

低开销的网络流量检测监测机制

余林琛^{1,2}, 赵保全¹

(1 中国地质大学(武汉) 计算机学院, 湖北 武汉 430074; 2 华中科技大学 计算机学院, 湖北 武汉 430074)

摘 要: 网络流量监测是网络管理的重要组成部分, 数据包的捕获是网络流量监测的前提。但是, 传统监测技术缺乏对网络流量监测并提供预警和预测功能。文中提出了低开销的网络流量监测机制。在捕获网络流量数据包后, 对其分析, 并以友好界面的形式将分析结果呈现给用户。基于此, 实现了流量统计和流量预警机制。当网络流量出现异常时, 系统可以及时地向网络管理员反馈, 从而让管理员了解网络的运行状态及可能存在的危机, 并做好相应的应对策略。实验表明, 结果符合设计需求。

关键词: 网络安全; 流量监测; 效率

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2012)09-0015-03

Network Traffic Monitoring Scheme with High Efficiency

YU Lin-chen^{1, 2}, ZHAO Bao-quan¹

(1 School of Computer Science, China University of Geosciences, Wuhan 430074, China;

2 School of Computer Science, Huazhong University of Science and Technology, Wuhan 430074, China)

Abstract: Network traffic monitoring system is the main part of network management, and how to capture the network data packets is the base, but traditional schemes lack of forecasting and alarm functions on network traffic with low efficiency. This paper proposes a network traffic monitoring scheme with high efficiency. In the scheme, we will analyze the captured data packets and put the good user interface on the analysis results to users. Based on the capturing and analyzing, the network traffic statistics and alarm has been implemented. When the network flow is abnormal, the systems can timely feedback to the network administrator, which allows the network administrator to understand the network running status and crisis and make some corresponding coping strategies. Experimentations show that the performance is good.

Key words: network security; traffic monitoring; efficiency

1 引言

网络信息数据分析与监测技术^[1]是监测网络流量状态、检测网络入侵行为以及排查网络安全隐患的一项有效技术。网络流量监测系统采用有效的网络监听技术, 在固定的网络节点上用抓包工具进行数据包采集, 并对抓到的数据包逐层分析, 并向网络管理员提供预警机制。这对于简化网络管理工作, 增强网络的可用性发挥了重要作用。

当前, 主要的网络流量监测方法包括: 基于流量

镜像协议分析^[2]、基于硬件探针的监测技术^[3]、基于 SNMP 的流量监测技术^[4]、基于 Netflow 的流量监测技术^[5-6]等。

流量镜像(在线 TAP)协议分析方式是把网络设备的某个端口(链路)流量镜像给协议分析仪, 通过 7 层协议解码对网络流量进行监测。与其他 3 种方式相比, 协议分析是网络测试的最基本手段, 特别适合网络故障分析。缺点是流量镜像(在线 TAP)协议分析方式只针对单条链路, 不适合全网监测。

硬件探针是一种用来获取网络流量的硬件设

备,使用时将它串接在需要捕捉流量的链路中,通过分流链路上的数字信号而获取流量信息.一个硬件探针监视一个子网的流量信息.对于全网流量的监测需要采用分布式方案,在每条链路部署一个探针,再通过后台服务器和数据库,收集所有探针的数据,做全网的流量分析和长期报告.

基于 SNMP 的流量信息采集,实质上是测试仪表通过提取网络设备 Agent 提供的 MIB(管理对象信息库)中收集一些具体设备及流量信息有关的变量.基于 SNMP 收集的网络流量信息包括:输入字节数、输入广播包数、输入包丢弃数、输入包错误数、输入未知协议包数等.与其他的方式相比,基于 SNMP 的流量监测技术受到设备厂家的广泛支持,使用方便,缺点是信息不够丰富和准确.

Netflow 流量信息采集是基于网络设备(Cisco)提供的 Netflow 机制实现的网络流量信息采集. Netflow 由路由器、交换机自身对网络流量进行统计,并且把结果发送到第三方流量报告生成器和长期数据库.一旦收集到路由器、交换机上的详细流量数据后,便可为网络流量统计、网络使用量计价、网络规划、病毒流量分析,网络监测等应用提供计数根据.

