NBOS:一个基于流技术的精细化网管系统

张维维，龚俭，丁伟，张孝国
(东南大学 计算机科学与工程学院,江苏 南京 211189)

摘要：在高速主干网环境下，为了实现对整个被管网络的实时流量监测和精细化网络管理，本文基于流抽样技术获取主干链路上的流数据，通过合理的测度设计和计算为用户提供网络服务质量、热点报告，让用户可以了解目前被管网络的宏观运行状态和需要被关注的微观细节信息，并进一步提供网络安全态势报告以及攻击细节。目前该系统正式运行于CERNET华东北地区网络中心，在20Gbps流量负载下成功实现了对江苏各高校网络的实时流量监测和精细化管理。

关键字：高速主干网，流技术，流量监测，网络管理，服务质量，热点，网络安全

中图分类号：TP393.0　　文献标识码：A

0 引言

近年来，网络规模迅速扩大，网络应用日益多样化，给流量监测分析和网络管理带来新的挑战。一方面，主干网带宽逐渐向40Gb/s甚至更高速率过渡，网络流量急剧增加，流量监测需要消耗大量的存储和计算资源，其性能难以赶上流量增长。为此，如何在保证低丢包率甚至不丢包的前提下，及时高效地接收和分析流数据是一个需要解决的难点问题。另一方面，基于网元设备的传统网管系统不能继续满足当前需求，网络管理员为下一代网管系统提出了新的精细化管理需求：从被管对象看，需要基于不同粒度的用户群对网络进行监测和管理（整个被管网可以构成一个用户群，一个学校网可以构成一个用户群，甚至单个用户也可以看成一个用户群）；从统计内容看，除提供一般的流量统计、性能分析、服务质量评估外，还需尽可能详细地挖掘流量背后隐藏的通信行为、应用模式、流量异常和安全事件；从观测时间看，不仅需要实时监测网络流量，了解当前运行状态，及时发现流量异常和网络攻击，还需要存储网络流量，分析其长期变化趋势；从监管视角看，不但需要感知整个被管网宏观运行情况和安全状态，预测未来发展趋势，而且在流量异常和攻击时还需提供微观细节信息，确定引起事故的服务端口和主机。

在高速主干网环境下，为了实现对整个被管网络的实时流量监测和精细化网络管理。传统基于SNMP和基于包嗅探的流量分析技术，无法同时满足数据精度和性能需求。SNMP技术无法分析具体的用户流量和协议组成；包嗅探技术需要依靠专用“探针”，设备费用昂贵，全报文采集分析资源开销大，其性能也难以适用于高速主干网环境。当前流技术的发展为主干网流量监测和分析提供了可能。一方面，流技术提供了丰富的基于端到端连接的流量信息；另一方面，需要处理的流数据量不到全报文的5%，有效减少了高速大容量网络状况下数据采集和分析对网络设备的要求。目前有许多开源流数据采集分析工具，其中最具代表的是Flow-tools[1]和Nfdump[2]，但是它们大多是基于串行化程序设计，在突发流量下会出现很高的丢包率，无法满足主干网环境下流量监测的性能要求。此外目前基于NetFlow流数据的网管系统基本上是基于流记录内容的直接统计，缺乏深度处理，实用性都不够理想。

本文给出了一个新型精细化网管系统NBOS的设计。首先基于网络流技术，并通过多缓冲区和多线程并发机制，能够在保证低丢包率甚至不丢包的前提下，及时高效地采集高速主干网上的流数据；其次根据IP地址所属行政单位和地理位置，提出了一种因特网区域划分的方法；然后基于网络区域的划分，以适当的时间粒度为单位，从被管单位（或者个人）的角度出发，通过合理的测度设计，计算一般流量行为，评估网络服务质量，挖掘流量热点，发现流量异常，分析网络安全威胁；最后在上述统计分析的基础上感知整个被管网的宏观运行情况和安全状态，并在出现流量异常或网络攻击时，进一步提供微观细节的数据分析。目前该系统正式运行于CERNET华东北地区网络中心，基于主干链路上获取的流数据，实现对江苏各高校网络的实时流量监测和精细化网络管理。经过性能测试发现，本文提出的方案能在20Gbps的流量负载下保证较好系统性能。

