Mastering Network Flow Traffic Analysis

Implementing and analyzing flow data across network topologies for threat detection

Gilberto Persico

www.bpbonline.com

First Edition 2025

ISBN: 978-93-65890-266

LIMITS OF LIABILITY AND DISCLAIMER OF WARRANTY

The information contained in this book is true and correct to the best of author's and publisher's knowledge. The author has made every effort to ensure the accuracy of these publications, but the publisher cannot be held responsible for any loss or damage arising from any information in this book.

All trademarks referred to in the book are acknowledged as properties of their respective owners but BPB Publications cannot guarantee the accuracy of this information.

Dedicated to

*My partner **Tin**, my daughter **Flora**,*
*my father **Carlo** and my mother **Grazyna***

About the Author

Gilberto Persico is a Unix system, networking, and security engineer with over 30 years of experience in the IT world, working as a programmer, architect, security auditor, and systems and network engineer. He worked for IBM, Sun Microsystems, Oracle, and Huawei both as an employee and as a freelancer, designing, developing, and supporting production-ready enterprise-grade architectures for dozens of very important customers. He also conceived, designed, and developed Fl0wer, a new generation network flow analysis product, and deployed it in a big setup, successfully controlling two big data centers. He plays cello in his free time and raises a daughter when not hacking things in his lab. He also loves resurrecting old systems, retro-computing, and is currently works as a NOC team leader in Econocom.

About the Reviewer

Md Nahidul Kibria is currently staff engineer at HelloFresh. With over a decade of experience in software development and cloud infrastructure, he specializes in migrating legacy systems to cloud-native environments. He focuses on improving the synergy between DevOps and SecOps processes and enhancing infrastructure scalability, data streaming technologies, and security.

He has worked with companies of various sizes, designing and developing microservices-based platforms, implementing service mesh strategies, and leading cloud migration initiatives. His expertise includes cloud technologies such as AWS and Kubernetes, infrastructure as code, data streaming technologies, and application security.

He is also an active member of the global cybersecurity community, serving as a red team member and community lead, and has presented at prestigious conferences. He believes that learning is a lifelong journey and enjoys sharing his insights through writing and public speaking on topics such as data streaming, application scaling, and advanced threat hunting.

He holds a bachelor's degree in computer science and is passionate about building resilient, scalable, and secure systems. Outside work, he enjoys exploring emerging technologies and finding innovative ways to simplify and enhance complex operations.

Acknowledgement

I want to express my deepest gratitude to my family and friends for their unwavering support and encouragement throughout this book's writing, especially my partner Tin and my daughter Flora. I want to thank my father for teaching me the meaning of patience and determination, and my mother for helping me in my darkest moments. I love you all.

I am also grateful to BPB Publications for their guidance and expertise in bringing this book to fruition. It was a long journey of revising this book, with valuable participation and collaboration of reviewers, technical experts, and editors.

I would also like to acknowledge the valuable contributions of my colleagues and co-worker during many years working in the tech industry, who have taught me so much and provided valuable feedback on my work.

Finally, I would like to thank all the readers who have taken an interest in my book and for their support in making it a reality. Your encouragement has been invaluable.

Preface

Managing the enterprise network security is a complex task that requires a comprehensive understanding of the latest technologies. On one side, passive network traffic analysis still makes sense for several reasons, but what is going to provide a more scalable approach to network security is the analysis of network traffic flows.

The book aims to familiarize the readers with network traffic flows analysis technologies, giving a deep understanding on the difference between active and passive network traffic analysis, the advantages and disadvantages of each methodology, with a special focus about network flow traffic analysis, which due to its scalability, privacy, ease of implementation and effectiveness, is beginning to play a leader role in the field of network security. The book allows a reader to dive deep into tools and technologies that can be used and leveraged to effectively deploy a scalable and affordable network monitoring solution capable of giving a clear idea of all internal traffic flows and providing an effective, almost-real-time data breach detection mechanism.

Throughout the book, you will learn how common network infrastructures are built, how the flow protocols work, what kind of data is managed, and how you can effectively take advantage of it.

The book targets professionals with job roles such as incident responder, forensic investigator, SOC analyst, network administrator, or a student seeking to extend their knowledge on network flow analysis. The book assumes readers know network topologies, the OSI and TCP/IP models, and have a basic understanding of capturing network data. The book heavily relies on Linux knowledge.

The reader will learn to set up their infrastructure to obtain flow traffic data and make the most of this information. The reader will also learn to understand how to relate to normal and unknown traffic, how to set up tasks to automate network controls, and how to assess their network in a passive but proactive way. The reader will acquire the knowledge and skills that will allow them to be untied by limitations of traditional analysis tools and embrace a new and scalable way to improve security on high-speed networks.

