Architektura systému pro zákonné odposlechy

FIT VUT Technický report

Libor Polčák, Petr Kramoliš, Michal Kajan, Tomáš Martínek

Technický report č. FIT-TR-2011-008
Fakulta informačních technologií, Vysoké učení technické v Brně

Last modified: 9. ledna 2012

Architektura systému pro zákonné odposlechy

Libor Polčák, Petr Kramoliš, Michal Kajan, and Tomáš Martínek

Vysoké učení technické v Brně, email:
{ipolcak,xkramo00,ikajan,martinto}@fit.vutbr.cz

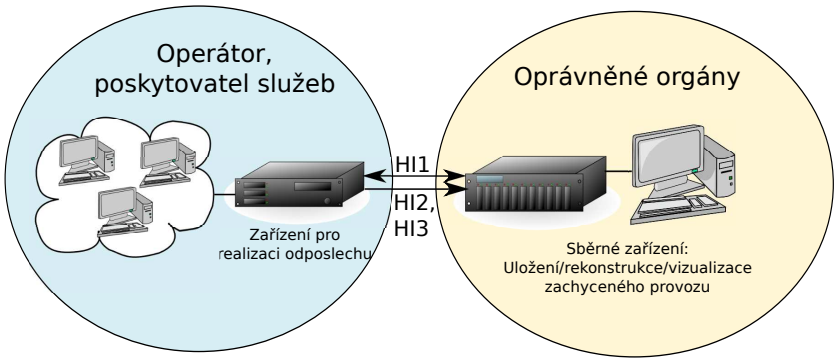
Abstrakt Tato technická zpráva popisuje návrh architektury prototypu systému pro sběr dat pro zákonné odposlechy vyvíjeného v rámci projektu *Moderní prostředky pro boj s kybernetickou kriminalitou na Internetu nové generace*. Tento prototyp dále slouží jako základní prostředí pro vývoj nových technik dynamické identifikace uživatele v prostředí IPv6 sítí, pro sběr dat a vývoj nových metod v oblasti rekonstrukce a vizualizace zachyceného provozu a také jako základní testovací prostředí pro vývoj mikro-sondy a vysokorychlostní sondy. Součástí zprávy je podrobný popis architektury navrhovaného systému a způsob komunikace mezi jednotlivými bloky. Zdůrazněny jsou problémy, které bylo potřeba při návrhu zohlednit, včetně způsobů jejich řešení. Zvláštní pozornost je věnována problematice dynamické identifikace cíle v prostředí IP sítí a návrhu architektury bloku IRI-IIF, který je za tuto funkci v systému pro sběr dat pro zákonné odposlechy zodpovědný.

1 Úvod

Systém pro zákonné odposlechy je nástroj, který umožňuje oprávněným orgánům sledovat aktivitu podezřelých osob využívajících veřejných komunikačních prostředků jako jsou telefonní sítě nebo Internet. Jelikož nejvíce informací o aktivitě sledovaného cíle je k dispozici na straně poskytovatele služeb (telefonní operátor, poskytovatel Internetového připojení), je systém rozdělen do dvou hlavních částí (viz obrázek 1): 1) Zařízení pro realizaci odposlechu na straně poskytovatele, které je schopno veškerou komunikaci sledovaného cíle zachytávat a předávat ji oprávněným orgánům. 2) Sběrné zařízení na straně oprávněných orgánů, které je schopno zachycený provoz ukládat a podrobněji analyzovat resp. vyhodnocovat.

Komunikace mezi oběma částmi probíhá skrze standardizované rozhraní definované normou ETSI [6], které se skládá ze tří částí:

1. *Rozhraní HI1*: slouží pro předávání požadavků na odposlech ze strany oprávněných orgánů a dodatečných informování o průběhu odposlechu.
2. *Rozhraní HI2*: slouží pro přenos metainformací o aktivitě sledovaného cíle (začátek/konec hovoru, připojení/odpojení od sítě, změna identity, apod.)
3. *Rozhraní HI3*: slouží pro přenos obsahu zachycené komunikace sledovaného cíle (obsah telefonního hovoru, SMS zprávy, datového přenosu apod.)



Obrázek 1. Architektura systému pro zákonné odposlechy

Zařízení pro realizaci odposlechu na straně poskytovatele je zodpovědné za: a) příjem vstupních požadavků a ověření jejich platnosti, b) sledování aktivity zadaného cíle (připojení/odpojení od sítě) včetně případné dynamické změny jeho identity v rámci sítě poskytovatele (např. změna IP adresy) a informace o přiřazení zasílat oprávněným orgánům skrze rozhraní HI2 a c) zachycení veškeré komunikace zadaného cíle a její předání oprávněným orgánům skrze rozhraní HI3.

Sběrné zařízení na straně oprávněných orgánů: a) přijímá zachycenou komunikaci od jednoho nebo více poskytovatelů služeb, b) zajišťuje bezpečné a spolehlivé uložení získaných informací do vlastního datového úložiště, c) provádí rekonstrukci zachycené komunikace, která zahrnuje např. uspořádání komunikace do oddělených toků a jejich klasifikace (FTP, email, web, apod.). Obě tyto úlohy je nutné provádět s ohledem na přeuspořádání vstupních dat nebo dokonce na jejich neúplnost a d) případnou vizualizaci rekonstruovaných dat, které mohou dále sloužit např. jako důkazní materiál pro případné soudní líčení.

Tato technická zpráva je zaměřena na popis prototypu zařízení pro realizaci odposlechu na straně poskytovatele služeb navrženého v rámci projektu *Moderní prostředky pro boj s kybernetickou kriminalitou na Internetu nové generace*. Kapitola 2 se podrobněji věnuje architektuře systému a jeho rozdělení na dílčích bloky. Podrobný funkční popis jednotlivých bloků, včetně jejich vzájemné komunikace je součástí kapitoly 3. Problematika spolehlivého přenosu zachycených dat v rámci LI systému je podrobněji popsána v kapitole 4. Technikám detekce dynamické identifikace odposlouchávaného cíle je věnována samostatná kapitola 5 se zaměřením na protokoly DHCPv4 a SLAAC.

Podrobnější popis dalších částí systému pro zákonné odposlechy vytvořených v rámci tohoto projektu je umístěn v samostatných technických zprávách. Mezi tyto zprávy patří:

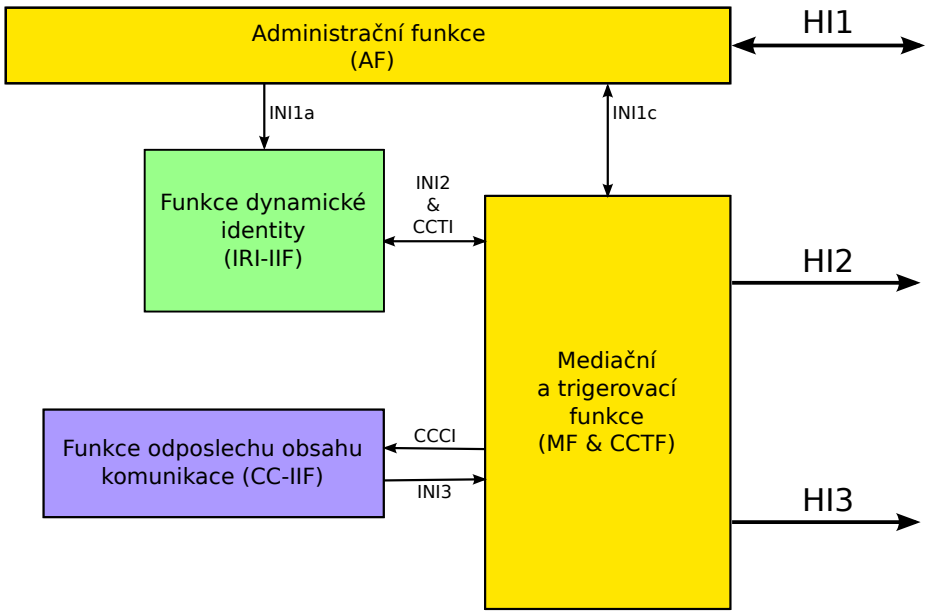
– Návrh prototypu mikro-sondy a vysokorychlostní sondy pro monitorování síťového provozu [17].

– Rekonstrukce a vizualizace datových toků zachycené komunikace [16].

2 Architektura systému

Jedním z cílů projektu *Moderní prostředky pro boj s kybernetickou kriminalitou na Internetu nové generace* je vytvoření prototypu systému pro sběr dat pro zákonné odposlechy (Lawful Interception System – LIS), který bude schopen komunikovat s vytvořenou vysokorychlostní sondou a mikro-sondou a poskytnout data pro rekonstrukci a klasifikaci datových toků. Tato kapitola je zaměřena na popis architektury vyvíjeného LIS, která vychází z doporučení ETSI [6,7,9,11,8,12,10,13] a je také inspirována architekturou LIS firmy CISCO publikovanou v RFC3924 [2].

Všimněte si, že cílem této aktivity není vytvořit zcela nový LIS, který by dokázal konkurovat současným komerčním řešením, ale vytvořit zjednodušený prototyp LIS, který by měl tvořit základní prostředí pro : a) vývoj nových technik dynamické identifikace uživatele v prostředí IPv6 sítí, b) sběr dat a vývoj nových metod v oblasti rekonstrukce a vizualizace zachyceného provozu a c) jako základní testovací prostředí pro vývoj mikro-sondy a vysokorychlostní sondy. Z těchto důvodů jsou některé části vyvíjeného LIS zjednodušeny oproti standardu ETSI. Tato zjednodušení jsou v textu podrobněji popsána.



