VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Sieťové aplikácie a správa sietí
PCAP NetFlow v5 exportér

2024 Boris Hatala (xhatal02)

Obsah

1 Úvod

Táto technická správa popisuje implementáciu PCAP NetFLow v5 exportéru vo forme konzolovej aplikácie napísanej v jazyku C++. Aplikácia je kompilovaná pomocou g++ a je určená pre Linux.

Problematika

V rámci tohto projektu sa zameriavame na analýzu sieťovej prevádzky a jej prevod do formátu NetFlow v5, ktorý predstavuje štandard pre monitorovanie a analýzu sieťových tokov. Sieťová prevádzka je zvyčajne zachytávaná do PCAP súborov, ktoré obsahujú detailný záznam o jednotlivých sieťových paketoch. Z týchto záznamov je možné rekonštruovať a analyzovať toky, ktoré predstavujú agregované informácie o komunikácii medzi dvoma sieťovými bodmi, charakterizované napríklad IP adresami, portmi a protokolmi.

Cieľom projektu je vytvoriť nástroj `p2nprobe`, ktorý dokáže:

- Načítat sieťové pakety zo súboru vo formáte PCAP
- Extrahovať informácie o sieťových tokoch výhradne pre protokol TCP
- Agregovať tieto informácie do formátu NetFlow v5
- Odosielať agregované toky na vzdialený kolektor prostredníctvom UDP

2 Návod na použitie

Program je potrebné preložiť pomocou prikázu `make`. Následne je možné ho spustiť z príkazového riadku s nasledovnými argumentami:

```
./p2nprobe [-a <akt. t.>] [-i <neak. t.>] <pcap súbor> <kolektor_ip:kolektor_port>
```

- `-a <aktívny timeout>` - Nastaví aktívny timeout v sekundách
- `-i <neaktívny timeout>` - Nastaví neaktívny timeout v sekundách
- `<pcap súbor>` - Cesta k PCAP súboru, ktorý sa má analyzovať.
- `<kolektor_ip:kolektor_port>` - Adresa a port NetFlow kolektora.

Argumenty `-a` a `-i` sú voliteľné, ak nepoužijete tieto argumenty, tak aktívny aj neaktívny timeout bude automaticky nastavený na 60 sekúnd. Argumenty v príkazovom riadku môžu byť v hocijakom poradí.

Príklad spustenia

```
./p2nprobe large.pcap 127.0.0.1:2055 -a 10 -i 3
```

Program spracuje súbor `large.pcap`, ktorého toky vo formáte NetFlow v5 pošle na kolektor s ip adresou `127.0.0.1` a portom `2055`. Aktívny timeout bude nastavený na 10 sekúnd a neaktívny na 3 sekundy.

3 Návrh

Aplikácia je rozdelená na zdrojový súbor `p2nprobe.cpp` a hlavičkový súbor `p2nprobe.h`. V tejto kapitole sú popísané jednotlivé časti implementácie a technické detaily.

Knižnice

Program využíva knižnice:

- `pcap` na spracovanie PCAP súborov.
- Siete (`netinet` a `arpa/inet`) na prácu s IP adresami a TCP/UDP hlavičkami.
- Štandardná knižnica C++ (`map`, `string`) na ukladanie aktívnych a čakajúcich tokov.

Štruktúry

`Flow` Štruktúra reprezentujúca jednotlivý tok, obsahuje:

- Zdrojová a cieľová IP adresa (`struct in_addr`).
- Zdrojový a cieľový port (`uint16_t`).
- Počet paketov a bajtov v toku (`uint32_t`).
- Čas začiatku a konca toku (`uint32_t`).

`NetFlowV5Packet` Štruktúra obsahujúca hlavičku pakety NetFlow v5 a maximálne 30 záznamov typu NetFlow v5.

```
    struct NetFlowV5Packet {
        struct NetFlowV5Header header;
        struct NetFlowV5Record records[30];
};
```

- `NetFlowV5Header` je štruktúra, ktorá predstavuje hlavičku NetFlow v5 pakety, nachádza sa v nej verzia pakety, doba prevádzky exportéra v milisekundách pred odoslaním, počet záznamov v pakete atď.
- `NetFlowV5Record` je štruktúra, ktorá obsahuje informácie o jednotlivých záznamoch v NetFlow v5 pakete ako sú zdrojová a cieľová ip adresa a port, začiatok a koniec záznamu v milisekundách vzhľadom na dobu prevádzky exportéra atď.