基金项目：国家科技支撑计划课题“新一代可信任互联网安全和网络服务”（2008BAH37B04）
交稿日期：2012年9月12日
作者简介：张维维（1984年8月），男，江苏南通，博士研究生，主要从事网络行为学研究，Email:wwzhang@njnet.edu.cn

1 系统设计

1.1 设计目标

作为国家科技支撑计划课题“新一代可信任互联网安全和网络服务”的组成部分，NBOS (Network Behavior Observation System)是用于监控和管理 CERNET2 网络服务质量和网络安全状态的新型精细化网管系统。按照下一代精细化网管系统的要求，该系统需要基于主干链路上被动测量获取的流数据，通过合理的网络区域划分，以适当的时间粒度为单位，准实时的分析被管单位不同网络互联级别的流量行为，评估被管单位和全网的服务质量，发现流量▪点和流量异常，▪▪网▪安全事件，限制异常流量和网络攻击对正常网络服务的影响。此外，该系统还需进一步保存当前的流量分析结果，提供历史信息检索和流量特征▪化▪▪分析，并基于被管▪位的流量行为和安全状况，感知整个被管网络宏观的运行状态和发展趋势。

为了提供丰富的网管内容，还需设计相▪的网▪▪度，表1 为描述不同的网管内容，分别给出 NBOS 所需的测度及其语义。

1.2 总体结构设计

按照数据采集、分析、▪示的一般流程，NBOS 系统应该包含数据采集模块、数据分析模块和可视化模块。考虑整个数据分析模块功能的复杂多样性，根据测度计算内容的不同进一步分成基本流量行为计算模块、服务质量评估模块、热点和异常检测模块、以及安全威胁分析模块。这四个测度计算模块根据当前时间片内采集的流数据，各自从实时和历史两个方面计算和分析相应的测度及变化趋势，从微观和宏观两个角度分析并提供总体概况和微观细节描述。为了进一步精简各测度计算模块的功能，提取出它们相同的数据处理部分，统一由数据预处理模块完成。此外，作为一个完整的系统，还需要增加系统管理模块来保证整个系统正常运行。

综上所述，整个 NBOS 系统包含数据采集、数据预处理、基本测度计算、服务质量评估、热点及异常检测、安全威胁分析、可视化和系统管理八个模块，如图 1 所示。数据采集模块获取网络流数据，经过预处理模块统一的数据处理生成 NBOS 中间流数据并

表 1 测度定义

测度分类	测度名称	测度语义
基本流量行为	应用分布	观测周期内 www、p2p、邮件、多媒体等九种应用分别所占流量比例
	协议分布	观测周期内 TCP、UDP、ICMP 三种协议分别所占流量比例
	端口分布	观测周期内，占流数排名前二十的端口及其所占流数比例
服务▪量评估测度	全网出/入带宽	观测周期内，流出/流入被管网络的带宽总和
	全网逻辑往返时延	观测周期内，被管网和外部子网间所有端到端逻辑往返时延的平均值。端到端▪▪往返▪延通▪流持▪▪▪除以▪流发送报文轮数来估计[7]
	全网逻辑丢包	观测周期内，被管网和外部子网间所有端到端逻辑丢包的平均值。端到端逻辑丢包通过丢包率和往返报文数之间的关系进行推测[8]
热点和异常检测测度	源/宿流量热点	观测周期内，所占出/入流量超过总流量一定阈值的主机
	源/宿活动热点	观测周期内，所占出/入流数超过总流数一定阈值的主机
	源/宿访问热点	所占源/宿关联 IP 数超过总关联 IP 数一定阈值的主机
	热源/宿前缀	观测周期内，所占流量超过总流量一定阈值的源/宿前缀
	热连接	观测周期内，所占流量超过总流量一定阈值的连接
安全威胁分析测度	全网威胁类型	当前年内各威▪▪型的子▪型数、威▪源数、波及网▪数、▪化▪▪
	全网威胁子类型	观测周期内各威胁子类型的威胁源 IP 数量，感染单位数，感染主机数，感染数量▪化▪▪
	被管单位威胁状态	观测周期内各被管单位内部威胁源数量，恶意代码类型数量，感染主机数，危险程度

