VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV INFORMAČNÍCH SYSTÉMŮ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF INFORMATION SYSTEMS

MODELOVÁNÍ L2 PROTOKOLŮ ZAJIŠŤUJÍCÍCH BEZSMYČKOVOST

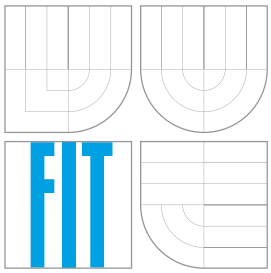
DIPLOMOVÁ PRÁCE
MASTER'S THESIS

AUTOR PRÁCE
AUTHOR

Bc. MATEJ HRNČIŘÍK

BRNO 2012

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV INFORMAČNÍCH SYSTÉMŮ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF INFORMATION SYSTEMS

MODELOVÁNÍ L2 PROTOKOLŮ ZAJIŠŤUJÍCÍCH BEZSMYČKOVOST

MODELLING OF L2 LOOP-PREVENTING PROTOCOLS

DIPLOMOVÁ PRÁCE
MASTER'S THESIS

AUTOR PRÁCE
AUTHOR

Bc. MATEJ HRNČIŘÍK

VEDOUCÍ PRÁCE
SUPERVISOR

Ing. VLADIMÍR VESELÝ

BRNO 2012

Zadání diplomové práce/14268/2011/xhrnci00

Vysoké učení technické v Brně - Fakulta informačních technologií

Ústav informačních systémů Akademický rok 2011/2012

Zadání diplomové práce

Řešitel: **Hrnčiřík Matej, Bc.**
Obor: Počítačové sítě a komunikace
Téma: **Modelování L2 protokolů zajišťujících bezsmyčkovost**
Modelling of L2 Loop-Preventing Protocols
Kategorie: Počítačové sítě

Pokyny:

1. Zjistěte aktuální stav a nasazení protokolů zajišťujících bezsmyčkovost sítě na vrstvě L2 - TRILL, SPB, aj.
2. Analyzujte existující implementace TRILLu a SPB - ať už SW, tak i HW (Cisco, HP, Juniper).
3. Podle doporučení vedoucího implementujte podporu jednoho z výše uvedených protokolů v prostředí OMNeT++ a rozšiřte obecný XML konfigurační soubor.
4. Ověřte fungování modelu vůči reálné topologii a analyzujte výsledky.

Literatura:

- J. Touch and R. Perlman, "RFC 5556 - Transparent Interconnection of Lots of Links (TRILL)", IETF, 2010.
- Z. Kraus, "Modelování a analýza počítačové sítě VUT", DP, FIT 2010.
- M. Matuška, "TRILL", Lupa.cz, roč. 2010, č. 10, Praha, CZ, s. 40, ISSN 1213-0702.

Při obhajobě semestrální části diplomového projektu je požadováno:

- Body 1) a 2).

Podrobné závazné pokyny pro vypracování diplomové práce naleznete na adrese http://www.fit.vutbr.cz/info/szz/

Technická zpráva diplomové práce musí obsahovat formulaci cíle, charakteristiku současného stavu, teoretická a odborná východiska řešených problémů a specifikaci etap, které byly vyřešeny v rámci ročníkového a semestrálního projektu (30 až 40% celkového rozsahu technické zprávy).

Student odevzdá v jednom výtisku technickou zprávu a v elektronické podobě zdrojový text technické zprávy, úplnou programovou dokumentaci a zdrojové texty programů. Informace v elektronické podobě budou uloženy na standardním nepřepisovatelném paměťovém médiu (CD-R, DVD-R, apod.), které bude vloženo do písemné zprávy tak, aby nemohlo dojít k jeho ztrátě při běžné manipulaci.

Vedoucí: **Veselý Vladimír, Ing.**, UIFS FIT VUT
Datum zadání: 19. září 2011
Datum odevzdání: 23. května 2012

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
Fakulta informačních technologií
Ústav informačních systémů
612 66 Brno, Božetěchova 2

L.S.

doc. Dr. Ing. Dušan Kolář
vedoucí ústavu

Abstrakt

Práce popisuje současně používané technologie, které zajišťují bezsmyčkovost sítí na linkové vrstvě. Objasňuje základní problematiku Ethernetových sítí, následně popisuje protokoly zajišťující bezsmyčkovost - TRILL a SPB. Pro každý z vybraných protokolů jsou uvedeny výhody i nevýhody a jsou porovnány s ostatními technologiemi. V praktické části je popsána prvotní implementace protokolu IS-IS v prostředí OMNeT++, který slouží jako základní stavební kámen TRILLu. Důležitým tématem je ověření správnosti implementace.

Abstract

This thesis informs about currently used technologies, which provide loop protection on data link layer of computer networks. It clarifies issues of Ethernet networks. Chosen protocols are then closely described. There are presented advantages and disadvantages of chosen protocols and they are compared to other technologies. Practical section describes the initial implementation of IS-IS in OMNeT++ environment, which serves as the basic building block of TRILL protocol. An important issue is to verify correctness of implementation.

Klíčová slova

STP, TRILL, SPB, počítačové sítě, přepínač, linková vrstva, simulace, OMNeT++, Ethernet, IS-IS, ANSA, směrovač

Keywords

STP, TRILL, SPB, computer networks, switch, data link layer, simulation, OMNeT++, Ethernet, IS-IS, ANSA, router

Citace

Matej Hrnčiřík: Modelování L2 protokolů zajišťujících
bezsmyčkovost, diplomová práce, Brno, FIT VUT v Brně, 2012

Modelování L2 protokolů zajišťujících bezsmyčkovost

Prehlásenie

Prehlasujem, že som túto diplomovú prácu vypracoval samostatne pod vedením pána Ing. Vladimíra Veselého.