I hope you will find this book informative and helpful.

Chapter 1: Foundation of Network Flow Analysis - This chapter lays the base for conducting network analysis using flow protocols, with a strong bias towards network security. The reader will learn about the essentials, important concepts of types of network analysis types, advantages, scalability, and sustainability of different types of analysis. Additionally, in this chapter, the reader will learn about the proper tools to get effective results in each of the different analysis types, focusing on the network flow analysis. The reader will familiarize himself with the differences between packet and flow analysis while learning the advantages and disadvantages of statistical flow analysis.

Chapter 2: Fixed and Dynamic Length Flow Protocols - This chapter will discuss both the fixed length flow protocols and the dynamic length flow protocols, their advantages, and drawbacks. The chapter describes NetFlow v1, NetFlow v5, NetFlow v9, sFlow v5, and IPFIX. By the end, the chapter will discuss case studies on the protocols' benefits or misuse and their identification.

Chapter 3: Network Topologies - This chapter primarily focuses on various network topologies found in companies and ways to implement proper flow analysis in different contexts, from classical flat infrastructure to frontend/backend/DMZ to Virtual Private Clouds and ways to discover blind points.

Chapter 4: Implementing Flow Export on Layer 2 Devices - This chapter will guide the reader to implement flow data export on the most widespread Layer 2 devices (switches and access points) from most vendors on the market, and will also describe a solution to get NetFlow/IPFIX data from a switch using port mirroring.

Chapter 5: Implementing Flow Export on Layer 3 Devices - This chapter will guide the reader to implement flow data export on the most widespread Layer 3 devices like firewalls, routers, load balancers, and wireless gateways from most vendors on the market.

Chapter 6: Implementing Flow Export on Servers - This chapter focuses on solutions for implementing flow export on servers, which may be required in contexts where you want to see the flow traffic but cannot manage network infrastructure, like cloud environments or hosting services.

Chapter 7: Implementing Flow Export on Virtualization Platforms - This chapter focuses on solutions for implementing flow export on virtualization systems like VMware and Proxmox, which can give you network visibility in traffic not crossing the network infrastructure (imagine traffic between different virtual machines on the same hypervisor).

Chapter 8: Ingesting Data into Clickhouse and Elasticsearch - This chapter shows the user how to ingest raw flow data into more usable and structured analysis platforms like Elasticsearch and Clickhouse (open-source high-performance OLAP).

Chapter 9: Flow Data Analysis: Exploring Data for Fun and Profit - This chapter will discuss how we can do interesting analysis of the flow data we are getting from the network, and will teach the reader to understand better what is happening inside their network infrastructure, by showing a lot of examples. It will also give the reader further in-depth knowledge about identifying patterns and anomalies, and how to detect security threats.

Chapter 10: Understanding the Flow Matrix - This chapter introduces an often too underestimated concept, the matrix of flows happening inside the company network. A deep dive into the concept will allow the reader to take advantage of it to improve the security posture of the whole network.

Chapter 11: Firewall Rules Optimization Use Case - This chapter describes a real use case of NetFlow data to approach a quite complex problem of firewall optimization rules in a complex (but now becoming quite common) environment.

Chapter 12: Simple Network Anomaly Detection System Based on Flow Data Analysis - This chapter focuses on how to identify network anomalies and data breaches by using the flow matrix and some Python scripting using Pandas. It will show the reader how to automate continuous checking, trying to address the problem of the slowness in identifying a breach.

Code Bundle and Coloured Images

Please follow the link to download the
Code Bundle and the ***Coloured Images*** of the book:

https://rebrand.ly/46sqwki

The code bundle for the book is also hosted on GitHub at
https://github.com/bpbpublications/Mastering-Network-Flow-Traffic-Analysis.
In case there's an update to the code, it will be updated on the existing GitHub repository.

We have code bundles from our rich catalogue of books and videos available at **https://github.com/bpbpublications**. Check them out!

Errata

We take immense pride in our work at BPB Publications and follow best practices to ensure the accuracy of our content to provide with an indulging reading experience to our subscribers. Our readers are our mirrors, and we use their inputs to reflect and improve upon human errors, if any, that may have occurred during the publishing processes involved. To let us maintain the quality and help us reach out to any readers who might be having difficulties due to any unforeseen errors, please write to us at :

errata@bpbonline.com

Your support, suggestions and feedbacks are highly appreciated by the BPB Publications' Family.