Obrázek 2. Architektura prototypu systému pro zákonné odposlechy

Základní blokový diagram navrhovaného LIS je uveden na obrázku 2. Vstupní požadavky k odposlechu jsou přijímány skrze rozhraní HI1 a zpracovávány tzv. *Administrační funkcí* (*Administration Function* - AF). Blok AF nejdříve provádí kontrolu správného vyplnění povinných a volitelných položek – je zkontrolována unikátnost *Lawful Interception IDentifier* (LIID) [11,12], korektnost časových údajů a přítomnost odposlouchávající agentury v systému. Pokud je vše v pořádku, je odposlech zařazen do fronty čekajících odposlechů. AF je následně zodpovědná za korektní inicializaci a ukončení odposlechu, tj. konfiguraci ostatních částí systému (skrze rozhraní INI1a a INI1c) tak, aby bylo zajištěno, že budou zachycena všechna data přenášená v povoleném intervalu pro odposlech a zároveň nebudou zaznamenána žádná data mimo platnost odposlechu. Veškeré požadavky na přidání, či odebrání odposlechu jsou navíc z bezpečnostních důvodů v AF zaznamenávány.

Aktuální verze vyvíjeného LIS předpokládá manuální verzi rozhraní HI1 [11]. Oprávněný orgán (*lawful enforcement agency* - LEA) zasílá požadavky v papírové formě dopisem, či faxem, nebo elektronicky pomocí datových schránek. Za kontrolu pravosti příkazu k odposlechu a jeho korektní vložení do LIS je odpovědný pověřený zaměstnanec síťového operátora.

Identifikace jednotlivých uživatelů sítě (např. jejich IP adresa) se obecně může v průběhu času měnit. S ohledem na tuto skutečnost je v LIS k dispozici blok označený jako *Funkce dynamické identity* (*Intercepted Related Information - Internal Interception Function* - IRI-IIF). Úlohou bloku IRI-IIF je detekovat v síťovém provozu zprávy, které se vztahují ke změně identity uživatelů (např. protokoly DHCP, RADIUS apod.) a udržovat informace o aktuální identitě odposlouchávaných cílů. Každá změna identity odposlouchávaného cíle je pak neprodleně signalizována ostatním částem LIS (podrobnější informace o způsobu předávání zpráv jsou uvedeny v kapitole 3). Blok IRI-IIF dále vytváří informační zprávy (metadata) sledující začátky a konce odposlouchávaných spojení (např. uživateli byla přidělena IPv4 adresa protokolem DHCP nebo platnost adresy vypršela) a odesílá je skrze rozhraní INI2.

Aktuální verze vyvíjeného LIS podporuje odposlouchávání uživatele na základě MAC adresy nebo rozsahu IPv4 a IPv6 adres definovaného pomocí IP adresy sítě a masky (odposlech jedné IP adresy je speciální případ rozsahu). Při sledování identity uživatele jsou podporovány protokoly DHCPv4 a SLAAC. V současné době také připravujeme podporu protokolu DHCPv6 a možnost identifikace uživatele na základě uživatelského jména používaného pro autentizaci v rámci protokolu PPPoE a protokolu RADIUS. Více informací o dynamické identifikaci uživatele naleznete v sekci 5.

Blok označený jako *Funkce odposlechu obsahu komunikace* (*Content of Communication - Internal Interception Function* - CC-IIF) sleduje síťový provoz operátora a kopíruje veškerý obsah komunikace vztahující se k některé IPv4 nebo IPv6 adrese sledovaného cíle. Vstupní požadavky na započetí, či ukončení odposlechu jsou zasílány rozhraním CCCI. Zachycená data jsou odesílána rozhraním INI3.

Aktuální verze CC-IIF bloku je realizována softwarově, přičemž v následujících etapách projektu by měla být tato realizace nahrazena hardwarově akcelerovanou variantou v podobě mikro-sondy nebo vysokorychlostní sondy.

Centrální správa zachycených dat aktuálně probíhajících odposlechů je úlohou *Mediační funkce* (*Mediation Function* - MF). MF zpracovává jak metainformace od bloku IRI-IIF, tak i obsah zachycené komunikace od bloku CC-IIF. Oba tyto typy dat navzájem kombinuje a zasílá oprávněným orgánům skrze rozhraní HI2 resp. HI3. Z důvodu zjednodušení správy odposlechů je MF ve vyvíjeném LIS kombinovaná s trigerovací funkcí (*Content of Communication Trigger Function* - CCTF), která je zodpovědná za konfiguraci CC-IIF sond. Toto rozhodnutí bylo inspirováno architekturou LIS firmy CISCO [2].

Rozhraní HI2 slouží k předávání metadat o detekovaných síťových spojeních. Z důvodu zjednodušení nejsou ve vyvíjeném LIS tato metadata odesílána oprávněným orgánům přímo ve formě proudového zpracování, ale jsou ukládaná do textových souborů. Každý textový soubor obsahuje metadata pro jeden odposlech identifikovaný pomocí LIID. Na každém řádku souboru s meta informacemi je uložena jedna událost. U každé události je evidován čas výskytu, typ (*IRI begin*, *end*, *report*, nebo *continue*), *Communication identifier* (CID) [11,12], typ dynamické detekce (sekce 5), protokol síťové vrstvy a používaná adresa na síťové a linkové vrstvě.

Rozhraní HI3 slouží pro přeposílání obsahu komunikace podezřelých osob. Z důvodu zjednodušení se kopie zachycených dat ukládá do souboru ve formátu PCAP [1]. V každém vytvořeném souboru jsou uložena data patřící k jednomu odposlechu identifikovaném pomocí LIID [11].

3 Popis činnosti

V této kapitole se podrobněji zaměříme na činnost jednotlivých bloků LIS, na obsah předávaných zpráv a používané identifikátory.

V rámci LIS se používají následující identifikátory:

- *Lawful Interception IDentifier* (LIID): Řetězec alfanumerických znaků, který jednoznačně identifikuje odposlech [11]. LIID je dodáván LEA a je unikátní nejen v rámci LEA, ale také na mezinárodní úrovni [12] (řetězec obsahuje dvoupísmennou zkratku státu definovanou ISO 3166-1 [15]). Všechna data přenášena rozhraními HI2 a HI3 musí být tímto identifikátorem označena.
- *Network IDentifier* (NID): Identifikátor používaný v síťových protokolech k označení účastníků komunikace. V současnosti LIS podporuje MAC adresu, statickou IPv4 a IPv6 adresu nebo obecně rozsah adres definovaný adresou sítě a maskou.
- *HI1 IDentifier* (HI1ID): Jednoznačná identifikace odposlouchávané osoby, která je součástí požadavku na odposlech ze strany LEA. Tento identifikátor může být tvořen např. MAC adresou, IPv4/IPv6 adresou nebo zcela obecně např. rodným číslem, adresou trvalého bydliště apod. V případě obecné identifikace je úkolem pověřeného pracovníka na straně operátora tuto identitu

převést na takovou, jakou je možné použít ve zbytku systému. Pověřený pracovník obvykle využije interní databáze obsahující seznam zákazníků a převede HI1ID na NID.

- *Communication IDentifier* (CID): Identifikátor, který jednoznačně identifikuje komunikaci [11,12]. CID se skládá z:
  - unikátního ID operátora přiděleného LEA,
  - NID, kterého se odposlech týká,
  - *Communications Identity Number* (CIN) Identifikuje sezení, nebo komunikaci v rámci jednoho odposlechu, který je identifikován pomocí LIID,
  - *Delivery Country Code* (DCC) Dvoupísmenné označení země, kde se nachází MF.
- *System IDentifier* (SID): 32-bitový celočíselný identifikátor označující množinu odposlechů (LIID). SID jsme do LIS zavedli kvůli minimalizaci dat přenášených ze sond IRI-IIF a CC-IIF do MF&CCTF. Mapování SID na množinu LIID se uchovává v MF&CCTF, které také spravují přidělování nových SID. Výhodou použití SID je jeho pevná velikost, což umožňuje vytvářet jednodušší hlavičky pro data zasílaná z IRI-IIF a CC-IIF do MF&CCTF.

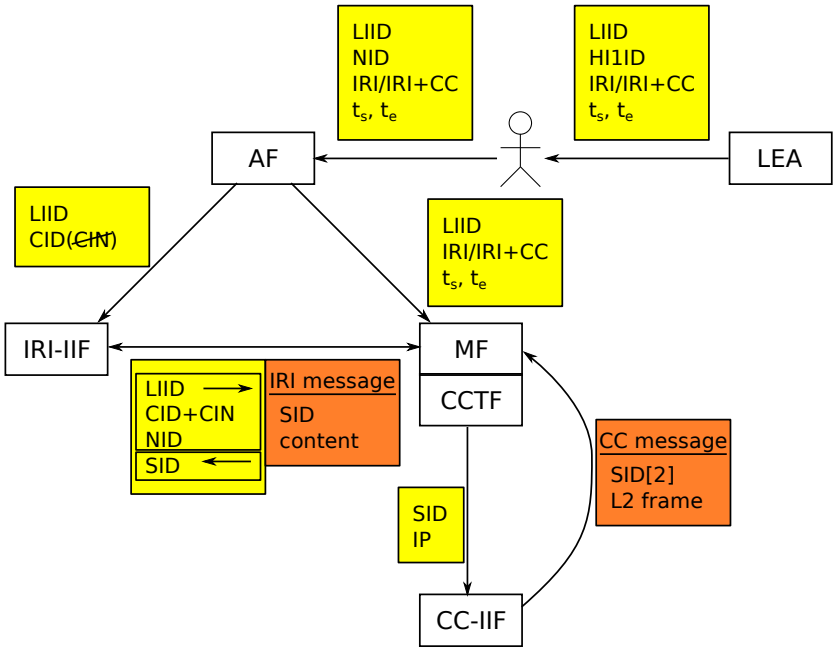
Komunikace a přenos jednotlivých identifikátorů mezi bloky LIS je znázorněna na obrázku 3. Následující část textu je věnována podrobnějšímu popisu této komunikace společně s činností jednotlivých bloků.

