng2002@hotmail.com

Email: rickyqiu1024@gmail.com

Survey Result

4. 请选择您正在或者最近做的产品的类型（可多选）

MIS类	2
OA类	1
Windows客户端软件	2
手机等移动设备上的软件	1
服务器软件	5
互联网产品	9
通信设备	1
其他	GPS, SaaS, Test tool

5. 在您参与的项目中，系统稳定性测试包含在测试范围之内吗？

包含	14
不包含	4
不清楚	

6. 您自己参与过系统稳定性测试的设计和执行吗？

设计	1
执行	1
设计和执行	5
均未参与	11

Warm up

测试的目的是为了更好的质量。

你觉得一个**高质量的空调**应该具备的特性包括哪些？

关于一个高质量的空调

好的特性	分类
1. 制冷/制热很快	性能很好
2. 噪音很低	对环境的影响小
3. 很省电	高效的使用资源
4. 持续长时间（>24小时？）的稳定工作	稳定性好
5. 在寒冷（< -20°）的冬天也能启动，正常工作	环境的适应能力
6. 在高温（>40 °）的情况下也能正常工作	环境的适应能力
7. 电压不稳时，能工作，至少不损坏	稳定性好
8. 能持续使用很多年（> 8年？）	稳定性好

不只是功能！

关于软件

“在人类历史上，从来没有一个社会像现在这样完全依赖于一种经常具有严重缺陷的产品。”

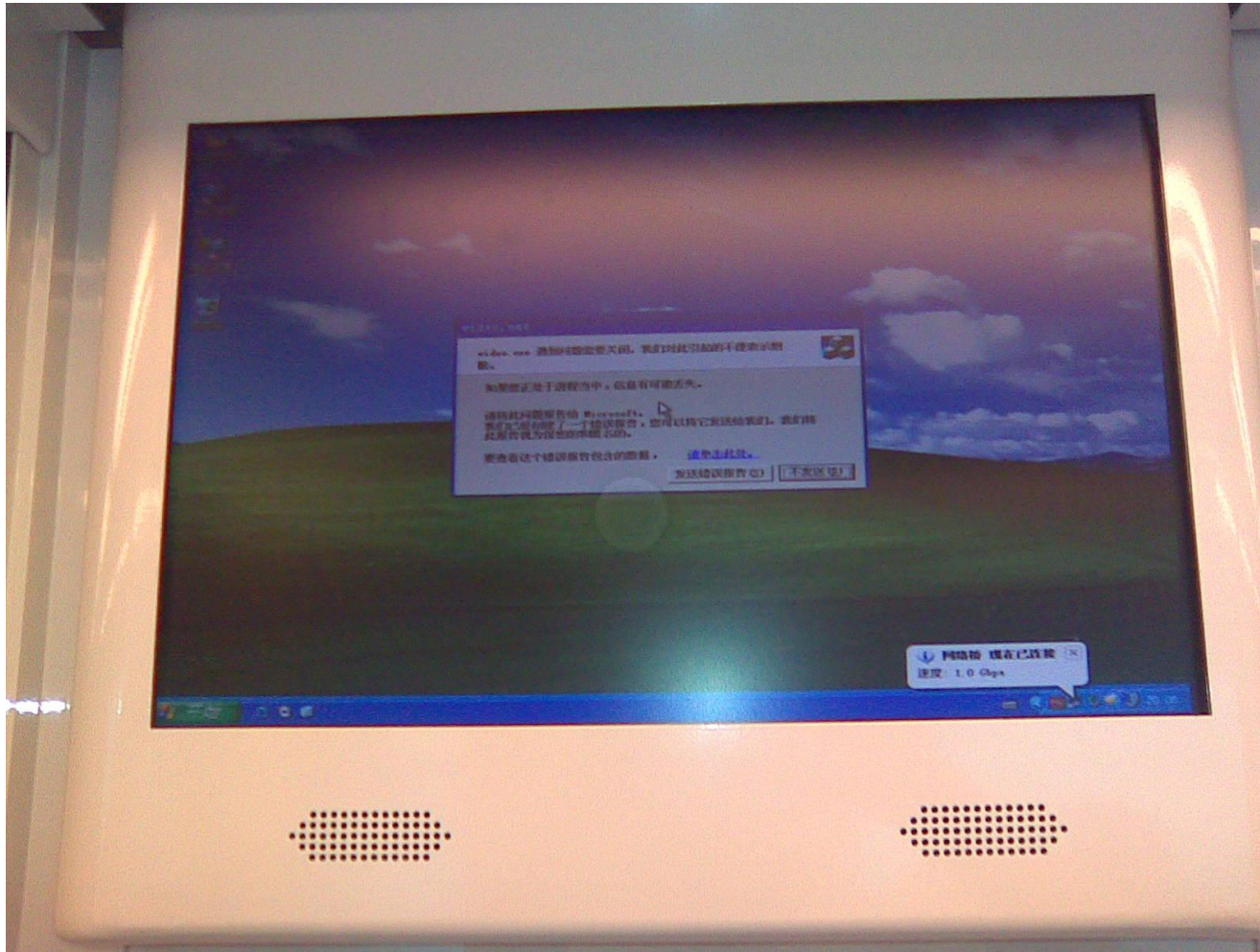
-- 《探索式软件测试》 by James Whittaker



1996年6月4日，阿丽亚娜5型运载火箭的首航，运送4颗卫星。升空37秒后发生爆炸。

5型火箭沿用4型的导航系统SRI，但速度是其5倍。所需计算的数据在4号火箭上从未出现，火箭点火升空后测算所得的水平速度以64位浮点数值表示，SRI将其转换为16位的整数值，但因为16位整数值并不足以表示火箭的水平速度，最终导致系统的内存溢出，并复写了部分数据。

Something around us



“工业界利用软件测试来达到创新性和可靠性之间的平衡。”

我们每个功能都测到了，为什么到用户那里还是有那么多问题？



lab vs field

还需要系统稳定性测试！

系统稳定性测试，Why?

系统稳定性的bug可能只占产品所有bug的5%*，但是每一个bug都可能导致高昂的技术支持/维护代价，甚至用户退货或者产品召回。

* 内部统计数据

**Q: 我们应该从哪些方面来设计
稳定性测试呢？**

1. 时间

产品能持续长时间工作吗？

可能出现的问题

1. 内存或者句柄泄漏
2. 队列堆积
3. 存储等慢速设备持续跟不上
4. **Cache, Buffer**累积性的问题

...



10.204.148.112-Win2k8-lab
10.204.168.112 - Win2k8-64-ESX
10.204.168.203
10.204.168.166
10.204.150.88
10.204.168.114
10.204.168.110
10.204.168.191-Win2k3
10.204.148.213
10.204.168.207-Perf
10.204.168.35
10.204.168.174
10.204.168.123
10.204.168.179

Windows Task Manager

File Options View Help

Applications Processes Services Performance Networking Users

Image ...	PID	User Name	CPU	Memory (Private Working ...)	Handles	Threads	I/O Read Bytes	I/O Write Bytes	Description
csrss.exe	496	SYSTEM	00	848 K	1,255	13	490,046	0	Client Server
csrss.exe	544	SYSTEM	00	80 K	76	7	116,872	0	Client Server
csrss.exe	4020	SYSTEM	00	340 K	359	10	2,107,439	0	Client Server
dwm.exe	4164	Administ...	00	100 K	77	3	276	348	Desktop Wind
explorer.exe	4188	Administ...	00	2,744 K	888	21	16,493,627	3,745,019	Windows Expl
httpd.exe	6608	SYSTEM	00	2,041,796 K	528	66	293,773,742	1,254,500	Apache HTTP
httpd.exe	7616	SYSTEM	00	288 K	88	4	14,466	1,394	Apache HTTP
inssausched.exe	3116	SYSTEM	00	6,780 K	233	2	78,613	17,469	Trend Micro Ir
TMSSChild.exe	2380	SYSTEM	00	138,888 K	304	17	97,725,824	20,285,321	Trend Micro Ir

