VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
FACULTY OF INFORMATION TECHNOLOGY

ÚSTAV INFORMAČNÍCH SYSTÉMŮ
DEPARTMENT OF INFORMATION SYSTEMS

PCAP NETFLOW V5 EXPORTÉR

SEMESTRÁLNÍ PROJEKT
TERM PROJECT

AUTOR PRÁCE **MAKSIM DUBROVIN**
AUTHOR

VEDOUCÍ PRÁCE **Ing. KAMIL JEŘÁBEK, Ph.D.**
SUPERVISOR

BRNO 2024

Abstrakt

Projektový popis programu, který převádí soubory zachycených paketů ve formátu PCAP do jednosměrné sekvence paketů formátu NetFlow v5 a poté odesílá tyto pakety do kolektoru.

Abstract

Project description of a program that converts captured PCAP packet files into a unidirectional sequence of NetFlow v5 packets and then sends these packets to a collector.

Klíčová slova

Síť, Správa sítí, Sledování sítí, PCAP, NetFlow v5, Exportér, Programovací jazyk C.

Keywords

Network, Network Management, Network Monitoring, PCAP, NetFlow v5, Exporter, C Programming Language.

Citace

DUBROVIN, Maksim. *PCAP NetFlow v5 exportér*. Brno, 2024. Semestrální projekt. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Ing. Kamil Jeřábek, Ph.D.

Obsah

Kapitola 1

Úvod

Projekt „PCAP NetFlow v5 exportér“ měl za cíl vytvořit program s názvem p2nprobe, který extrahuje informace o síťových tocích ze souboru PCAP a odesílá je kolektoru ve formátu NetFlow v5. Nástroj čte pakety ze souboru PCAP, který je zadán jako argument příkazového řádku, a agreguje je do jednotlivých toků. Tyto toky následně odesílá protokolem UDP do kolektoru NetFlow v5, kde jsou přijaty a dále zpracovány. Moje úloha spočívala v implementaci exportéru, přičemž jsem se zaměřil výhradně na export TCP toků, zatímco ostatní typy provozu byly ignorovány.

Kapitola 2

Popis řešení projektu

2.1 Základní informace o programu

Program byl vytvořen v programovacím jazyce C s využitím knihovny libpcap.

2.2 Návrh aplikace

V rámci tohoto projektu jsem strukturoval provádění programu do několika hlavních kroků:

1. Inicializace a zpracování parametrů zadaných v příkazovém řádku.
2. Čtení a analýza souboru s paketovými daty a jejich seskupení do jednotlivých toků.
3. Export zpracovaných toků ve formátu NetFlow v5.

Algoritmus vyžaduje specifické datové struktury pro správu informací o toku, jako je například dynamicky alokované pole toků.

2.3 Popis implementace

Pro implementaci hlavních kroků programu byly vytvořeny veřejné funkce v jednotlivých modulech. Každá funkce řeší specifickou část zpracování dat, od získání informací z příkazové řádky až po export toků v požadovaném formátu.

get_exporter_info_from_cla

Tato funkce zpracovává argumenty zadané v příkazovém řádku. Ukládá přijaté parametry do struktury typu exporter_info, která obsahuje informace potřebné pro export toků (např. IP adresu a port kolektoru). Funkce ověřuje, zda jsou zadány všechny požadované argumenty, zda mají správný formát, a zajišťuje inicializaci nastavení exportéru podle zadání uživatele. V případě chybného formátu nebo chybějících argumentů funkce vrátí chybový kód.

create_flow_array

Funkce create_flow_array vytváří dynamické pole toků, známé jako flow_array. Toto pole slouží k uložení jednotlivých záznamů toků, které jsou postupně agregovány při zpracování paketů. Funkce se stará o alokaci paměti pro toto pole a připravuje datové struktury tak, aby bylo možné efektivně přidávat a vyhledávat toky během zpracování.

packet_handler

Funkce packet_handler slouží jako obslužná rutina pro každé nové zachycené paketové data, která jsou zpracována pomocí pcap_loop. Je zodpovědná za analýzu každého příchozího paketu, identifikaci jeho příslušnosti k existujícímu toku nebo vytvoření nového toku, pokud takový neexistuje. Tato funkce také aktualizuje statistiky toku, jako jsou počty paketů a přenesených bajtů, a přidává informace do dynamického pole flow_array.

send_netflow

Funkce send_netflow zajišťuje odesílání zpracovaných toků ve formátu NetFlow v5 na kolektor. Tato funkce připraví zprávu v požadovaném formátu, včetně hlaviček a seskupených toků, a odesílá ji pomocí protokolu UDP na zadaný kolektor. Funkce je navržena tak, aby odesílala data spolehlivě a v odpovídajících časových intervalech podle požadavků formátu NetFlow.

Kapitola 3

Vysledný program

3.1 Spuštění programu

./p2nprobe <host>:<port> <pcap_file_path> [-a <active_timeout> -i <inactive_timeout>]

Pořadí parametrů je libovolné. Popis parametrů:

- <pcap_file_path> - zde bude uvedena cesta k souboru PCAP, který se má zpracovat;
- <host> - IP adresa nebo doménové jméno kolektoru;
- <port> - port kolektoru, kam budou zprávy odesílány;
- -a <active_timout> - počet sekund pro nastavení aktivního timeoutu exportu flow (defaultní hodnota při nepoužití parametru 60);
- -i <inactive_timeout> - počet sekund pro nastavení inaktivního timoutu exportu flow (defaultní hodnota při nepoužití parametru 60).

3.2 Testování a výsledky testů

Testování programu bylo provedeno pomocí nástroje `nfcapd`, který slouží jako NetFlow kolektor. Pro ověření správnosti exportovaných dat byla jako referenční výstupní data použita data generovaná nástrojem `softflowd`, který vykonává obdobné funkce jako vyvíjený program. Výstupy obou programů byly následně porovnány, aby se zjistilo, zda můj exportér p2nprobe zachovává klíčové vlastnosti NetFlow exportu.

Během analýzy dat byly porovnávány základní parametry, jako jsou celkový počet toků, objem přenesených bajtů a počet paketů. Získaná data ukázala, že program p2nprobe spolehlivě zachovává všechny tyto základní charakteristiky. Během testování se však projevily drobné časové odchylky mezi časovými značkami u jednotlivých toků, které lze přičíst zaokrouhlovacím chybám při agregaci časových údajů. Tyto odchylky jsou malé, obvykle se pohybují v řádu desetin milisekund, a nemají zásadní vliv na přesnost sledovaných toků.

Výsledky testů jsou uvedeny ve dvou souborech pro snadné porovnání. Výstup generovaný nástrojem `softflowd` je uložen v souboru `tests/softflowd_output.txt`, zatímco výstup z mého programu p2nprobe se nachází v souboru `tests/p2nprobe_output.txt`. Tyto soubory obsahují záznamy o všech klíčových parametrech toku a slouží jako podklad pro detailní porovnání obou výstupů.

Kapitola 4

Bibliografie

- IBM NetFlow v5 format – Zdroje, z nichž jsem získal strukturu formátu NetFlow v5, nutnou pro správný export toků do kolektoru.
- TCPdump and libpcap Homepage – Dokumentace k knihovně `pcap.h`, poskytující informace o funkcích pro práci se síťovými pakety.
- softflowd, a flow-based network monitor
- *Programming with Libpcap - Sniffing the network from our own application* by Luis MartinGarcia (Hakin9 Magazine, issue 2/2008) – Článek poskytující detailní návod k použití knihovny Libpcap při vytváření vlastních aplikací pro zpracování síťových paketů.