.....
Matej Hrnčiřík
22. mája 2012

Poďakovanie

Rád by som poďakoval predovšetkým vedúcemu svojej práce, ktorý mi poskytoval v dobe vzniku morálnu podporu a podnecoval moju tvorivosť. Odvďačiť by som chcel formou predania úžasného receptu, ktorý som vymyslel a postupne zdokonaľoval počas môjho štúdia na tejto škole. Zoberieme 0,5 kg cestovín a uvaríme na spôsob al dente. Zvyšok receptu sa týka omáčky, ktorej treba venovať obzvlášť veľkú pozornosť, keďže sa jedná o delikátnu záležitosť. Tajomstvo spočíva v samotnom zložení Univerzálnej Hnedej Omáčky (UHO), ktorá dostala názov podľa obvyklej farby a jej univerzálneho využitia (môže byť podávaná k čomukoľvek). Jej zloženie je však zakaždým unikátne. Metódou Monte Carlo vyberieme n ingrediencií (kde n je prirodzené číslo z intervalu (3, 9)) z chladničky a všetky ich zmiešame a opatrne vložíme spolu s uvarenými cestovinami do hrnca. Medzi povinné ingrediencie patrí cibuľa a bravčový lunchmeat obsahujúci minimálne 60% tuku. Po zmiešaní varíme na miernom ohni po dobu 15-20 minút. Postupne pridávame papriku, čierne korenie a soľ v takých pomeroch, aby sme čo najviac priblížili odtieňu hnedej #903620 v RGB. Po dovarení podávame s archívnym ročníkom vína Château de bougie, ktoré je v týchto končinách skôr známe pod názvom Hradná svieca.

Obsah

Kapitola 1

Úvod

Počítačové siete patria medzi najrýchlejšie sa rozvíjajúce technológie posledných desaťročí. Technologický pokrok vývoja sieťových zariadení z hľadiska hardwaru je nutné reflektovať novými a optimálnejšími protokolmi. Charakter počítačových sietí sa mení obrovskou rýchlosťou a preto nové protokoly musia byť schopné sa adaptovať. Moderná doba vyžaduje komplexné a škálovateľné siete. Výrazný je rozvoj cloudových technológií.

Staré protokoly je obtiažne nahradiť novými predovšetkým z dôvodu spätnej kompatibility. Je treba počítať s tým, že nasadzovanie nových technológií bude inkrementálne a sieť musí byť funkčná aj v prípade, kedy dôjde k interakcii starých protokolov s novými. Dôležitá je preto simulácia sieťovej prevádzky v simulačnom prostredí akým je napríklad OMNeT++. Výhodou simulácie je cena prístrojov, ktorá je rovná cenám počítačov, na ktorých simulácia prebieha. Simulácia je efektívnym a flexibilným prostriedkom ako otestovať nové technologické postupy pred ich reálnym nasadením.

Práca sa zameriava, na bližší popis súčasných L2 protokolov, ktoré zaisťujú bezslučkovosť v prepínaných sietiach. Zameriava sa ako na staré, tak aj na nové protokoly a zvýrazňuje kontrasty medzi nimi. Prvá časť dokumentu je zameraná hlavne na nový protokol TRILL, ktorý sa javí ako potencionálny náhradník STP protokolu. V druhej časti sa venujem protokolu IS-IS.

Práca je štrukturovaná nasledovne: kapitola 2 popisuje základné princípy Ethernetových sietí, ktoré sú potrebné pre pochopenie nasledujúcich kapitol. Konkrétne pojednáva o mechanizme prepínaných sietí, štruktúre MAC adries a Ethernetových rámcov.

V kapitole 3 sa nachádza zhrnutie existujúcich protokolov zaisťujúcich bezslučkovosť v sietiach na druhej vrstve ISO/OSI modelu. Popísaný je protokol TRILL v kontraste s nevýhodami STP protokolu. V závere kapitoly sa nachádza popis protokolu SPB a rôznych proprietárnych protokolov.

Kapitola 4 obsahuje špecifikáciu smerovacieho protokolu IS-IS. Objasňuje základné mechanizmy výmeny informácií medzi smerovačmi, systém voľby tzv. Designated IS a podrobne popisuje všetky typy paketov.

V kapitole 5 nájdete popis simulačného prostredia OMNeT++ a výskumnej skupiny ANSA. Podrobnejšie sa venujem návrhu a implementácii modulu simulujúceho činnosť IS-IS protokolu. Nasleduje popis aktivity modulu a jeho konfigurácia. Zároveň sa v tejto kapitole nachádza zhrnutie dosiahnutej funkcionality.

Nasledujúca kapitola 6 opisuje proces simulácie protokolu IS-IS. V prvej časti kapitoly je zobrazená modelová topológia. Nasleduje analýza udalostí v čase a porovnanie dosiahnutých výsledkov s hodnotami získanými z fyzických zariadení zapojenými v identickej topológii.

Kapitola 7 zhrňuje dosiahnuté výsledky, obsah práce a popisuje možný budúci vývoj.

Kapitola 2

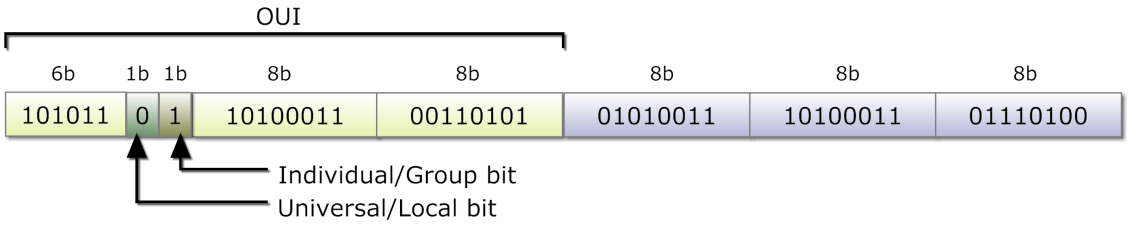
Prepínanie v L2 sietiach

Kapitola je zameraná predovšetkým na objasnenie základných princípov Ethernetových sietí. Zameriava sa popis na štruktúru MAC adresy a Ethernetových rámcov. Informácie v tejto kapitole boli čerpané zo zdrojov [14, 5, 15, 20, 23].

Počítačové siete si môžeme v základe predstaviť ako uzly, ktoré sú navzájom poprepájané dátovými linkami. Každý uzol má svoju špecifickú adresu, pomocou ktorej sa ostatné uzly navzájom rozpoznávajú. Počítačová sieť umožňuje sieťovým prvkom komunikovať medzi sebou zasielaním správ. Pod pojmom L2 sieť je myslená sieť založená na technológiách 2. vrstvy (*data link layer*) ISO/OSI modelu. Rámec tohto dokumentu sa ale týka výhradne problematiky Ethernetových sietí.