EUQ Test Plan 2.jmx (E:\Tools\jakarta-jmeter-2.3.2\script\EUQ Test Plan 2.jmx) - Apache JMeter (2.3.2 r665936)

文件 编辑 运行 选项 帮助

500 / 500

EUQ Test Plan

- HTTP Header Manager
- HTTP Cookie Manager
- HTTP Request Defaults
- User Parameters
- Thread Group
- G-Login
- G-View
- G-Delete
- G-Order
- G-NotSpam
- Monitor Results
- Aggregate Report
- View Results Tree

工作台

聚合报告

名称: Aggregate Report

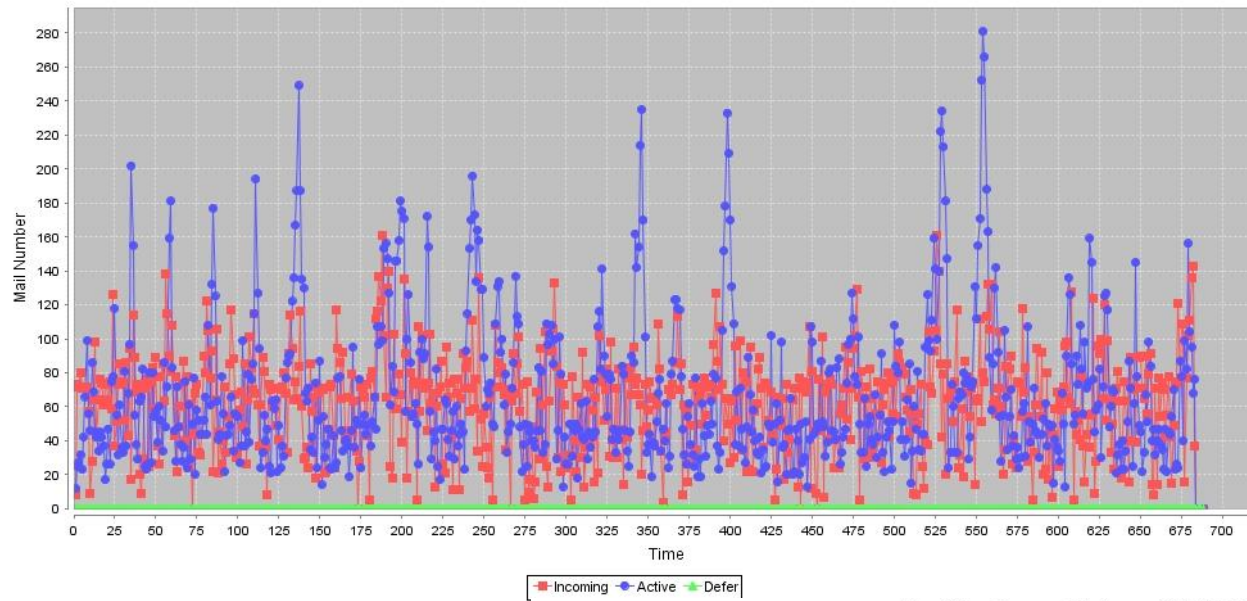
注释:

所有数据写入一个文件

文件名: aaa 浏览... Log/Display Only: ☐ 仅日志错误 ☐ Successes

Label	# Samples	Average	Median	90% Line	Min	Max	Error %	Throughput	K
G-Login:T-Login	14729	1249	397	3277	4	22300	15.94%	9.5/sec	
G-Order:T-Login	12706	1061	205	3276	4	21673	10.77%	8.2/sec	
G-Delete:T-Login	12536	1128	206	3278	4	21773	10.13%	8.1/sec	
G-View:T-Login	12945	1140	215	3281	4	21732	12.04%	8.3/sec	
G-NotSpam:T-Login	12716	1109	212	3278	4	21730	11.31%	8.2/sec	
G-Login:T-Logoff	14683	713	110	1141	1	20764	16.22%	9.5/sec	
G-Order:T-OrderBy	12655	1234	162	3324	2	21779	8.87%	8.2/sec	
G-View:T-View	12897	1327	256	3607	2	21917	14.83%	8.3/sec	
G-Delete:T-Delete	12491	1455	354	3998	2	21954	11.86%	8.1/sec	
G-NotSpam:T-NonSpam	12671	1481	387	4014	2	22433	13.02%	8.2/sec	
G-Order:T-Logoff	12620	584	44	1071	0	20856	10.92%	8.2/sec	
G-View:T-Logoff	12860	598	49	1070	1	20828	12.19%	8.3/sec	
G-NotSpam:T-Logoff	12634	648	55	1077	1	20221	11.38%	8.2/sec	
G-Delete:T-Logoff	12454	619	50	1073	1	20428	10.25%	8.1/sec	
总体	181597	1024	173	3064	0	22433	12.22%	117.0/sec	

Postfix Queue Status - 10.204.168.207 - Parent



Postfix Queue Status - 192.168.21.167 - Parent



2. 大流量

产品能处理这么大的流量吗？

1. 系统崩溃 crash/hang
2. 性能严重下降
3. 资源耗尽
4. DB或者磁盘空间满
5. 触发系统的一些行为
 - 流量控制
 - 进/线程动态增长
 - 文件purge/rotate
 - ...



奥运门票预售暂停5天 系统开工半小时即瘫痪

http://news.xinhuanet.com/sports/2007-10/31/content_6977044.htm

“北京奥运会官方票务网站的浏览量在第一个小时达到800万次，每秒钟从网上提交的门票申请超过20万张；票务呼叫中心热线的呼入量超过了380万人次。由于瞬间访问数量过大，技术系统应对不畅，造成很多申购者无法及时提交申请。至上午11点，各个销售渠道仅售出门票约 9000张”



Dynamic Process Number



3. 并发

单独运行是好的，那么一起运行呢？

1. 竞争条件
2. 功能之间的冲突
3. 后台定期任务
 - 数据的统计和汇总
 - heartbeat状态查询
 - 模块/数据文件更新
 - 日志的purge和rotate