但是,传统方法缺乏对于流量监测数据的高效率预测与告警等功能.本文将对此展开研究,并提出了一种新的网络流量监测机制.

2 高效低开销的网络流量监测机制

本系统所采用的流量监测模型^[7]可以表述为如下.考虑网络受监测端口总体发出/进入的数据量($N_i(\text{out})(T_j)$, $N_i(\text{in})(T_j)$),表示了该端口流出和流进的字节数总量.基于此,可以求出该端口流进、流出的速率,即单位时间内的数据量,设置为 $F_i(\text{out})$, $F_i(\text{in})$. 则如下式表达:

$$F_i(\text{out}) = (N_i(\text{out})(T_j) - N_i(\text{out})(T_k)) / (T_j - T_k)$$

$$F_i(\text{in}) = (N_i(\text{in})(T_j) - N_i(\text{in})(T_k)) / (T_j - T_k)$$

但是,现实网络中大数据量的传输网络往往存在较大传输与处理的延迟.因此,为了增加数据流量获取的真实性,可以采用康托三分集的分形理论.该集合中有无穷多个点,所有的点处于非均匀分布状态,此点集具有自相似性,其局部与整体是相似的,所以是一个分形系统.基于康托三分集的意义在于,既然不能一次取定两个时间点来计算网络流量,那么,可以基于分形理论取定一批自相似的点集,通过该点集的集合来计算网络流量的速率,并基于该点

集集合还可以预测未来流量变化,从而降低了监测开销,提高了准确度.需要特别说明的是,在计算中,所取得的每一个时间点,都是康托三分集中的一个点.

本系统由用户或网络管理员设定相应的过滤规则(如选择的协议、目的地址、流量的限制等),然后开启抓包进程,系统将根据选择的规则捕获数据包,在对数据包分析处理后以友好界面的形式将分析结果呈现出来.系统的基本模型如图 1 所示.

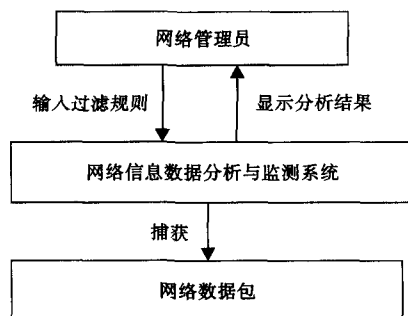


图 1 系统的基本模型

本系统共设计了以下几个功能:数据包过滤规则模块:该模块的设计允许管理员根据特定的需求输入相应的过滤规则.其中过滤规则包括:协议选择、地址选择、网卡选择、流量的上下限设置等.数据包捕获控制模块:该模块主要用于数据包的捕获与停止.显示模块:该模块显示的内容主要有:已选择的过滤规则、捕获数据包的内容分析等.流量统计模块:该模块由两个子模块组成.第一个是将数据包的当前流量等信息的统计并显示在流量统计窗口;第二个模块是讲实时的流量信息以图形的形式显示在流量图中.保存模块:该模块用于将数据包的分析内容保存在文本文档中.

本系统的具体工作流程如下:

(1) 数据包过滤.在该模块中,有协议选择、过滤规则和流量设置三个子模块.在协议选择中,本系统并未全部囊括所有的协议,只是将网络协议(混杂模式)、以太网协议、ARP 协议、IP 协议、ICMP、TCP、UDP 等常用的几个协议列举出来,当双击协议的时候,系统在捕获数据包的时候只捕获与所选协议相匹配的数据包;如果没有选择协议,系统将以默认的方式进行数据包捕获.在过滤规则模块中,网络管理员可以输入源 IP 地址或目的 IP 地址,系统在捕获数据包时指捕获与输入的地址相匹配的数据包;如果没有输入,系统将以默认方式进行捕获数据包.在流量设置模块中,网络管理员可以对流量上