2.1 MAC adresa

Adresovanie zariadení v Ethernetových sietiach je možné na základe unikátnych adries zvaných *MAC adresy*. MAC adresa je reprezentovaná 48 bitmi (obrázok 2.1).



Obr. 2.1: Príklad MAC adresy, vypracované z [20]

Prvých 24b unikátne reprezentuje výrobcu zariadenia (prípadne typ protokolu). Táto časť MAC adresy nesie názov OUI - *Organizationally Unique Identifier* a je prideľovaná organizáciou IEEE. Špeciálny význam má tzv. *Inidividual/Group* (I/G) bit a *Universal/Local* (U/L) bit. I/G bit udáva, či sa jedná o adresu jednej stanice (hodnota 0) alebo skupinovú adresu (hodnota 1). U/L bit oznamuje, či sa jedná o adresu pridelenú univerzálne (hodnota 0) alebo lokálne (hodnota 1). Príklad lokálneho pridelenia je napríklad špeciálna adresa `FF-FF-FF-FF-FF-FF` adresujúca všetky stanice v sieti (*broadcast*).

2.2 Ethernetový rámec

V telekomunikačnej terminológii sa balíky informácií putujúce cez sieť nazývajú *protocol data units* (PDU). PDU vyskytujúci sa na L2 (data link) vrstve je označovaný termínom *rámec* (frame). Štruktúra Ethernetového rámcu transferovaného medzi jednotlivými zariadeniami po sieti je popísaná na obrázku 2.2.

Preamble	Start of frame delimiter	destination MAC	source MAC	802.1Q tag	Ethertype	Payload	FCS	Interframe gap
7 octets 10101010	1 octet 10101011	6 octets	6 octets	4 octets	2 octets	46-1500 octets	4 octets	12 octets

Obr. 2.2: Štruktúra Ethernetového rámcu, vypracované z [20]

Preamble - 7 oktetov, striedavo 0 a 1. Slúži na synchronizáciu hodín príjemcu.

Start of frame delimiter - Označenie začiatku rámcu (oktet 10101011).

Destination MAC - MAC adresa príjemcu.

Source MAC - MAC adresa odosielateľa.

802.1Q tag - Volitelná položka rámca jednoznačne identifikujúca VLAN, do ktorej rámec patrí.

Ethertype - Identifikuje protokol, ktorý sa nachádza v užitočnom obsahu (payload).

Payload - Dáta, ktorých dĺžka je minimálne 46 a maximálne 1500 oktetov (minimálna dĺžka je nutná pre správnu detekciu kolízií v rámci segmentu).

FSC - Frame check sequence označuje kontrolný súčet, ktorý sa počíta pomocou CRC32 zo všetkých položiek rámcu okrem preambule a FCS samotného. Slúži ku kontrole správnosti dát.

Interframe gap - Potom čo je zaslaný rámec, vysielač musí vyslať minimálne 96b (12 oktetov) dát a na to slúži toto pole.

2.3 Problematika prepínania rámcov

V minulosti boli v sietiach rožšírené sieťové mosty a rozbočovače (*bridge* a *hub*). Preposielanie rámcov v týchto zariadeniach fungovalo z dnešného pohľadu primitívnym spôsobom - rámec prijali na jednom rozhraní (porte) a preposlali na všetky ostatné porty, na ktorých mali pripojené iné zariadenia. Dnes tento spôsob rozosielania označujeme termínom *broadcast*. Dochádzalo tak k zbytočnému plytvaniu dostupného pásma a vyťažovaniu zariadení.

Časom bolo vyvinuté inteligentné zariadenie dnes známe ako prepínač (*switch*). Ethernetový prepínač je zariadenie, ktoré zasiela prijaté rámce iba do tej časti siete, kde sa nachádza koncové zariadenie a zbytočne nezaťažuje sieť broadcastovaním rámcov. Prepínač je nevyhnutným aktívnym sieťovým prvkom a základom *prepínaných sietí*. Adresovanie zariadení prebieha na základe MAC adries. Každý prepínač si buduje vlastnú *CAM tabuľku* (Content addressable memory), do ktorej ukladá jednotlivé MAC adresy zariadení spárované s fyzickým rozhraním, na ktorom je dané zariadenie prepojené.

Zakaždým keď prepínač prijme rámec, pozrie sa na zdrojovú MAC adresu a v prípade, že danú MAC adresu ešte nemá v CAM tabuľke obsiahnutú, asociuje ju s portom, na ktorom bol rámec prijatý. Posielanie paketu ďalej (*forwarding*) je realizované vyhľadávaním cieľovej adresy v CAM tabuľke. Ak danú MAC adresu nájde, prepošle rámec na daný port. V opačnom prípade je rámec replikovaný a šírený všetkými smermi, kde ešte nie je sieť preskúmaná. Vykonávaním týchto krokov sa prepínač postupne naučí umiestnenie všetkých pripojených zariadení.

Bez použitia protokolu, ktorý zabraňuje vytváranie slučiek na sieti by sa mohlo stať, že sa rámec vráti nejakou cestou na ten istý prepínač. Akonáhle by vznikla takáto slučka, rámce by sa začali replikovať exponenciálne, čo by po istej dobe viedlo k zahlteniu celej siete. L2 rámce sami o sebe neposkytujú žiadny ochranný mechanizmus proti zacykleniu ako je to u L3. Preto je prítomnosť protokolu na prepínačoch, ktorý túto problematiku rieši, aboslútnou nevyhnutnosťou.

Kapitola 3

Prehľad existujúcich protokolov

Kapitola popisuje základnú funkčnosť L2 protokolov, používaných na zabezpečenie sietí proti vzniku slučiek. Zameriava sa predovšetkým na protokoly STP, TRILL a SPB.