功能之间

- Status Monitor Tested by QA_1

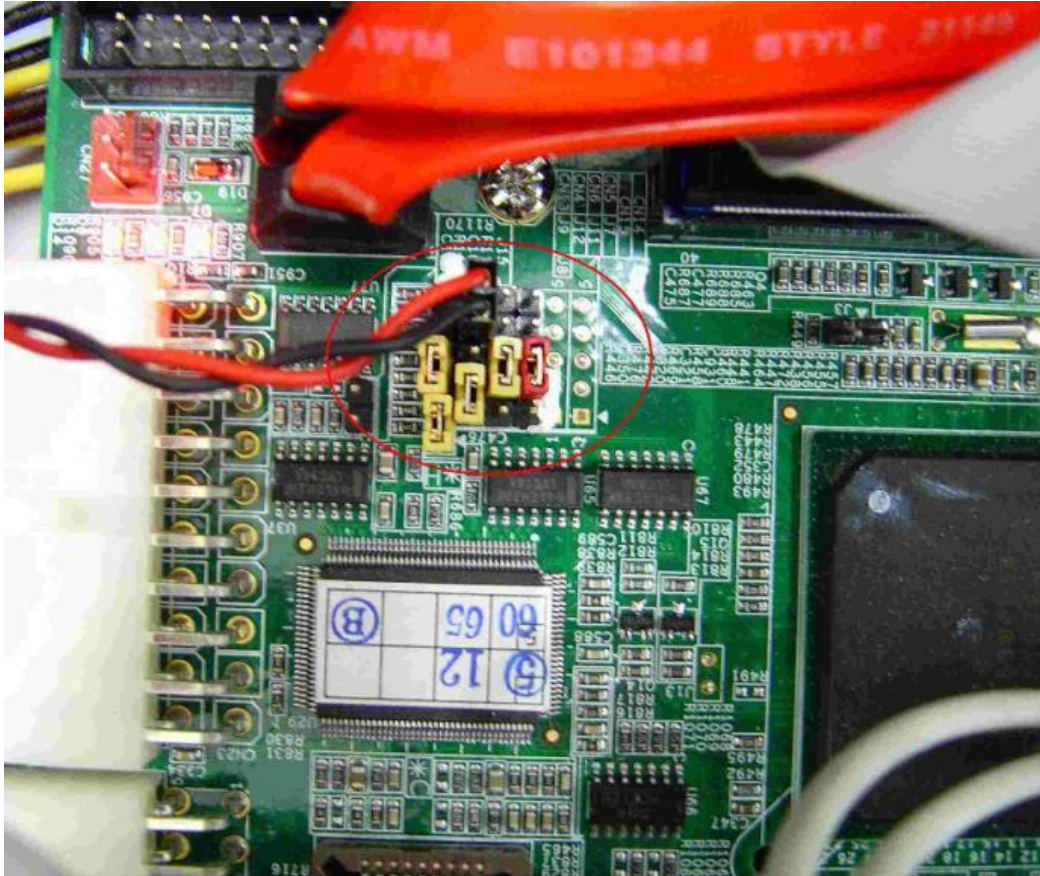
检查SMTP服务是否可用，每5秒

- Telnet [IP_SMTP_Server] 25
- If get 2xx, means OK. 否则重启服务

- SMTP Server Tested by QA_2

- 防攻击的功能
- 如果来自某个IP的客户端连上了，什么都不做，就断开，那么超过一定次数，会被列入黑名单。

功能和性能



HW watchdog

`#/sbin/wd`

`#insmod
/lib/modules/panic.ko`

4. 使用环境

在不同用户的环境下都能正常工作吗？

1. 平台

硬件, OS (Win2008, RHEL 5.3)

2. 网络

低带宽, 丢包, 延迟

Proxy, Firewall, DMZ

3. 其它基础架构

DNS, LDAP, 虚拟化...

4. 和其它应用的共存

...



First bug?

9/9

0800 Andam started
 1000 " stopped - andam ✓


1300 (032) MP-MC { 1.2700 9.037 847 025
 1.982647000 9.037 846 995 correct
~~2.130476415~~ 4.615925059(-2)

(033) PRO 2 2.130476415
 correct 2.130676415

Relays 6-2 in 033 failed special speed test
 in relay " 10.000 test.

Relays changed

1100 Started Cosine Tape (Sine check)
 1525 Started Multi-Adder Test.

1545  Relay #70 Panel F
 (moth) in relay.

First actual case of bug being found.

~~1630~~ 1630 Andam started.
 1700 closed down.

Relay 3145
 Relay 3370

Mark II Computer (Relay Calculator), Harvard University, 9 September 1947.

运营商的DNS重定向

```
C:\WINDOWS\system32\cmd.exe - nslookup
D:\>
D:\>nslookup
*** Can't find server name for address 10.204.16.18: Non-existent domain
*** Default servers are not available
Default Server: UnKnown
Address: 10.204.16.18

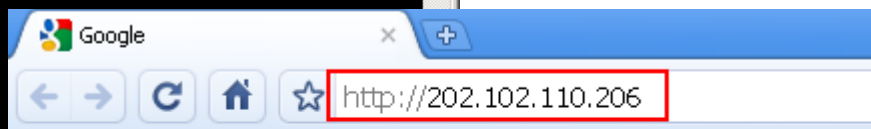
> sdfsdaf.sdafsa.wer.com
Server: UnKnown
Address: 10.204.16.18

Non-authoritative answer:
Name: nj.redirect.local
Address: 202.102.110.206

>
> sdf.wer234sdfa.wer.cn
Server: UnKnown
Address: 10.204.16.18

Non-authoritative answer:
Name: sdf.wer234sdfa.wer.cn
Address: 220.250.64.25

>
```



5. 用户使用方法

如果用户这样使用的话...

1. 添加1万个白名单/黑名单
2. 配置200条内容过滤的规则
3. 删除正在使用的配置
4. 同时执行两套程序

...

测试的期望结果:

1. 检查或者禁止
2. 能正常工作



6. 极端情况



不能控制它不会发生，那么用户的期望是什么？

车可以毁，人不能亡！

软件也会遇到极端的情况，用户的期望？

软件可以**Crash**，但是数据不能丢！

不能正常工作，但是重启后可以恢复！

...

从6个方面系统化的考虑

时间

- 哪些功能需要不间断的运行？
- 那些部分会有累积的数据产生？

大流量

- 哪些模块是承受压力的地方？
- 哪些地方累积的数据和流量有关？

并发

- 考虑所有的schedule task
- 可能相互冲突的任务、操作

使用环境

- 产品所依赖的硬软件平台，以及周边的服务
- 如果外部环境出问题的时候，产品的行为

用户使用方式

- 在实际中用户可能会使用的复杂配置

极端情况

- 可能发生的极端情况，产品的行为和期望
- 如何在lab模拟这样的情况？

Q: 如何来实施稳定性测试?

稳定性测试的挑战

- 使用怎样的测试环境？
- 不知道该测试哪些功能？
- 该如何设置产品的参数和配置？
- 如何产生大流量的请求？
- 如何长时间自动的运行？
- 不知道**Pass**还是**Fail**？
- 系统的压力有多大？
- 问题出在哪儿？

产品的配置和使用

测试流量的产生

产品和资源使用的监控

案例分析

一个电子邮件处理系统，主要功能特性如下：

1. 负责对进出公司的email进行扫描
过滤病毒，垃圾邮件，关键词，附件类型和大小， etc
2. 支持SMTP和POP3, 支持TLS加密
3. 支持白名单，黑名单
4. 运行在Linux平台
5. 有一个基于Web的管理界面，提供产品配置，日志，报表等功能。

I. 产品的配置和使用

- 硬件

- 最小配置
- 推荐配置

- OS

- RHEL 5.3?