`activeFlows` Typ `std::map` obsahujúci aktívne toky, kde kľúčom je zdrojová a cieľová IP adresa a port.

`flowsBuffer` Typ `std::map` obsahujúci toky, ktoré čakajú na export.

4 Implementácia

Vstupným bodom programu je funkcia `main`, ktorá spracuje argumenty z príkazového riadka a následne sa pokúsi otvoriť PCAP súbor.

Logika spracovania paketov

Zavolaním funkcie `pcap_loop` sa začne spracovávať každý paket pomocou callback funkcie `callback`, kde sa zisťuje, či je paket TCP. Ak áno, tak sa z neho extrahujú informácie o toku a aktualizuje sa záznam v `activeFlows`. Teda, buď sa vytvorí nový tok, alebo sa aktualizuje už existujúci. Toky sú ukladané do mapy `activeFlows` podľa zdrojovej a cieľovej IP adresy a portu.

Timeouty

Po každom spracovanom pakete sa prejde mapa `activeFlows` a zistí sa, či nejaký tok nepresiahol timeout. Tento výpočet je vykonávaný na základe časového údaju z poslednej načítanej pakety. Ak tok presiahol aktívny timeout alebo neaktívny, tak sa presunie do mapy `flowsBuffer` na export.

Inicializácia NetFlow v5 pakety

Inicializácia NetFlow v5 pakety sa vykonáva vo funkcii `initNetFlowV5Packet`. Táto funkcia nastaví všetky potrebné polia hlavičky NetFlow v5 pakety a zabezpečí, že paketa je pripravená na naplnenie záznamami o tokoch. Funkcia `initNetFlowV5Packet` vynuluje celú štruktúru pakety pomocou `memset`, nastaví verziu pakety, vypočíta dobu prevádzky exportéra v milisekundách od začiatku programu a nastaví pole `sysUptime`, nastaví `unixSecs` a `unixNsecs`, a sekvenciu toku do poľa `flowSequence`.

NetFlow v5 záznam

Záznamy sú postupne pridávané do NetFlow v5 pakety v cykle, kde sa prechádza mapa `flowsBuffer`. Položka first a last sú vypočítané v milisekundách na základe časového údaja pakety a času začiatku programu.

Export

Po každom spracovanom pakete sa prejde mapa `flowsBuffer` a overí sa či nejaký tok čaká na export. Ak áno, vytvárajú sa záznamy v NetFlow v5 formáte, zapĺňa sa paketa a záznamy z `flowsBuffer` sa postupne mažú. Paketa sa odošle v momente, keď je naplnená, teda obsahuje 30 záznamov, alebo keď program skončí spracovávanie PCAP súboru. V takom prípade sa všetky zvyšné toky exportujú. Po exportovaní pakety sa incializuje nová a proces sa opakuje.

5 Testovanie

Na testovanie boli použité PCAP súbory generované pomocou programu `Wireshark` alebo `tcpdump`, kolektor `nfcapd` z nástroja `nfdump`, referenčný exportér `softwflowd` a program `Wireshark`. Pri testovaní a debugovaní mi rovnako pomáhali aj rôzne print funckie.

Pomocou `nfcapd` a `softwflowd`

Kolektor `nfcapd` som spustil na lokálnej adrese a porte 2055. Exportoval som rovnaké PCAP súbory pomocou `softwflowd` a `p2nprobe` na kolektor `nfcapd` a následne som pomocou nástroja `nfdump` porovnal výstupy. Analyzoval som počet exportovaných tokov a správnosť exportovaných informácií.

Pomocou `Wireshark`

Po spustení kolektoru `nfcapd` som exportoval PCAP súbor pomocou `p2nprobe` a `softwflowd` sledoval som správnosť exportovaných informácií pomocou programu `Wireshark`.

Zdroje

[1] Cisco Systems: *NetFlow Export Datagram Format*. Dostupné z: https://www.cisco.com/c/en/us/td/docs/net_mgmt/netflow_collection_engine/3-6/user/guide/format.html#wp1006108.

[2] Lucas, M. W.: *Network Flow Analysis*. No Starch Press, 2010.

[3] Cisco Systems: *RFC 3954: Cisco Systems NetFlow Services Export Version 9*. Dostupné z: https://datatracker.ietf.org/doc/html/rfc3954.

[4] Claise, B., Trammell, B., Aitken, P.: *RFC 7011: Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information*. Dostupné z: https://datatracker.ietf.org/doc/html/rfc7011.