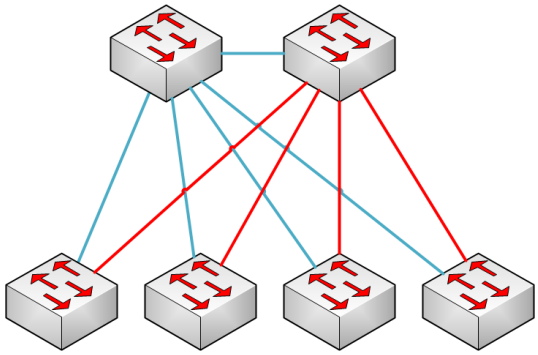
Takmer všetky existujúce protokoly riešiace danú problematiku však majú pevný základ - zaužívaný štandard **Spanning Tree Protocol (STP)**. Jeho prítomnosť je na každom dnešnom prepínači samozrejmosťou. Svoje prvenstvo si udržiava najmä vďaka tomu, že je tak masívne rozšírený a jeho nahradenie nebude jednoduchý proces. Pre dnešné komplexné siete je stále použiteľný, avšak zďaleka nie efektívny. Má množstvo nežiadúcich vlastností, ktoré redukujú v určitých pohľadoch výkon celej siete. Za posledné roky bolo vyvinutých viacero protokolov, ktoré tieto negatíve vlastnosti STP minimalizovali. Vo väčšine prípadov však platí, že buď špecifikácia týchto protokolov vychádza zo STP a tým pádom sú vylepšené len niektoré vlastnosti, alebo sa jedná o úplne nové protokoly, ktoré sú ale proprietárne, a tým pádom nie je známa ich špecifikácia. Existuje minimum otvorených L2 protokolov zabezpečujúcich bezslučkovosť siete, ktoré majú šancu na reálne rozšírenie. Jedná sa predovšetkým o protokoly **TRILL** (Transparent Interconnect of Lots of Links) a **SPB** (Shortest Path Bridging).

3.1 STP

Počítačovú sieť si môžeme predstaviť ako graf, kde aktívne prvky predstavujú uzly a linky medzi nimi ako hrany. V takomto grafe sa prirodzene môžu nachádzať slučky. V praxi je úplne bežný výskyt takýchto sieťových topológií, kde sa mimo iných dôvodov vyskytujú slučky z dôvodu prítomnosti redundantného zapojenia, aby sa zvýšila odolnosť voči poruchám. Mechanizmus prepínania rámcov neobsahuje žiadnu metódu starnutia rámcov (napr. TTL na L3 vrstve) ani iný mechanizmus na elimináciu slučiek. Tento problém rieši *Spanning Tree Protocol* - STP. Radia Perlman navrhla jeho prvú verziu v roku 1990 a bol vydaný ako súčasť štandardu IEEE 802.1d, ktorý sa postupne dočkal viacerých aktualizácií. [15, 17, 5]

Princíp fungovania STP spočíva v naučení sa topológie danej lokálnej siete. Následne sa zvolí jeden z uzlov za koreň, od ktorého sa vytvára aktívna topológia bez cyklov - kostra. Tá je potom použitá k prepínaniu rámcov po sieti. Voľba koreňa prebieha na základe identifikátora uzlu - *BridgeID*, ktorý sa skladá z MAC adresy a modifikovateľnej priority. Vytvorená kostra spája aktívne uzly práve jednou cestou, takže nemôže dôjsť k zacykleniu. [15]

Na obrázku 3.1 je zobrazená modelová topológia, v ktorej sú prepínače redundantne poprepájané a tým pádom by sa bez použitia STP vyskytli slučky. STP vytvorí kostru:



Obr. 3.1: Demonštrácia činnosti STP

modrou farbou sú vyznačené aktívne spoje a červenou deaktivované. Takáto sieť nie je zďaleka optimálna. Celá spodná vrstva je pripojená na jeden vrchný prepínač, čo spôsobuje nerovnomerné rozloženie záťaže. Hlavným problémom je ale skutočnosť, že množstvo spojov je nevyužitých a tým pádom je zbytočne znížená priepustnosť siete.

Ďalšia činnosť STP spočíva v aktivovaní náhradných spojov v prípade výpadku určitého segmentu siete. Ak nastane výpadok na jednej linke a existuje redundantný spoj, je potrebné ho v čo najkratšom čase aktivovať a presmerovať ním tok dát. Pri zmene topológie je nutné, aby bola celá sieť čo najrýchlejšie skonvergovaná. Doba konvergencie STP je väčšinou do 30 sekúnd (v extrémnom prípade až 50 sekúnd), čo je v dnešnej dobe neprijateľné. Okrem toho má STP charakter *fail-open*, čo znamená, že pri poruche môže nastať situácia, kedy sa aktivuje v jednej chvíli hlavný aj záložný spoj a zahltí sa tým sieť. Zisťovanie zdroju chýb pri zahltení ako aj iných problémoch je v L2 sietiach neľahká činnosť, keďže nie sú dostupné nástroje známe z L3. [15, 17]

Zrýchlenie doby konvergencie je možné riešiť použitím vylepšenou variantou - **Rapid Spanning Tree Protocol** (IEEE 802.1w). Bol navrhnutý tak, že ponecháva mnoho parametrov nezmenených a funkčne vychádza z pôvodného STP. Vďaka tomu je spätne kompatibilný s prepínačmi, ktoré nevedia RSTP. Zlepšila sa predovšetkým doba konvergencie, ktorá je maximálne 6 sekúnd. RSTP je v súčasnosti začlenený do novšieho štandardu IEEE 802.1d z roku 2004. [15, 17, 5]

So zavedením virtuálnych sietí vznikli nové problémy a požiadavky na aktívne topológie. Problém sa objaví v sieti využívajúcej STP, ktorá je rozdelená pomocou VLAN na segmenty, ktoré ale nepokrývajú celú sieť. Nemusí byť vhodné šíriť citlivé dáta cez zariadenia, ktoré sú v správe iného subjektu. STP síce zabezpečí bezslučkovosť siete aj pri rozdelení siete pomocou VLAN, ale vypočítaná kostra grafu nemusí byť optimálna. Preto bol vyvinutý **Multiple Spanning Tree Protocol** (IEEE 802.1s), ktorý pre každú VLAN alebo skupiny VLAN vypočíta rôzne kostry grafu, takže dôjde k optimálnejšiemu rozloženiu. MSTP je dnes súčasťou štandardu 802.1Q. [15, 17]

Existujú ďalšie deriváty STP od firmy Cisco. Jedná sa o protokoly **Per-VLAN Spanning Tree** (PVST a PVST+) a **Rapid-PVST+**, ktoré sú ale proprietárnymi protokolmi a je preto nutné použiť zariadenia firmy Cisco. Základným mechanizmom fungovania PVST je vytváranie oddelených aktívnych topológií pre kažú VLAN. [17, 15]