Scénář odposlechu

LEA požádá soud o odposlech specifického uživatele identifikovaného HI1ID. Žádost o odposlech musí mít definované časové období, ve kterém je možné data podezřelého odposlouchávat, tj. počátek odposlechu (t_s) a konec odposlechu (t_e). Součástí specifikace odposlechu je i požadavek, zda má LIS zachytávat pouze metainformace o spojeních realizovaných podezřelou osobou (IRI), nebo zda bude povolen sběr veškeré komunikace uživatele (IRI+CC). Po schválení odposlechu jsou jeho parametry předány pověřenému pracovníku poskytovatele připojení k Internetu (ISP), ke kterému je podezřelý připojen.

Pověřený pracovník ISP je taková osoba, která má pověření ke konfiguraci LIS. Po přijetí požadavku k odposlechu zkontroluje platnost odposlechu. S využitím interní databáze převede dodaný HI1ID na NID (dále v textu označovaném jako $\mathrm{NID_{In}}$) a provede zadání odposlechu do LIS. Vyvíjený LIS umožňuje zadávat odposlechy skrze webové rozhraní nebo pomocí specializovaných nástrojů z příkazové řádky.

Abychom si blíže ukázali činnost AF i dalších modulů systému, uvažujme příklad, ve kterém LEA_1 zažádá o IRI zprávy označené LIID = X týkající se konkrétního uživatele identifikovaného jménem a adresou. Uvažujme, že pověřený zaměstnanec na základě databáze uživatelů zjistí, že tento uživatel se do sítě připojuje pomocí zařízení s MAC adresou *00:11:22:aa:bb:cc*. Oproti tomu LEA_2 má zájem o odposlech veškeré komunikace počítače s konkrétní IPv4 adresou *10.0.0.1* a přeje si zachycená data označovat LIID = Y. Pověřený zaměstnanec



Obrázek 3. Výměna zpráv v rámci LIS vyvíjeného v rámci projektu *Moderní prostředky pro boj s kybernetickou kriminalitou na Internetu nové generace*

LEA	LIID	NID_{In}	t_s	t_e	typ
LEA_1	X	MAC: 00:11:22:aa:bb:cc	1.1.2011	1.1.2012	IRI
LEA_2	Y	IPv4: 10.0.0.1	2.1.2011	15.5.2012	IRI+CC

Tabulka 1. Příklad konfigurace LIS, kterou zadává pověřený pracovník ISP

nakonfiguruje LIS na odposlech adresy *10.0.0.1*. Parametry obou odposlechů, které zaměstnanec zadal do LIS jsou uvedeny v tabulce 1.

AF konfiguruje IRI-IIF a MF&CCTF 10 sekund před stanoveným začátkem odposlechu. Ke konfiguraci dalších částí LIS dochází co nejpozději, abychom minimalizovali dobu, po kterou je možné získat informace o odposlechu ze sond LIS. Na druhou stranu probíhá konfigurace LIS s krátkým časovým předstihem, aby se zamezilo ztrátě dat přenášených těsně po začátku odposlechu, ke které by došlo v průběhu konfigurace LIS. Při ukončení odposlechu čeká AF 10 sekund, než odebere odposlech ze zbytku systému. Toto zpoždění bylo zavedeno, aby bylo zaručeno bezproblémové zpracování dat zachycených těsně před povoleným koncem odposlechu. Blok MF&CCTF je zodpovědný za odstranění dat nasbíraných sondami mimo povolené rozmezí odposlechu.

AF posílá MF&CCTF pro každý odposlech jeho LIID, období platnosti (t_s a t_e) a informaci, zda odposlech povoluje odchycení celého obsahu komunikace. MF&CCTF si tyto informace ukládá do tzv. *tabulky LIID*. Podle této tabulky MF&CCTF ukládá odchycená data do správných souborů (určených různým LEA) a zahazuje data nasbíraná mimo interval platnosti odposlechů. Příklad obsahu tabulky LIID je znázorněn v tabulce 2.

LIID	t_s	t_e	typ	Soubor pro HI2	Soubor pro HI3
X	1.1.2011	1.1.2012	IRI	X.hi2	X.pcap
Y	2.1.2011	15.5.2012	IRI+CC	Y.hi2	Y.pcap

Tabulka 2. Příklad obsahu tabulky LIID uvnitř bloku MF&CCTF

Současně s konfiguraci bloku MF&CCTF posílá AF do bloku IRI-IIF pro každý odposlech jeho LIID a šablonu pro CID (CID, ve kterém není vyplněn CIN). Tabulka 3 ukazuje konfigurační data, která jsou poslána pro odposlechy *X* a *Y*.

LIID	Šablona pro CID (LEA, NID_{In}, CIN, DCC)
X	(LEA_1, MAC: 00:11:22:aa:bb:cc, -, CZ)
Y	(LEA_2, IPv4: 10.0.0.1/32, -, CZ)

Tabulka 3. Příklad konfigurace zasílané z AF do IRI-IIF

CC-IIF sondy pracují na síťové vrstvě ISO/OSI modelu [14] a zachytávají pakety odposlouchávaných osob na základě IP adresy (verze 4 i 6). Pokud je jako NID_{In} vstupního požadavku použit jiný identifikátor než IP adresa (např. MAC adresa), potom je tento identifikátor potřeba převést na IP adresu nebo rozsah IP adres (dále označovány jako NID_{CC}), které je CC-IIF sonda schopna zpracovávat. Tuto převodní funkci zajišťuje v LIS blok IRI-IIF, který sleduje komunikaci protokolů používaných pro přidělování IP adres (DHCP, SLAAC apod.) a udržuje si aktuální tabulku IP adres odposlouchávaných osob. Protože může být současně používáno několik možností pro dynamickou konfiguraci, navíc některé z nich umožňují přiřazení několika různých adres (např. SLAAC [19]), je nutné uvažovat případy, kdy se jeden odposlech vztahuje na několik různých IP adres.

Jeden podezřelý může být předmětem odposlechů několika různých LEA. Aby nedocházelo k duplicitnímu přeposílání stejných dat z IRI-IIF do MF&CCTF a především z CC-IIF do MF&CCTF (označených pouze různými LIID), přiřazuje vyvíjený LIS každému odposlechu nově zavedený identifikátor SID, který zastupuje množinu odposlechů identifikovaných pomocí LIID. Díky tomu je možné zachycená data označit jedním SID a jejich kopírování a rozesílání různým LEA

realizovat až v rámci bloku MF&CCTF. Komunikace mezi bloky CC-IIF resp. IRI-IIF a MF&CCTF se tak minimalizuje a vyžaduje jen nezbytně nutnou šířku pásma.

Jelikož se případná duplikace zachycených dat a jejich zasílání jednotlivým LEA provádí až MF&CCTF, je v témž bloku také realizován samotný algoritmus vytvářející relaci přiřazující LIID k SID a uložena tzv. *tabulka SID*, která tuto relaci reprezentuje. Pokaždé, když je do systému přidán nový požadavek na odposlech statické IP adresy nebo dojde k dynamické změně IP adresy odposlouchávaného cíle (např. skrze protokol DHCP), může obecně dojít ke změně relace přiřazující LIID a SID. Blok IRI-IIF, který tyto události sleduje, informuje blok MF&CCTF. MF&CCTF upraví odpovídajícím způsobem relaci přiřazující LIID a SID a o provedené změně informuje oba bloky IRI-IIF a CC-IIF, které novou nebo upravenou hodnotu SID potřebují pro označení zachytávaných dat nebo vytvářených metainformací.

Pro názornou ukázku uvedeného principu uvažujme následující posloupnost událostí, která by mohla nastat v případě našeho demonstračního příkladu:

1. Do LI systému jsou vloženy dva požadavky na odposlech X a Y definované výše.
2. Dne 1.1.2011 zahájí AF odposlech požadavku X a předá příslušné zprávy blokům IRI-IIF a MF&CCTF.
3. Sledováním protokolu DHCP detekuje blok IRI-IIF, že počítači s MAC adresou *00:11:22:aa:bb:cc* byla přiřazena IPv4 adresa *10.0.0.1*. Dále bylo zjištěno, že počítač začal používat IPv6 adresu *2001:db8::5* (např. sledováním protokolu DHCPv6).
4. IRI-IIF informuje MF&CCTF, že v odposlechu s LIID = X došlo k přiřazení IPv4 adresy *10.0.0.1* a že bude tato komunikace označována CID = *(LEA_1, MAC: 00:11:22:aa:bb:cc, 1, CZ)*.
5. MF&CCTF přiřadí pro tento odposlech SID_1 a zašle tuto informaci zpět bloku IRI-IIF, který si ji poznačí do svých vlastních tabulek.
6. Dále IRI-IIF ohlásí MF&CCTF, že v odposlechu s LIID = X došlo k přiřazení IPv6 adresy *2001:db8::5* a že bude používán CID = *(LEA_1, MAC: 00:11:22:aa:bb:cc, 2, CZ)*.
7. Jelikož nedochází ke shodě s některou z již existujících IP adres v tabulce SID, přiřadí MF&CCTF tomuto odposlechu SID_2 a zašle tuto informaci bloku IRI-IIF, který si aktualizuje interní tabulky.
8. Dne 2.1.2011 zahájí AF odposlech požadavku Y a předá příslušné zprávy blokům IRI-IIF a MF&CCTF.
9. IRI-IIF může okamžitě ohlásit odposlech IPv4 adresy *10.0.0.1* s CID např. *(LEA_2, IPv4: 10.0.0.1/32, 5, CZ)*.
10. Jelikož dochází ke shodě s již odposlouchávanou IPv4 adresou z požadavku X, přidělí MF&CCTF pro tento odposlech SID_1.