- 周边的服务器

- DNS: 域名记录
- AD: 帐号
- Mail server: 帐号

- 产品配置

- Email 路由
- 过滤规则
- 黑白名单
- 打开log, report等功能

使用/测试

- 测试工具在使用产品
- 既是测试人员，也是用户



Stability Test
Check list

II. 测试流量的产生

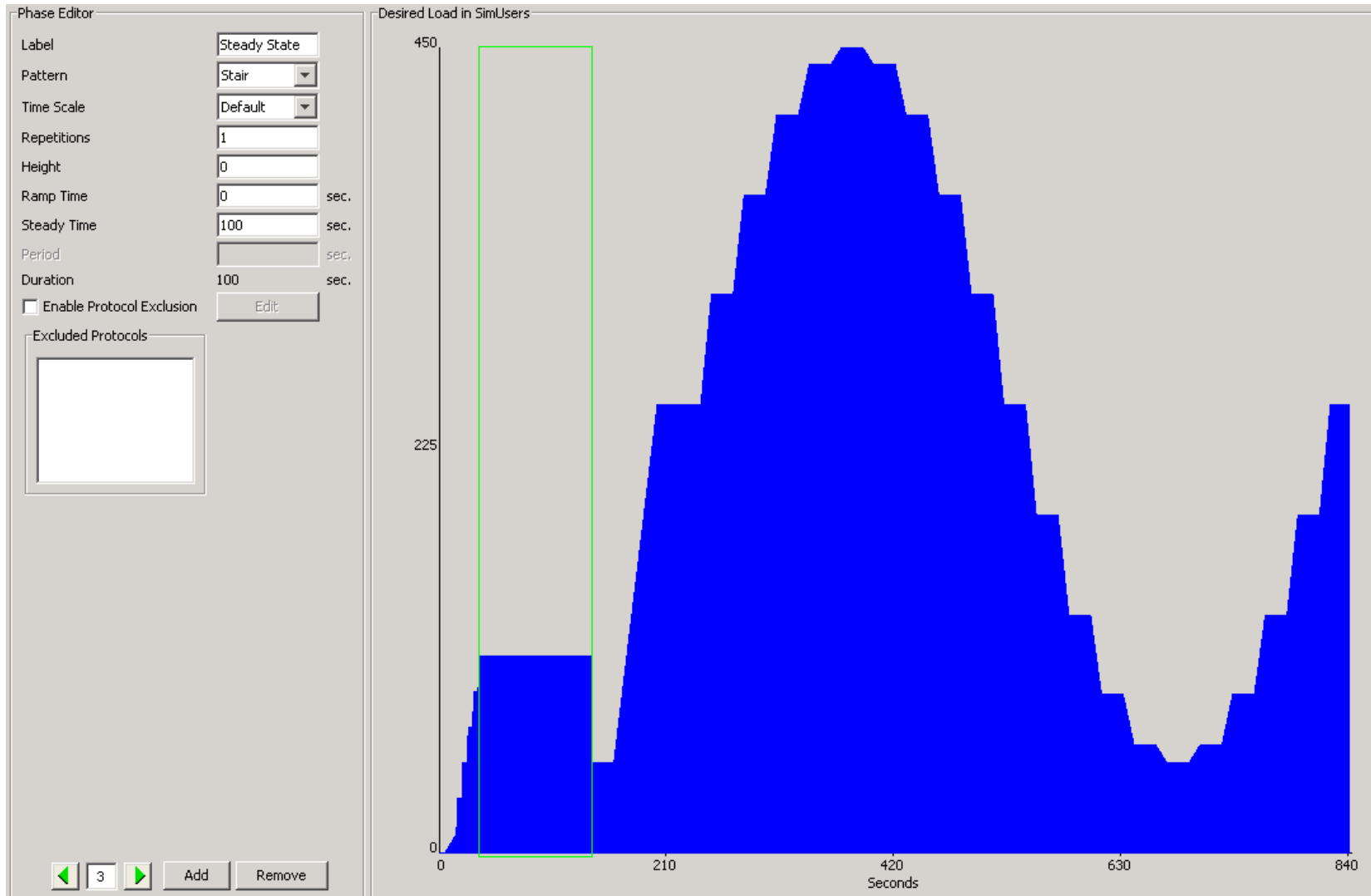
- 数据的类型

- 不同类型的样本
 - 正常文本邮件、正常带附件邮件、病毒、垃圾邮件、关键词，大附件、敏感附件类型，etc
- TLS加密流量
- POP3流量
- 发给不同的用户组
- ...
- 按比例混合

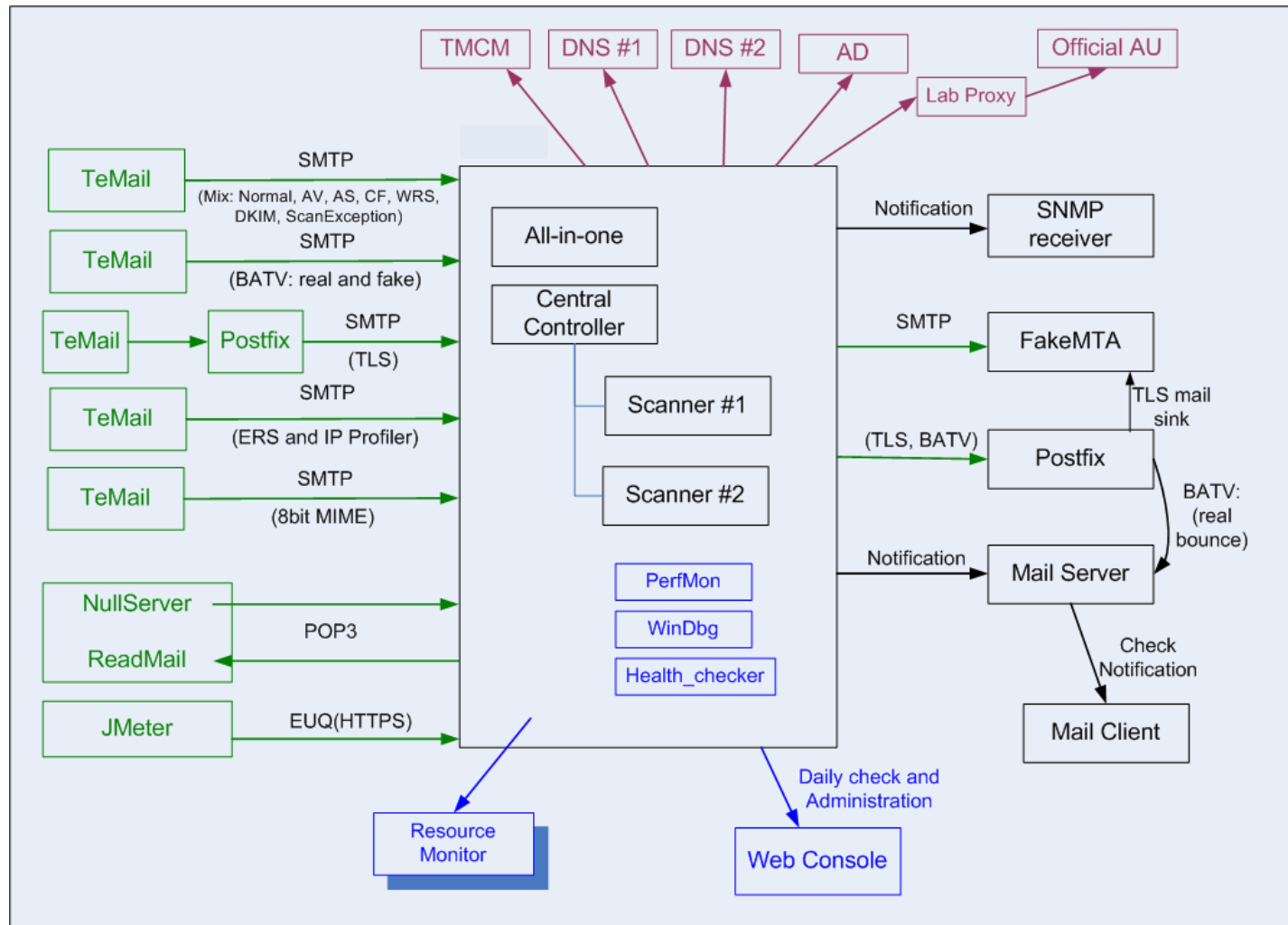
- 压力的设定

- 多大的流量比较合适？
- 流量的模型

Load curve



混合流量



How long?

讨论：稳定性测试应该执行多长时间？

24小时？1天？1周？还是一个月？ ...

