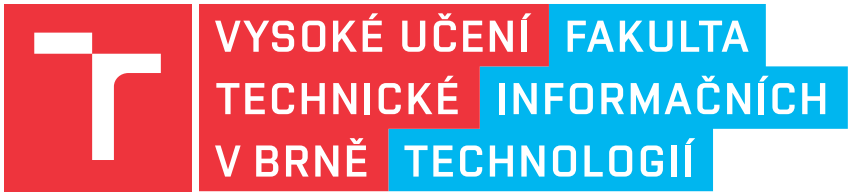

ISA 2022 - Dokumentácia k projektu: **Generovanie NetFlow dát zo zachytenej sieťovej komunikácie**

14. novembra 2022

Marián Backa (xbacka01)

Obsah

1 Uvedenie do problematiky

Cieľom projektu je vytvoriť program na získanie tokov rekordov z jednotilivých paketov .pcap súboru. Ide teda o analýzu zachyteného dátového toku. Jednotlivé pakety z pcap súboru združujeme do takzvaných tokov na základe spoločných vlastností a udržujeme o nich užitočné informácie. Ako programátorom sledovanie tokov nám umožňuje omnoho rýchlejšie a jednoduchšie pochopiť, čo sa na sieti dialo ako keby sme sa snažili analyzovať jednotlivé pakety ručne. Program posiela vygenerované toky na kolektor.

2 Návrh aplikácie

Program najprv spracuje argumenty. Následne sa prechádza .pcap súbor paket po pakete. Z hlavičiek každého packetu získa program potrebné informácie pre toky. Na základe získaných informácí sa zistí, či paket patrí do už existuhúceho toku, alebo vytvára nový. Po získaní informácí z hlavičiek sa kontroluje neaktívny časovač včetkých tokov. Pri vytváraní nového toku sa kontroluje veľkosť cache a pri aktualizácí existujúceho toku sa kontroluje aktívny časovač. Toky ktoré prekračujú limit časovačov sú exportované na kolektor a odstránené z cache. Program takto pracuje do spracovania všetkých paketov z .pcap súboru. Nakoniec program exportuje všetky toky z chache na kolektor, vyprázdni cache a ukončí sa.

3 Popis implementácie

Program je napísaný v jazyku C++. Najprv spracuje argumenty pomocou funkcie `getopt()`.[4] Potom program inicializuje socket na komunikáciu s kolektorom pomocou UDP spojenia. Následne sa prechádza .pcap súbor paket po pakete, podobne ako bolo ukázané na prednáške v programe readpcap.c[6], z hlavičiek každého paketu získa program potrebné informácie[11][12][5]. Toky sú uložené v mape[1] a ich klúčom je tuple[2] zložená z pätice: zdrojová ip, cieľová ip, zdrojový port, cieľový port a protokol. Po získaní informácí z paketu medzi ktoré patrí aj najnovší čas programu skontroluje či niektoré uložené toky neexpirovali z dôvodu neaktívneho časovača. Ak áno, odošle ich na kolektor[9][8][7] a odstráni z cache. Následne program spracováva daný packet a buď ho zaradí do exisujúceho toku, kde potom prebieha aktualizácia dát a kontrola aktívneho časovača, alebo vytvorí nový, kde pred vložením do cache prebieha kontrola na zaplnenie cache, ak je cache plná tak sa najstarší tok exportuje, aby sa vytvorilo miesto pre nový. Proces sa opakuje do vyčerpania paketov z .pcap súboru. Nakoniec program exportuje všetky existujúce toky z cache[10].

4 Základné informácie o programe

Pre zistenie náležitosti packetu do toku sa používa pätica[3]: src ip, dst ip, src port, dst port a protokol. Pokiaľ by cache presiahla limit pri vkladaní nového toku, exportuje sa najstarší tok - tok ktorý nedostal nový paket najdlhšiu dobu a vytvorí sa tak miesto pre nový. Program neimplementuje ukončovanie TCP tokov na základe flagov. Tokom protokolu ICMP sa dosadzuje miesto čísla portov 0, keďže ich protokol túto informáciu neobsahuje, následne s nimi program pracuje tak isto ako s UDP či TCP tokmi.

5 Zaujímavejšie pasáže implementácie

Použitie tuple ako klúč do mapy pre získanie netflow záznamu.

```
typedef std::tuple<uint32_t, uint32_t, uint16_t, uint16_t, uint8_t> tuple_key;

   std::map<tuple_key,struct flowrecord> flow_map;
```

6 Návod na použitie

Preloženie programu sa vykoná príkazom `make`.
Spustenie programu:
`./flow [-f <file>] [-c <netflow_collector>[:<port>]] [-a <active_timer>] [-i <inactive_timer>] [-m <count>]`

Popis argumentov:

`-f <file>`	Meno analyzovaného súboru, ak nezadané, tak STDIN.
`-c <netflow_collector>[:<port>]`	IP adresa, alebo hostname NetFlow kolektoru, Voliteľne i UDP port, (127.0.0.1:2055, ak nezadané).
`-a <active_timer>`	interval v sekundách, po kterom sa exportujú aktívne záznamy na kolektor, (60, ak nie je zadané).
`-i <inactive_timer>`	interval v sekundách, po ktorom vypršení sa exportujú neaktívne záznamy na kolektor (10, ak nie je zadané)
`-m <count>`	veľkosť flow-cache. Při dosažení max. veľkosti dôjde k exportu najstaršieho záznamu v cachi na kolektor (1024, ak nie je zadané).

Všetky argumenty sú volitelné.

Referencie

[1] *C++ map*. [online]. URL: https://www.geeksforgeeks.org/map-associative-containers-the-c-standard-template-library-stl/.

[2] *C++ tuple*. [online]. URL: https://www.geeksforgeeks.org/tuples-in-c/.

[3] *Formát netflow datagramov*. [online]. URL: http://www.cisco.com/c/en/us/td/docs/net_mgmt/netflow_collection_engine/3-6/user/guide/format.html#wp1003394.

[4] *Getopt manual*. [online]. URL: https://www.gnu.org/software/libc/manual/html_node/Getopt.html.

[5] *ICMP packet wikipédia*. [online]. URL: https://en.wikipedia.org/wiki/Internet_Control_Message_Protocol.

[6] *ISA read-pcap.c*. [online]. URL: https://moodle.vut.cz/pluginfile.php/504654/mod_folder/content/0/pcap/read-pcap.c?forcedownload=1.

[7] *man libpcap*.

[8] *man nfcapd*.

[9] *man nfdump*.

[10] *Netflow wikipedia*. [online]. URL: https://en.wikipedia.org/wiki/NetFlow.

[11] *TCP packet wikipédia*. [online]. URL: https://en.wikipedia.org/wiki/Transmission_Control_Protocol.

[12] *UDP packet wikipédia*. [online]. URL: https://en.wikipedia.org/wiki/User_Datagram_Protocol.