Obsah tabulky MF&CCTF bloku vyjadřující mapování identifikátorů LIID na SID po provedení výše uvedených kroků je znázorněn v tabulce 4. Podobně transformace identifikátorů $\mathrm{NID_{In}}$ na $\mathrm{NID_{CC}}$ a přiřazení odpovídajících SID prováděné uvnitř bloku IRI-IIF jsou uvedeny v tabulce 5.

SID	NID_{CC}	LIID	CID
1	IPv4: 10.0.0.1/32	X	(LEA_1, MAC: 00:11:22:aa:bb:cc, 1, CZ)
		Y	(LEA_2, IPv4: 10.0.0.1/32, 5, CZ)
2	IPv6: 2001:db8::5/128	X	(LEA_1, MAC: 00:11:22:aa:bb:cc, 2, CZ)

Tabulka 4. Příklad obsahu tabulky SID uvnitř bloku MF&CCTF. Tabulka vyjadřuje vztah mezi SID a odpovídajícím NID_{CC}. Pro každý SID je navíc v tabulce uložena množina LIID a pro každý LIID odpovídající CID, který byl konkrétní komunikaci přiřazen.

LIID	NID_{In}	NID_{CC}	SID	CID
X	MAC: 00:11:22:aa:bb:cc	IPv4: 10.0.0.1	1	(LEA_1, MAC: 00:11:22:aa:bb:cc, 1, CZ)
		IPv6: 2001:db8::5	2	(LEA_1, MAC: 00:11:22:aa:bb:cc, 2, CZ)
Y	IPv4: 10.0.0.1	IPv4: 10.0.0.1	1	(LEA_2, IPv4: 10.0.0.1/32, 5, CZ)

Tabulka 5. Příklad transformace NID_{In} na NID_{CC} a přiřazení SID uvnitř IRI-IIF

Zobecnění algoritmu mapování LIID na SID pro rozsahy

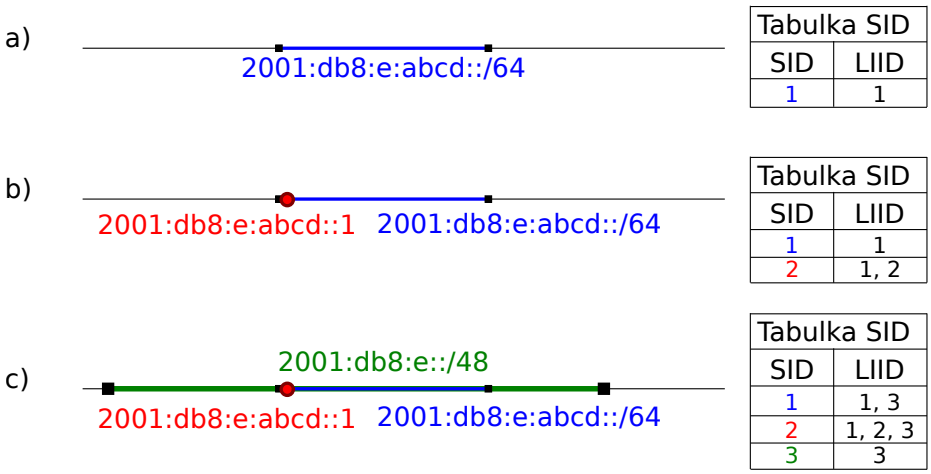
V předchozím příkladě jsme demonstrovali použití SID identifikátorů na příkladě, kdy je sledovaný cíl předmětem různých požadavků na odposlech z různých LEA. Pro jednoduchost jsme uvažovali situaci, kdy dochází k překryvu konkrétních IP adres. Při reálném nasazení LIS však může docházet i k situacím, kdy se překrývají nejen jednotlivé adresy, ale i rozsahy IP adres. Může se tak například stát, že jeden požadavek na odposlech sleduje rozsah IP adres, který je podrozsahem IP adres jiného požadavku. V následující části textu je proto relace mapování LIID na SID zobecněna na rozsahy a následně demonstrována na příkladě.

Při požadavku ze strany IRI-IIF na odposlech nového rozsahu adres R_N se kontroluje, zda již v tabulce SID neexistuje odpovídající větší nebo menší rozsah, a podle potřeby se vygeneruje nový SID a upraví informace uložené v tabulce SID. Konkrétně se přiřazení SID řídí následujícími pravidly:

- Pokud není žádná adresa z rozsahu adres R_N odposlouchávána, pak je vygenerován nový SID a odposlech je uložen do tabulky SID.
- Pokud je R_N vlastní podmnožinou jiného rozsahu adres, např. rozsahu R_V (tzn. R_N obsahuje nějakou část adres obsažených v R_V, ale ne všechny), pak se vygeneruje nový SID a do tabulky SID se k nově přidávanému odposlechu uloží i odposlechy vztahující se k R_V.
- Pokud se již odposlouchává stejný rozsah adres, pak se pouze nový odposlech uloží do tabulky SID ke stávajícím odposlechům pro daný rozsah a všechny podrozsahy adres.
- Při existenci odposlechů pro menší rozsahy adres se vytvoří nový SID pro nově vkládaný rozsah a zároveň se informace o nově přidávaném odposlechu uloží i ke všem odposlouchávaným podrozsahům adres.

LIS podporuje pouze rozsahy specifikované maskou, tzn. pro všechny platné rozsahy adres A, B platí $A \cap B = \emptyset \vee A \subseteq B \vee B \subseteq A$ a nemůže tedy nastat jiný případ než zmíněné. Mechanismus přidělování SID a správy *tabulky SID* je demonstrován na následujícím příkladu:

1. Předpokládejme například, že je v systému nejdříve aktivován odposlech $LIID_1$ vztahující se na adresový rozsah *2001:db8:e:abcd::/64*. IRI-IIF požádá o přidělení nového SID a MF&CCTF přidělí např. SID_1. Do tabulky SID MF&CCTF uloží, že data označená SID_1 se vztahují k odposlechu $LIID_1$. Situaci znázorňuje část a) obrázku 4.
2. V případě, že je následně aktivován odposlech $LIID_2$ vztahující se na jedinou IPv6 adresu *2001:db8:e:abcd::1*, pak je přidělen nový SID. Předpokládejme, že má nově přidělený SID pro odposlechnutá data vztahující se k této IPv6 adrese hodnotu SID_2. Pak MF&CCTF uloží do tabulky SID, že data označená SID_2 se vztahují k odposlechům $LIID_1$ a $LIID_2$. Změna je zachycena v části b) obrázku 4.
3. Pokud by byl následně aktivován odposlech $LIID_3$ vztahující se na celý rozsah *2001:db8:e::/48* a uvažujme, že pro odposlechnutá data vztahující se k tomuto IPv6 rozsahu byl přidělen SID_3, pak MF&CCTF uloží do tabulky SID, že data označená SID_3 se vztahují k odposlechu $LIID_3$. MF&CCTF dále doplní, že data označená SID_1 i SID_2 se vztahují také k odposlechu $LIID_3$. Tento stav je zachycen v části c) obrázku 4.



Obrázek 4. Příklad postupného přidávání odposlechů a vytváření vazeb mezi rozsahy IPv6 adres, SID a LIID

Pokud IRI-IIF vytváří metainformace nebo pokud sonda CC-IIF zachytí paket vztahující se k některé sledované IP adrese, je potřeba zjistit, zda tato adresa

patří do některého z odposlouchávaných rozsahů. Pokud je nalezeno dokonce několik rozsahů (navzájem vnořených) s různými SID, potom je vždy nutné nalézt ten nejmenší z nich (nejvíce zanořený) a jeho SID použít pro označení zachycených dat. V uvedeném příkladě jsou data z rozsahu *2001:db8:e::/48* označována SID_3, až na podrozsah adres *2001:db8:e:abcd::/64*. Data vztahující se ke komunikaci stroje s IPv6 adresou *2001:db8:e:abcd::1* jsou označována SID_2. Ostatní data vztahující se ke komunikaci s alespoň jednou IPv6 adresou z rozsahu *2001:db8:e:abcd::/64* jsou označována SID_1.

Sondy CC-IIF jsou konfigurovány pomocí IP adres (příp. rozsahu IP adres), které mají odposlouchávat a SID, kterým odposlechnutá data označovat. U sond CC-IIF se předpokládá zapojení za tapem, hubem, nebo SPAN portem. Z kopie provozu síťové linky sondy odstraňují pakety, které nejsou předmětem některého z odposlechů. Na svůj výstup propustí pouze rámce, které obsahují IP datagramy z některého z rozsahů, které jsou aktuálně odposlouchávány. Všimněte si, prosím, že odchycená data mohou sondy CC-IIF označit až dvěma identifikátory SID, protože se odposlech může vztahovat jak ke zdrojové, tak k cílové IP adrese.

