VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Síťové aplikace a správa sítí

Export DNS informací pomocí protokolu Syslog

Dokumentace

16. listopadu 2018

Petr Marek

Obsah

1 Zadání

Cílem projektu bylo vypracovat aplikaci, která bude umět zpracovávat data protokolu DNS[5] (Domain Name System) a vybrané statistiky exportovat pomocí protokolu syslog[3] na centrální logovací server.

2 Úvod do problematiky DNS

Jelikož počítače pracují na úrovni binárních dat, i přenos dat po síti musí probíhat touto cestou. Nicméně pro člověka je náročné a nepraktické pamatovat si sekvenci několika číslis pro každou stránku, kterou chce na internetu navštívit. Těmto sekvencím se říká IP adresy a identifikují každé zařízení v internetu. Proto vznikla potřeba mapovat IP adresy na lehce zapamatovatelná doménová jména. To byl základní kámen při vzniku systému DNS, který tato aplikace využívá. Zprávy mezi koncovým zařízením a DNS serverem lze odchytávat a analyzovat. To bude také jedna z hlavních funkcí tohoto programu.

3 Uvedení do problematiky logování

Každý programátor měl někdy potřebu sbírat informace o tom, co jeho aplikace dělá. Ať už z čistě informativního hlediska, či pro nalezení původu nějakého problému. K tomu účelu je nutné mít informace o tom, co se v daném systému děje. K tomu slouží logování, které přijímá potřebné informace a ukládá je. Naše aplikace k tomu využívá protokol syslog. Používá se zejména proto, že umožňuje oddělení softwaru, který generuje zprávy, systému, který je uchovává a softwaru, který je zaznamenává a analyzuje.

4 Základní informace o aplikaci

Aplikace naslouchá buď na síťovém rozhraní, nebo zpracovává daný pcap soubor. Pokud aplikace naslouchá na síťovém rozhraní, jsou statistiky periodicky odesílány na syslog server po vypršení definované doby dané přepínačem -t. Při zpracování pcap souboru jsou statistiky odeslány po jeho zpracování. Při obdržení signálu SIGUSR1 vypíše aplikace statistiky na standardní výstup. Tento signál lze vyslat programově. K tomu účelu aplikace při spuštění odposlouchávání vypíše informaci o ID programu (PID). Ten je pro usnadnění odfiltrování uveden prefixem "INFO:".

4.1 Formát statistik

```
domain-name rr-type rr-answer count
```

- domain-name : doménové jméno
- rr-type : typ odpovědi
- rr-answer : odpověď
- count : počet shodných odpovědí

5 Návrh aplikace

Po zapnutí aplikace a zpracování argumentů příkazové řádky je nutné získat data k analýze. K tomu slouží tzv. sniffer. Je založen na pcap API sloužící k odchytávání provozu na síti. V aplikaci je nastaven pro odchytávání paketů zdrojového portu 53, tedy ty obsahující odpověď serveru DNS. Z této zprávy jsou vybrány potřebná data. Z nich se vytvoří statistika, která se zobrazí uživateli, či odešle na centrální logovací server.

6 Popis implementace

6.1 Zachycení paketů

Nejprve je potřeba odchytit samotný paket, k čemuž slouží již zmiňované pcap API. Aplikace odchytává pouze pakety z ethernetového rozhraní. Po odchycení paketu je třeba z něj extrahovat data. Pokud se jedná o UDP paket, celá DNS zpráva bude v jednom paketu. Jestliže je zpracováván TCP paket[2], je třeba rozlišovat, zda-li je DNS zpráva již úplná či nikoliv. K tomu účelu slouží příznaky PUSH a SEQ v TCP hlavičce. Jestliže je příznak PUSH nastaven, znamená to, že data mají být ihned zpřístupněna aplikaci. Příznak SEQ zase umožňuje sledovat jednotlivé řetězce zpráv.

Tento mechanismus je v aplikaci implementován pouze částečně. Pokud přichází TCP pakety ve správném pořadí, program je skládá za sebe. Tyto části označujeme jako segmenty. Jakmile přijde PUSH příznak, jsou takto poskládaná data jako jedna zpráva zpracována. Jestliže se vyskytne neočekávaný segment, je dosud rozpracovaná zpráva zahozena a začíná skládání segmentů od nově příchozího paketu.

Během odchytávání paketů lze vysláním signálu SIGUSR1 vypsat statistiky. Další signál SIGALRM je použit kvůli nastavení časovače, který zajišťuje periodické odesílání statistik na logovací server.

6.2 Zpracování DNS odpovědi

Formát DNS zpráv je specifikován v *RFC 1035*[5]. Implementováno je i bezpečnostní rozšíření protokolu DNS, DNSSEC. Tento protokol je popsán v *RFC 4035*[1].

Z každého příchozího DNS paketu je zpracována pouze první vyskytující se odpověď *RR Answer*. Navíc podporovány jsou jen vybrané typy odpovědí. Pakety neobsahující žádnou odpověď, či obsahující nepodporované typy záznamů jsou ignorovány. K dekódování některých typů (např. *RRSIG*) bylo zapotřebí zakódovat *base64* formát dat, jelikož jeho původní binární forma není tisknutelná. K tomu zde slouží volně dostupný modul *base64* od René Nyffeneggera.

Z takto zpracovaných dat jsou vždy před jejich tisknutím či odesláním na server vytvořeny statistiky. Statistiky jsou po jedné zprávě odesílány na logovací server formou UDP paketů.

6.3 Statistiky

Statistiky se vytváří z řetězce, ke kterému program po každém celém zpracovaném paketu připojuje další data. Jeden řádek těchto dat může vypadat například takto:

```
google.com A 172.217.23.238
```

Celý tento řetězec se prochází a unikátní záznamy kopírují do pomocného řetězce. Dále se prochází tento pomocný řetězec a ke každému jeho záznamu se připojí číslo, označující počet jeho výskytů v původních datech. Takto vytvořený řetězec již je výsledná statistika.

6.4 Syslog zpráva

Syslog zprávy jsou ve speciálním formátu. Ten je popsán v RFC 5424[3]. Formát časové známky je pak blíže specifikován v *RFC 3339*[4]. Tyto informace jsou doplněny ke každému řádku statistiky, který se odesílá na logovací server. je tedy implementována zjednodušená varianta, která odesílá data po jednom řádku. V aplikaci je použit formát:

<PRI>VERSION TIMESTAMP HOSTNAME APP–NAME MSG

- PRI: Priorita zprávy
- VERSION: verze specifikace protokolu syslog
- TIMESTAMP: časová známka - GMT čas
- HOSTNAME: IP adresa zařízení, které zaslalo syslog zprávu
- APP-NAME: Název aplikace
- MSG: Zpráva

Celá zpráva může vypadat například takto:

<134>1 2018-09-20T22:14:15.003Z 127.0.1.1 dns-export - - - google.com A 172.217.23.238 68

Pokud se při výpisu objeví tento název aplikace "dns-export[- - -]", je třeba nastavit šablonu pro syslog, která zobrazuje pouze přijatou zprávu, bez dalšího formátování.

7 Návod na použití

```
dns-export [-r file.pcap] [-i interface] [-s syslog-server] [-t seconds]
```

- -r: zpracuje daný pcap soubor
- -i: naslouchá na daném síťovém rozhraní a zpracovává DNS provoz
- -s: hostname/ipv4/ipv6 adresa syslog serveru
- -t: doba výpočtu statistik, výchozí hodnota 60s

Všechny parametry jsou nepovinné, nicméně platí určitá omezení. Odposlouchávat lze buď ze síťového rozhraní nebo ze souboru. Tedy možnosti -r a -i se vzájemně vylučují, ale alespoň jedna z nich musí být použita. Také parametr -t nemá při zpracování pcap souboru opodstatnění, statistiky jsou odeslány ihned po jeho zpracování. Aplikace se ukončuje zasláním signálu SIGINT nebo SIGTERM. Při použití jiného signálu může dojít k úniku paměti.

Reference

[1] ARENDS, R., AUSTEIN, R., LARSON, M. et al. *Resource Records for the DNS Security Extensions* [online]. Duben 2005 [cit. 24. října 2018]. Dostupné na: <https://www.rfc-editor.org/rfc/rfc4034.txt>.

[2] DICKINSON, J., DICKINSON, S., BELLIS, R. et al. *DNS Transport over TCP - Implementation Requirements* [online]. Duben 2016 [cit. 24. října 2018]. Dostupné na: <https://www.rfc-editor.org/rfc/rfc7766.txt>.

[3] GERHARDS, R. *The Syslog Protocol* [online]. Duben 2009 [cit. 24. října 2018]. Dostupné na: <https://www.rfc-editor.org/rfc/rfc5424.txt>.

[4] KLYNE, G. a NEWMAN, C. *Date and Time on the Internet: Timestamps* [online]. Červen 2002 [cit. 24. října 2018]. Dostupné na: <https://www.rfc-editor.org/rfc/rfc3339.txt>.

[5] MOCKAPETRIS, P. *Domain names - implementation and specification* [online]. Listopad 1987 [cit. 24. října 2018]. Dostupné na: <https://www.rfc-editor.org/rfc/rfc1035.txt>.