一些参考：

1. 你的产品的类型和特性？

手机？PC软件？服务器？基础服

2. 按数据量来估算

处理一个月的邮件量

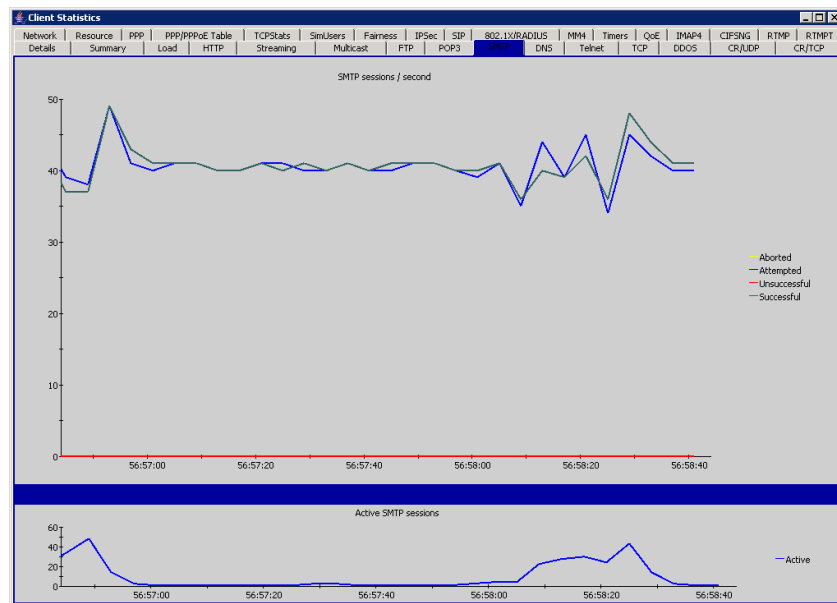
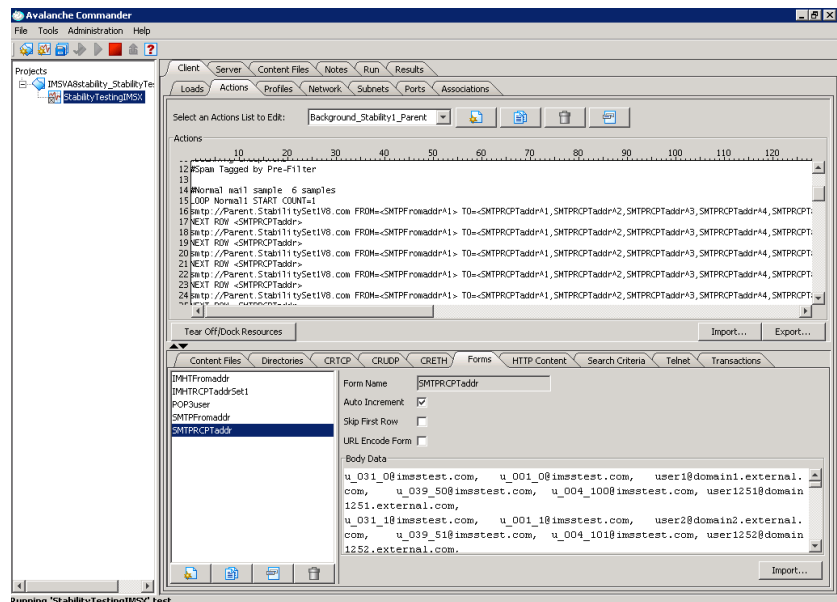
保留100G的垃圾邮件

...

3. 尽量长



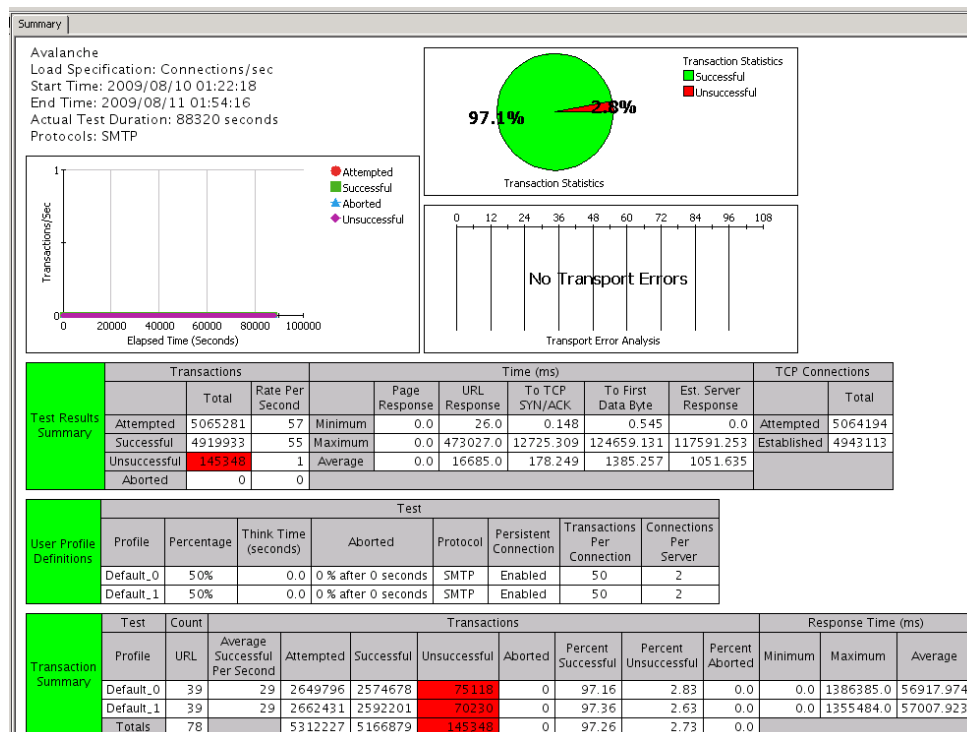
商业工具 - Avalanche



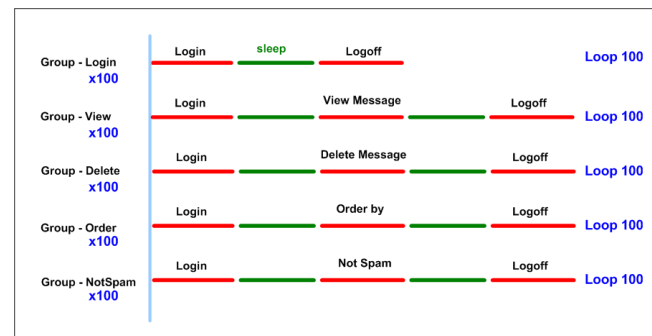
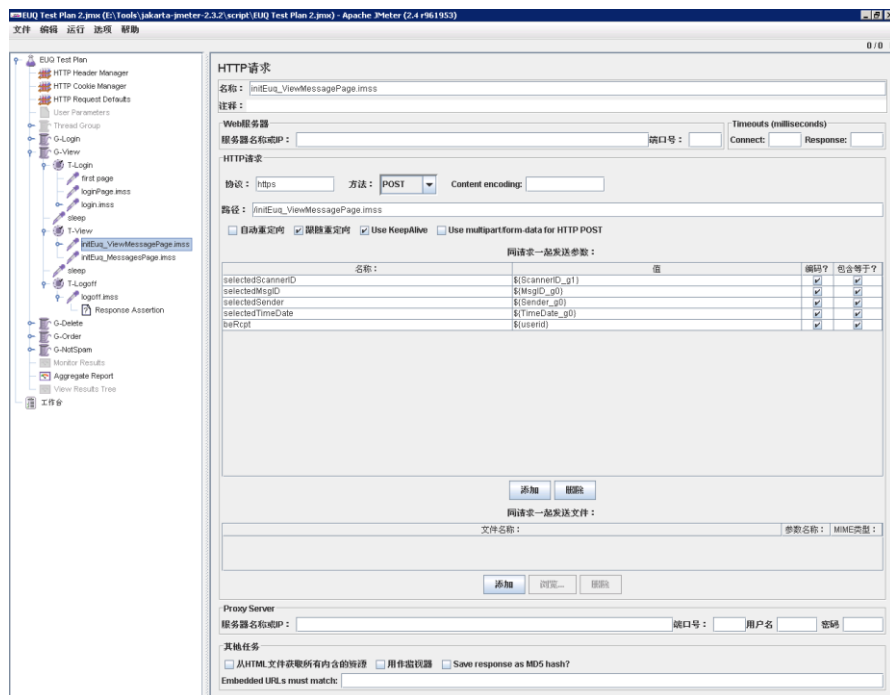
流量的配置

流量的统计

实时流量的状况



免费工具 - JMeter



聚合报告

名称:

Aggregate Report

注释:

所有数据写入一个文件

文件名

浏览...

