



基于用户体验的信息系统主动运维平台

吴 舜¹, 许大卫¹, 魏 征², 胡成臣², 向万红³, 郭 前⁴, 唐亚哲²

(1. 国网冀北电力有限公司信息通信分公司 北京 100053; 2. 西安交通大学 西安 710049;
3. 远光软件股份有限公司 珠海 519085; 4. 南京云帕信息技术有限公司 南京 211100)

摘 要:传统的运维监控只是监控系统各个组件的健康工作情况,对用户级体验的细粒度监控尚缺乏有效手段。针对此问题,研究基于用户体验的信息系统主动运维平台,从业务的网络流量大数据分析出发,提出监测用户网络行为及其体验耗时。系统实施的关键在于实时识别用户的行为,测量每个行为的具体响应时间,监测用户级应用的性能。测试表明,该平台实现了预期的功能,系统吞吐率较高,内存占用较少,实现了用户应用级性能实时监测,为提高用户体验提供了有效手段。

关键词:系统监控;性能;用户体验;主动运维

doi: 10.11959/j.issn.1000-0801.2015059

Active Operation and Maintenance Platform for Information System Based on User Experience

Wu Shun¹, Xu Dawei¹, Wei Zheng², Hu Chengchen², Xiang Wanhong³, Guo Qian⁴, Tang Yazhe²

(1.State Grid Jibei Information & Telecommunication Company, Beijing 100053, China; 2. Xi'an Jiaotong University, Xi'an 710049, China; 3. Ygsoft Inc., Zhuhai 519085, China; 4. Nanjing Yinspect Co., Ltd., Nanjing 211100, China)

Abstract: The traditional monitoring maintenance system only watches the running metrics. However, these metrics is not sufficient to reflect the status of the whole system for fine-grained monitoring of the system at users and application level. An active operation and maintenance platform based on user experiences was proposed. Through the mining of big network data, the platform can recognize a variety of user behaviors, and measure the user experience time along with the system performance at user's level. Tests and experiments show that the platform works well with large throughput and low memory occupation. The application level performance monitoring, which provides an efficient way to improve user experience, was realized.

Key words: system monitoring, performance, user experience, active operation and maintenance

1 引言

在电网企业信息系统运维实践中,即使配置、部署了各种各样的监控系统,也常有业务系统出现问题却怎么也

找不到故障点的情况。究其原因,主要是当前大部分监控系统是分离地监控业务系统组成部分的工作状态(如网络、服务器、数据库等),并且这些监控的级别很低,没有区分用户和业务,也没有从业务系统本身的工作状态寻找原

收稿日期:2014-09-15;修回日期:2015-02-10

论文引用格式:吴舜,许大卫,魏征等. 基于用户体验的信息系统主动运维平台. 电信科学, 2015059

Wu S, Xu D W, Wei Z, *et al.* Active Operation and Maintenance Platform for Information System Based on User Experience. Telecommunications Science, 2015059



因。尤其在网络环境下,仅知道某个交换机端口的吞吐率或者分组丢失率,很难发现某个业务运行不正常的原因。本文的工作目标是解决上述问题,通过监控业务系统尤其是网络中运行的业务系统的通信行为,实现基于用户体验的主动运维。

通常,一个业务系统具有针对某个具体业务的多种功能,如财务系统通常会有各种报账、查询等。作为终端用户的财务统计人员,直接操作在业务系统上,便发生了对应于各种功能的操作,定义这些操作为相应的此业务系统上的用户行为。

在业务系统的实际运行中,行为的响应速度直接反映了用户业务运行是否正常,也影响着终端用户的用户体验,过长的响应时间(出现故障时甚至导致响应时间无限长)将直接导致用户使用系统时的不适。例如,用户点击“已完成单据查询”,在多长时间之后业务系统做出响应并展示出相应的结果,直接影响着业务系统的用户体验满意度。

作为业务系统的运维人员,需要在用户体验下降时及时发现终端用户在使用业务系统时存在的问题,这就要求运维人员能通过某种方式实时了解用户对业务系统的使用情况,并且对用户满意度进行评估,出现问题时及时解决。