存入中间共享缓存，然后各测度计算模块借助读接口函数读取当前时间片内流数据计算各被管单位以及全网不同网络互联级别的测度并写入测度数据库，最后可视化

模块基于 web 页面显示各测度计算结果。

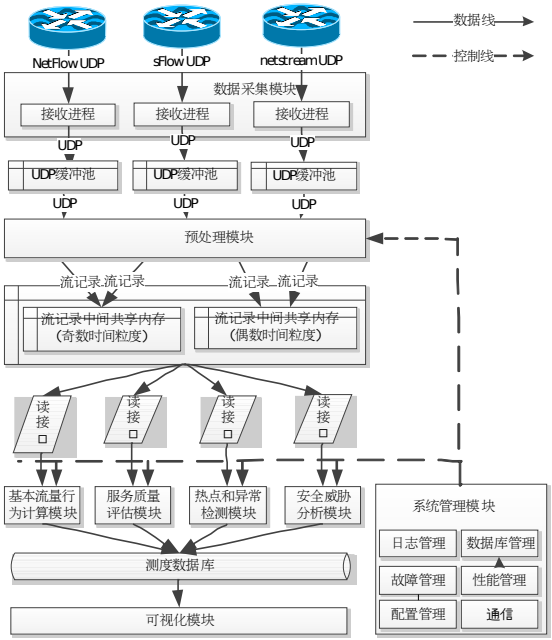

图 1 NBOS 系统总体结构图

从总体看，各测度独立成单独的计算模块，将可变的测度计算部分从稳定的数据处理流程中分离出来，提高了系统的稳定性和新测度计算的可扩展性；从局部看各模块功能彼此独立，基于合适的中间缓存可以实现系统的并行化运行，提高系统的处理效率。

2 网管数据的采集与维护

数据采集模块基于被动测量的流技术获取主干链路上的流数据 UDP 报文，并将采集到的 UDP 报文缓存到 UDP 缓冲池中。预处理模块同时从 UDP 缓冲池中读取 UDP 报文，将所有的数据格式统一转换为 NetFlow 格式，再经过一系列的数据预处理（时间片划分，数据过滤、短流合并、往返流合并、网络区域划分、流应用类型识别），得到中间流记录数据，并存放到中间共享缓存，供测度计算模块读取。关于数据采集和预处理模块的详细设计可见文献[3]。下面针对数据采集和存储维护涉及到的关键问题进行介绍。

2.1 多数据源问题

目前主流的流技术包括 Cisco NetFlow 技术[4]、华为 Netstream 技术[5]、以及 Juniper Sflow 技术[6]等，为了支持这 3 种数据源在不同抽样比条件下的并发工作，系统统一将接收到 SFlow 和 NetStraem 数据格式转换成 Netflow 格式。值得注意的是，SFlow 提供的其实是报文 trace 数据，而非真正意义上的流数据，因此需要仿照 Netflow 机制重新组流[4]；而 NetStream 则与 Netflow 比较接近，直接提取流记录中的对应字段即可。

2.2 数据采集性能问题

主干链路上流数据的实时采集，需要同时满足数据精度和采集性能两方面的要求。为此，本系统选用信息丰富的流数据作为数据源，基于并行化设计理念对数据和功能进行分解，并在此基础上通过多缓冲区和多线程并行化机制有效地提高数据采集效率；为了进一步提高整个系统后期数据处理的效率，统一对原始数据进行数据处理，通过聚合数据减少系统开销，通过减少冗余操作提高处理效率。经过性能测试，该解决方案能够在20Gbps的流量负载下保证较好的系统性能：丢包率小于10^{-6}，数据采集和预处理内存开销小，约8MB-76MB，具体见文献[3]。

2.3 时间片长度选取

为了兼顾测度计算模块的实时数据统计和历史趋势分析，采集数据需要在时间维度上进行划分和存储，因此需要合理地选取时间片长度，同时满足数据分析的实时性和数据处理的性能。太长的时间片影响数据统计分析的实时性，而太短的时间片又会影响数据采集和处理的效率，因此一个合理的时间片长度，应该同时满足数据分析的实时性和数据处理的效率两方面的需求。

$T \le MAX\_REAL\_TIME$ （公式 1）

$MAX\_REVERSE\_TIME \ge Pre(D(T)) + P(D'(T))$（公式 2）

公式1约束数据分析实时性要求，限定取 ▪ 上界；公式2满足数据处理效率需求，限定取值下界。MAX_REVERSE_TIME[3]为报文最大错序时间；Pre(D(T))为一个时间片数据预处理时间；P(D'(T))为一个时间片测度计算分析时间；MAX_REAL_TIME为出于实时性考虑能够容忍的最大滞后时间。通过实测数据源，MAX_REVERSE_TIME保持在30秒以内，为此本文设定时间片长

度为5分钟，可以同时满足实时性和处理性能需求。

2.4 网络区域划分

精细化网络管理的核心思想是能够基于不同粒度的子网对网络进行监测和管理既能将被管网看成一个整体，提供宏观的网络运行状况和安全状态报告，又能将被管网划分成一个个被管单位（或者用户个人），提供微观的细节信息说明。为此需要提供一种合理有效的方法，能够由粗到细逐层对被管网进行更细粒度的子网划分。

考虑到 IP 地址分配时，常常将一块连续的地址空间分配给同一个用户单位。因此一个合理的网络区域划分方法就是根据 IP 地址所属行政单位（ISP 或者用户单位）或者地理位置（国家、地区、城市）进行子网划分。本文首先将整个因特网划分成被管网内和被管网外，被管网内细分到各被管单位网，被管网外国际互联细分到各国家网，国内互联细分各 ISP 网；CERNET 内细分到各省网。

2.5 测度结果存储维护

各测度计算模块的测度计算结果需存入数据库中进行维护。测度具体存放时保存三种数据状态：当前周期数据、周数据、年数据。为了便于数据管理和提高数据检索效率，NBOS 设计了“基本数据库”和“历史测度数据库”两类数据库，分别用于存放稳定的数据表格和动态增长的数据表格。其中，基本数据库名为 NBOS_Base，存放当前周期测度，便于首页快速显示。历史测度数据库按“年份_NBOS”命名，每年建一个年数据库，存放各测度的周数据表和年摘要数据表。

3 系统应用

作为新一代精细化网管系统，NBOS 能够管理更大范围的被管网络，提供更加丰富的网管内容，综合运用实时流量监测和历史趋势分析，有效结合宏观总体概括和微观细节描述。目前该系统正式运行于东南大学的 CERNET 华东北地区网络中心，基于 JSERNET 和 CERNET 主干网互联节点的 Cisco 路由器的 NetFlow 流数据，以及 JSERNET 和电信网互联节点 Juniper 路由器的 sFlow 流数据，实现对江苏各高校网络的实时流量监测和精细化网络管理。下面通过截取 NBOS 系统的部分运行界面，阐述该系统在日常网络管理，流量异常发现和安全事件检测方面的应用。

3.1 日常网络管理

图 2 是从 NBOS 系统首页截取的部分图片和表格，基于这些信息网络用户可以高效的进行日常网络管理。首先，系统不再局限于单个校园网或者企业网的管理，而是负责

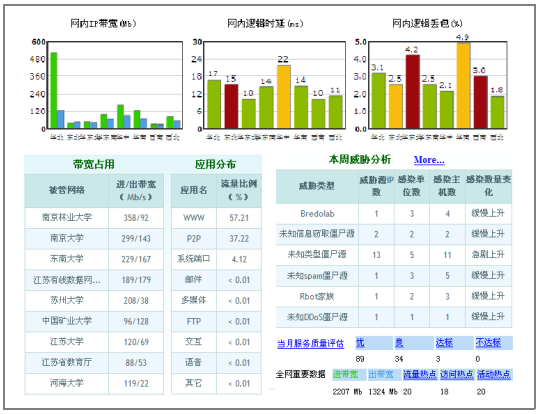

图 2 NBOS 系统首页部分截图

监管整个江苏教育网的所有单位。其次，系统提供了丰富的网管信息：基于提供的全网应用分布，用户可以了解网络应用的业务类型；通过被管网到教育网各地区宏观带宽、逻辑时延、逻辑丢包的计算，以及全网服务质量的评估，用户可以掌握整个网络当前性能状况，便于流量规划和高质量网络通信服务的开展；基于发现的热点主机，用户可以有目的监控服务器，以及方便地检测到因网络攻击而流量膨胀的主机通过查看网络安全威胁分析报告，用户不但可以知道网络存在的威胁和隐患，还可以准确定位攻击源，了解攻击范围和造成的影响。再次，该系统的实时流量监测能够让用户及时发现网络攻击事件，进一步的历史趋势分析又能为用户预测威胁的传播范围。最后，网管用户仅通过扫描一眼首页，就能了解整个网络的运行现状，只在出现问题时，用户才需点进去查看细节信息，有效的提供用户的工作效率。

