网络安全视角下全流量智能采集监测方案研究

汤迎春[1] 陈晓军[1] 张建光[1]
TANG Yingchun CHEN Xiaojun ZHANG Jianguang

摘 要 网络攻击手段日益复杂多变，对关键信息基础设施的安全防护提出了更高要求。网络态势感知与监测作为保障系统稳定运行的关键环节，其重要性不言而喻。针对关键信息基础设施网络安全监测中流量采集精度低、动态适应性不足的问题，文章提出了一种具备自适应性和协同特性的分布式全流量采集动态方法，旨在高效、准确地捕获和传输网络流量，适用于关键信息基础设施的网络安全监测。通过部署具备协同反馈机制的流量传感器集群，结合基于网络状态的动态捕获比率调整算法，实现流量采集效率与资源占用的平衡。研究表明，该方案在千兆级网络环境下，其多级部署架构可灵活适配关键信息基础设施的网络监测需求。

关键词 网络安全；流量镜像；分布式采集；自适应

doi：10.3969/j.issn.1672-9528.2025.04.028

0 引言

针对关键信息基础设施中安全相关数据采集面临的核心问题。首先，需要精确识别并收集安全相关数据，这在大数据环境下尤为挑战，因为安全相关数据具有5V特性[1-2]。其次，数据收集过程必须高效且准确，不能对关键信息基础设施的性能造成负担，同时随着基础设施的复杂性增加，设计这样的收集方案变得更加困难。再者，关键信息基础设施的异构性和高速性为数据收集带来了额外挑战，需要构建自适应、上下文感知的收集机制。最后，必须实施严格的安全保护措施，确保数据在收集、存储和访问过程中的安全。

随着关键信息基础设施网络规模扩大，传统抽样采集方法难以满足 APT 攻击溯源、零日漏洞检测等高精度需求。现有流量采集方案存在三方面矛盾：全流量存储与计算资源消耗、静态采集策略与动态网络负载、单点采集盲区与分布式协同效率。本文设计一种支持动态反馈、资源感知的分布式流量采集架构，解决流量过载丢包、攻击特征漏检等核心问题。首先简要回顾了流量采集领域的相关研究工作。其次，深入探讨了全流量采集系统的实际应用场景和关键要素，并在此基础上详细阐述了本文所提出的方法。最后，对研究结果进行了全面总结。

1 流量采集设计要求

1.1 功能性和安全性

安全数据采集器的核心功能需涵盖数据智能识别、异构网络适配与轻量化资源管理，通过机器学习算法动态筛选目标数据类型（如流量特征、异常行为），并基于容器化技术实现跨平台部署[3-4]。安全性设计需采用分层防御策略：传输层使用 TLS/SSL 加密防止窃听，存储层通过 SHA-256 哈希校验保障数据完整性，访问层实施 RBAC 权限控制与双因素认证，同时引入实时审计日志追踪数据流向，构建全生命周期防护体系。

1.2 采集节点

网络数据采集节点按功能分为四类。

（1）传输节点（如核心路由器、边缘交换机），利用 NetFlow/sFlow 协议镜像流量[5]。

（2）安全节点（如下一代防火墙、AI 驱动的 IDS），深度解析 DDoS 攻击、APT 攻击等威胁数据[6]。

（3）专用采集节点（如 Prometheus 监控服务器、ELK 日志代理），实现应用层数据抓取与预处理[7]。

（4）终端节点（移动设备、IoT 传感器），通过 SDK/Agent 采集端侧行为数据，形成分布式数据源矩阵[8]。

1.3 采集工具

硬件工具（如 FPGA 加速的 DPDK 抓包设备）提供 10 Gbit/s+ 高吞吐采集能力，但存在部署僵化问题；软件工具（如 Wireshark、tcpdump）支持动态规则配置，却受限于操作系统内核性能；协议工具（如 SNMP Trap、IPFIX）依赖标准化数据模板，难以覆盖私有协议场景。通用型采集器需融合软硬件优势，采用 eBPF 技术实现内核级数据过滤，并通过插件架构兼容 OpenConfig/YANG 模型，最终达成性能、成本与灵活性的三角平衡[9-10]。