本文提出的主动运维平台通过对业务系统的实时监控,以用户发生各种行为时产生的流量为输入,借助细粒度网络流量分析方法识别出每个用户行为,并且给出用户发生此行为的响应时间,也就是用户行为性能,帮助运维人员及时发现系统存在的问题,提供优化业务系统的依据,提高用户体验。

2 系统设计

针对引言中提出的功能需求,本平台采用自顶向下的设计思路,要得到用户发生某个行为的响应时间,就需要知道在用户发生此行为对应的流量中,最开始和最末尾部分流量的时间点,二者之间的差值就是用户在此行为上的体验时间。而在通常情况下,系统获取到的网络流量是各个用户各个行为产生流量的混合体,所以在上述层次之下便是如何识别出特定用户行为的流量,故系统最终设计为一种上下双层结构。

2.1 行为性能测量

本文提出的主动运维平台以监听的方式实施,通过监听用户流量,识别分析用户行为,度量行为的时间,最终实现对业务系统的实时监控。

运维平台在网络中部署的典型拓扑如图 1 所示,平台串接在网络之中,以用户发生各种行为时的流量和内置规则库作为输入,如图 2 所示,最终得到的输出便为识别出的用户行为以及每个行为的用户体验时间。

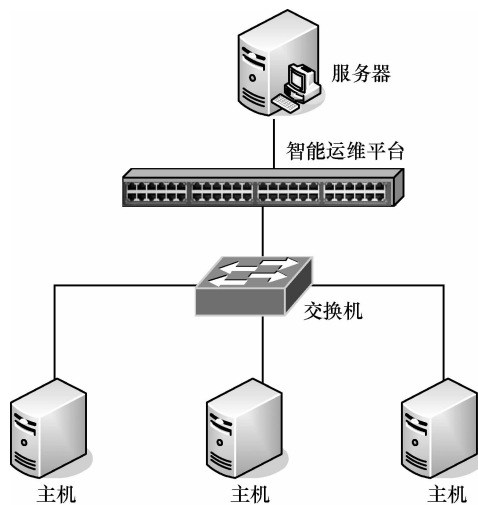


图 1 智能运维平台在网络中的位置

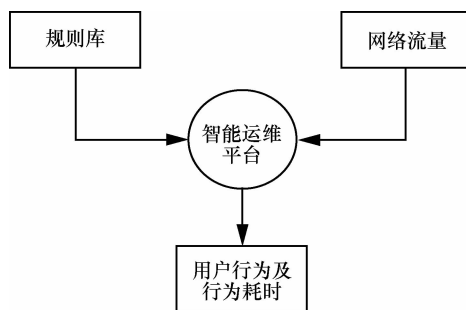


图 2 平台输入输出示意图

本平台测量的目标系统是电网企业信息系统典型的 B/S 架构的客户端,特点是使用 HTTP 请求和响应与服务器进行交互。需要注意的是,在本平台中最后得到的是某个行为的用户体验时间,而不是行为发生时某个 HTTP 请求和其响应之间的耗时,故在本平台对于用户发生某个行为时产生的一大批 HTTP 请求和响应中,取最开始和最后的请求响应对,并且计算第一个请求响应对的请求时间点到最后一个请求响应对的响应时间点的时间差,从而判断用户从最开始发生此行为一直到行为完成的耗时,也就是用户体验时间。例如,当用户在某财务系统上点击“已完成单据查询”链接时,会产生一系列 HTTP 请求与响应,在本系统的研究中,只取具有先后顺序的 HTTP 请求和响应,如图 3 所示。

为描述方便,本文使用字母表示发生的 HTTP 请求响应对,图 3 中最开始的圆 A 代表第一个请求响应对,最后

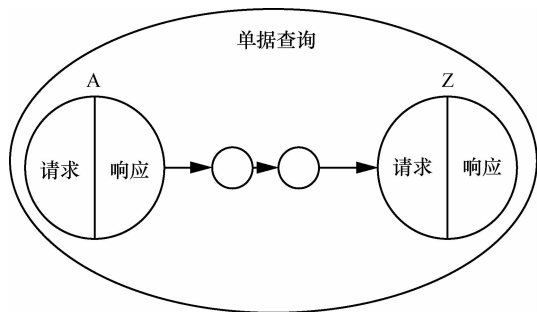


图 3 一个行为产生的 HTTP 请求和响应

的圆 Z 代表最后一个请求响应对,本运维平台首先识别出圆 A,记录下圆 A 的请求和响应时间戳,之后只需要识别出期望的圆 Z,并且记录下圆 Z 的请求时间戳和响应时间戳,用圆 Z 的响应时间戳减去圆 A 的请求时间戳就是此行为的用户体验时间,如式(1)所示。