4 Komunikační protokoly mezi částmi systému

Součástí vývoje prototypu LIS (jak softwarové, tak hardwarové verze) je i otázka návrhu a implementace komunikačního protokolu mezi jednotlivými částmi systému. Komunikačním protokolem se v rámci návrhu LIS myslí protokol na transportní a vyšší vrstvě ISO/OSI síťového modelu [14]. Samotná norma ETSI nedefinuje výběr protokolu, pouze poskytuje příklady možných způsobů implementace komunikační vrstvy s ohledem na její jednoduchost a rychlost. V zásadě však norma definuje dva základní požadavky, které musí komunikace v LIS splňovat:

- spolehlivost
- bezpečnost

Tyto požadavky jsou směrovány hlavně na rozhraní INI2 od IRI-IIF a INI3 od CC-IIF směrem k MF. Spolehlivost doručování dat je vyžadována z důvodu zajištění nezpochybnitelnosti zachycených dat, která by mohla být použita i v případném soudním líčení. Zachycená data je nutno přeposílat bez modifikace do MF pro uložení a další zpracování do vhodnější podoby (již zmíněna rekonstrukce a vizualizace těchto dat).

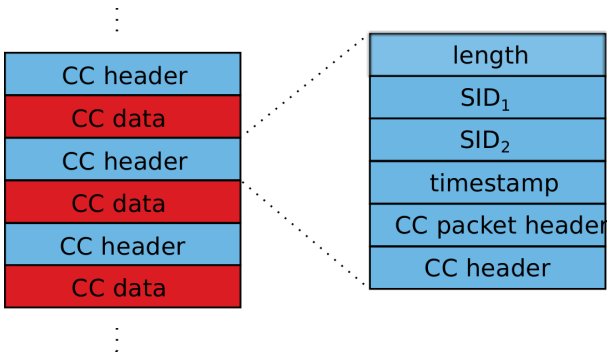
Sonda může být připojena k MF zařízení dvojím způsobem: *dedikovaným spojem* nebo *zapojením do infrastruktury operátora*. V případě dedikovaného spoje se jedná o přímé spojení sondy se zařízením MF. Data neproudí přes další síťová zařízení operátora. V druhém případě se jedná o zapojení sondy jako dalšího síťového zařízení do síťové infrastruktury operátora. Data se v tomto případě posílají přes síť operátora, než se dostanou k zařízení MF. V prvním případě odpadá nutnost zajistit ochranu dat šifrováním. To vede i na jednodušší zpracování zachycených dat na výstupu sondy pro jejich přeposlání ven směrem k MF. Ve druhém případě jsou data vystavena riziku odchycení a možného odhalení odposlechu, protože jsou přenášena přes síť operátora, která je mimo kontrolu LIS.

Z těchto důvodů (a protože se jedná o citlivá data uživatelů) je nutno je chránit pomocí šifrování. V obou případech je však nutno řešit spolehlivost přenosu dat, aby nedocházelo k jejím ztrátám.

Norma ETSI [8] uvádí příklady zapouzdření zachycených dat pro účely jejich dalšího přeposlání. Tyto příklady jsou uvedeny především s ohledem na jednoduchost implementace jak na straně CC-IIF, tak MF. Jedná se o jednoduché příklady užití protokolu UDP [21] na transportní či protokolu RTP [23] na aplikační vrstvě. V případě UDP se skutečně jedná o nejjednodušší možnost volby transportního protokolu, protože účelem UDP je pouze diferenciace aplikace naslouchající na určitém portu v příslušném komunikačním bodu a volitelně kontrola integrity pomocí kontrolního součtu. UDP sám o sobě nezahrnuje žádnou detekci ztráty, mechanismus znovu zasílání ani žádné jiné pokročilé techniky použity v komplexnějších protokolech (např. TCP).

ETSI [8] uvádí RTP jako příklad protokolu pro potřeby přeposílání dat na vyšší než transportní vrstvě, kvůli využití určitých položek hlavičky pro potřeby LIS a také pro nenáročnou implementaci. RTP pro komunikaci využívá služby protokolu nižší vrstvy – a to protokol UDP.

V rámci řešení způsobu zapouzdření je potřebné se zabývat i problémem identifikace zachycených a přenášených dat po jejich příjmu na straně MF. V MF musí proběhnout identifikace a spárování přijatých dat s odposlechem již dříve uloženým v tabulce SID (viz sekce 3). Z těchto důvodů je nutno k odposlechnutým paketům přidávat před odesláním z CC-IIF další identifikační informace. Sonda samotná bude zachytávat data už od linkové vrstvy, ke kterým se přidá nová hlavička pro potřeby následné identifikace v MF. V CC-IIF vznikne proud dat, který obsahuje zachycené pakety (*CC data*) a před každým z nich pomocnou hlavičku (*CC header*), jak je znázorněno na obrázku 5. Alternativně můžeme tato data považovat za posloupnost zpráv, kde se každá zpráva skládá z *CC header* a *CC data*.



Obrázek 5. Struktura zapouzdření zachycených rámců

Struktura *CC header* pro jeden paket zahrnuje:

1. délku přenášeného bloku (2B)
2. 2x SID identifikátor (8B)
3. časovou značku (4B)
4. směr toku zprávy přes CC-IIF (1B)
5. IP adresy, porty a transportní protokol zachyceného paketu (16B, 16B, 2B, 2B, 2B)

Při přenášení dat do MF je nutné uvažovat i maximální velikost IP datagramu, který je možné počítačovou sítí přenést, aniž by překročil maximální povolenou velikost (MTU). MTU není ve všech sítích stejná, ale závisí na použité technologii na linkové vrstvě ISO/OSI modelu [14]. Použití protokolu IPv6 však zaručuje MTU alespoň 1280 B [4]. Pro zjištění MTU na cestě od zdroje komunikace k cíli (PMTU) je definován algoritmus využívající zprávy ICMPv6 [18]. Díky záruce minimální MTU a tím pádem i PMTU mohou jednoduchá zařízení odesílat všechny IPv6 datagramy o velikosti 1280 B nebo menší a nemusí implementovat protokol zjišťování PMTU. Toho je možné využít v počáteční fázi implementace mikro-sondy.

Při výběru vhodného transportního protokolu je potřeba zohlednit i charakter přenášených dat. Například UDP přenáší data ve formě zpráv. Cílová aplikace vždy obdrží data ve stejných blocích, jako je zdroj odeslal. Nicméně není zaručeno doručení ve stejném pořadí, jako byly zprávy odeslány. Některé zprávy se mohou ztratit a některé na cestě duplikovat. Pro usnadnění návrhu protokolu pro přenos dat v rámci LIS je vhodné nedělit jeden odposlechnutý paket do několika UDP zpráv. Pro zajištění spolehlivosti přenosu by bylo nutné *CC header* doplnit o položku umožňující znovu seřazení zpráv do původního pořadí a také detekci ztráty a duplikace zprávy.

Naproti tomu TCP přenáší data ve formě proudů a cílová aplikace obdrží data obvykle v takové velikosti bloků, o jaký si požádá, nezávisle na tom, v jakých blocích byla data odeslána. Pomocí TCP je tedy možné přenášet zachycená data jako proud a efektivně využívat šířku pásma tím, že se několik kratších paketů odešle v rámci jednoho TCP segmentu, případně se jeden odposlechnutý paket rozdělí do více TCP segmentů.

Pro návrh sondy vyvíjeného LIS proběhla analýza protokolů použitelných pro komunikační vrstvu. Analyzovali jsme protokoly UDP a nadstavby nad ním, protokol RTP a probíhá studium detailních vlastností protokolu TCP. Pro implementaci je snaha vyjít ze schválených standardů RFC pro zajištění kompatibility komunikace.

V první fázi procesu oživování sondy se soustředíme na jednoduchou implementaci komunikačního protokolu nad UDP z důvodu ověření funkcionality komunikace. Pro další vývoj je však snahou komunikaci zajistit způsobem známým z implementací protokolu TCP. Protože mikro-sonda je hardwarové zařízení, proces implementace bude nutné rozdělit do hardwarové a softwarové části. Cílem je navrhnout a implementovat tzv. *offload engine*, který by umožňoval realizovat spolehlivou komunikaci.

5 Dynamická identifikace uživatele

Pro účely zákonných odposlechů se komunikace uživatele obvykle identifikuje a zaznamenává na základě IP adresy. Uživateli může být IP adresa přiřazena staticky nebo dynamicky. V případě statické adresy stačí do LIS přidat odposlech konkrétní IP adresy. Pokud se však v síti přidělují IP adresy dynamicky, může být uživateli během platnosti odposlechu přiřazeno dokonce několik různých IP adres. Z toho důvodu je součástí LIS blok IRI-IIF pro dynamickou identifikaci uživatele, který analyzuje komunikaci založenou na protokolech pro přidělování IP adres jako jsou např. DHCPv4 [5], SLAAC [24], DHCPv6 [3], RADIUS [22] apod.

Vyvíjená IRI-IIF požívá dvě rozhraní pro komunikaci s ostatními bloky LIS: 1) Rozhraní INI1a sloužící pro administraci odposlechů (přidávání nebo odebírání odposlechů apod.); 2) Rozhraní INI2 pro komunikaci s blokem MF&CCTF, které slouží pro přenos metainformací a servisních informací, jejichž podrobný popis lze nalézt v kapitole 3). Metainformace jsou organizovány do tzv. *IRI zpráv*, jejichž podrobný seznam je uveden v tabulce 6.