1. 国家计算机网络应急技术处理协调中心安徽分中心
安徽合肥 230041

2 流量镜像

流量镜像技术通过复制物理服务器或网络设备的接口流量，实现关键数据监测。其核心功能包括流量过滤、数据包截断及跨网段转发，可精准筛选安全相关流量并转发至监测设备，提升分析效率[11]。部署时需定义源设备（如 A、B）与目标设备（D），并在网络设备（交换机 / 路由器）上配置镜像会话，如图 1 所示。

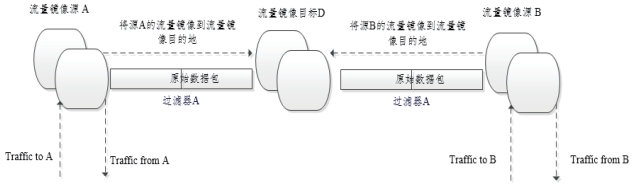

图 1 镜像流量示例

确定源设备 A：选定用于流量镜像的物理服务器或网络设备 A，其网络接口流量将被复制。

确定源设备 B：选定另一台用于流量镜像的物理服务器或网络设备 B。

配置目标设备 D：配置一台物理服务器、网络设备或安全监测设备 D，用于接收来自源设备 A 和 B 的镜像流量。设置流量过滤器：在物理网络中，通过交换机、路由器等网络设备配置流量镜像规则，筛选特定流量。

配置源设备 A 的流量镜像会话：在交换机等设备上配置会话，指定源设备 A 网络接口、过滤器条件和目标设备 D 接收接口，使符合条件的流量从 A 复制到 D。

配置源设备 B 的流量镜像会话：重复上一步，为源设备 B 配置会话，将符合条件的流量从 B 复制到 D。

当会话创建后，任何满足过滤器 A 中配置规则的流量都会被封装在以太网帧头部，并发送至目标。

3 全流量镜像采集方案

关键信息基础设施（关基）系统的网络规模庞大、结构复杂，随着云计算、软件定义网络（SDN）等新技术的应用，设计一个分布式、动态、灵活且可扩展，专注于数据采集和传输的系统很有必要[12]。

本文聚焦于高效、准确地捕获和传输网络流量，不涉及数据的进一步分析。当然，可将其他分析程序或系统与本文研究集成，如集成 Suricata 作为入侵检测和防御系统（IDS/IPS）。

镜像流量技术是网络监控与分析的关键手段，能按实际任务需求灵活应用于多种场景。在服务器入口或数据中心出口等单个关键位置，配置网络设备的镜像功能，可将流量复制至专门的镜像流量传感器或分析工具，实现对流量的深度剖析与实时监控，且不干预原始网络流量，保障了分析的快速、灵活和功能丰富[13]。但单个镜像点存在资源限制，难以全面捕获所有重要流量。因此，在网络多个关键位置设置镜像点，将流量复制至多个传感器协同作业，提升了流量捕获的完整度与精确度，克服了资源限制问题。

针对资源受限的挑战，本文提出自适应流量采样技术。该技术可依据当前网络负载和传感器性能，动态调整采样率，在流量高峰时减轻传感器压力，流量低谷时捕获更多细节信息。融入镜像流量分析后，在资源有限的情况下也能最大限度地捕获并分析流量。

为降低静态协作镜像点分析可能导致的数据包捕获非确定性，设计了自适应协作数据包捕获方法。正常网络状态下，由一个协作组的单个镜像流量传感器捕获所有数据；当传感器接近过载阈值时，系统动态启用其他传感器参与捕获，防止过载，减少数据在多个传感器上的分散，简化后续分析流程。

考虑到数据包捕获和转发对 CPU、内存及网络吞吐量的资源消耗，构建了动态可配置系统。该系统能根据网络实际情况，自动调整采样率、捕获时长等关键参数，实现资源优化配置与高效利用。应用于镜像流量分析，既保持了监控高效性，又最大限度减轻了对网络性能的影响。

