



基于 Tileria 平台的网络细粒度应用行为识别

吴 舜¹, 苏 丹¹, 吴 佳¹, 李 坤¹, 许大卫¹, 刘 昀¹, 魏 征²

(1. 国网冀北电力有限公司信通分公司 北京 100053; 2. 西安交通大学 西安 710049)

摘 要:网络管理员需要根据用户对应用的使用情况决定如何合理地分配带宽,提供区别化服务,进行安全监测和管理,以进行更加细致和准确的 IT 运维。研究了网络细粒度应用行为识别系统,能根据应用产生的流量精准识别用户在应用上的行为。与 x86 架构相比,为了实现更高的吞吐率和更低的功耗,应着重在 Tileria 平台上对应用识别进行并行设计和实现。最后对两种平台下系统的各种性能测试结果进行了对比,并得出结论。

关键词:网络行为识别;Tileria 平台;并行化

doi: 10.3969/j.issn.1000-0801.2013.11.016

Fine-Grained Network Traffic Identification Based on Tileria Platform

Wu Shun¹, Su Dan¹, Wu Jia¹, Li Kun¹, Xu Dawei¹, Liu Yun¹, Wei Zheng²

(1. Information & Telecommunication Company of State Grid Jibei Electric Power Co., Ltd., Beijing 100053, China;

2. Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: It is essential for network managers to decide how to use a reasonable allocation of bandwidth, providing differentiated services, as well as conducting security monitor and management so as to do fine-grained and precise IT operation. The fine-grained traffic identification method proposed was able to identify the user's behavior accurately according to the traffic that the application generates. In addition, compared with the traditional x86 platform, in order to achieve higher throughput and lower power consumption, the method was redesigned and implemented in the Tileria platform with parallelization. Finally, a variety of system performance test results were compared under the two platforms.

Key words: network traffic identification, Tileria platform, parallelization

1 引言

传统 x86 平台上的 CPU 是一维总线架构,当核的个数超过 8 个或 16 个时,在性能上会产生较大的瓶颈,而 Tileria 平台的处理器基于二维网状架构,把重点放在核与核之间的通信上,从而在提供高性能应用的同时也能大大节约功耗。

本文提出了一套高效的网络细粒度应用行为识别系统,根据应用产生的流量可以精准识别用户在应用上的各

种行为,满足网络管理员的相关需求。首先,用户行为识别(也可称为应用行为识别)可以为企业网络带来更高的网络安全等级。当某种应用行为被识别出后,服务提供商可使用白名单或黑名单手段保证网络的安全运行;其次,用户行为的精确识别可以帮助管理员优化网络,可针对不同的用户行为、用户群体,提供服务质量不同的网络,从而实现整个网络系统的最优化。

本系统运行在核心网关等大型设备上,而核心网关等大型设备的吞吐率都在万兆级别,所以研究重点在于:如

何在可接受的资源消耗下,使系统的吞吐率达到万兆级别或者更高,以满足未来几年的需要。为实现这个目标,从系统并行化入手,建立多个用于识别数据分组的单元,识别单元根据规则库对数据分组进行匹配,匹配到哪条规则就表示匹配到哪个行为,故也称为匹配引擎。多个匹配引擎同时对数据流量进行识别,从而充分发挥系统的并行性。在 Tilera 平台下,多核可编程智能数据分组引擎(multi-core programmable intelligent packet engine, mPIPE)作为数据分组调度器,通过散列方法,保证同一条流中的所有数据均在同一个匹配引擎上进行处理,同时将这些不同的流分配到不同的匹配引擎上,保证匹配结果的准确性和负载均衡。

2 系统模块及并行化结构

2.1 数据分组数据提取——FlowSifter

匹配引擎是将数据分组中某个域的值与规则库中的相应域进行比对,从而进行匹配。而在这之前,系统需要一个中间件,从原始的数据分组中提取数据并转换为合适的数据结构,供匹配引擎使用。在本系统中,FlowSifter 完成这样的功能。比如,它可以将一个 HTTP 数据分组转换为如下形式,存储在自定义的数据结构中:user-Agent: Mozilla/4.0(compatible;MSIE 5.01;Windows NT 5.0;DigExt)、Host:www.www.icconcept.com:8080 等格式化的数据,如图 1 所示。

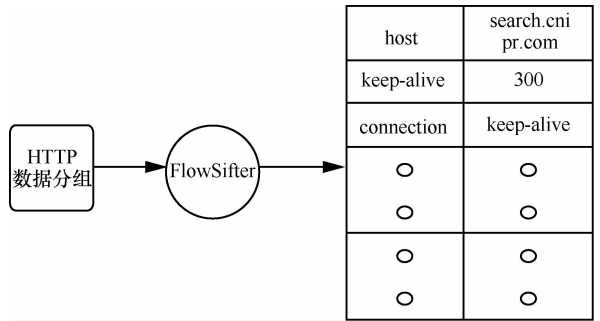


图 1 FlowSifter 的功能

2.2 匹配引擎——ROOM

数据分组被 FlowSifter 解析成格式化的数据后,就要被匹配引擎处理,从而识别出用户行为。匹配引擎要进行匹配功能,需要两种数据,一种是提前组织好的规则库,另一种是需要处理的格式化的数据,将这两种数据作为输入,便能输出匹配结果。

本文使用的匹配引擎为 ROOM(rule organized optimal matching engine)。现有技术中,大部分匹配引擎的规则组

织方法是基于正则表达式的,而 ROOM 的规则库是基于语义组织的。

一般来说,将规则集中的规则按照基于语义匹配的方法进行组织,即在每个域上建立一个基于正则表达式的匹配节点。举例来说,对于表 1 的规则集,利用基于语义的匹配方法会建立两个匹配节点,分别针对 host 域和 keep-alive 域,在每个域上基于正则表达式建立一个 DFA(definite finite automate),用于识别格式化数据在这个域上匹配的一条或多条规则。以同样的方法匹配第 2 个域,之后对每个域的匹配结果取交集,匹配格式化数据所有域的一条或多条规则就是所得的结果。

表 1 有 3 条规则的规则集

编号	host 域	keep-alive 域
1	Search.cni pr.com	300
2	Web.qq.com	400
3	Web.qq.com	500

然而,这种匹配方法存在一个问题:当规则数目十分巨大,每个域都需要对全部规则进行匹配且需要对各个域的匹配结果取交集时,极大地减慢了系统的速度。考虑到这种缺点,ROOM 采用了一种去除匹配冗余性的结构——树型匹配,加快了匹配速度。

2.3 并行化结构

在 Tilera 平台上设计如图 2 所示的系统架构。其中,数据分组流量通过 4 个 XGBE XAUI 口进入系统,被 mPIPE 以线速解析。mPIPE 按流划分数据分组,将这些流轮询地调度到每个处理器上,即 worker tile。在 worker tile 上,FlowSifter 模块和 ROOM 模块完成对数据分组的处理。

3 实验结果

本文分别在两种平台下测试系统的性能,x86 平台的配置环境为 Intel Xeon E5-2620 12 核(2.0 GHz,2×6 核)处理器,32 GB 内存;Tilera 平台的配置环境为 TILEGx8036,36 tiles(1.1 GHz),16 GB 内存。本次实验的测试流量从网络截取,流量大小为 4.5 GB,持续时间为 82 min,数据分组平均长度为 618 byte。在 Tilera 平台下,系统需要从 XGBE 口接收流量,而另一台设备或者通过其他方式能给 XGBE 口发送数据分组,测试中,将 TILEGx8009 作为流量产生设备通过光口 XGBE1、XGBE2 以合计 20 Gbit/s 的速度向系统发送数据分组流量。而 TILEGx8009 自身将流量文件转换为数据分组从光口发送出去,故两个平台可使用

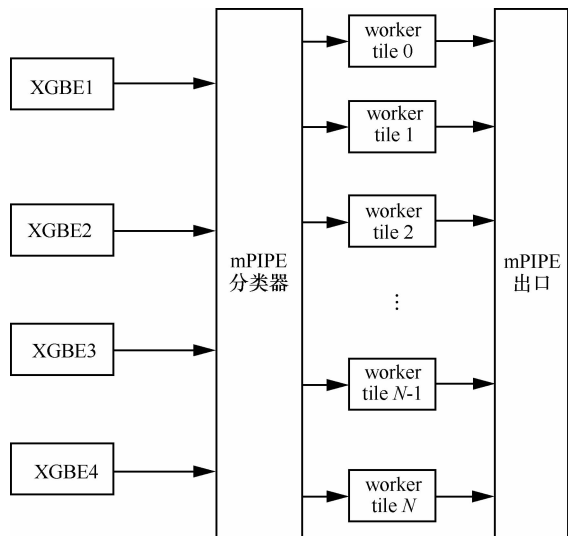


图2 Tilera 平台上的系统架构

同一个流量文件。

3.1 案例测试

本系统根据用户使用应用产生的流量达到对用户行为进行识别的目的。在两种典型应用案例下测试系统,证实结果的精准和有效性:检测互联网的浏览市场份额;检测在线视频流量所属应用的种类及所占比例。

首先,检测用户使用的浏览器市场份额。浏览器的市场份额对于用户和投资商有重要的参考价值,需要一个精准的测试系统测试浏览器的市场份额。以上述测试流量作为输入,根据基于语义的规则对流量中的 HTTP 数据分组进行匹配,方法是:提取出属于每种浏览器的页面视图的数量,并计算每种页面视图所占的比例,从而得到浏览器市场份额。图3是使用这种方法对测试流量进行匹配得到的结果。

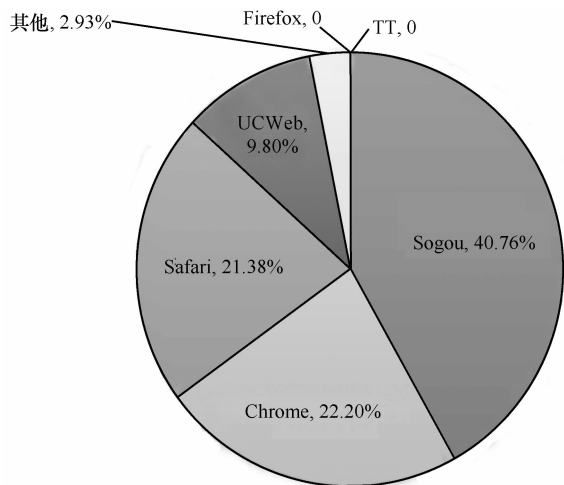


图3 使用页面视图比例得到的浏览器市场份额

从图3中可以看出,由于搜狗(Sogou)浏览器使用了代理加速,从而使其占据了最多的浏览器市场份额,火狐(Firefox)和TT浏览器在浏览器竞争中基本没有市场份额,同时会有更多其他种类的浏览器在这方面进行竞争。

其次,检测在线视频流量所属应用的种类及所占比例。对于想提高网络服务质量的管理员,将视频流量与其他流量区分开是十分重要的。但新的网络技术的发展,如HTML5,使得纯文本内容和视频内容都使用HTTP进行传送,识别二者变得更加困难。在此次实验中,试图检测出HTTP流量中的视频流量,这些视频流量分别来自中国排名前6的在线视频应用,如土豆、音乐台、优酷、PPTV、LeTV、iQiyi。所得的结果见表2。

表2 各个视频流量所占大小

在线视频应用	所占的流量大小(MB)
土豆	122.4
音乐台	9.2
优酷	217.6
PPTV	52.9
LeTV	33.4
iQiyi	29.2

从测试结果可以看出,土豆和优酷产生大部分的视频流量,验证了它们是中国最大的两家在线视频提供商,同时出现了越来越多的在线视频提供商,如音乐台、LeTV。

3.2 吞吐率测试

在x86平台上,用数据分组流量文件的大小(单位为MB)除以处理时间(单位为s)即可得到x86平台上的吞吐率。在Tilera平台上,使用9核TILEGx8009处理器将上述流量文件中的数据分组从光口以20 Gbit/s的速度打入,之后分别在核数为1、2、4、8、16、32、35的情况下进行吞吐率测试。每个核每处理200万个数据分组,计算一次吞吐率(单位为Mbit/s),计算式为:单次处理分组数的总字节大小 $\times 8 \text{ byte}/(1\,000 \times 1\,000 \times \text{处理时间})$ 。x86平台和Tilera平台的吞吐率如图4所示。

从图4可以看出,x86平台在10核时达到36.1 Gbit/s的系统吞吐率,10核表示运行5个FlowSifter+ROOM对、1个InsertFlowTable线程和1个主线程;Tilera平台在32核时达到19.5 Gbit/s的吞吐率,这是因为受到线速的限制,如果没有线速的限制,吞吐率会更高。