Log/Display Only:

☐ 仅日志错误

☐ Successes

Configure

Label	# Samples	Average	Median	90% Line	Min	Max	Error %	Throughput	KB/sec
T-Login	12313	892	673	1731	10	30349	34.45%	88.3/sec	2022.5
T-Logoff	11908	359	11	426	1	54162	35.01%	86.9/sec	350.8
T-Delete	1974	1511	897	3682	2	54509	52.99%	14.4/sec	252.3
T-OrderBy	2068	1148	663	2519	2	54559	36.61%	15.1/sec	425.0
T-View	2095	1163	627	3311	2	25518	54.75%	15.3/sec	296.4
T-NonSpam	1949	1517	865	3848	2	44809	54.59%	14.2/sec	246.2
总体	32307	805	369	1786	1	54559	38.46%	231.7/sec	3564.7

常用的测试工具

- 商业的
 - LoadRunner / SilkPerformer
 - Visual Studio Team System, test edition
 - Avalanche / Smartbits / TestCenter
 - IXIA
 - WebLoad (Web)
- 免费的
 - Jmeter (HTTP(S), SMTP, POP3, DNS)
 - OpenSTA
 - WAS
- 内部开发

针对私有的客户端或者协议

Reference: 性能测试（1）：测试工具

<http://blog.csdn.net/superqa/archive/2010/08/30/5850991.aspx>

III. 监控：产品运行和资源使用的状况

挑战：

不知道Pass还是Fail？

系统的压力有多大？

问题出在哪儿？

目标：

1. 观察系统的运行情况，有无异常
2. 发现潜在的问题
3. 判断目前所加的压力是否合适
4. 发现系统运行的性能瓶颈

进程状态监控

```
[root@Prefiltertest home]#  
[root@Prefiltertest home]# ./health_checker_imsva  
health_checker v0.4.5   Jan 31, 2007
```

pid=0 means process does not exist.

```
----- original pid -----  
2010-08-18 19:28:18    IMSSAU:      is not running!  
2010-08-18 19:28:18    MsgAction:    11411  
2010-08-18 19:28:18    NTPD:      is not running!  
2010-08-18 19:28:18    HWMonitor:  is not running!  
2010-08-18 19:28:18    MailTraffic: 4562  
2010-08-18 19:28:18    Realtime:   4520  
2010-08-18 19:28:18    NAMED:      4428  
2010-08-18 19:28:18    IMSSD:      14609  
2010-08-18 19:28:18    FoxDNS:     4395  
2010-08-18 19:28:18    Monitor:    30587  
2010-08-18 19:28:18    FoxProxy:   is not running!  
2010-08-18 19:28:18    Postgre:    30623  
2010-08-18 19:28:18    PolicyServer: 29721  
2010-08-18 19:28:18    EUQUI8447:  is not running!  
2010-08-18 19:28:18    EUQUI8446:  is not running!  
2010-08-18 19:28:18    AdminUI:    17798  
2010-08-18 19:28:18    IMSSTASK:   5308  
2010-08-18 19:28:18    IMSSMANAGER: 25963  
2010-08-18 19:28:18    Postfix:    28694  
2010-08-18 19:28:18    CMAgent:    is not running!
```

初始进程状态及PID

```
2010-08-18 19:28:18    /var/coredumps: no core file!
```

```
2010-08-18 19:29:08    IMSSD, new pid:      17923  
2010-08-18 19:29:52    PolicyServer, new pid:      19370  
2010-08-18 19:30:03    IMSSAU, new pid:      19768
```

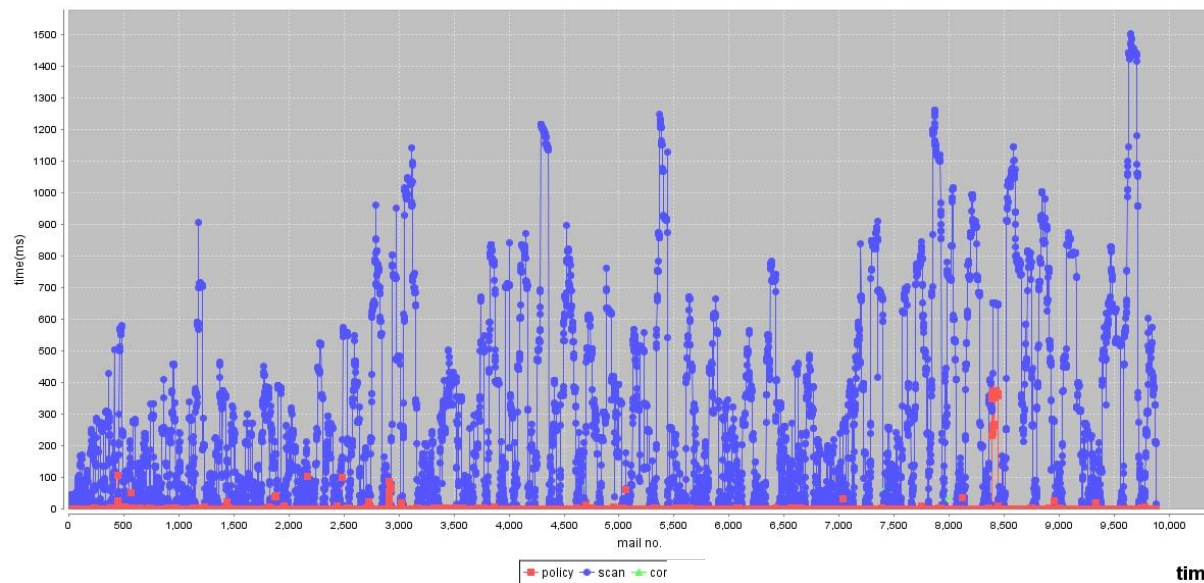
进程重启

```
2010-08-18 19:30:36    Postfix, new pid:      0
```

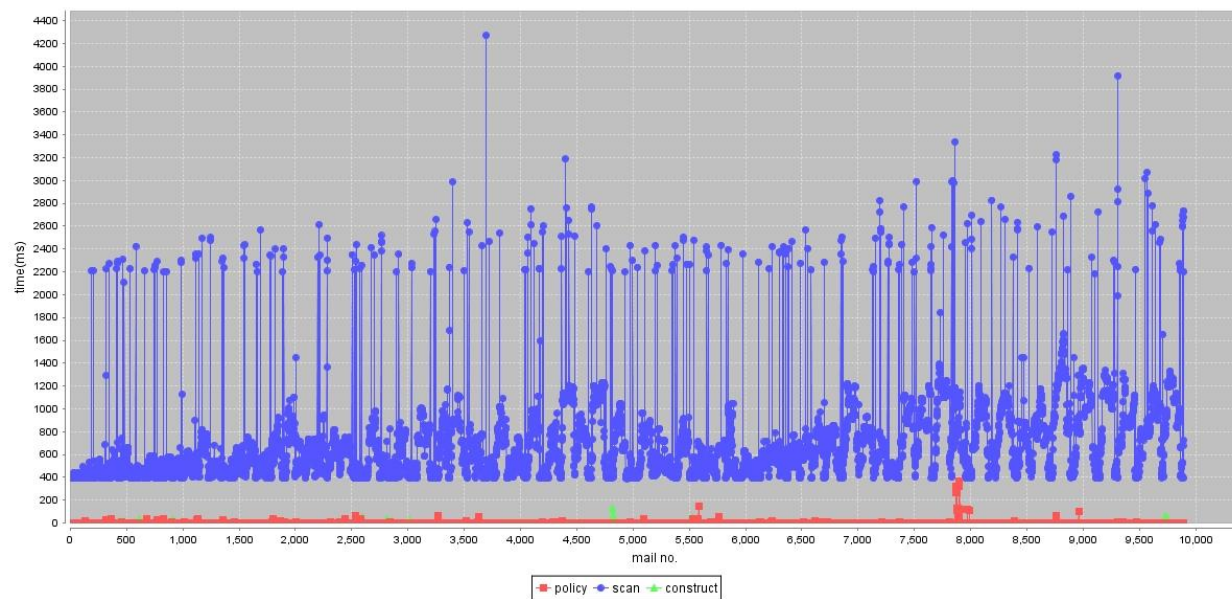
进程停止

响应时间

time usage for every mail



time usage for every mail



对产品状况的监控

- 数据流量的状况
 - 异常中断
 - 不期望的波动
- 产品指标的监控
 - 事务/任务的响应时间
 - 并发用户数
 - 队列的状况
 - 数据库/web 服务器等模块的状况
 - ...
- 核心模块的监控
 - 进（线）程个数
 - 异常重启
 - Crash dump
 - 内存泄漏
 - ...
- 外部工具
 - Perfmon
 - Top/ps
 - tcpview
- 产品自身
 - Dashboard等功能
 - Log分析
- OS
 - Linux /proc file
 - 接口 (Windows WMI)