综上所述，镜像流量技术通过灵活的镜像点配置、高效的协作分析、自适应的采样策略和资源优化配置，为网络监控与分析提供了全面、高效、可靠的解决方案，满足不同场景的多样化需求。

3.1 采集架构

镜像流量系统由分布式传感器、通信枢纽（CI）及智能处理模块构成多层次架构。远端传感器实时捕获并加密回传流量至 CI 中枢，合并器将多源数据流整合为时序统一队列。系统通过可视化界面实现集中监控与指令下发，内置自适应控制器基于网络状态、负载均衡动态调优运行参数。数据经 CI 处理后通过先进先出（FIFO）管道推送至命名管道，支持外部系统无损对接，形成集采集、聚合、调控、输出于一体的智能流量分析平台，兼具实时性、扩展性和跨平台兼容性。

通过各类传感器捕获网络流量，利用稳健的 CI 进行数据传输，通过合并器整合数据流，前端界面提供直观控制，自适应控制器实现智能调整，以及 FIFO 输出机制促进外部集成，共同构建了一个全面、高效、灵活的镜像流量监控与分析系统。

3.2 协作流量捕获

通过联合和协作的方式使用多个传感器来提高捕获性能的一般方法，是捕获流量的不相交子集 C_k。

$$C_k \cap C_l = \emptyset,\ 1 \le k < \mathrm{l} \le n \tag{1}$$

在每个传感器（从 1 到 n）上捕获流量的不相交子集，使得这些捕获的子集的并集。

$$T = \cup_{k=1}^{n} C_k \tag{2}$$

各传感器按预设规则采集互斥流量子集 $C_1, C_2, \cdots, C_n$，其并集覆盖总流量 T。基于包头字段（如源 IP 哈希值）对 2^x 取模实现流量划分（$x \leq$ 字段位数）[14]，利用 2 的幂特性简化过滤器规则生成（二进制掩码 AND 运算替代复杂计算）。当单节点触发带宽配额告警时，系统执行动态重分配：（1）本地降级，停止捕获预设低优先级流量（P2P 等），对高优先级流量（HTTP 等）启动采样；（2）协同接管，合并器基于最新流量特征生成新哈希参数（如调整 x 值扩展剩余类数量），通知空闲节点接管溢出流量子集；（3）状态同步，新参数通过 CI 实时下发至全部传感器，确保各节点持续捕获互斥子集。

3.3 自适应实现机制

为实现自适应性，采用了一个与文献 [15] 类似的反馈循环。通过算法展示了单个传感器自适应性的反馈循环执行方法。在协作场景中，自适应性的工作原理类似，但不同之处在于，会协调地使用多个反馈循环实例，每个传感器对应一个。

监控数据以可配置的固定间隔发出。监控数据包含诸如正在捕获、传输或丢弃的数据包数量等信息。

使用最大发送速率（MSR）来确定传感器的最大性能。MSR 描述了传感器每秒可以捕获并发送到系统中的数据包数量。在运行时动态地确定 MSR，采用一种能够检测丢弃并根据需要调整 MSR 的策略，以避免在后续操作中出现丢弃。最初，没有设置 MSR。为确定 MSR，也可以使用其他动态策略或静态配置。

伪代码如下：

```
算法 adaptive_sensor_system(monitoring_data)
    输入 : monitoring_data ( 监测数据 )
    输出 : max_send_rate ( 最大发送速率 ), capture_ratio ( 捕获比率 ), filter_rules ( 过滤规则 )
    // 步骤 1：初始化
    monitor_sensor = 初始化 MonitorSensor 实例
    max_send_rate_calculator = 初始化 MaxSendRateCalculator 实例
    capture_ratio_calculator = 初始化 CaptureRatioCalculator 实例
    pcap_filter_generator = 初始化 PcapFilterGenerator 实例
    // 步骤 2：设置监测数据调用 monitor_sensor.set_monitoring_data(monitoring_data)
    // 步骤 3：计算最大发送速率
    max_send_rate = 调用 max_send_rate_calculator.calculate_max_send_rate(monitoring_data)
    // 步骤 4：计算捕获比率
    capture_ratio = 调用 capture_ratio_calculator.calculate_capture_ratio(monitoring_data)
    // 步骤 5：生成过滤规则
    filter_rules = 调用 pcap_filter_generator.generate_filter_rules(capture_ratio)
    // 步骤 6：返回结果
    返回 (max_send_rate, capture_ratio, filter_rules)
结束算法
```