Priamočiarym riešením ako minimalizovať nedostatky STP je rozdeliť sieť do čo najmenších L2 segemtnov (broadcast domén) a navzájom ich prepojiť L3 prvkami. V prípade poruchy, ktorá môže viesť k zahlteniu siete, bude nefunkčná práve postihnutá doména a ostatné segmenty siete ostanú "nedotknuté". L3 vrstva poskytuje sofistikovanejšie možnosti

pre stavbu rýchlo konvergujúcich, väčších a stabilnejších sietí. Broadcast domény by nemali byť geograficky príliš odľahlé a mal by sa obmedzovať počet koncových staníc na doménu (v dnešnej dobe je za doporučené maximum považované množstvo 1000 koncových staníc). Niekedy je však potrebné budovať veľké L2 segmenty, čo sa týka predovšetkým prípadov datacentier a poskytovateľov služieb.

3.2 TRILL

Informácie použité v tejto kapitole sú vypracované zo štandardov [9, 10, 11], doplňujúce informácie sú čerpané z [17].

Na základe vedomostí o nedostatkoch STP bolo potrebné navrhnúť nový protokol. Výskumná skupina na čele s Radia Perlman vyvinuli protokol **TRILL** - *Transparent Interconnection of Lots of Links*, ktorý bol v júly 2011 vydaný ako IETF RFC štandard. TRILL je popísaný nasledujúcimi dokumentami:

- **RFC 6325** – Základné definície a formáty protokolu TRILL.
- **RFC 6326** – Použitie protokolu IS-IS v TRILLe.
- **RFC 6327** – Susedstvo TRILL prepínačov (Adjacency).

Je nutné podotknúť, že TRILL môže byť okrem Ethernetu (IEEE 802.3) nasadený aj nad PPP(Point to Point Protocol) linkami (definované v IETF RFC 6361). V rámci tohoto dokumentu sa však budem sústrediť len na Ethernetovú implementáciu.

Protokol TRILL má nájsť využitie ako náhrada protokolov rodiny STP predovšetkým v datacentrách, u poskytovateľov služieb a všade tam, kde sú vyžadované škálovateľné L2 Ethernetové komunikačné služby. Tvorcovia protokolu pre dosiahnutie cieľa zvolili analogické postupy známe z problematiky smerovania sietí (L3 routing), ktoré umožňujú dosiahnuť vysokú škálovateľnosť (napr. pridaná smerovacia hlavička a hop-count) pri zachovaní vlastností Ethernetových sietí (napr. broadcast a absencia nutnosti explicitnej konfigurácie).

Medzi stanovené ciele pri vývoji protokolu patrilo:

- Využitie najvýhodnejšej cesty v danej topológii bez nutnosti sledovania kostry grafu.
- Využitie prenosovej kapacity všetkých liniek bez nutnosti rozloženia záťaže pomocou VLAN.
- Paralelné využitie liniek pre prenos od rovnakého zdroja k rovnakému cieľu (tzv. *multipathing*).
- Podpora ľubovoľnej L2 topológie.
- Možnosť analýzy chovania L2 siete formalizovanými nástrojmi.
- Protokol musí byť fail-safe tj. pri nejasnom stave musí byť linka radšej zablokovaná ako aktivovaná.
- Zabránenie zacykleniu L2 rámcov.
- Zachovanie kompatibility so staršími prepínačmi v jednej L2 infraštruktúre.

- Možnosť šírenia broadcastových, multicastových a unicastových paketov s neznámou cieľovou adresou po celej L2 doméne.
- Jednoduchosť konfigurácie.
- Transparencia voči koncovým staniciam.
- Vysoká rýchlosť konvergencie.

Cieľom pri návrhu rozhodne nebolo:

- Zabezpečenie siete voči útokom ako MAC flooding, MAC spoofing, ARP poisoning a pod. Rozhodne by však nemal pripievať k novým typom sieťových útokov.
- Zvýšenie počtu koncových staníc v jednej broadcast doméne.
- Zavedenie hierarchického smerovania ako to poznáme v L3.

3.2.1 Základný princíp

Klasické prepínače, ktoré poznajú iba STP (a z neho odvodené protokoly popísané v kapitole 3.1) a prepínače ovládajúce TRILL môžu koexistovať spolu na jednej sieti. RFC 6325 definuje názov **RBridge** (Routing Bridge) pre zariadenie implementujúce protokol TRILL. Porty RBridge je možné rozdeliť nasledovne:

1. porty, ku ktorým sú pripojené iné TRILL prepínače
2. porty, ku ktorým sú pripojené koncové stanice, klasické prepínače a iné zariadenia neovládajúce TRILL

Na základe typu pripojenia sa líši spôsob spracovania prijatých a odoslaných rámcov. V prípade že je RBridge pripojený k inému TRILL prepínaču, je nutné prijatý/odoslaný rámec dekapsulovať/enkapsulovať spôsobom definovaným v štandarde popisujúcom TRILL. V opačnom prípade sa na daných portoch RBridge tvári ako klasický prepínač a komunikácia prebieha spôsobom známym z konvenčných L2 sietí.

Elimináciu slučiek v sieti rieši TRILL zavedením mechanizmu známeho z L3 vrstvy - využíva smerovací *link-state* protokol **IS-IS** (Intermediate System To Intermediate System). Po pripojení TRILL prepínača do siete nie je nutná žiadna dodatočná konfigurácia (tzv. *zero configuration*). Po zapnutí, zariadenie automaticky vyhľadá svojich susedov a pomocou IS-IS vybuduje celú topologickú mapu prepínanej siete. Ďalším krokom je vypočítanie optimálnych ciest k ostatným TRILL prepínačom. Zároveň sa vypočíta zdieľaný distribučný strom, ktorý sa bude používať pre zasielanie *multi-destination* rámcov na portoch prvého typu. Na portoch druhého typu sa pre tento typ rámcov použije metóda *floodingu* známeho z klasických Ethernetových sietí.