限、流量下限、数据包上限、数据包下限进行设置,当在流量统计过程如果某时刻流量高于或低于设定值,系统将弹出相应的窗口提示网络管理员。

(2) 数据包捕获。该模块主要控制数据包的捕获与停止,并根据当前状态禁用或启用某些功能。当点击开始按钮时系统会弹出当前检测到的网络接口,网络管理员选择需要的网络接口开始数据包捕获。在未选择相应的网络接口的时候,本系统是不能进行数据包捕获的。开始捕获数据包后,可根据需要点击停止按钮随时停止对数据包的捕获。

(3) 数据分析与显示。在该模块中有显示配置信息、显示统计信息、显示流量图、显示数据包内容分析四个子模块。当数据包开始捕获后,这四个模块即被启用。显示配置信息模块将选择的网络接口、选择的协议、输入的 IP 地址以及流量的设置信息显示在系统的主界面上。显示统计信息模块又被分为两个子区域:数据包信息区和流量信息区。在数据包信息区显示当前数据包、平均数据包、最大数据包以及数据包的总量和各个协议的数据包数量;在流量显示区显示当前流量、平均流量、最大流量、抓包开始的时间、当前时间以及流量达到最大时的时刻。显示流量图模块将流量图以条形图的形式显示出来。数据包内容分析后将数据包的序号、版本号、源地址、目的地址、协议类型、首部长度、总长度、标识、标志、片偏移、生存时间、抓取时间显示在系统的主界面上。

(4) 结果保存与预测预警。当抓包停止后,可以选择保存,并开始执行统计分析与预测预警机制。

3 系统实现与测试分析

按照上述流程与设计,本文设计实现了网络流量监测与分析系统,它包括了规则配置、流量监测、统计分析等主要过程。

如图 2 所示,展示了网络流量监测与分析中的规则配置过程。从图中可以看到,系统提供了主要的配置规则与监测策略、对象。用户可以按照需求自行选择、设置,体现了系统的灵活性。

如图 3 所示,网络流量中的数据包取到之后还需将数据格式化才能在主界面的列表框中显示出来,以下语句可以将编号、版本号、源地址、目的地址、头部长、总长度、标识、标志、片偏移、生存时间以及数据包的抓取时间依次格式化。格式过的内容

按顺序显示在主界面的列表框中。

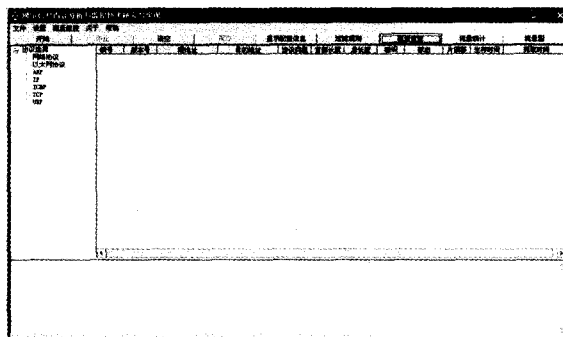


图 2 规则配置与分析

图 3 抓包分析示意

如图 4 所示,展示了网络流量统计与分析的结果。该部分的基本思想是,首先记录下开始抓包的时刻并传送到流量统计模块中的起始时刻的编辑框中,作为流量统计的基准时刻。通过时间函数动态获取系统的当前时间并实时的将时间传送到流量统计模块中的当前时刻。如果要计算当前流量和当前数据包,只需要统计出每秒内抓取的数据包个数以及这些数据包的总长度之和。通过单位换算后的值即为当前数据流量和当前数据包流量。从图中可以看到,结果符合预期。

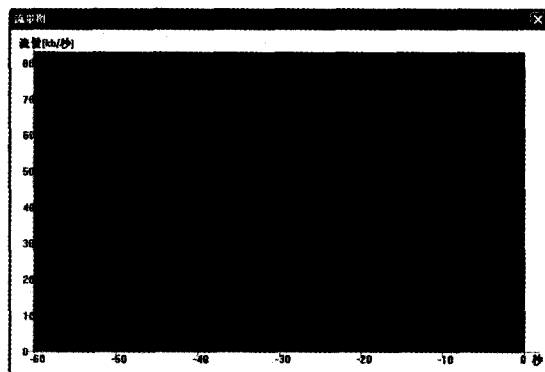


图 4 流量统计与分析

(下转第 22 页)

节点间达到平衡。

5 结束语

云计算的出现也为互联网带来了更丰富的应用,云数据存储技术是云计算中的核心领域,云计算及云存储的快速发展对于数据的负载均衡提出了很高的要求。本文对于 Hadoop 的数据负载均衡算法进行了深入的研究,并提出了改进算法,通过大量的实验验证了优化算法比原算法具有更高的效率。

参考文献:

- [1] 张建勋,古志民,郑超. 云计算研究进展综述[J]. 计算机应用研究,2010, 27(2):429-433.
- [2] 陈康,郑纬民. 云计算:系统实例与研究现状[J]. 软件学报,2009,20(5):1337-1348.
- [3] Dhruba Borthakur. The hadoop distributed file system: architecture and design [EB/OL]. (2008-09-02) [2010-08-25]. http://hadoop.apache.org/common/docs/r0.16.0/hdfs_design.html.

- [4] Jeffrey Dean, Sanjay Ghemawat. MapReduce: simplified data processing on large clusters[C]// Proceedings of the 6th Symposium on Operating System Design and Implementation. New York: ACM Press, 2004:137-150.
- [5] Hadoop HDFS[EB/OL]. (2011-10-18) [2011-10-25]. <http://hadoop.apache.org/hdfs/>.
- [6] Caibinbupt. Hadoop 源代码分析(重读 GFS 的文章) [EB/OL]. (2009-01-29) [2010-8-25]. <http://caibinbupt.javaeye.com/blog/318949>.
- [7] Tom White. Hadoop: the definitive guide[M]. United States of America: O'Reilly Media, Inc. 2009.
- [8] 分布式基础学习[EB/OL]. (2009-02-22) [2010-08-25]. <http://www.cnblogs.com/duguguuiyu/archive/2009/02/22/1396034.html>.

作者简介:

刘 琨 女,(1980—),讲师.研究方向为分布式系统、云计算。

(上接第 17 页)

4 结束语

本文基于三分集的分形理论,提出了网络监测与分析系统的设计思路,实现了网络流量的数据获取、监测、分析与预测等功能。实验结果表明,该系统具有较好的监测准确度,开销也较小,符合实际应用的需求。未来,还可继续实现面向不同应用的流量监测与分析,从而可以提取不同的监测模型。

参考文献:

- [1] Chang-Yong Lee, Hwan-Kuk Kim, Kyoung-Hee Ko, et al. A study on structure for monitoring and detecting VoIP abnormal traffic [C]//FGCN 2008. Washington DC, 2008: 178-182.
- [2] Cîrnei A, Boboc S, Leordeanu C, et al. Netpy: advanced network traffic monitoring [C] //INCOS 2009, IEEE Computer Society. Romania, 2009: 253-254.
- [3] Ning Zhang, Hong Bao. Research on traffic technology in communication network [C]//ICNDS 2009. China: GuiLin, IEEE Computer Society, 2009: 52-55.
- [4] Wenjian Xing, Yunlan Zhao, Tonglei Li. Research on the defense against ARP spoofing attacks based on winpcap [C] // Proceedings of Second International

Workshop on Education Technology and Computer Science (ETCS). IEEE Computer Society, Wuhan, China, 2010: 762-765.

- [5] Wang Xuren, He Famei. An implement of broadband network monitoring system based on libnids and winpcap [C]// Proceedings of International Conference on New Trends in Information and Service Science (NISS). Beijing, China, 2009: 812-814.
- [6] Xiaofan Lu, Weijia Sun, Huiping Li. Design and research based on WinPcap network protocol analysis system [C]// Proceedings of 2010 International Conference on Computer. Mechatronics, Control and Electronic Engineering (CMCE). Changchun, China, 2010: 486-488.
- [7] Yuhui Liu, Jinshan Sun, Rui Sun, et al. Next generation internet traffic monitoring system based on netFlow [C]// Proceedings of 2010 International Conference on Intelligent System Design and Engineering Application (ISDEA). Chengdu, China, 2010: 1006-1009.

作者简介:

余林琛 女,(1977—),讲师,博士研究生.研究方向为对等计算。