常见的资源使用的指标

- CPU
 - Usage (user, sys, iowat)
 - CPU load
 - Context switch
- Memory
 - Total / free / cache
 - Process private
- 磁盘
 - IO/s
 - Bytes/s
- 网络
 - Bandwidth
 - TCP连接数（建立, timewait, etc）
- Others

性能瓶颈

The screenshot displays the Windows Reliability and Performance Monitor interface. The left-hand navigation pane shows a tree structure with 'System' expanded, revealing 'System Diagnostics' and 'System Performance'. Under 'System Performance', three reports are listed: '20090724-0001', '20090728-0002', and '20090728-0003'. The main pane is divided into several sections:

- Diagnostic Results:** Contains a 'Warnings' section with a single warning: 'The system is experiencing excessive paging'. The warning details indicate that available memory is low and suggest upgrading physical memory or reducing system load.
- Performance:** Includes a 'Resource Overview' table showing the status and utilization of system components.
- CPU:** A section for monitoring CPU usage.
- Process:** A section for monitoring running processes.
- Services:** A section for monitoring system services.
- System:** A section for monitoring system-level performance.
- Network:** A section for monitoring network performance.
- Disk:** A section for monitoring disk performance.
- Hot Files:** A section for monitoring files that are causing most disk I/Os.

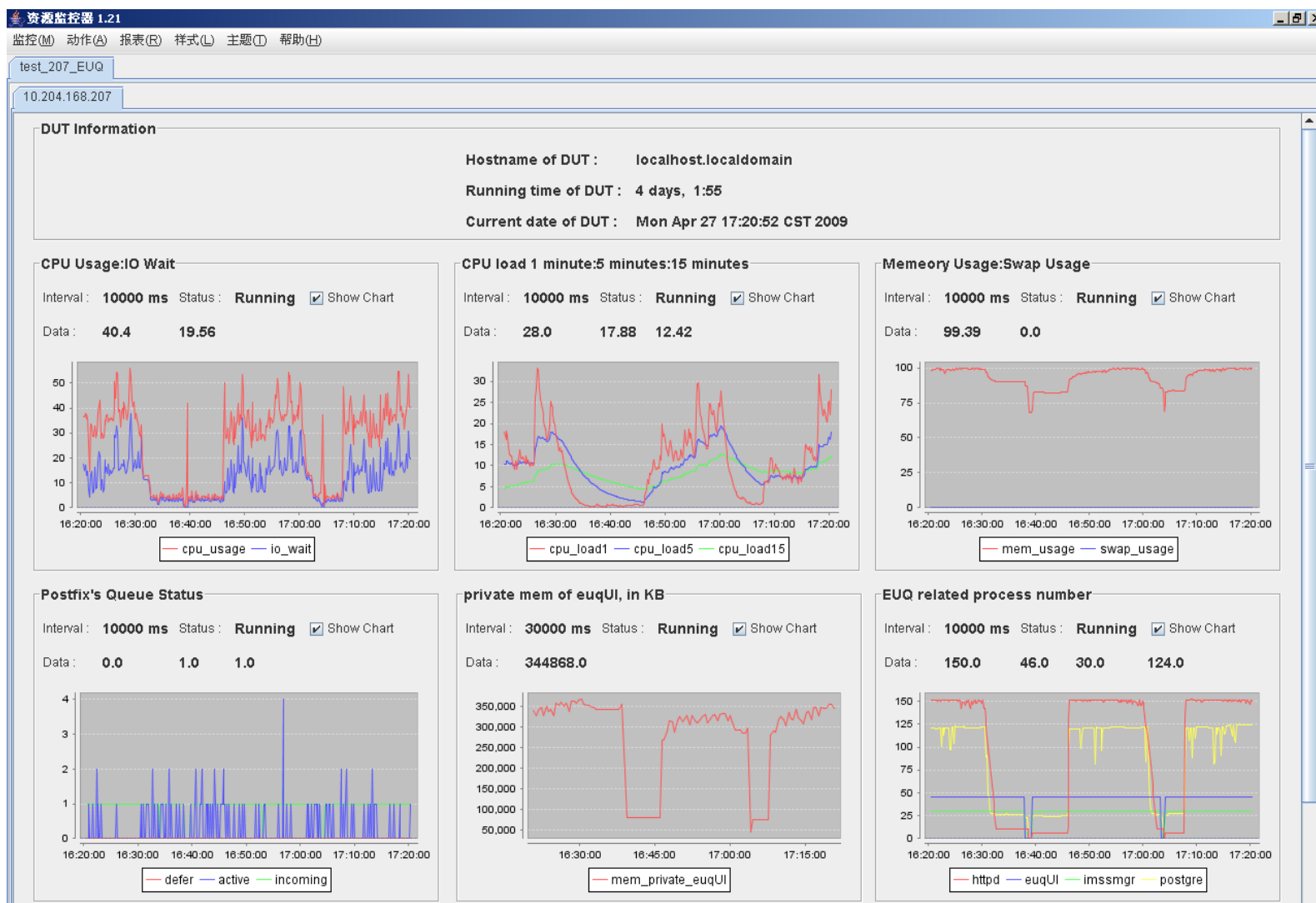
The 'Resource Overview' table is as follows:

Component	Status	Utilization	Details
CPU	Normal	23 %	Normal CPU load.
Network	Busy	62 %	Busiest network adapter exceeded 60%.
Disk	Normal	245 /sec	Disk I/O is between 100 and 500 (read/write) per second on disk 0.
Memory	Busy	96 %	304 MB Available.

