Ingetstor metadat z network flow do Apache Kafka

Patrik Krajč, Peter Grofčík

December 4, 2020

Exportér

1. Zdrojové súbory exportéra
   - packet_parser.c/h
   - flow_list.c/h
   - my_kafka.c/h
2. Ukočovanie toku dát
   - Príznakom FIN
   - Časovač
3. Spustenie a ukončenie programu
   - Spracovanie signálu ctrl+c

Packet parser

- struct Flow
  - struct flow_id (src/dst addr, protocol, src/dst port)
  - struct packet_info_buffer (isExported, index)
    - Array of packet_info (packet payload, timestamp)
- Funkcia :
  - Získavanie dát z paketov IPv4 a IPv6
  - Kontrola obojsmerného toku
  - Prevod štruktúry na reťazec

Flow list

- List, ktorého položky tvoria štruktúry Flow
- Funkcia:
  - Uchovávať všetky aktuálne toky data
  - Vkladanie/mazanie dátových tokov
  - Na základe časovača vyexportovať akutálne dáta

Kafka

- Inicializácia producenta
- Registrácia zdroja/topic voči apache kafka
- Posielanie správ na broker

Spustenie programu a ukončovanie toku dát

- Spustenie
  - Inicializácia zookeepr a apache kafka docker-compose up -d
  - sudo ./producer server timer
  - python3 consumer.py
- Ukončenie
  - Detekcia príznaku FIN v TCP hlavičke
  - Časovač, po jeho uplynutí vyexportuje časť dát z listu
  - Spracovanie signálu ctrl + c