$$behavior_elapsed_time = Z.response_timestamp - A.request_timestamp \quad (1)$$

在图 3 中,本平台对于“已完成单据查询”行为,只需要依次捕捉请求响应对 A 和 Z 的到来,对于其他如请求响应对 B、C,放过即可。通过可配置的规则来决定对于某个行为,平台需要捕获到哪两个请求响应对。换句话说,规定平台需要关联哪两个请求响应对才能识别一个行为,并且得到这个行为的用户体验时间,将这种可配置的规则称为关联分析规则,规则样例见表 1。

表 1 关联分析规则样例	
关联规则	行为名称
A,Z	已完成单据查询
X,Y	凭证查询

综上所述,本平台以请求响应对 A、B、C 等作为待关联数据输入,以可配置的关联分析规则作为规则输入,最终得到用户发生的行为以及用户体验时长。也就是说,请求响应对 A、B、C 是底层提供给上层的服务数据,同 TCP/IP 网络的分层原理一样,将这一层称为关联分析层。显然,像请求响应对 A、B、C 这样的请求响应对,并不是直接可以从网络中不做任何处理就可以获取到的,在第 3 节中将叙述如何获得请求响应对 A、B、C。

2.2 请求响应对识别

像 A、B、C 这样的请求响应对,在本平台中一般代表某个特定行为的开始或结束,因此是具有特定语义的,称本层为语义抓取层。语义抓取层为关联分析层提供具有语义特性的请求响应对,供上层进行关联分析,也就是为上

层提供服务数据。

如前所述,用户在点击“已完成单据查询”时,会产生一系列 HTTP 请求和响应,该层应该如何区分不同的 HTTP 请求和响应是一个问题。在本层使用语义抓取规则,规定满足特定条件的 HTTP 请求和响应才是特定的请求响应对 A 或者其他。例如表 2 中的语义抓取规则,其中第一条规则规定:HTTP 请求中 URI 的字段值与正则表达式“Ywdjywc.*”匹配,HTTP 请求中 referer 的字段值与正则表达式“index.jsp”匹配,同时服务器对这个 HTTP 请求的响应数据与正则表达式“ywc_data”匹配,如果这 3 个正则表达式都可以匹配到,则认为捕获到了 A 请求响应对,之后将其输出到一张按用户 IP 地址分开的表中,见表 3。表 3 作为语义抓取层的输出、关联分析层的输入,由语义分析层提供给关联分析层,之后由关联分析层进行关联分析。

表 2 语义抓取规则样例

语义抓取规则	请求响应对标识码
URI: Ywdjywc.* refer: index.jsp response_data:ywc_data	A
URI:\friends callname:setlogin response_data:friends_name	X

表 3 语义抓取层输出样例

用户 IP 地址	请求响应对标识码列表
192.168.0.1	A,Z,...
192.168.0.2	X,A,Y,Z,...

综上所述,语义抓取层的输入为用户操作业务系统时产生的一系列 HTTP 请求和响应,也就是网络中的原始流量以及一张内置的语义抓取规则表,输出为提供给上层的关联分析数据,也就是请求响应对。

因此,整个平台最终分为两层:关联分析层和语义抓取层。两层的输入、输出及相互间的关系可用图 4 表示。自底向上,首先语义抓取层将用户产生的流量与语义抓取层规则库进行匹配,得到的输出为每个用户产生的不同请求响应对;而语义抓取层的输出作为关联分析层的输入,关联分析层将下层提供的输入数据与关联分析规则库进行关联分析,最终得到识别出的用户行为及每个行为的用户体验满意度。

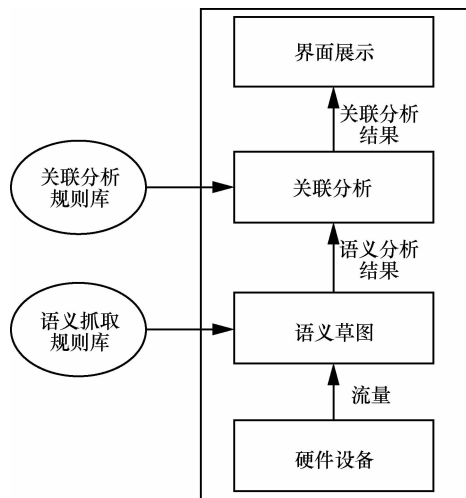


图4 系统分层架构设计

3 关键技术实现

如第2节所述,要得到用户发生某个行为的响应时间,就需要知道在用户发生此行为对应的流量中,最开始和最末尾部分流量片段的时间点,二者之间的差值就是用户在此行为上的体验时间。在系统的实现过程中,同一个用户在一段时间内会发生多个行为,而每个行为由行为开始和行为结束两部分组成,因此,多个行为之间的开始和结束在时间上会发生覆盖,记行为1的开始时间为 t_1 ,结束时间为 t_2 ,行为2的开始时间为 t_3 ,结束时间为 t_4 ,上述问题可以描述为 $t_1 < t_3 < t_2 < t_4$,即行为时间跨度发生交叉。

针对上述情况,本平台在实现中采用了一种保存用户行为上下文的办法,即用户行为候选表。这张表中保存了用户一段时间内可能会发生的所有行为,例如,当系统在 t_1 时刻检测到行为1的开始时,便在用户行为候选表中加入行为1,并且记录时间点 t_1 。同样,随后系统会在用户行为候选表中加入行为2,并记录时间点 t_3 。对于后面 t_2 时刻到来的行为1结束点,只需要查找用户行为候选表便可以命中行为1,并且得到行为1的用户体验时间。同理,可继续命中行为2并且得到其用户体验时间。系统的关键实现在于用户行为体验时间测量算法,具体计算过程如下。

输入:用户请求响应对标识码表(表3)、关联分析规则表(表1)。

输出:用户行为体验耗时。

1:从用户请求响应对标识码表中取出一个IP地址和其对应的请求相应对标识码,如(192.168.0.1,A);

2:使用3中的标识码在用户行为候选表中匹配用户

的候选行为;

3:if命中某个候选行为

4:then输出得到的用户行为并且计算此行为的体验耗时;

5:删除此候选行为;

6:else

7:使用3中标识码查找关联分析规则表;

8:if查找到可能的用户行为

9:then将查找到的用户行为加入用户行为候选表;

10:记录行为的开始时间等信息;

11:else

12:丢弃此次的请求相应对标识码;

13:end if

14:end if

15:返回至步骤1

4 实验结果

4.1 功能测试

本平台按照如图1所示的拓扑在某电网企业财务系统进行了测试,根据最终得到的输出统计出用户体验时间最长的10种行为,并且给出了每种行为的用户体验时间,如图5所示。

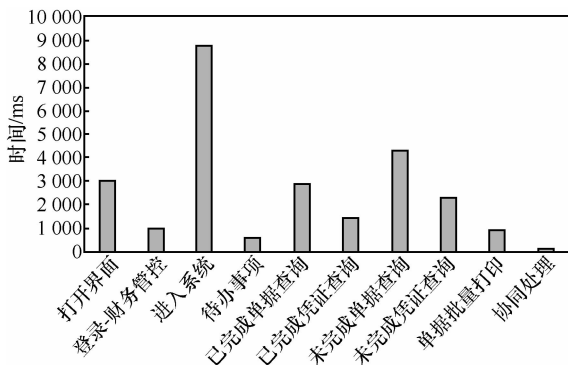


图5 某财务系统行为性能统计

从图5可以看出,在这个财务系统中最慢的行为是“进入系统”行为,这说明这个行为的用户体验较差,用户登录后进入系统可能从服务器加载了大量数据,从而导致服务器响应过慢;或者网络负载过重,使得数据传输过慢;也可能是使用了不合理的SQL查询、服务器负载过重等。总之,运维人员可以通过这种方式实时监控系统中用户在业务系统上各个行为的体验,对有问题行为进行专项优化。

4.2 吞吐率测试

对吞吐率进行测试,在平台上使用离线流量进行测试。具体来说,针对不同大小、不同比例的混合流量进行处理,根据最终处理完整个离线流量的时间来计算平台的吞吐率。由于本平台的目标流量是某财务管控系统流量,故将纯目标流量与其他无关流量进行不同比例的混合,最终得到 5 种离线流量,见表 4。测试结果即在不同比例流量下的不同吞吐率。

表 4 不同比例的混合流量

名称	大小/byte	目标流量所占比例	平均分组长长度/byte
流量 1	8 649 566	90.62%	674.8
流量 2	1 1036 926	71.02%	694.8
流量 3	15 046 514	52.10%	685.2
流量 4	24 885 785	31.50%	725.9
流量 5	73 617 106	10.65%	760.0

图 6 为在 5 种不同混合流量下测试得到的系统吞吐率。可以看出,从流量 1 到流量 5 吞吐率逐渐变高,这是因为无关流量的增加并不会增加本系统的负荷,本系统只对流量中的目标流量进行处理,从而表现出系统吞吐率的升高。一般情况下,网络中的流量不会总是无关流量或者目标流量,所以采取混合流量的方法进行测试是十分合理的。

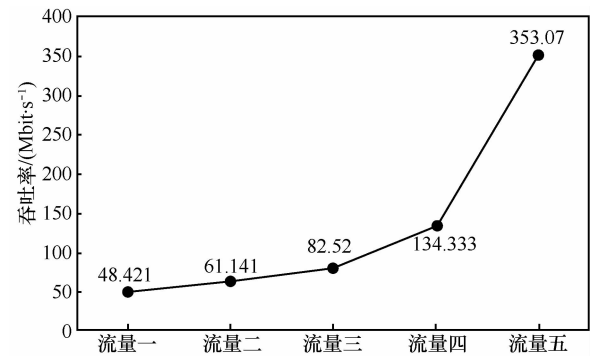


图 6 不同混合流量下测试得到的系统吞吐率

4.3 内存测试

对设备进行长时间的模拟在线环境测试,将目标流量从另外一台设备打入本平台,监控系统的内存使用情况,测试系统的稳定性,图 7 是系统运行 72 h 内存的变化情况。图 7 中有两条线是因为本系统运行时需要两个例程。可以看出,系统的内存开销很低,并且运行稳定。

5 结束语

本文提出了一种识别用户行为并且测量用户行为体

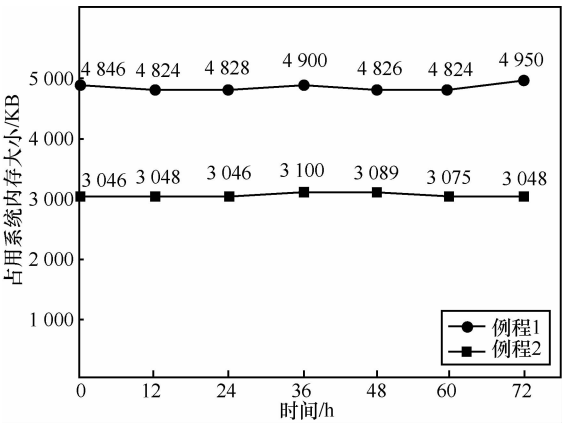


图 7 系统运行 72 h 内存占用情况

验的方法,同时将其平台化,帮助运维人员定位业务系统中存在的缺陷。使得运维人员在用户体验下降时,及时发现终端用户在使用业务系统时存在的问题,运维人员通过这种方式实时监控系统中用户在业务系统上各个行为的体验,对有问题的行为进行专项优化。

本文的创新主要有以下几点。

(1) 本平台相比于传统的服务器端监控,创新性地实现了与之相对应的用户端体验监控,多维度的监控提供了更有力的运维信息。

(2) 本平台将用户体验监控与系统组件监控有机结合起来,更细粒度地分析出发生故障时是在哪个组件上、因为哪种用户行为导致组件运行出现问题。

(3) 本平台不但进行用户行为识别,并且测量每个用户行为的用户体验时间,为运维人员提供强有力的故障追踪及优化依据。

(4) 本平台以旁路的方式监听网络流量,无需在用户一端安装任何软件等,对用户使用业务系统完全无干扰。

(5) 本平台的设计采用典型的分层设计,极大地减少了系统各模块之间的耦合性,从而使得开发与开发、开发与测试、测试与测试之间都可以互不干扰地进行,极大推动了项目进度。