3.2.2 Zapúzdrenie

Pri použití protokolu TRILL sa k originálnemu Ethernetovému rámcu pridá TRILL hlavička a následne nová Ethernetová hlavička. Pôvodná hlavička a obsah rámca ostane nezmenený. Na koniec rámcu sa namiesto pôvodnej položky *Frame Check Sequence* (FCS) vypočíta nová hodnota. Štruktúra nového enkapsulovaného rámcu je znázornená na obrázku 3.2.

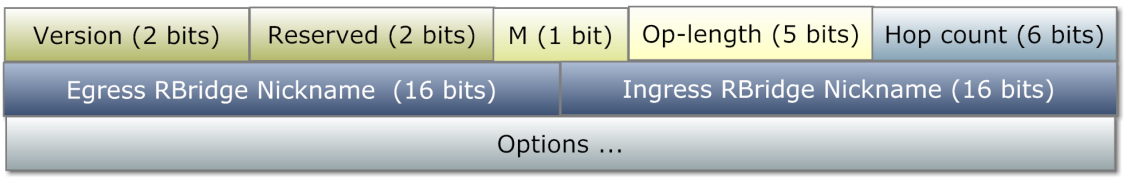
Obr. 3.2: Enkapsulácia Ethernetového rámca protokolom TRILL [9]

Pôvodná zdrojová a cieľová MAC adresa je týmto spôsobom ukrytá. Vonkajšia pridaná Ethernetová hlavička slúži pre prenos rámcu na susedný RBridge. Od klasickej Ethernetovej hlavičky sa líši identifikátorom *Ethertype*, ktorý má hodnotu `0x22F3` a označuje tým rámec, v ktorom sa nachádzajú TRILL dáta.

Pri využití VLAN je vo vonkajšej Ethernetovej hlavičke k dispozícii 802.Q tag. Hodnota tagu identifikuje VLAN, ktorou sa transportujú enkapsulované rámce medzi jednotlivými TRILL prepínačmi. Vnútorný 802.1Q tag tak ostáva nemodifikovaný a určuje VLAN, v ktorej sa nachádza koncová stanica. Jedna transportná VLAN môže prenášať rámce pre viacero koncových VLAN.

3.2.3 TRILL hlavička

TRILL obaľuje pôvodný Ethernetový rámec špecifickou hlavičkou, ktorej štruktúra je zobrazená na obrázku 3.3.



Obr. 3.3: Hlavička protokolu TRILL podľa [9]

Version - 2 bitová hodnota udávajúca verziu TRILL protokolu. Aktuálna je verzia 0. Každý RBridge musí kontrolovať verziu a pokiaľ danú verziu nerozpozná, musí daný rámec zahodiť.

Reserved - 2 bity sú rezervované pre eventuálne budúce použitie v nasledujúcich verziách. V aktuálnej verzii je táto položka nastavená na hodnotu 0.

Multi-destination - Bit určujúci, či má byť daný rámec doručený určitej skupine koncových staníc.

- 0 (FALSE) – Položka Egress RBridge Nickname obsahuje ID TRILL prepínača pre známu unicastovú MAC adresu.
- 1 (TRUE) - Položka Egress RBridge Nickname obsahuje ID označujúce distribučný strom. Toto ID je vybrané vstupným TRILL prepínačom pre daný rámec.

Options length - Veľkosť poľa *Options* na konci hlavičky. Veľkosť udáva koľko 4-oktetových chunkov bude mať položka *Options*. Ak je veľkosť 0, žiadne dodatočné nastavenia sa na konci hlavičky nenachádzajú.

Hop count - 6 bitová hodnota nastavená vstupným TRILL prepínačom určujúca, koľkými ďalšími TRILL prepínačmi môže daný rámec prejsť, kým bude zahodený. Ide o položku ekvivalentnú TTL známu z L3. Hodnota je pri každom prechode prepínačom znížená o 1. Keď hodnota dosiahne 0, rámec sa zahodí. Ide o mechanizmus zabraňujúci zacykleniu rámca.

Určenie hodnoty Hop count nie je jasne definovaná. Základná hodnota by mala byť nastavená medzi očakávanou dĺžkou trasy a najdlhšou možnou dĺžkou trasy. Je vhodné pripočítať ešte nejakú vhodne zvolenú hodnotu v prípade chyby v smerovaní. Nastavená hodnota samozrejme musí reflektovať aktuálnu topológiu a hodnotu prenastaviť pri zmene topológie, aby bolo zaručené bezproblémové doručovanie rámcov k cieľu.

Egress RBridge Nickname - 16 bitová hodnota udávajúca ID TRILL prepínača, ktorým má daný rámec opustiť TRILL sektor. TRILL umožňuje mať adresovaných 2^{16} TRILL prepínačov, pričom hodnota `0x0000` je rezervovaná a indikuje, že ID nie je špecifikované. Hodnoty od `0xFFC0` po `0xFFFE` sú zatiaľ rezervované pre budúce použitie a hodnota `0xFFFF` je permanentne obsadená.

Hodnota ID výstupného Rbridge je nastavená vždy vstupným TRILL prepínačom a počas celej cesty nesmie byť zmenená a to platí ako v prípade unicastového, tak aj multi-destination doručovania. Zvolené ID musí byť samozrejme v danom TRILL sektore unikátne. ID môže byť nastavené manuálne alebo automaticky. V prípade automatického zvolenia sa ID počíta pomocou hashu z hodnôt ako System ID, čas a dátum a ďalších zdrojov entropie aby sa s čo najväčšou pravdepodobnosťou predošlo kolíziám. Jednotlivé RBridge medzi sebou komunikujú a vyjednávajú svoje ID. Každý RBridge má nastavenú určitú prioritu k svojmu ID a v prípade kolízie vyhráva ten, ktorý ju má vyššiu. Po reštarte sa RBridge vždy snaží použiť ID, ktoré mal naposledy nastavené. Pri postupnom pridávaní prepínačov do TRILL sektoru je pravdepodobnosť kolízie ID minimálna. Väčšia pravdepodnbnosť je pri premostení viacerých TRILL sektorov do jedného. Doba vyjednávania ID musí byť preto čo najmenšia.

Ingress RBridge Nickname - 16 bitová hodnota udávajca ID RBridge, ktorým daný rámec vstupuje do TRILL sektoru. Ostatné vlastnosti ostávajú rovnaké ako v prípade položky *Egress RBridge Nickname*.