4 典型应用

全流量采集监测安全事件方案在某个应用系统，对于具有跨省或地市级分支应用系统的场景具有应用价值，如图 2 所示。

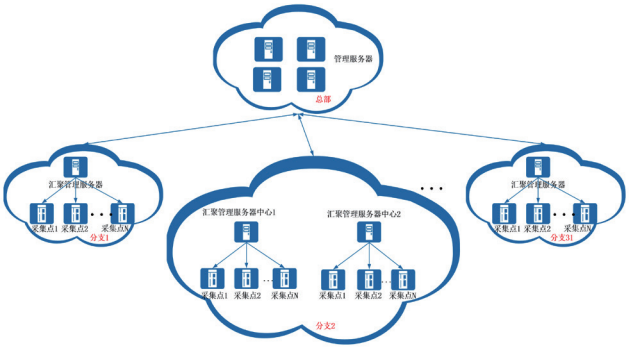

图 2 有分支的应用系统典型网络场景

全流量采集技术虽然能够捕获全面的流量数据，但通常这些数据的流量规模较大。若直接将采集到的数据通过网络传输至主管理机，会给网络带宽带来沉重负担，可能导致网络拥塞。因此，在面对大型被采系统时，通常采用多级部署的方式来实施全流量采集技术的采集监测软件。具体而言，第一级部署负责将采集到的全流量数据传输至汇聚服务器，在那里进行初步筛选和重组，以减少数据量并优化传输效率。随后，经过处理的数据再通过网络传输至第二级主管理服务器进行进一步处理。而对于规模较小的被采系统，由于其流量相对较小，可以简化部署流程，仅采用单级部署方式即可满足需求。此外，全流量采集技术还支持横向扩展，采用开放式架构，能够随着被采集点的增加灵活地扩展系统规模，如图 3 所示。

（1）交换机镜像：通过配置交换机实现 N:1 端口流量复制，支持商业设备原生镜像功能或分光器物理旁路部署，需规避双向端口同时镜像导致的数据冗余。

（2）流量采集服务器：基于多网卡架构实现动态速率适配与资源优化，通过多传感器协同机制划分数据子集，保障数据完整性与接口标准化。

（3）汇聚服务器：作为数据处理的核心，采用动态负载均衡与高效聚合算法整合多源数据流，支持并行化处理满足秒级查询需求。

（4）主管理服务器：集成 Web 智能管理平台，支持参

数动态调优，统一管理各个组件，协调任务分发和调度，提升系统响应与运行稳定性。

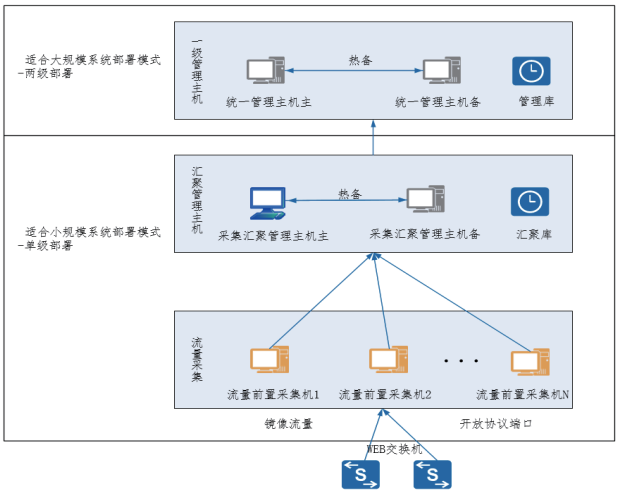

图 3 软件部署架构

5 结论

全流量采集监测技术是一种高效、全面且易于部署的网络安全监测方案，适用于客户服务层的安全事件监测。通过该技术，企业可以实现对 Web 流量的全面监测和分析，为网络安全防护提供有力的技术支持。

全流量采集监测安全事件方案在等保三级以上应用系统这一典型场景中的应用具有实际价值。通过实施该方案，可以有效提升这一类系统的网络安全防护能力，降低安全风险。同时，也需要关注数据处理量、隐私保护和技术更新等挑战，并采取相应的应对策略，以确保方案的可持续性和有效性。

参考文献：

[1] BEJTLICH R. The practice of network security monitoring understanding incident detection and response[J]. The practice of network security monitoring: understanding incident detection and response,2013(8):376.