本文的关键技术在于通过监听用户流量、识别分析用户行为、测量行为的体验时间,最终实现对业务系统的实时监控。本平台以用户发生各种行为时的流量和内置规则库作为输入,计算第一个请求响应对的请求时间点到最后—个请求响应对的响应时间点的时间差,从而计算用户体验时间,最终得到的输出便为识别出的用户行为以及每个行为的用户体验时间。

本平台下一步还需要对除了 HTTP 之外的其他协议进



行功能方面的扩展,以满足不同的应用场景。在性能方面,本文已经探究出更快速的语义抓取算法,目前正在整合中,之后可以支持更多用户、更复杂的系统监控。

参考文献

- 1 Chen Y, Mahajan R, Sridharan B, *et al.* A provider-side view of web search response time. Proceedings of the ACM SIGCOMM, Hong Kong, China, 2013: 243~254
- 2 Cherkasova L, Gardner R. Measuring CPU overhead for I/O processing in the Xen virtual machine monitor. Proceeding of USENIX Annual Technical Conference, Boston, MA, USA, 2005: 387~390
- 3 Hassidim A, Raz D, Segalov M, *et al.* Network utilization: the flow view. Proceedings of IEEE INFOCOM, Turin, Italy, 2013: 1429~1437
- 4 Meyer T, Wohlfart F, Raumer D, *et al.* Measurement and simulation of high-performance packet processing in software routers. Proceedings of Leistungs-, Zuverlässigkeits- und Verlässlichkeitsbewertung von Kommunikationsnetzen und Verteilten Systemen, GI/ITG-Workshop MMBnet, 2013
- 5 Balachandran A, Voelker G M, Bahl P, *et al.* Characterizing user behavior and network performance in a public wireless LAN. ACM SIGMETRICS Performance Evaluation Review, 2002, 30(1): 195~205
- 6 Li H, Hu C C. ROOM: rule organized optimal matching for fine-grained traffic identification. Proceedings of IEEE INFOCOM, Turin, Italy, 2013
- 7 吴舜, 苏丹, 吴佳等. 基于 Tiler 平台的网络细粒度应用行为识别. 电信科学, 2013, 29(11): 94~98
Wu S, Su D, Wu J, *et al.* Fine-grained network traffic identification based on tilera platform. Telecommunications Science, 2013, 29(11): 94~98
- 8 朱凯, 张超, 张凯等. 基于 TCP 数据包层分析的移动互联网用户体验的评估方法. 北京邮电大学学报, 2014, 37(s2): 40~45
Zhu K, Zhang C, Zhang K, *et al.* Assessment of user's QoE for mobile internet based on TCP packet layer analysis. Journal of Beijing University of Posts and Telecommunications, 2014, 37(s2): 40~45
- 9 Viswanath B, Bashir M A, Crovella M, *et al.* Towards detecting anomalous user behavior in online social networks. Proceedings of the 23rd USENIX Security Symposium, San Diego, CA, USA, 2014
- 10 Barford P, Plonka D. Characteristics of network traffic flow anomalies. Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement, San Francisco, USA, 2001: 69~73

[作者简介]



吴舜,男,国网冀北电力有限公司信息通信分公司工程师、高级技师,主要研究方向为电力企业信息系统运检技术、用户体验监测技术、软件定义网络技术。



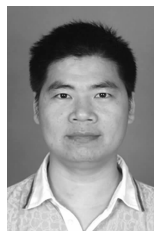
许大卫,男,国网冀北电力有限公司信息通信分公司助理工程师、副主任,主要从事信息系统运维管理工作。



魏征,男,西安交通大学硕士生,主要从事网络测量、计算机网络相关研发工作。



胡成臣,男,博士,西安交通大学电信学院计算机系副教授、博士生导师,主要研究方向为计算机网络通信系统,包括软件定义网络、网络测量等。



向万红,男,远光软件股份有限公司副总裁,主要研究方向为电力企业信息系统架构设计。

郭前,男,南京云帕信息技术有限公司联合创始人,主要从事移动互联网、云计算的开发与研究。



唐亚哲,男,博士,西安交通大学电信学院计算机系副教授、博士生导师。主要研究方向为计算机网络、网络管理和网络安全。