Zpráva	Popis zprávy
IRI report	Zpráva je zasílána v okamžicích, kdy klient žádá o přiřazení IP adresy nebo mu byla žádost o přiřazení z jakéhokoliv důvodu odmítnuta.
IRI begin	Zpráva je zasílána v okamžiku, kdy byla klientovi přiřazena IP adresa.
IRI continue	Zpráva je zasílána, je-li klientovi prodloužena doba, po kterou může používat danou IP adresu.
IRI end	Zpráva oznamuje, že klientovi byla odebrána IP adresa nebo se jí sám vzdal.

Tabulka 6. IRI zprávy zasílané skrze rozhraní INI2

Architektura bloku IRI-IIF je znázorněna na obrázku 6 a skládá se z jádra *IRI-IIF Core* a modulů pro analýzu jednotlivých protokolů. Moduly analyzují provoz na odposlouchávané lince, zpracovávají podporované protokoly pro přiřazení IP adres (DHCPv4, SLAAC, apod.) a předávají do jádra *IRI-IIF Core* informace, na jejichž základě lze snadno generovat výstupní IRI zprávy. Podrobnější struktura zprávy předávané do jádra *IRI-IIF Core* a její parametry jsou uvedeny v tabulce 7.

Jádro *IRI-IIF Core* se stará o komunikaci IRI-IIF s okolními bloky LIS tj. a) přijímá vstupní požadavky na odposlechy a udržuje si tabulku právě probíhajících odposlechů, b) přijímá zprávy od jednotlivých modulů, provádí jejich filtraci na základě aktuálně probíhajících odposlechů a generuje výstupní IRI zprávy. Důvodem pro modulární architekturu bloku IRI-IIF je to, aby pro přidání dal-

šího podporovaného protokolu stačilo vytvořit pouze daný modul a nebyly nutné zásahy do ostatních částí IRI-IIF.

Parametr	Popis parametru
Module ID	Identifikator modulu, který vygeneroval danou zprávu pro *IRI-IIF Core*.
IRI type	Typ zaznamenané události, který je významově schodný s typem IRI zprávy. Může tedy nabývat hodnot *report*, *begin*, *continue* a *end*.
Timestamp	Časová značka udává čas události, která způsobila vygenerovaní zprávy. Nejčastěji se jedná o čas odposlechu paketu, který vyvolal danou událost. V jiných případech se může jednat například o čas vypršení platnosti přiřazené IP aresy apod.
IP address	IPv4 nebo IPv6 adresa spjatá s událostí. Například IP adresa přiřazená klientovi v rámci zprávy IRI begin. (Může se jednat i o rozsah adres.)
Identifiers	Množina všech identifikátorů spjatých s danou událostí. V současné verzi se zde vkládá pouze MAC adresa, předpokládá se však rozšíření i o další identifikátory jako je uživatelské jméno pro RADIUS apod.

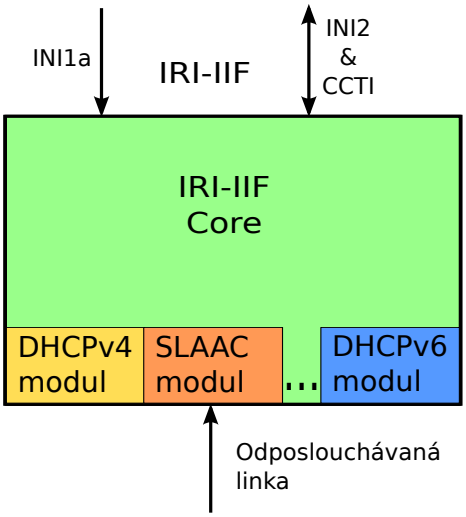
Tabulka 7. Parametry zprávy předávané mezi moduly a jádrem IRI-IIF bloku

Přidělování dynamické IP adresy může probíhat v intervalech několika hodin nebo dokonce dní, přičemž požadavek na odposlech může přijít kdykoli. To představuje problém pro LIS, který by nemusel být schopen odposlouchávat komunikaci hned od začátku odposlechu, ale až po dalším přiřazení IP adresy odposlouchanému uživateli. Z toho důvodu se ve vyvíjeném *IRI-IIF Core* udržuje tabulka všech přiřazených IP adres. V současné době IRI-IIF využívá dvou tabulek, jednu pro IPv4 a druhou pro IPv6. Do tabulek se ukládají dvojice IP adresa a MAC adresa. Zatímco u tabulky pro IPv4 se předpokládá vazba jedna IPv4 adresa na jednu MAC adresu, v tabulce pro IPv6 může být jedné MAC adrese přiděleno i více IPv6 adres.

Současná verze IRI-IIF bloku obsahuje dva moduly, první pro DHCPv4 a druhý pro SLAAC. Oba jsou podrobněji popsány v následujících podkapitolách.

5.1 DHCPv4

Protokol DHCPv4[5] se používá pro dynamické přidělování IPv4 adresy. Toto přidělení adresy je založeno na komunikaci mezi klientem a DHCP serverem. DHCP server (dále jen server) je zpravidla umístěn ve stejné podsíti jako klient a často je obsažen přímo ve směrovači. Server má k dispozici určitý rozsah adres, ze kterého přiděluje jednotlivým klientům adresy. Přidělení IPv4 adresy pomocí protokolu DHCP probíhá obvykle v následujících krocích:



Obrázek 6. Architektura IRI-IIF bloku

1. Klient pomocí všesměrového vysílání (broadcast) požádá o přidělení adresy (zasláním zprávy *DHCP discover*).
2. Server zašle klientovi odpověď s navrhovanou adresou (zasláním zprávy *DHCP offer*).
3. Klient všesměrovým vysíláním požádá o navrhovanou adresu (zasláním zprávy *DHCP request*).
4. Server zašle klientovi potvrzení o přidělení adresy (zasláním zprávy *DHCP ack*).

Platnost přidělené IPv4 adresy je časově omezena na hodnotu uvedenou v položce *lease time*. Po vypršení této doby již nesmí klient přidělenou adresu používat, pokud si před jejím vypršením úspěšně nepožádal o její prodloužení. Žádost o prodloužení adresy probíhá v následujících krocích:

1. Klient všesměrovým vysíláním požádá o prodloužení adresy (zasláním zprávy *DHCP request*). Klient jinými slovy žádá o stejnou adresu, kterou měl doposud přidělenu.
2. Server zašle klientovi buď potvrzení o přidělení adresy (zasláním zprávy *DHCP ack*) nebo odmítnutí tohoto požadavku (zasláním zprávy *DHCP nack*).

Uvedený scénář týkající se prodloužení platnosti adresy se může opakovat vícekrát. V případě odmítnutí požadavku o prodloužení adresy (*DHCP nack*) nemá klient jinou možnost, než znovu požádat server o nabídku dostupných adres (zasláním zprávy *DHCP discover*) a na základě nové nabídky (*DHCP offer*) si vybrat adresu jinou. Pokud již klient nebude adresu dále používat, může

(volitelně) sám předat serveru informaci o uvolnění adresy (zasláním zprávy *DHCP release*), čímž se adresa vrátí zpět mezi nepřiřazené adresy a DHCP server jí může přidělit ostatním klientům.

Činnost IRI-IIF

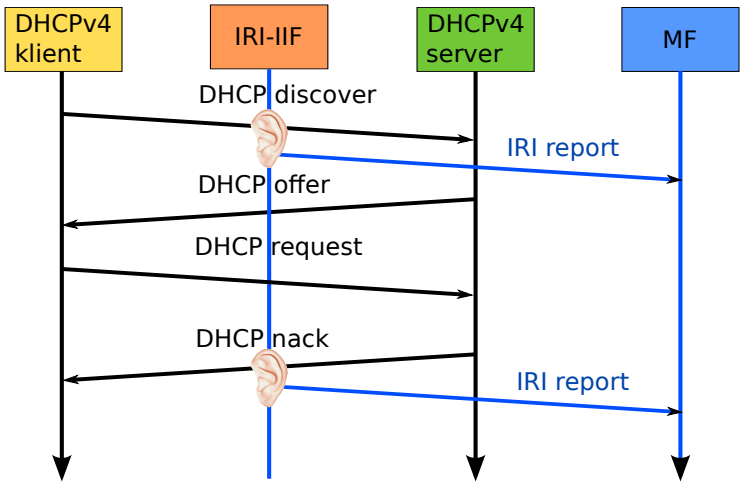
Blok IRI-IIF analyzuje uvedené DHCP zprávy a na jejich základě udržuje aktuální tabulku přidělených IPv4 adres společně s jejich dobou platnosti a generuje pro odposlouchávané adresy příslušné IRI zprávy. Činnost bloku IRI-IIF se řídí podle následujících pravidel:

1. Při zachycení žádosti *DHCP discover* odeslané z odposlouchávané MAC adresy generuje blok IRI-IIF zprávu typu *IRI report*.
2. Blok IRI-IIF páruje žádost *DHCP request* s odpovědí *DHCP ack*, či *DHCP nack*. Pokud server odpoví kladně (*DHCP ack*), může se jednat buď o přidělení nové adresy nebo o prodloužení předchozí adresy. IRI-IIF je schopen tyto dva stavy rozlišit tak, že zkontroluje obsah tabulky přidělených IPv4 adres a provede následující operace:
   (a) Pokud není nově přidělená adresa v tabulce obsažena, generuje IRI-IIF zprávu typu *IRI begin* a přidá adresu do tabulky společně s její dobou platnosti.
   (b) V opačném případě se jedná o prodloužení platnosti adresy, blok IRI-IIF generuje zprávu typu *IRI continue* a aktualizuje v tabulce dobu platnosti této adresy.
3. Pokud je požadavek o přidělení adresy zamítnut zasláním odpovědi *DHCP nack*, generuje IRI-IIF zprávu typu *IRI report*.
4. Platnost přiřazené IP adresy končí zasláním zprávy *DHCP release* ze strany klienta, popř. vypršením doby platnosti. Po ukončení platnosti, již nesmí být IP adresa klientem dále používaná. Úlohou IRI-IIF bloku je platnost adres periodicky kontrolovat a v okamžiku ukončení některé z nich vygenerovat zprávu typu *IRI end*.
5. V případě, že v době zahájení odposlechu již sledovaný cíl komunikuje s přidělenou IP, blok IRI-IIF tuto adresu dohledá v tabulce přidělených adres a generuje zprávu typu *IRI begin*.