Options - Veľkosť položky je definovaná pomocou hodnoty *Options length*. Pole Options vôbec nemusí existovať (nulová veľkosť). Maximálna veľkosť je 128 oktetov. V aktuálnej verzii protokolu sú v poli definované hodnoty ako *Critical Hop by Hop* alebo *Critical Ingress to Egress*. Zvyšné miesto je zatiaľ určené pre budúce použitie.

3.2.4 Smerovanie

Ako bolo spomenuté v kapitole 3.2.1, TRILL využíva narozdiel od STP koncept smerovania na elimináciu slučiek v sieti a optimálne doručovanie rámcov. Pre smerovanie a výpočet najvýhodnejších ciest využíva protokol IS-IS. Je však nuté poznamenať, že TRILL IS-IS nie je kompatibilný s L3 IS-IS. Keďže sa jedná o link state protokol, každý RBridge vie

o všetkých ostatných TRILL prepínačoch v danom TRILL sektore a pozná teda kompletnú topológiu. Použitie IS-IS má nasledujúce výhody:

- Beží priamo nad L2 a nevyžaduje dodatočnú konfiguráciu (nemusia byť priradené IP adresy).
- Je jednoducho rozšíriteľný pomocou definovania nových TLV (*Type Length Value*) elementov pre prenos TRILL informácií.

Jednotlivé RBridge sa o sebe navzájom dozvedia zasielaním Hello rámcov, ktoré sú doručované pomocou multicastu všetkým ostatným TRILL prepínačom. Tieto rámce sú preposielané klasickými prepínačmi ďalej, zahodené koncovými stanicami a spracované sú len ostatnými TRILL prepínačmi. TRILL *Hello rámce* sú odlišné od L3 IS-IS *Hello paketov* tým, že majú menšiu veľkosť a podporujú fragmentáciu určitých informácií. Na základe prijatých *Hello rámcov* sa na každej multiaccess linke v sieti zvolí *Designated RBridge* (DRB), ktorý má za úlohu:

- Zvoliť VLAN ID na ktorej bude prebiehať komunikácia medzi susednými TRILL prepínačmi. Zvolené VLAN ID je použité vo vonkajšej Ethernetovej hlavičke ktorá obaľuje TRILL hlavičku + pôvodný rámec.
- Zvoliť tzv. *Appointed Forwarder* (AP) pre každú VLAN na linke (môže byť aj on sám). AP má za úlohu spracovať všetky natívne rámce na konkrétnej linke danej VLAN. Zapúzdruje prijaté rámce do TRILL dátových rámcov v prípade, že je vstupným TRILL prepínačom a analogicky opačnú operáciu vykonáva v prípade, že je výstupným TRILL prepínačom.

Viac o protokole IS-IS sa môžete dozvedieť v kapitole 4.

3.2.5 Distribučný strom

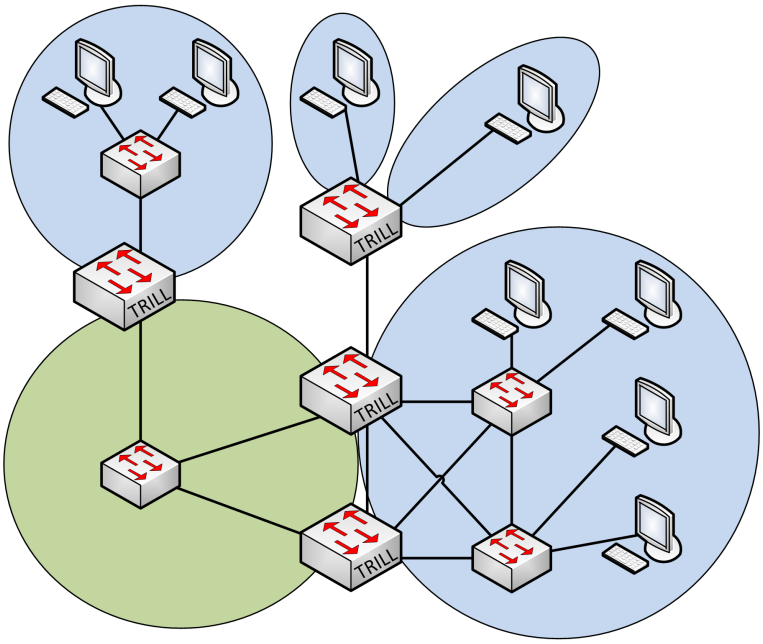
Distribučné stromy (Distribution trees) slúžia pre doručovanie multi-destination rámcov. Stromy sú obojsmerné. Teoreticky by pre jeden TRILL sektor stačil jeden strom ako pri použití STP. Pri vypočítaní viacerých stromov sme ale schopní využívať multipathing pri multi-destination rámcoch a umožňuje nám vybrať najoptimálnejšiu cestu. Pre výpočet stromov sa nevyužíva STP ale sú určené na základe link state informácii. Každému distribučnému stromu sa priradí koreňový RBridge. TRILL využíva ďalej množstvo optimalizačných techník pre „vylepšenie“ stromov. Každý distribučný strom by mal byť orezaný na základe každej prítomnej VLAN tým spôsobom, že sa odrežú vetvy, na ktorých nie sú žiadni potencionálny príjemci.

Aby sa zabránilo vytvoreniu dočasných slučiek pri prenose multi-destination rámcov, využíva TRILL mechanizmus zvaný *Reverse Path Forwarding Check*. Jeho funkcionalita spočíva v kontrole multi-destination rámcov na základe použitého distribučného stromu. Ak sa takýto rámec nevráti očakávanou linkou, musí byť zahodený.

3.2.6 Kompatibilita s klasickými prepínačmi

Jedným z cieľov pri návrhu bola maximálna kompatibilita so súčasnými prepínačmi. Klasické prepínače môžu fungovať úplne kdekoľvek (na okraji, v jadre, na spojoch medzi TRILL prepínačmi). Nie je potrebné vytvárať ucelené oblasti prepínačov nového alebo starého typu. Klasické prepínače sú z pohľadu TRILL prepínačov transparentné - sú súčasťou služby na

nižšej vrstve. Klasický prepínač pošle TRILL rámec správnou cestou na základe vonkajšej Ethernetovej hlavičky i napriek tomu, že vnútornému obsahu nerozumejú.