3.3 内存测试

对于内存,测试了两种类型的内存消耗:一种是静态

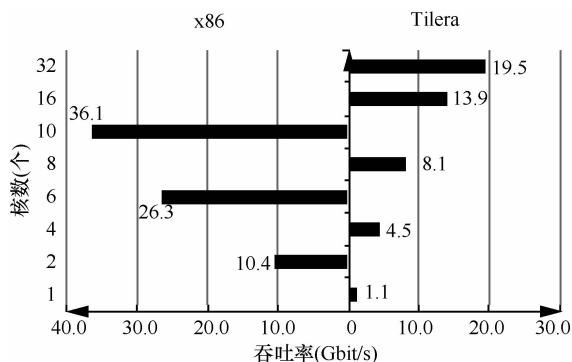


图4 两种平台下的系统吞吐量

内存,指系统初始化时一些必需的数据结构所占用的内存,如建立匹配引擎所需的树型匹配结构等;另一种是动态内存,指系统处理数据分组期间产生的一些中间数据所占用的内存。两种平台下的静态内存消耗如图5所示,动态内存消耗如图6所示。

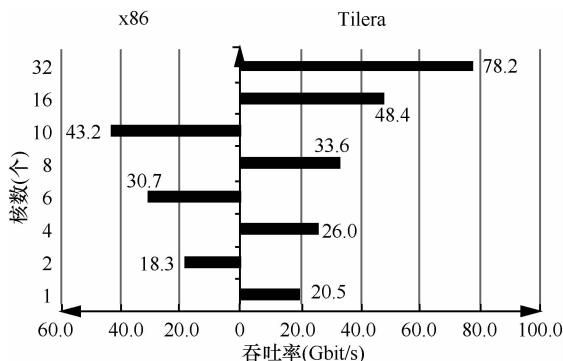


图5 两种平台下的静态内存消耗

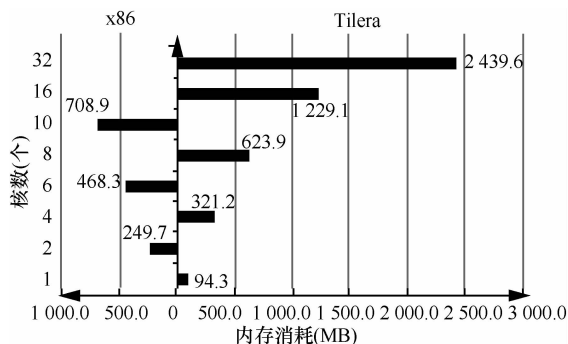


图6 两种平台下的动态内存消耗

可以看出,静态内存消耗随着核数的增加而增加,这是因为在静态内存上每个匹配线程都需要一个自己的树型匹配结构,从而避免数据竞争;在动态内存上,匹配产生的中间数据也随着核数的增加而增加。在Tilera平台上的内存消耗中,静态内存包括一些额外的消耗,如mPIPE队列等,且在运行过程中,每个tile需要将数据分

组缓存起来进行处理,使得动态内存随着核数的增加而增加。可以看到,系统在两种平台上的内存消耗基本呈线性的,是可以接受的。

4 结束语

本文研究了如何更好、更快、更有效地基于Tilera平台在互联网中对用户在应用上的行为进行细粒度识别,其中,FlowSifter模块使得匹配引擎只关心匹配,而不必再提取每一个数据分组的数据,从而变串行为并行。基于语义的规则,匹配方法和新型树型匹配结构使得匹配引擎的匹配速度获得比较好的提升,同时,在内存占用方面也有所优化。最后,与x86平台实现的系统进行对比实验,得出结论:相对于单核性能强劲的x86系列处理器,Tilera平台的处理器在核数较多的情况下性能功耗比更高。