3.2 异常流量检测

图 3 热点连接发现

NBOS 热点页面能够提供当前占全网总流量超过 1%的热点连接，如图 3 所示，2011 年 3 月 16 日 16 点 05 分，发现热点连接（南京大学 219.219.127.4/32，APNIC Debogon Project 202.38.3.253/32，未知服务，未知服务，ICMP）。通过了解 APNIC Debogon Project 是 APNIC 的一个和改善 AS 有关的项目，当时南京大学正在参与此项目。

3.3 安全事件检测

"CERNET华东(北)地区网络中心" 的安全状态

感染主机IP	威胁源IP	威胁源类型	威胁源子网归属	首次感染时间	最近感染时间
219[redacted].66	130.237.188.216	Bredolab	瑞典	2012-07-19 03:12:56	2012-09-11 16:45:38

图 4 安全事件发现

NBOS 安全分析页面能够提供各个被管单位的安全事件列表，如图 4 所示，2012 年 7 月 19 日到 9 月 11 日期间，华东北网络中心服务器 219.XX.XX.66 受到瑞典主机 130.237.188.216 的 Bredolab 攻击。根据该攻击事件的详细信息，网络管理员在该主机上发现攻击源码并采取了保护措施。

4 结论

为了对高速主干网进行实时流量监测和精细化网络管理，本文围绕主干链路上流数据实时采集的性能问题以及精细化网管设计和实现的关键技术进行讨论。针对流数据实时采集面临的数据精度和性能问题，选用信息丰富的流数据作为数据源，通过采用多缓冲区和多线程并行化机制有效地提高数据的采集性能(20Gbps 的流量负载下，丢包率小于 10^{-6}，数据采集和预处理较小的内存开销，约 8MB-76MB)；此外，为满足精细化网管需求，本文提出了一种网络区域划分方法，根据 IP 地址所属行政单位和地理位置进行子网划分；然后基于网络区域划分，以适当的时间粒度为单位，通过合理的测度选择和计算，提供网络服务质量评估、热点报告及安全事件告警，让用户可以了解被管网络的运行状态和需要被关注的信息，并提供网络安全态势报告。最后该系统成功运行于 CERNET 华东北地区网络中心，进行日常网络管理，以及发现流量异常和安全事件。

文献：

[1] http://www.splintered.net/sw/flow-tools/

[2] http://nfdump.sourceforge.net/

[3] ZHANG Weiwei, GONG Jian. NetFlow-Based Network Traffic Monitoring. APNOMS 2011

[4] http://www.cisco.com/en/US/docs/ios/solutions_docs/netflow/nfwhite.html

[5] http://www.huawei.com/en/

[6] http://www.juniper.net/us/en/

[7] 张晓宇,龚俭,吴桦. 一种基于 NetFlow 特定流记录的平均往返时延估计方法. 计算机应用与软件,2010,27(5)

[8] Hua Wu,Jian Gong. Packet Loss Estimation of TCP Flows Based on the Delayed ACK Mechanism. APNOMS 2009, LNCS 5787, pp. 540–543, 2009

NBOS： A Fine-Grained Network Management System

ZHANG Wei-Wei, GONG Jian, DING Wei, ZHANG Xiao-Guo

(School of Computer Science and Engineering, Southeast University, Nanjing Jiangsu 211189)

Abstract: In order to achieve real-time traffic monitoring and fine-grained network management on high-speed backbone link, this paper used flow sampling technology to collect data, designed and calculated metric to provide QoS evaluation, hot-spot report, network macroscopic running state and microscopic detail information, network security situation report and attack details. The

system has been deployed in CERNET network center of northeast China, to monitor and manage Jiangsu universities networks in the 20Gbps backbone network environment.
Keywords: backbone network, flow technology, traffic monitoring, network management, QoS, hotspot, network security