**题目5：基于网络嗅探的网站用户体验数据监测、存储和分析**

|  |  |
| --- | --- |
| **试题简介** | 目前企业对自身客户服务网站的用户访问体验质量非常关心，希望能主动获知用户网站访问的体验数据，包括服务影响管理数据和服务性能管理数据两类。网络嗅探用户数据体验分析主要架构在网络交换机层面，通过交换机端口镜像，抓取网络协议数据包，分析网站Client IP，Server IP，URL，Cookie/Session ID，以及响应时间，数据大小，用户访问的错误，以及网络失败，或者网络超时等等这些数据，然后进行分析，将分析结果以商业智能图表的形式实时提供给用户。 http://www.chinasoft.org.cn/rjds/UploadFiles/20131028174126888.jpg |
| **业务场景** | 针对用户网站体验，能实现基于5个纬度的网络包数据采集、存储及分析： 　　Who(谁) 　　　·完整的会话sessions？ 　　　·IP地址？用户名？所属群组？或匿名？ 　　　·哪些是最不满意的用户？ 　　　·哪些是最勤奋的用户？ 　　　·哪些人关闭窗口？ 　　　·哪些人违规(IP、时间)？ 　　　·满意度问题通知谁？ 　　What(什么) 　　　·访问什么业务？调用什么服务或页面？ 　　　·什么服务访问最多？ 　　　·什么服务/页面最慢？资源开销最大？ 　　　·系统发生了什么错误？因为什么触发错误？错误提示给用户什么？ 　　　·用户使用的设备/版本？ 　　　·人工与系统时间开销多少？ 　　How(如何) 　　　·如何快速准确定位错误？ 　　　·如何知道哪些用户受到不良影响？ 　　　·如何提升访客转换率？ 　　　·如何改善优化业务效率？ 　　　·如何优化应用减少bug？ 　　Where(哪里) 　　　·用户主要来自哪里？ 　　　·哪里导致满意度问题？ 　　　·用户在哪里离开业务？ 　　　·哪里出错(网络、服务器)？ 　　　·哪些地方流量大？ 　　　·哪些地方用户多？ 　　When(何时) 　　　·不同时间尺度的访问曲线？ 　　　·什么时间出的错误？ 　　　·什么时间做的动作？ 　　　·历史完整回溯？ |
| **功能性需求** | 1、可以实时地集中监控任何B/S架构(J2EE、.Net、WAP)的业务系统，实现跟踪所有用户，在所有时间里的、所有应用操作的、所有体验障碍（如系统错误、性能差等）解析导致不良用户体验的主要系统资源或应用服务瓶颈的根源；可回朔任何时间点的任何用户的任何操作，甚至是用户屏幕错误重放。 　　2、能评估用户体验的满意度和趋势度 　　3、改进以交易漏斗分析业务流程的效率（如访客转换率等） 　　4、监听指标： 　　　(1)旁路监听用户发出请求 　　　　·Client-IP, Server-IP, Gateway-IP 　　　　·Requested URL (Page,Object) 　　　　·Cookie/Session-id 　　　　·Referrer 　　　　·GET & POST requests 　　　　·User agents (phone, browser type)  　　　(2)系统响应与服务质量 　　　　·Response? 　　　　·Server error, web site error, Hit OK 　　　　·Content of page (i.e order details, functional errors) 　　　(3)是否送达用户 　　　　·Delivered? 　　　　·Aborted? 　　　　·Network timeout |
| **非功能性 需求** | 在实际场景中要考虑实时数据存储和分析的效率问题，选用适合场景的数据库进行数据存储。要考虑到抓取数据后的存放问题，如存放多久，存储空间多大，如何清空过期的历史数据等。 |
| **其他限制 条件** | 代码可采用JAVA方式实现，B/S结构。无嵌入代码，不重启服务器，无需修改应用；被动的网络侦听，不影响应用性能。数据收集后以关系或NoSQL数据库形式存储，数据库类型不限制。商业智能图表展现可以利用API组件自己开发或利用商业产品成品。嗅探测试网站可使用已运行网站或自己搭建。网站访问行为可自行模拟。 　　验证方式： 　　1、将嗅探程序置于实际运行的网站服务器中进行部署，实时性监控，100%无遗漏的客户体验记录； 　　2、要求采用旁路侦听或端口拷贝方式，不干扰网站运行、不影响性能、不使用探针（Adaptor）或插件、不修改网站应用（与开发框架、数据库、中间件及底层平台无关）、不需要重起应用即能监测。 　　判断依据（数据表或图表展现）： 　　1、可跟踪系统所有用户在测试期内的所有应用操作，以及所有客户体验障碍（如系统错误等） 　　2、可解析导致不良用户体验的主要系统资源或应用服务瓶颈的根源 　　3、可回朔任何时间点的任何用户的任何操作，甚至是用户屏幕错误重放 　　4、可评估客户体验的满意度和趋势 |
| **测试数据 或平台** | 自选 |
| **其他要求** | 无 |
| **作品交付** | 1. 源代码和可执行程序 　　2. 设计文档，包括算法中的数据结构、索引机制、查询和写入设计、测试结果 　　3. 测试网站说明,可复制，可展示 |

**元信息**

**(1)旁路监听用户发出请求**

**•Client-IP, Server-IP, Gateway-IP**

**•Requested URL (Page,Object)**

**•Cookie/Session-id**

**•Referrer**

**•GET & POST requests**

**•User agents (phone, browser type)**

**(2)系统响应与服务质量**

**•Response?**

**•Server error, web site error, Hit OK**

**•Content of page (i.e order details, functional errors)**

**(3)是否送达用户**

**•Delivered?**

**•Aborted?**

**•Network timeout**

删除cookie影响

**网站概况**

**流量分析-趋势分析**

**流量分析-对比分析**

**流量分析-当前在线**

**流量分析-访问明细**

**来源分析-来源分类**

**来源分析-搜索引擎**

**来源分析-搜索词**

**来源分析-最近搜索**

**来源分析-来路域名**

**来源分析-来路页面**

**来源分析-升降榜**

**受访分析-受访域名**

**受访分析-受访页面**

**受访分析-升降榜**

**访客分析-地区分布**

**访客分析-终端详情**

**访客分析-新老访客**

**访客分析-忠诚度**

**访客分析-活跃度**