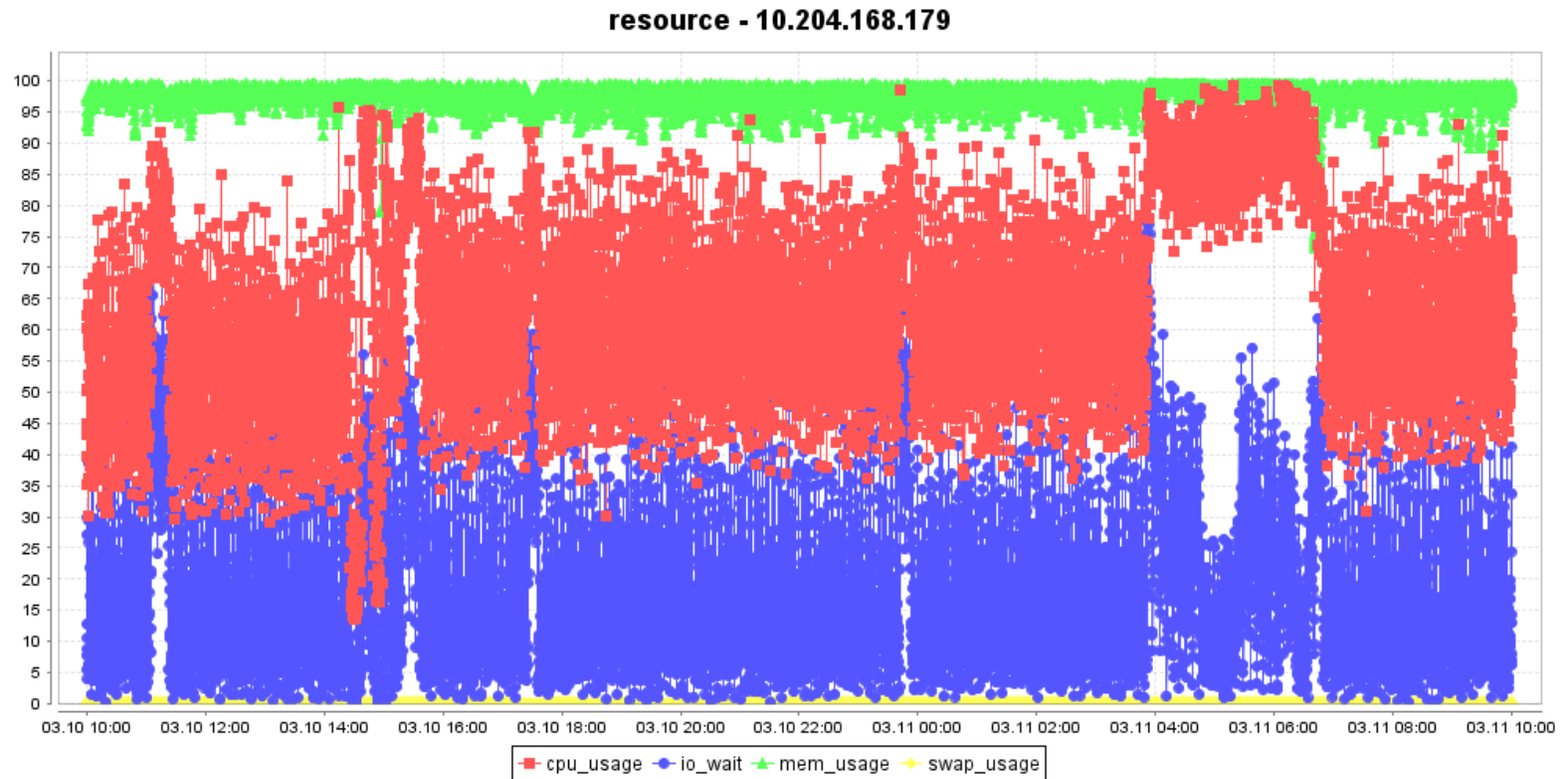
The 'Files Causing Most Disk I/Os' table is as follows:

Disk	File	Reads/sec	Kb/Read	Writes/sec	Kb/Write
0	D:\Program Files\Microsoft SQL Server\MSSQL\1\MSSQL\Data\imss_log.LDF	0.0	0	136.9	2
0	D:\Program Files\Trend Micro\SMSS\lib\patternwsapi\lpt\vpn.306	23.6	59	0.0	0
0	C:\PerfLogs\System\Performance\20090728-0003\NtKernel.etl	0.0	0	9.9	63

Resource Monitor



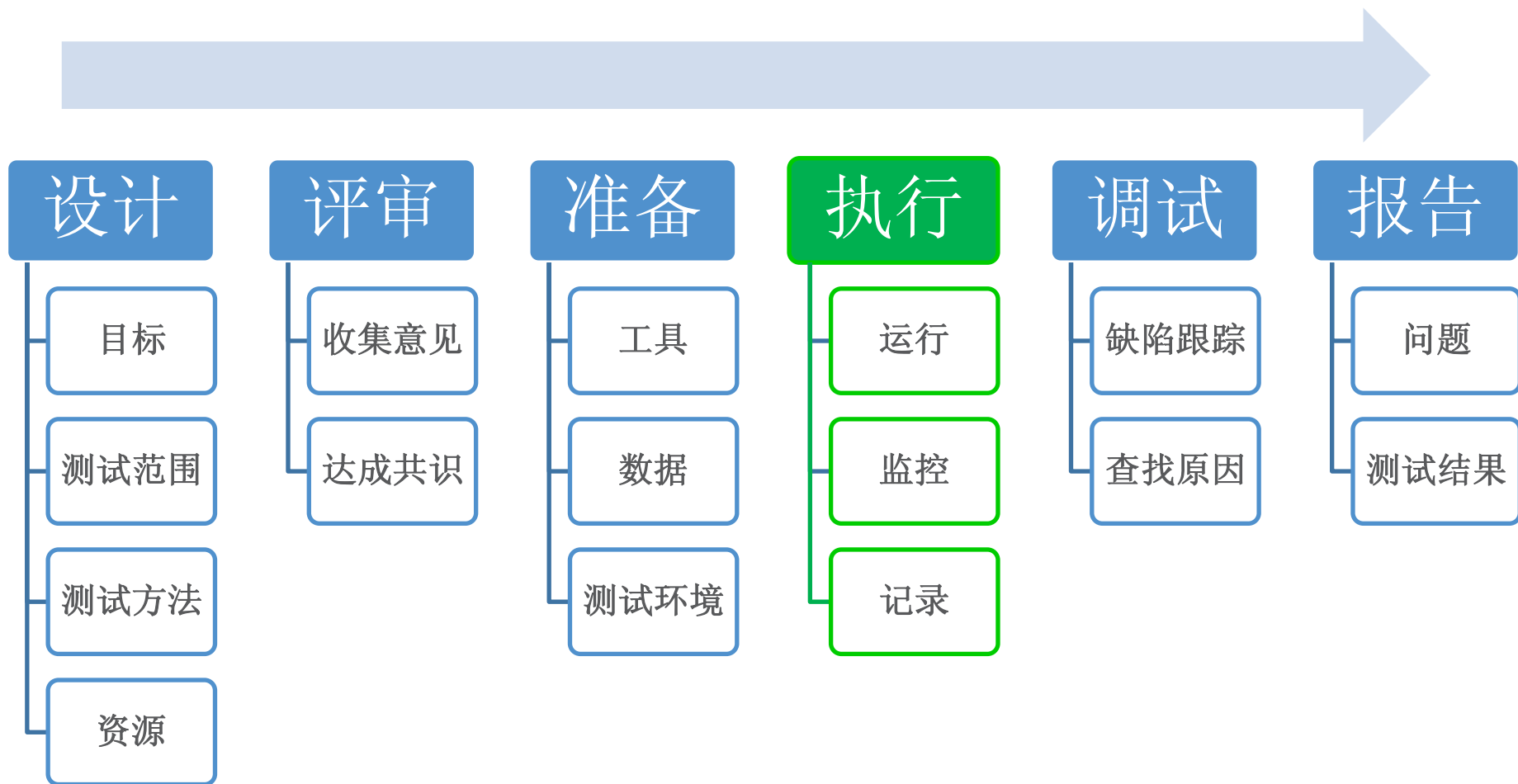
CPU usage



常见的工具

- OS自带的
 - Perfmon
 - Top/iostat/vmstat
- 自己开发的
 - script
 - Resource Monitor
- 业界的
 - Gnaglia
 - Nagios
 - SysInternals (tcpview, procexp, nmap)
 - IOMeter (disk, network)
 - 监控宝

一套完整的流程



一些心得

- 对产品深入的了解
- 相关的技术，OS, DB, network, Debugging, etc。
- 工欲善其事，必先利其器
- 团队的协作，不同功能的测试人员、不同角色之间。
- 反思、调整和持续改进

参考资料

1. 二十大软件灾难
http://atlives.blog.hexun.com/39323241_d.html
2. http://en.wikipedia.org/wiki/Software_bug
3. Performance Testing Guidance for Web Applications, by J.D. Meier, etc
4. How We Test Software at Microsoft, by Alan Page, etc.
5. <http://jakarta.apache.org/jmeter/>
6. <http://www.perftestplus.com/pubs.htm>
7. 软件性能测试过程详解与案例剖析, 段念
8. Web性能测试实战, 陈绍英等

Q&A