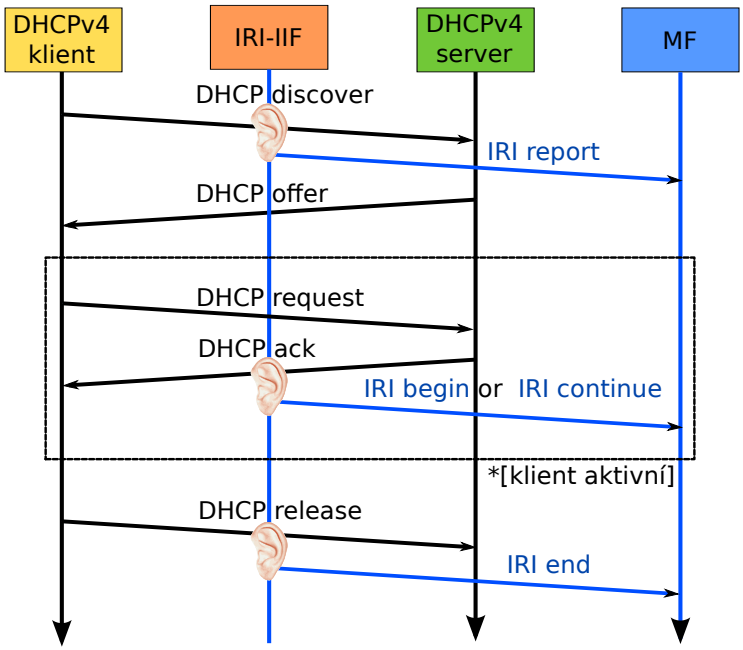
Uvedené IRI zprávy se generují pouze v případě, že daná událost souvisí s některým z odposlouchávaných cílů. Názorně jsou tato pravidla zachycena na obrázcích 7 a 8.

5.2 Bezstavová autokonfigurace adres (SLAAC)

Bezstavová autokonfigurace adres (SLAAC) slouží pro automatické přidělení IPv6 adresy koncové stanici [24]. Přidělení probíhá mezi klientem a směrovačem pomocí protokolu ICMPv6. Na rozdíl od jiných protokolů pro dynamické přiřazení adresy není klientovi přidělena adresa přímo SLAAC serverem, ale klientovi jsou zaslány pouze konfigurační udaje na jejichž základě si klient sám vygeneruje



Obrázek 7. Generování IRI zpráv při zamítnutí přidělení IP adresy pomocí DHCP



Obrázek 8. Generování IRI zpráv při přidělení IP adresy pomocí DHCP

IPv6 adresu. Mezi tyto konfigurační údaje patří zejména prefix dané sítě (horní

část IPv6 adresy) a doba platnosti tohoto prefixu. Postup pro přidělení IPv6 adresy probíhá v následujících krocích:

1. Klient si nejprve vygeneruje tzv. *lokální* IPv6 adresu tak, že si sám zvolí (např. náhodně) identifikátor rozhraní (spodní část IPv6 adresy o velikosti 64 bitů) a horní část nastaví na prefix fe80::/64. Následně je nutné ověřit, zda někdo takto vygenerovanou adresu již nepoužívá. K ověření unikátnosti použije klient protokol *Neighbor Discovery* [20], jehož podrobnější popis je uveden níže.
2. Aby si klient mohl přiřadit také globální IPv6 adresu, musí nejprve znát prefix podsítě, ve které se nachází. O tuto informaci může požádat nejbližší směrovač(e) opět skrze protokol *Neighbor Discovery* (podrobnější popis je uveden níže).
3. Jakmile má klient k dispozici prefix podsítě, vygeneruje si globální adresu buď na základě lokální adresy - záměnou prefixu lokálního adresy za prefix dané podsítě nebo zcela nezávisle na lokální adrese. V prvním případě se identifikátor rozhraní (spodní část adresy) globální adresy shoduje s identifikátorem rozhraní lokální adresy. Ve druhém případě se identifikátory rozhraní navzájem liší. Pro takto vytvořenou globální adresu je nezbytné opět ověřit, zda není používána jiným klientem sítě.

Ověření lokální i globální IPv6 adresy (Duplicate Address Detection – DAD) pomocí protokolu *Neighbor Discovery* probíhá v následujících krocích:

1. Klient vloží vygenerovanou adresu do zprávy s označením *Neighbor solicitation* a tuto zprávu zašle na skupinovou (multicast) adresu odvozenou od vygenerované IPv6 adresy.
2. Pokud klientovi do určitého časového intervalu nepřijde odpověď ve formě zprávy *Neighbor advertisement* oznamující, že vygenerovaná adresa je již používána, považuje klient adresu za unikátní a začne ji používat.

Získání informace o prefixu sítě pomocí protokolu *Neighbor Discovery* probíhá v následujících krocích:

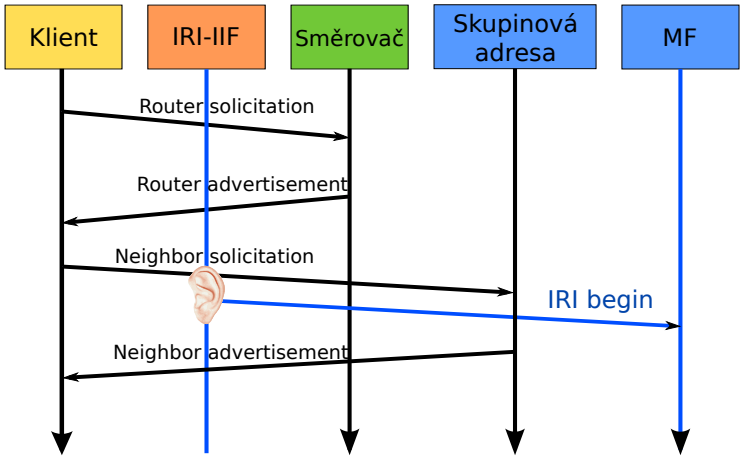
1. Klient vyšle pomocí skupinového vysílání (multicast) žádost o konfigurační údaje skrze zprávu *Router solicitation*. Tato zpráva je odeslána na skupinovou (multicast) adresu ff02::2, kde jsou zařazeny všechny směrovače v dané podsíti.
2. Směrovač(e) odpovídají zprávou *Router advertisement* obsahující informace o prefixu podsítě a době jeho platnosti.

Uvedeným způsobem si může klient vygenerovat i více globálních IPv6 adres. Doba platnosti přiřazené globální adresy je omezena na hodnotu zaslanou v rámci zprávy *Router advertisement* (RA) v položce *Valid lifetime* [20]. Klient si může prodlužovat platnost přiřazené adresy po dobu, kdy směrovače periodicky zasílají RA. Mechanismus, kterým by klient oznámil ostatním síťovým uzlům, že již adresu nebude používat není definován.

Činnost IRI-IIF

Z pohledu realizace bloku IRI-IIF je nezbytné uvažovat následující problémy:

1. Odposlouchávaný cíl může komunikovat jak s použitím lokálních adres (v rámci interní sítě), tak s využitím globálních adres. Těchto lokálních i globálních adres si klient může vygenerovat libovolné množství.
2. Výše uvedený postup je implementován korektně jen na určitých systémech. V průběhu testování byly např. nalezeny systémy zcela ignorující jak DAD zprávu *Neighbor solicitation* (klient nereaguje na požadavek o ověření adresy), tak zprávu *Neighbor advertisement* (klient nereaguje na informaci o již používané adrese). Pro IRI-IIF se jedná o nepříjemný problém, neboť nelze spolehlivě párovat požadavek a odpověď na ověření adresy.
3. Pokud bylo pomocí DAD jednou úspěšně ověřeno, že se vygenerovaná IPv6 adresa v síti nevyužívá, platnost adresy se automaticky prodlužuje, pokud je v síti nadále inzerován použitý prefix. Doba platnosti lokálních adres je obecně neomezená. V rámci protokolu SLAAC není přítomen mechanismus, kterým by koncová stanice informovala o uvolnění adresy. IRI-IIF proto nedokáže určit, kdy se daná adresa přestala používat a nemůže ji jednoduše odstranit z tabulky přiřazených adres.



Obrázek 9. Generování IRI zprávy při bezestavové autokonfiguraci IPv6 adresy.