[2] HAN L C, GUO Z C , HUANG X Y ,et al.A multifunctional full-packet capture and network measurement system supporting nanosecond timestamp and real-time analysis[J/OL].IEEE transactions on instrumentation and measurement,2021[2024-06-19].https://ieeexplore.ieee.org/document/9431234.DOI:10.1109/TIM.2021.3080375.

[3] BUMGARDNER V K N, MAREK V W. Scalable hybrid stream and hadoop network analysis system[C]//Proceedings of the 5th ACM/SPEC international conference on Performance. NewYork: ACM, 2014:219-224.

[4] MUNZ G, CARLE G. Distributed network analysis using TOPAS andwireshark[C/OL]//NOMS Workshops 2008-IEEE Network Operations and Management Symposium Workshops.Piscataway:IEEE,2008[2024-09-01].https://ieeexplore.ieee.org/abstract/document/4509943.DOI:10.1109/NOMSW.2007.27

[5] CALYAM P, DHANAPALAN M, SRIDHARAN M,et al.Topology-aware correlated network anomaly event detection anddiagnosis[J]. Journal of network and systems management, 2014,22:208-234.

[6] HOFSTEDE R, CELEDA P, TRAMMELL B, et al.Flow monitoring explained: from packet capture to data analysis with netflow and IPFIX[J].IEEE communications surveys & tutorials, 2014,16(4):2037-2064.

[7] OTOMO K, KOBAYASHI S, FUKUDA K,et al.Finding anomalies in network system logs with latent variables[C]// Proceedings of the 2018 Workshop on Big Data Analytics and Machine Learning for Data Communication Networks. NewYork: ACM, 2018: 8-14.

[8] SHERRY J, LAN C, POPA R A, et al. BlindBox: deep packet inspection over encrypted traffic[J].ACM SIGCOMM computer communication review, 2015,45(4):213-226.

[9] 朱俊芳 , 李彦泽 , 郭超 , 等 . 面向全流量的网络安全大数据系统研究 [J]. 网络空间安全 , 2023,14(5):10-16.

[10] 马超 , 张典 , 石小川 . 基于旁路技术的工控网络流量分析安全检测系统及方法 :CN202211636451.5[P].2023-01-20.

[11] 黄超 . 基于深度学习的网络流量异常检测研究 [D]. 上海 : 华东师范大学 ,2022.

[12] 马亚洲 , 龚俭 , 杨望 . 面向应急响应的高速网络流量采集设计与实现 [J]. 通信学报 ,2014:46-51.

[13] 王俊健 , 冯诚波 . 基于日志和流量采集的网络系统事件溯源方法及系统 :CN201910494641.X[P]. 2019-10-29.

[14] GAD R, KAPPES M, MUELLER-BADY R, et al.Header field based partitioning of network traffic for distributed packet capturing and processing[C]//2014 IEEE 28th International Conference on Advanced Information Networking and Applications (AINA). Piscataway:IEEE, 2014:866-874.

[15] BRUN Y, DESMARAIS R, GEIHS K,et al.A design space for self-adaptive systems[J]. Software engineering for self-adaptive systems II, 2013,7475:33-50.

【作者简介】

汤迎春（1985—），男，硕士研究生，工程师，研究方向：信息系统集成与运维、系统安全保障技术等。

陈晓军（1987—），男，硕士研究生，高级工程师，研究方向：信息系统集成与运维、系统安全保障技术等。

张建光（1984—），男，硕士研究生，工程师，研究方向：信息系统集成与运维、系统安全保障技术等。

（收稿日期：2024-12-04）