Piracy

If you come across any illegal copies of our works in any form on the Internet, we would be grateful if you would provide us with the location address or website name. Please contact us at **business@bpbonline.com** with a link to the material.

If you are interested in becoming an author

If there is a topic that you have expertise in, and you are interested in either writing or contributing to a book, please visit **www.bpbonline.com**. We have worked with thousands of developers and tech professionals, just like you, to help them share their insights with the global tech community. You can make a general application, apply for a specific hot topic that we are recruiting an author for, or submit your own idea.

Reviews

Please leave a review. Once you have read and used this book, why not leave a review on the site that you purchased it from? Potential readers can then see and use your unbiased opinion to make purchase decisions. We at BPB can understand what you think about our products, and our authors can see your feedback on their book. Thank you!

For more information about BPB, please visit **www.bpbonline.com**.

Join our book's Discord space

Join the book's Discord Workspace for Latest updates, Offers, Tech happenings around the world, New Release and Sessions with the Authors:

https://discord.bpbonline.com

Table of Contents

Chapter 1

Foundation of Network Flow Analysis

Introduction

This chapter introduces you to network flow analysis, ways of performing it, techniques and technologies involved in network security, and their advantages. Nowadays, technology is playing a pivotal role in the success of most companies. Moreover, computer-based technology relies on a strong foundation: computer networks. A famous computer slogan used by *Sun Microsystems*[1] once said *The network is the computer*. When you browse the Internet, your computer uses the company's network to connect to various websites. Maybe you can reach only some websites, in which case it means that some sort of network control is already in place, or maybe you can go anywhere (which does not suggest that controls are not in place). In any case, if you are in your office, you are connected to the company's internal network unless you are doing smart work. A network is simply made by computers with network interfaces, cables, and networking devices (like hubs, switches, routers, firewalls, etc.), each performing exactly the task they were developed for.

But what is network security? Simple, with the advent of computer networks first and the Internet later, protecting internal data and users from theft and fraud has become increasingly complicated than before. And in current times, if you read about network security news, there is always a company that has been hacked or was the victim of malware or data theft.

1. https://spectrum.ieee.org/does-repurposing-of-sun-microsystems-slogan-honor-history

The interesting fact is that normally, everyone assumes that the internal network of the company is a safe place without risks or menaces. It is no surprise that searching zero trust online yields numerous articles; we will get back to it in the following chapters, but let us just assume that to keep a network perimeter safe, you must work on it in some way!

Structure

In this chapter we will discuss the following topics:

- Computer network
- Computer network analysis
- Common network security threats to company networks
- Network security traffic analysis
- Techniques for performing network security traffic analysis
- Packet-inspection network traffic analysis
- Network flow-based traffic analysis
- Basics of network protection
- Firewalls and packet filters
- Network proxies
- Intrusion detection systems
- Intrusion prevention systems
- Pros and cons of packet-inspection network traffic analysis
- Open-source and commercial solutions
- Pros and cons of network flow-based traffic analysis
- Open-source and commercial solutions
- Challenge of analyzing 800Gbps networks

Objectives

This chapter will introduce the user to the world of corporate network security, corporate networking, and the different types of network traffic analysis, as well as introduce the network flow traffic analysis.

Computer network

A computer network is a collection of interconnected computers and devices that can communicate with each other, share resources, and exchange data. These networks can be as small as a few devices at home or office or as large as the global Internet, connecting billions of devices worldwide. Computer networks enable the sharing of information, resources, and services, facilitating communication and collaboration between users and systems.

Usually, depending on the size of the company, inside networks are split and deployed in different ways, both on the physical and logical levels. Normally there are distinct designs for logical and physical, because modern network devices allow this split distinction. This has several benefits from the perspective of security and availability of the network service.

An example of a simple logical split can be in terms of frontend, backend, and employee networks. These networks can be split into Layer 2 (switches or VLANs) and interconnected by routers or firewalls in Layer 3.

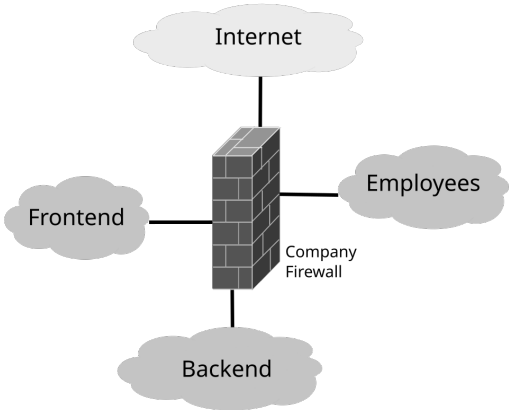

***Figure 1.1**: Logical network view*

The same network infrastructure, considered on the physical design, can have different ways to be deployed. In our example, implementing full redundancy (using proper protocols and configurations) can be as presented in *Figure 1.2*:

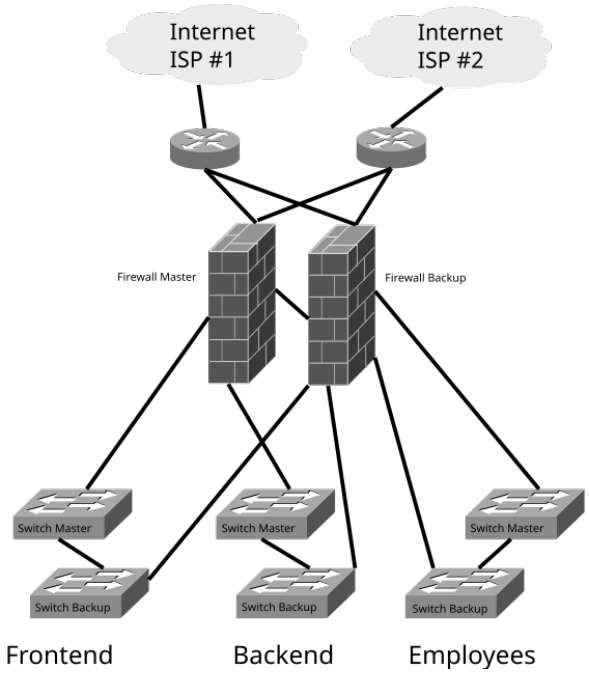

***Figure 1.2**: Physical network view*

Computer network analysis

Computer network analysis refers to examining and evaluating computer networks to understand their performance, security, efficiency, and overall functionality. It involves various techniques and tools to gain insights into network behavior and make informed decisions about network design, optimization, troubleshooting, and security.

Here are some key aspects of computer network analysis:

- **Performance monitoring**: Network administrators and analysts use various monitoring tools to track the performance of a network. This includes measuring bandwidth utilization, latency, packet loss, and network throughput. Performance analysis helps in identifying bottlenecks and optimizing network resources.
- **Security analysis**: Network analysis is crucial for identifying and mitigating security threats. It involves monitoring network traffic for suspicious activities, such as intrusion attempts, malware infections, and unauthorized access. Security analysts use **intrusion detection systems** (**IDS**), firewalls, and other security tools to analyze network traffic patterns and detect anomalies.
- **Troubleshooting**: When network issues occur, network analysis is used to diagnose and resolve problems. By examining network traffic, logs, and configuration settings, administrators can pinpoint the root causes of network

outages, connectivity problems, or performance issues. This process is essential for maintaining network reliability.

- **Optimization**: Network analysis helps in optimizing network resources and configurations. By studying traffic patterns and usage data, administrators can make informed decisions about network design, capacity planning, and load balancing. This ensures that the network operates efficiently and cost-effectively.
- **Capacity planning**: Analyzing network usage trends over time helps forecast future capacity requirements. This is important for ensuring that the network can handle increased traffic and new applications without degradation in performance.
- **Protocol analysis**: Network analysts often use packet sniffers and protocol analyzers to capture and analyze network traffic at a granular level. This is useful for diagnosing protocol-related issues and ensuring that network protocols are functioning as expected.
- **QoS analysis**: **Quality of service** (**QoS**) analysis involves assessing the network's ability to deliver different types of traffic with varying levels of priority. This is crucial for ensuring that real-time applications like voice and video conferencing receive the bandwidth and low latency to perform well.
- **Traffic engineering**: Network analysis is used to optimize traffic routing and distribution within a network. This is particularly important in large-scale networks to balance traffic loads and minimize congestion.
- **Network visualization**: Visualization tools and techniques are often used to represent network data graphically. This helps network administrators and analysts better understand network topologies, traffic flows, and dependencies.
- **Compliance and auditing**: Network analysis is also important for ensuring that a network complies with regulatory requirements and internal policies. It helps in auditing network activity and maintaining compliance records.

In summary, computer network analysis is a multidisciplinary field that involves the use of various tools and methodologies to gain insights into the performance, security, and efficiency of computer networks. It plays a crucial role in maintaining the reliability and integrity of modern networks in an ever-evolving technological landscape. In this book, we are focusing on the network security topics.

Common network security threats to company networks

Common network security threats, often referred to as cybersecurity threats or menaces, pose significant risks to computer networks and the data they contain. These threats can lead to data breaches, financial losses, and reputational damage for organizations. Here are some of the most common network security threats: