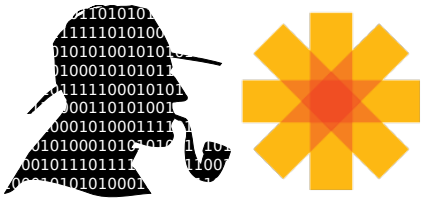
Excel@FIT 2015
http://excel.fit.vutbr.cz

Zákonné odposlechy v SDN

Barbora Franková*



Abstrakt
Tato práce se zabývá využitím softwarově definovaných sítí v oblasti zákonných odposlechů. Staví na implementaci systému pro zákonné odposlechy vyvinuté v rámci projektu Sec6Net. Přínosem práce je rozšíření tohoto systému v několika oblastech, ve kterých SDN nabízí potenciál ke spolehlivější identifikaci odposlouchávaných uživatelů a efektivnějšímu využití sítě. První zmíněný cíl je realizován prostřednictvím modulu funkce dynamické identity, druhý pak pomocí modulu pro dynamickou konfiguraci síťových sond.

Klíčová slova: Softwarově definované sítě — OpenFlow — OpenDaylight — Zákonné odposlechy — Sec6Net

Přiložené materiály: *N/A*

*xfrank08@stud.fit.vutbr.cz, *Fakulta informačních technologií, Vysoké učení technické v Brně*

1. Úvod

Softwarově definované sítě rozšiřují programovatelnost sítě. Zatímco zařízení v klasických sítích mají pevně danou funkcionalitu od výrobce a složitý operační systém, softwarově definované sítě umožňují získat globální pohled na síť a řídit chování sítě pomocí programovatelných síťových zařízení.

Tato práce je zaměřena na využití znalosti topologie ze softwarově definovaných sítí (SDN) v systému pro zákonné odposlechy. Cílem je spolehlivější identifikace odposlouchávaného užitatele, efektivnější využití sítě a dynamická konfigurace sond na základě aktuální topologie. Žádný ze současných systémů pro zákonné odposlechy (např. Cisco, Verint) ale nevyužívá výhod, které poskytuje propojení systému pro zákonné odposlechy s SDN.

Řešení je postaveno na implementaci systému pro zákonné odposlechy SLIS, který vznikl v rámci projektu Sec6Net. SLIS jsem v rámci projektu propojila s SDN kontrolerem OpenDaylight, který poskytuje rozhraní pro zjištění kompletní topologie sítě. Vytvořila jsem modul pro OpenDaylight, který získává informace o identitách koncových uživatelů a předává je dál systému pro zákonné odposlechy. Topologie se využívá při přidávání požadavků na odposlech tak, aby se konfigurovala pouze jedna sonda. Využití SDN umožňuje také směrování komunikace k sondám, které neleží přímo na lince, kterou odposlouchávaná data prochází, a vyvažování zátěže mezi sondami.

V sekci 2 jsou rozebrány teoretické pojmy. Podsekce 2.1 je věnována stručnému popisu systému SLIS, který vznikl v projektu Sec6Net. V podsekci 2.2 je vysvětlen princip softwarově definovaných sítí. V sekci 3 je popsán návrh rozšíření systému pro zákonné odposlechy a v sekci 4 implementace těchto rozšíření.

2. Teorie

2.1 Zákonné odposlechy

Rostoucí počítačová kriminalita je jedním z rysů současné společnosti. Z tohoto důvodu byly některé z útoků,

např. neoprávněný přístup, odposlouchávání, narušení systému nebo počítačové padělání, zaneseny do zákona jako trestná činnost. Přesné znění lze nalézt v zákoně o elektronických komunikacích (zákon č. 127/2005 Sb.) a v Evropském právu (Úmluva Rady Evropy o počítačové kriminalitě – Convention on Cybercrime, ETS No. 185).

Jednou z možností, jak bojovat proti počítačové kriminalitě, jsou systémy pro zákonné odposlechy. Doporučená architektura těchto systémů byla vytvořena úřadem ETSI pro všechny země Evropské unie [1]. Systémy pro zákonné odposlechy umožňují oprávněným orgánům sledovat komunikaci podezřelých subjektů v počítačové či telefonní síti. V rámci projektu Sec6Net[1] vznikla implementace systému pro zákonné odposlechy – Sec6Net Lawful Interception System (SLIS), který je určen pro nasazení v sítích poskytovatelů Internetu [2].

Součásti systému SLIS, které jsou důležité pro tuto práci, jsou následující:

- Funkce dynamické identity (IRI-IIF) – má za úkol dynamické zjišťování částečné identity sledováním probíhajících komunikací (relace, hovory, spojení apod.) a analýzou protokolů.
- Triggerovací funkce (CCTF) – konfiguruje jednotlivé sondy ve chvíli, kdy má být zahájen odposlech.
- Sondy pro odposlech (CC-IIF) – mají za úkol zachytávat obsah komunikace sledovaných uživatelů.

Odposlouchávaný uživatel musí být jednoznačně identifikovatelný v síti. Jeho *identita* se skládá z množiny *identifikátorů* používaných v síťovém prostředí (např. IP adresa, MAC adresa apod.). Některé identifikátory se mohou na straně poskytovatele dynamicky měnit např. skrze protokoly DHCP, RADIUS nebo SLAAC [3]. Již dostupné moduly pro IRI-IIF mají za úkol odesílat informace o tom, kdy a komu byly identifikátory přiděleny. IRI-IIF na základě těchto informací dynamicky propojí *částečné identity*, které patří jednomu uživateli nebo stroji [4, 5].

Jako příklad lze uvést požadavek na odposlech uživatele s určitou e-mailovou adresou. Pokud tento uživatel v průběhu odposlechu změní IP adresu svého zařízení nebo ke komunikaci použije jiné zařízení, IRI-IIF detekuje změnu a předá tuto informaci systému. CCTF pak může včas překonfigurovat připojené sondy CC-IIF a zachytit veškerý obsah komunikace.

[1] http://www.fit.vutbr.cz/~matousp/grants.php?id=517

2.2 Softwarově definované sítě

Koncept softwarově definovaných sítí je nový přístup k počítačových sítím, který se v poslední době prosazuje i v komerční sféře. SDN odděluje logiku sítě od samotného přeposílání paketů a tím umožňuje vytvářet programovatelné sítě. Tento přístup lze využít především ve velkých datacentrech a u poskytovetelů internetového připojení.

Výzkumu v této oblasti se věnuje například Google, který pomocí SDN propojil svá datacentra napříč kontinenty nebo McKeown aj., kteří zkoumali nasazení SDN v univerzitním kampusu [6, 7]. Princip SDN lze využít i v dalších aplikacích, jako je dynamická kontrola přístupu, vyvažování zátěže, virtualizace sítě nebo energeticky úsporné sítě.

Architektura softwarově definovaných sítí se skládá z vrstvy infrastruktury (data plane) a řídící vrstvy (control plane). Ve vrstvě infrastruktury se nachází jednotlivá zařízení, která mají na starosti rychlé přeposílání paketů. V řídící vrstvě se nachází oddělená kontrolní část ve formě kontroleru. Kontroler má přehled o topologii celé sítě a o prostředcích k směrování a přepínání paketů, které umožňují jednotlivá síťová zařízení. Nad řídící vrstvou lze vytvářet různé aplikace, upravující řízení sítě.

Komunikační rozhraní mezi řídící vrstvou a síťovými zařízeními může být například OpenFlow [7]. OpenFlow přepínač využívá koncept datových toků k identifikaci síťového provozu na základě pravidel, která jsou naprogramována staticky nebo dynamicky. Pravidla jsou uložena v tabulkách toků, které porovnávají procházející pakety se svými záznamy a zvolí akci na základě výsledku porovnání. Pravidla se vyhledávají sestupně od nejvyšší priority a je aplikováno první pravidlo, u kterého je nalezena shoda.

3. Návrh

Systém pro zákonné odposlechy SLIS je určen k nasazení u poskytovatelů internetového připojení. V klasických sítích mohou být CC-IIF sondy připojeny na lince s TAPem, který veškerá data zduplikuje. SLIS nastavuje všechny sondy stejně a sondy tak odposlouchávají veškerou zájmovou komunikaci. Jednou z možností, jak systém vylepšit, je použití softwarově definovaných sítí. Pokud systém rozšíříme o znalost topologie, můžeme jednoduše identifikovat cestu zájmové komunikace a přizpůsobit chování přepínačů.

Prvním rozšířením systému je *získávaní částečných identit* koncových stanic z kontroleru SDN. Částečná identita stroje zahrnuje identitu na síťové (L2) a linkové vrstvě (L3). Identifikátory získané z SDN kontroleru (MAC adresa, IP adresa a přepínač, ke kterému je

zařízení připojeno) poté IRI-IIF propojí s částečnými identitami uživatelů využívající detekované stroje (získaných z jiných modulů systému SLIS).

Druhým rozšířením je *dynamická konfigurace sond* na základě topologie sítě. Kombinací znalosti topologie z kontroleru a znalosti pozice sond v této topologii můžeme optimalizovat konfiguraci sond CC-IIF. Součástí tohoto řešení je i dynamická rekonfigurace přepínačů. Kombinace dynamické konfigurace sond a přepínačů umožní vyrovnávání zátěže a směrování toků k sondám, které neleží přímo na lince, kterou data prochází.

4. Implementace a testování

Pro implementaci navržených vylepšení jsem zvolila kontroler OpenDaylight. OpenDaylight je open source projekt, který má podporu i v komerční sféře (Cisco, HP a další). Tato práce využívá rozhraní REST [8] protokolem HTTP.

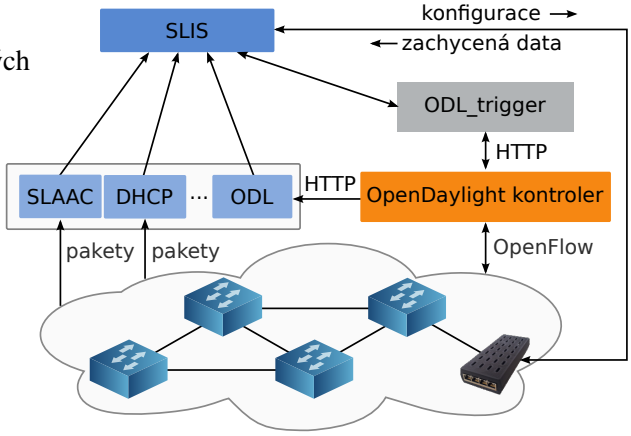
4.1 Získávání částečných identit

Získávání částečné identity z SDN je implementováno v jazyce Python jako modul *ODL* pro OpenDaylight kontroler. Na obrázku 1 je znázorněno zapojení *ODL* do systému.

Z kontroleru je možné získat tři typy identifikátorů pro dané zařízení: IP adresu, MAC adresu a identifikátor přepínače, ke kterému je toto zařízení připojeno. Funkci dynamické identity modul odesílá zprávy, když byl detekován začátek a konec spojení. Součástí zpráv je uvedená trojice identifikátorů. IRI-IIF odpovídající identifikátory propojí a tím rozšíří identitu tohoto zařízení.

4.2 Dynamická konfigurace sond

Koncept softwarově definovaných sítí umožňuje získat kompletní topologii a v systému pro zákonné odposlechy může znalost topologie zlepšit nastavování jednotlivých CC-IIF sond. Sondy jsou konfigurovány pomocí triggerovací funkce. Pokud přijde požadavek na zahájení odposlechu, musí tato funkce na základě topologie rozhodnout, kterou sondu nastaví a jaká data k ní bude přeposílat. Z toho důvodu byl vytvořen nový modul do systému pro zákonné odposlechy *ODL_trigger*. Tento modul byl implementován v jazyce Python a jeho zapojení do systému je znázorněno na obrázku 1. Modul *ODL_trigger* se v pravidelných intervalech dotazuje kontroleru na aktuální topologii, vkládá pravidla pro směrování kopií odposlouchávaných dat k sondám a triggerovací funkce ve SLIS na základě informací od tohoto modulu konfiguruje jednotlivé sondy.



Obrázek 1. Schéma zapojení systému pro zákonné odposlechy, SDN kontroleru, modulu pro zjišťování dynamické identity – *ODL* (4.1) a modulu pro sledování topologie – *ODL_trigger* (4.2).

Na rozdíl od funkce dynamické identity, která se zajímá o měnící se identifikátory koncových zařízení, je pro triggerovací funkci nezbytné znát kompletní topologii. Jedinou informaci, kterou nejsme schopni získat dynamicky, je pozice CC-IIF sond v síti. Součástí modulu proto musí být konfigurační soubor, který specifikuje, na kterém rozhraní jsou připojeny.

Kombinací topologie získané z kontroleru a pozice sond z konfiguračního souboru modul vytváří grafovou reprezentaci, kde vrcholy grafu jsou jednotlivá zařízení a hrany odpovádají linkám. Ve chvíli, kdy přijde požadavek na odposlech, začíná modul s konfigurací síťových zařízení.

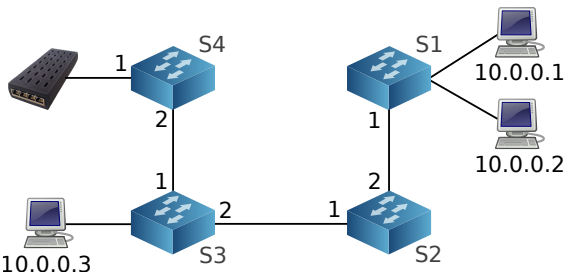
Konfigurace spočívá ve využití tří tabulek toků. Do první tabulky modul ukládá pravidla, která porovnávají procházející hlavičky paketů s IP adresou, která má být odposlouchávána. Pokud zdrojová nebo cílová adresa paketu odpovídají, je paket označen VLAN tagem a odeslán na výstupní port směrem k CC-IIF sondě. Následně je původní paket (bez VLAN tagu) předán třetí tabulce.

Druhá tabulka toků je na všech přepínačích stejná. Má za úkol porovnávat pakety s VLAN tagem a odesílat je směrem k CC-IIF sondě. Pravidla se prochází postupně od nejvyšší priority, proto musí být v první tabulce pravidlo s vysokou prioritou, které bude také porovnávat VLAN tag. Pakety, které budou takto označeny, pak nebude zpracovávat a pouze je předá druhé tabulce.

Třetí tabulka je plně pod správou kontroleru a přeposílá pakety k cílovým zařízením bez ohledu na pravidla v předchozích tabulkách.

Tímto způsobem se tedy vytvoří duplikát paketu s VLAN tagem a původní nezměněný paket se přepošle podle pravidel z kontroleru.

Modul zná aktuální topologii sítě a tak může jedno-



Obrázek 2. Ukázková topologie se zapojenou CC-IIF sondou.

duše zjistit, ke kterému přepínači je koncové zařízení s danou IP adresou přímo připojeno.

Uvažujme například topologii uvedenou na obrázku 2. Předpokládejme, že přišel požadavek na odposlech IP adresy 10.0.0.1. Zařízení s touto IP adresou je připojeno k přepínači S1. Na tento přepínač se vloží dvě pravidla s vysokou prioritou, která budou porovnávat danou zdrojovou a cílovou adresu v paketu. V případě, že jedna z těchto adres bude rovna 10.0.0.1, vloží se do paketu VLAN hlavička a odešle se na výstupní port 1. Ukázka pravidel je uvedena v tabulce 1 (porovnávání cílové IP adresy probíhá obdobně jako porovnávání zdrojové IP adresy). Na tomto i všech ostatních přepínačích se pak všechny pakety s VLAN hlavičkou budou přeposílat na rozhraní 1. Tato pravidla jsou uložena ve druhé tabulce a ukázka je uvedena v tabulce 2. Na přepínači S4 bude uloženo pravidlo, které ze všech paketů odesílaných na rozhraní 1 VLAN odstraní.

Sytém SLIS podporuje pravidla odposlechu konrétní IP adresy, trojice (IP adresa, port, protokol) a pětice (zdrojová IP adresa, port, cílová adresa, port a protokol). Přesnější požadavek jednoduše lze vyřešit pomocí přesnějších pravidel. U pětice pak můžeme libovolně rozhodnout, zda pravidlo pro vkládání VLAN tagů vložíme na přepínač, ke kterému je připojeno zařízení iniciátora komunikace nebo iniciovaného.

V případě, že je v topologii více CC-IIF sond, lze jednoduchým způsobem rozdělovat zátěž. Každá CC-IIF sonda bude mít vlastní VLAN tag. Při přidání odposlechu můžeme z grafu topologie zjistit, která sonda je nejblíž koncovému zařízení s danou IP adresou a při duplikování paketů vložíme VLAN tag nejbližší CC-IIF sondy. V druhé tabulce všech přepínačů pak budou pravidla, která pakety s VLAN hlavičkou odešlou směrem k odpovídající CC-IIF sondě.

V reálných zařízeních nemusí být k dispozici více tabulek toků. V takových případech je možné použít i alternativní přístupy. Jedním z nich je využití jednoho z fyzických portů přepínače, na který se bude duplikovat komunikace odposlouchávaného uživatele. Všechny pakety přijaté na tomto portu pak budou označeny a přeposlány směrem k sondě. K implementaci tohoto řešení stačí pouze jedna tabulka toků, ale nevýhodou je permanentní zablokování jednoho portu a nepřehlednost tabulky toků.

Tabulka 1. Ukázka pravidel pro odposlech v první tabulce toků. Porovnávání s hvězdičkou znamená, že na daném místě může být cokoliv. *Push/pop VLAN* značí přidání/odstranění VLAN tagu, *go-to table* znamená skoč do tabulky a *outport* odeslání paketu na výstupní port.

Prio	VLAN	IP zdroj	IP cíl	Ostatní	Akce
20	1	*	*	*	go-to tab 2
10	*	10.0.0.1	*	*	push VLAN outport 1 pop VLAN go-to tab 3
1	*	*	*	*	go-to tab 3

Tabulka 2. Ukázka pravidel v druhé tabulce toků.

Prio	VLAN	IP zdroj	IP cíl	Ostatní	Akce
10	1	*	*	*	outport 1

5. Závěr

Tato práce se zabývá rozšířením systému pro zákonné odposlechy tak, aby bylo možné využívat výhody softwarově definovaných sítí. Pro implementaci jsem zvolila kontroler OpenDaylight, který patří k nejpoužívanějším a má silné zastoupení v komerční sféře. Navrhla jsem dvě rozšíření:

- Modul *ODL* pro IRI-IIF, který je určen k získávání částečné identity. Modul se periodicky dotazuje kontroleru OpenDaylight na známé koncové stanice a změny hlásí IRI-IIF.
- Dynamická konfigurace CC-IIF sond. Jedná se především o rozšíření SLIS o modul *ODL_trigger*, který rozlišuje jednotlivé CC-IIF sondy a jejich pozice v topologii.

V klasických sítích je nutné umístit CC-IIF sondy přímo na linky, kterými bude procházet komunikace odposlouchávaného uživatele. Pokud by sonda byla připojena na jiné lince, bylo by velmi složité směrovat odposlouchávanou komunikaci přímo k sondě.

Využitím SDN v systému pro zákonné odposlechy je možné nastavit každou sondu jinak podle umístění v topologii. Díky tomu nedochází k odposlechu jednoho uživatele více sondami. Jednoduchým způsobem lze také směrovat k sondě zájmové pakety, které neprocházejí odposlouchávanou linkou. Je také možné

předejít zahlcení sondy přesměrováním toku odposlouchávaných dat k jiné, která bude v danou chvíli méně vytížená.

Součástí navazující práce bude důkladné otestování naimplementovaných částí a případně dalších rozšíření, jako je například využití OpenFlow přepínače jako CC-IIF sondy.

Poděkování

Ráda bych poděkovala Ing. Liboru Polčákovi za cenné rady, věcné připomínky a vstřícnost při konzultacích.

Literatura

[1] European Telecommunications Standards Institute: TR 101 943: Telecommunications security; Lawful Interception (LI); Concepts of Interception in a generic Network Architecture, 2001, v1.1.1.

[2] L. Polčák, T. Martínek, R. Hranický, S. Bárta, M. Holkovič, B. Franková, and P. Kramoliš. Zákonné odposlechy v moderních sítích. Technical report, FIT VUT v Brně, 2014.

[3] L. Polčák. Challenges in Identification in Future Computer Networks. In *ICETE 2014 Doctoral Consortium. Wien: SciTePress - Science and Technology Publications*, pages 15–24, 2014.

[4] L. Polčák and R. Hranický and T. Martínek. On Identities in Modern Networks. In *Journal of Digital Forensics, Security and Law*, volume 9, pages 9–22, 2014.

[5] A. Pfitzman and M. Hansen. A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management, August 2010. v0.34.

[6] J. Sushant, K. Alok, M. Subhasree, and O. Joon. B4: Experience with a Globally-Deployed Software Defined WAN. In *ACM SIGCOMM Computer Communication Review*, pages 3–14. ACM, 2013.

[7] N. McKeown and T. Anderson and H. Balakrishnan and G. Parulkar and L. Peterson. OpenFlow: enabling innovation in campus networks. In *ACM SIGCOMM Computer Communication Review*, pages 69–74. ACM, 2008.

[8] J. Medved, A. Tkacik, R. Varga, and K. Gray. OpenDaylight: Towards a Model-Driven SDN Controller architecture. In *15th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pages 1–6, 2014.