S ohledem na výše uvedené problémy byla první verze bloku IRI-IIF realizována následovně: IRI-IIF předpokládá, že ověřování unikátnosti pomocí zprávy *Neighbor solicitation* proběhne vždy v pořádku a klient ověřovanou IPv6 adresu začne používat. IRI-IIF proto sleduje především zprávy *Neighbor solicitation* a na jejich základě přidává IPv6 adresy do tabulky aktuálně přiřazených adres a

generuje zprávu typu *IRI begin*. Ostatní zprávy typu *IRI continue*, *IRI end* a *IRI report* nejsou prozatím podporovány vzhledem k jejich nejednoznačnému určení. Podrobněji se budeme uvedenými problémy a způsoby jejich řešení zabývat v další fázi výzkumu. Názorně je situace ilustrována na obrázku 9.

6 Shrnutí

Cílem této technické zprávy bylo popsat návrh architektury prototypu LIS vyvíjeného v rámci projektu *Moderní prostředky pro boj s kybernetickou kriminalitou na Internetu nové generace*. Tento prototyp by měl dále sloužit jako základní prostředí pro: a) vývoj nových technik dynamické identifikace uživatele v prostředí IPv6 sítí, b) sběr dat a vývoj nových metod v oblasti rekonstrukce a vizualizace zachyceného provozu a c) jako základní testovací prostředí pro vývoj mikro-sondy a vysokorychlostní sondy.

Architektura navrženého LIS je založena na standardech ETSI a je rovněž inspirována architekturou firmy CISCO publikovanou v RFC 3924 [2]. S ohledem na použití navrhovaného systému nebylo nutné implementovat všechna doporučení dle norem ETSI a rozhraní HI1, HI2 a HI3 byla proto v rámci prototypu zjednodušena. Podrobný popis navrhované architektury je uveden v kapitole 2. V průběhu návrhu architektury bylo nezbytné vyřešit několik důležitých problémů mezi které patří např. způsob předávání zpráv mezi jednotlivými bloky systému tak, aby nedocházelo k duplicitním přenosům odposlouchávaných dat. Důmyslná metoda založená na mapování požadavků na jednoznačné identifikátory typu SID je podrobněji popsána v kapitole 3. Dále jsme se zaměřili na problematiku spolehlivého a zabezpečeného přenosu zachycených dat mezi sondou a mediačním zařízením. Základní požadavky kladené na tento druh přenosu jsou shrnuty v kapitole 4. Zvláštní pozornost byla věnována problematice dynamické identifikace odposlouchávaného cíle v prostředí IPv4 a IPv6 sítí. Architektura bloku IRI-IIF, který tuto funkci v LI systému plní, byla navržena modulárně s ohledem na podporu různých protokolů pro přidělování IP adres. V současné verzi systému byly implementovány dva moduly: první pro analýzu protokolu DHCPv4 a druhý pro analýzu protokolu SLAAC. Podrobněji je tato problematika popsána v kapitole 5.

V dalších fázích projektu se zaměříme na rozšíření skupiny modulů pro dynamickou identifikaci uživatelů, pozornost bude věnována zejména protokolům používaným v prostředí IPv6 jako jsou DHCPv6 nebo RADIUS. Pokračovat budou také práce na vylepšování současné architektury, její modifikace s ohledem na připojení mikro-sondy resp. vysokorychlostní sondy a zajištěním zabezpečeného přenosu zachycených dat.

Reference

1. TCPDUMP/LIBPCAP public repository. [[online]], citováno 2011-10-24. URL http://www.tcpdump.org/

2. Baker, F.; Foster, B.; Sharp, C.: *RFC 3924 Cisco Architecture for Lawful Intercept in IP Networks*. 10 2004.
URL http://tools.ietf.org/html/rfc3924
3. Bound, J.; Volz, B.; Lemon, T.; aj.: *Dynamic Host Configuration Protocol for IPv6 (DHCPv6)*. 6 2003.
URL http://tools.ietf.org/html/rfc3315
4. Deering, S. E.; Hinden, R. M.: *RFC 2460 Internet Protocol, Version 6 (IPv6) Specification*. 12 1998.
URL http://tools.ietf.org/html/rfc2460
5. Droms, R.: *Dynamic Host Configuration Protocol*. 3 1997.
URL http://tools.ietf.org/html/rfc2131
6. European Telecommunications Standards Institute: *ETSI TR 101 943: Telecommunications security; Lawful Interception (LI); Concepts of Interception in a generic Network Architecture*. 7 2001, version 1.1.1.
7. European Telecommunications Standards Institute: *ETSI TR 101 944: Telecommunications security; Lawful Interception (LI); Issues on IP Interception*. 12 2001, version 1.1.2.
8. European Telecommunications Standards Institute: *ETSI TR 102 528: Lawful Interception (LI); Interception domain Architecture for IP networks*. 10 2006, version 1.1.1.
9. European Telecommunications Standards Institute: *ETSI TR 101 331: Lawful Interception (LI); Requirements of Law Enforcement Agencies*. 10 2009, version 1.3.1.
10. European Telecommunications Standards Institute: *ETSI TR 102 232-3: Lawful Interception (LI); Handover Interface and Service-Specific Details (SSD) for IP delivery; Part 3: Service-specific details for internet access services*. 10 2009, version 2.2.1.
11. European Telecommunications Standards Institute: *ETSI TR 101 671: Lawful Interception (LI); Handover interface for the lawful interception of telecommunications traffic*. 8 2010, version 3.6.1.
12. European Telecommunications Standards Institute: *ETSI TR 102 232-1: Lawful Interception (LI); Handover Interface and Service-Specific Details (SSD) for IP delivery; Part 1: Handover specification for IP delivery*. 8 2010, version 2.5.1.
13. European Telecommunications Standards Institute: *ETSI TR 102 232-4: Lawful Interception (LI); Handover Interface and Service-Specific Details (SSD) for IP delivery; Part 4: Service-specific details for Layer 2 services*. 8 2010, version 2.3.1.
14. International Organization for Standardization: *ISO/IEC international standard 7498-1:1994 Information Technology – Open Systems Interconnection – Basic Reference Model: The Basic Model*. 1994.
15. International Organization for Standardization: *International Standard ISO 3166-1, Codes for the representation of names of countries and their subdivisions - Part 1: Country codes, ISO 3166-1: 2006 (E/F)*. 2006.
16. Kajan, R.; Karpíšek, F.; Matoušek, P.; aj.: Rekonstrukce a vizualizace aplikačních dat z IP komunikace. Technical Report FIT-TR-2011-10, Faculty of Information Technology, Brno University of Technology, Brno, Czech Republic, 2011.

17. Kořenek, J.; Korček, P.; Kaštil, J.: Sondy pro monitorování provozu. Technical Report FIT-TR-2011-09, Faculty of Information Technology, Brno University of Technology, Brno, Czech Republic, 2011.
18. McCann, J.; Deering, S. E.; Mogul, J.: *RFC 1981 Path MTU Discovery for IP version 6*. 8 1996.
URL http://tools.ietf.org/html/rfc1981
19. Narten, T.; Draves, R.; Krishnan, S.: *Privacy Extensions for Stateless Address Autoconfiguration in IPv6*. 9 2007.
URL http://tools.ietf.org/html/rfc4941
20. Narten, T.; Nordmark, E.; Simpson, W.; aj.: *Neighbor Discovery for IP version 6 (IPv6)*. 9 2007.
URL http://tools.ietf.org/html/rfc4861
21. Postel, J.: *User Datagram Protocol*. 8 1980.
URL http://tools.ietf.org/html/rfc768
22. Rigney, C.; Rubens, A. C.; Simpson, W. A.; aj.: *Remote Authentication Dial In User Service (RADIUS)*. 6 2000.
URL http://tools.ietf.org/html/rfc2865
23. Schulzrinne, H.; Casner, S.; Frederick, R.; aj.: *RTP: A Transport Protocol for Real-Time Applications*. 7 2003.
URL http://tools.ietf.org/html/rfc3550
24. Thomson, S.; Narten, T.; Jinmei, T.: *IPv6 Stateless Address Autoconfiguration*. 9 2007.
URL http://tools.ietf.org/html/rfc4862

A Seznam zkratek

- AF – *Administration Function* – Administrační funkce
- CC – *Content of Communication* – Obsah komunikace
- CC-IIF – *Content of Communication - Internal Interception Function* – *Funkce odposlechu obsahu komunikace*
- CCTF – *Content of Communication Trigger Function* – Trigerovací funkce
- CID – *Communication identifier*
- CIN – *Communications Identity Number*
- DAD – *Duplicate Address Detection*
- DCC – *Delivery Country Code*
- DHCP – *Dynamic Host Configuration Protocol*
- DHCPv6 – *Dynamic Host Configuration Protocol for IPv6*
- HI *Handover Interface*
- HI1ID *HI1 IDentifier*
- IP – *Internet Protocol*
- IRI – *Intercept Related Information Function*
- IRI-IIF – *Intercept Related Information - Internal Interception Function* – Funkce dynamické identity
- ISP *Internet Service Provider* Poskytovatel připojení k Internetu (ISP)
- LEA – *Lawful Enforcement Agency* – Oprávněný orgán (odposlouchávající agentura)

– LIID – *Lawful Interception IDentifier*
– LIS – *Lawful Interception System* – Systém pro sběr dat pro zákonné odposlechy
– MF – *Mediation Function* – Mediační funkce
– MTU – *Maximal Transmission Unit* – Maximální velikost přenášeného paketu
– NID – *Network IDentifier*
– PMTU – *Path Maximal Transmision Unit* – Maximální velikost přenášeného na cestě
– RA – *Router Advertisment*
– RADIUS – *Remote Authentication Dial In User Service*
– SID – *System IDentifier*
– SLAAC – *Stateless Address Autoconfiguration*
– TCP – *Transmission Control Protocol*
– UDP – *User Datagram Protocol*