参考文献

- 1 Paxson V. Bro: a system for detecting network intruders in real-time. *Computer Networks*, 1999, 31(23/24): 2435~2463
- 2 Pang R, Paxson V, Sommer R, *et al.* Bnpac: a yacc for writing application protocol parsers. *Proceedings of ACM IMC*, Florida, USA, 2006
- 3 Li Z C, Xia G, Gao H Y, *et al.* NetShield: massive semantics-based vulnerability signature matching for high-speed networks. *Proceedings of the ACM SIGCOMM*, New Delhi, India, 2010
- 4 Li H, Hu C C. ROOM: rule organized optimal matching for fine-grained traffic identification. *IEEE INFOCOM 2013 Mini-Conference*, Turin, Italy, Apr 2013
- 5 Cisco IPS AIM and IPS NME for Cisco 1841 and Cisco 2800, 2900, 3800 and 3900 series ISR. http://www.cisco.com/en/US/prod/collateral/routers/ps5853/ps5875/product_data_sheet0900aecd806c4e2a_ps2641_Products_Data_Sheet.html, 2009
- 6 烽火网络 F-engine PL 系列 DPI 产品. <http://www.fhn.com.cn/product/listpro16/422.html>, 2011
- 7 RadwareANS G 系列产品. <http://www.radware.com/Products/ApplicationNetworkSecurity/default.aspx>, 2013
- 8 Snort: an open source network intrusion prevention and detection system, 2010
- 9 OpenDPI. <http://opendpi.org>, 2010
- 10 PACE: network analysis with layer-7 deep packet inspection, 2011



- 11 Tiler Inc. TILE-Gx 系列产品. http://www.tiler.com/products/processors/TILE-Gx_Family, 2013
- 12 Tiler Inc. Tiler 应用工具包. http://www.tiler.com/sites/default/files/productbriefs/NetLib-PB038-02_web_0.pdf, 2013
- 13 Tiler Inc. Tiler 软件平台. <http://www.tiler.com/products/software>, 2013



吴佳,女,国网冀北电力有限公司信通分公司数据中心工程师,主要从事信息系统运维和信息化项目管理工作。

[作者简介]



吴舜,男,国网冀北电力有限公司信通分公司网络管理工程师,主要从事信息化企业架构管理、信息通信专业同业对标和信息化项目建设管理工作。



苏丹,女,国网冀北电力有限公司信通分公司副主任,主要从事信息化项目管理及项目建设推进工作。

李坤,男,国网冀北电力有限公司信通分公司助理工程师、主任,主要从事信息系统运行管理、信息安全管理和信息化项目管理工作。

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

刘昀,女,国网冀北电力有限公司信通分公司信息安全主管,主要从事信息安全管理。

魏征,男,西安交通大学硕士研究生,主要从事计算机网络研发工作。

(收稿日期:2013-08-01)



·简讯·

“LTE 全球生态系统发展研讨会”成功开幕 FDD/TDD 协同发展引全球聚焦

2013 年 11 月 15 日,“LTE 全球生态系统发展研讨会”在北京国宾酒店召开。本次研讨会以“全球视野合作共赢”为主题,探寻 LTE 在全球市场的发展轨迹,探讨 LTE 的现状与未来。作为本次会议的焦点,LTE FDD 和 LTE TDD 的共同融合和混合运营无疑备受关注。本次研讨会特邀请政府主管部门、国内外运营商、电信设备商、芯片厂商以及第三方咨询机构进行了精彩演讲,充分展示了 LTE FDD 与 LTE TDD 在技术基础、网络演进和终端服务上的共通性以及两大模式融合发展的广阔前景。

来自 GSA 的数据显示,LTE 已经实现超过 1.26 亿的全球连接,83 个国家部署了 222 个 LTE 商用网络,LTE 的全球化扩张不断加速。同时根据第三方数据,目前全球已有约 23 个 LTE TDD 网络商用,另有 60 个

LTE TDD 网络正在建设或部署中。值得关注的是,23 个 LTE TDD 网络中有 11 个是由已经部署了 LTE FDD 的运营商所推出的,其中包括中国移动(香港)、软银(日本)及俄罗斯 MTS 等。在 LTE 全球化布局不断扩张的大背景下,LTE FDD/TDD 网络的协同发展格局正在不断清晰。

LTE 的融合发展走势引起了整个产业上下游的广泛关注。在本次研讨会上,工业和信息化部科技司副局级巡视员戴晓慧、中国移动研究院副院长黄宇红、GSA 主席 Alan Hadden、高通无线半导体技术有限公司大中华区总裁王翔以及 Signals Research 集团 CEO Mike Thelander 分别发表了相关的主题演讲。

本次会议由中国通信企业协会主办,人民邮电出版社协办,北京信通传媒有限责任公司承办,来自政府主管部门、电信运营商、设备厂商、芯片厂商、终端厂商、应用提供商、科研院所、新闻媒体等在内的近 200 人共同出席了本次研讨会。