Obr. 3.4: Príklad topológie s použitím klasických aj TRILL prepínačov

Aj napriek skutočnosti, že oba typy prepínačov je možné akokoľvek kombinovať, je vhodné ich umiestňovať tak, aby sa využila vlastnosť TRILL prepínačov rozdeľovať STP oblasti na menšie časti a skrývať MAC adresy koncových staníc pred transportnými prepínačmi, čím sa ušetrí miesto v CAM tabuľkách. Príklad takejto topológie je znázornený na obrázku 3.4. Nahradením niektorých klasických prepínačov TRILL prvkami došlo k rozdeleniu pôvodne jednej STP domény na päť (vyznačené elipsami). Zároveň je dodržaná celistvosť L2 siete (napr. keď ktorákoľvek koncová stanica vyšle *broadcastový* rámec, tak bude doručený všetkým ostatným staniciam). Zelenou farbou je vyznačený segment siete, ktorý slúži iba ako tranzitný. Napriek tomu že reprezentuje *backbone* siete, môže obsahovať aj klasické prepínače.

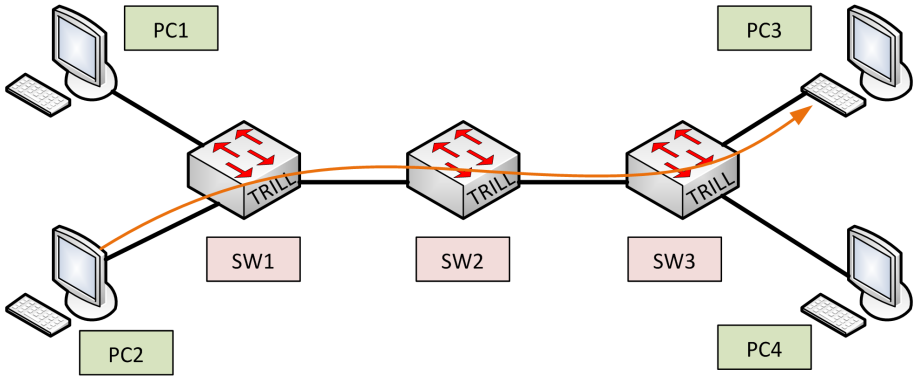
3.2.7 Demonštrácia komunikácie

Pre nasledujúci príklad uvažujme topológiu skladajúcu sa z 4 koncových staníc označenými PC1 - PC4 a TRILL prepínačmi označenými SW1 - SW3. Popísaná topológia je znázornená na obrázku 3.5.

Pre jednoduchosť príkladu predpokladáme, že všetky prepínače sa pomocou IS-IS už dozvedeli celú topológiu a sieť je skonvergovaná. Prepínače nevyužívajú VLAN a zatiaľ nepoznajú MAC adresy koncových staníc.

1. PC2 chce komunikovať s PC3. Vytvorí rámec, ktorý odošle prepínaču SW1.

2. SW1 zistí, že má k sebe pripojený PC2. Pozrie sa na hlavičku Ethernetového rámcu a zistí, že cieľovú MAC adresu ešte nepozná. Preto vykoná *flooding* na lokálnych portoch. Zistí, že na jednom z portov sa nachádza ešte PC1. Po floodingu rámec dorazí na SW1.



Obr. 3.5: Komunikácia v TRILL sieti

3. SW1 už vie, že na ďalšom porte má pripojený TRILL prepínač (RBridge). Preto pôvodný rámec obalí TRILL hlavičkou a následne ďalšou Ethernetovou hlavičkou, ako je znázornené na obrázku 3.2. Takto upravený rámec ďalej odošle na základe smerovacej tabuľky prepínaču SW2.

4. SW2 nemá ku sebe pripojené žiadne koncové zariadenie, preto nie je rámec dekapsulovaný z TRILL hlavičky. Hodnota *hop count* je dekrementovaná o 1. SW2 rámec ďalej odošle prepínaču SW3.

5. SW3 rámec kompletne dekapsuluje, pretože má k sebe pripojené koncové stanice a zároveň je výstupným TRILL prepínačom. Cieľovú adresu PC3 ešte nepozná, preto vykoná *flooding*. Medzičasom z prijatého rámcu zistí, že počítač PC2 je pripojený k SW1. Z floodingu sa dozvie, že má k sebe pripojený PC3 a PC4.

6. SW3 doručí rámec cieľu PC3.

7. Odpoveď poputuje obdobným spôsobom s tým rozdielom, že SW3 už bude vedieť, že PC2 je pripojený k SW1 a preto bude transferovaný najvýhodnejšou cestou priamo až k PC2.

TRILL routing table on SW1	
RBridge	Next hop
SW1	local
SW2	SW2
SW3	SW2

TRILL routing table on SW2	
Rbridge	Next hop
SW1	SW1
SW2	local
SW3	SW3

TRILL routing table on SW3	
RBridge	Next hop
SW1	SW2
SW2	SW2
SW3	local

Tabuľka 3.1: Smerovacie tabuľky jednotlivých prepínačov prislúchajúce príkladu z obrázku 3.5

Smerovacie tabuľky určujú smery, ktorými sa majú vyslať jednotlivé rámce aby dosiahli ktorýkoľvek RBridge v TRILL sieti. Keďže je využívaná implementácia IS-IS prokolu, ktorý

patrí medzi link state protokoly, je nutné aby každý RBridge poznal celú topológiu siete. Po skonvergovaní budú jednotlivé smerovacie tabuľky na prepínačoch vyzerať tak, ako je zhrnuté v tabuľke 3.1.

Po postupnom port floodingu a skonvergovaní sa jednotlivé prepínače naučia cesty k jednotlivým koncovým staniciam, ktoré ukladajú do prepínacích tabuliek. Tabuľka 3.2 zobrazuje obsah prepínacích tabuliek jednotlivých TRILL prepínačov po skonvergovaní.

SW1			
Target MAC address	location	output	encapsulation
PC1	directly connected	local port X	none
PC2	directly connected	local port Y	none
PC3	remotly connected	to SW3	outer Eth + TRILL
PC4	remotly connected	to SW3	outer Eth + TRILL

SW3			
Target MAC address	location	output	encapsulation
PC1	remotly connected	to SW1	outer Eth + TRILL
PC2	remotly connected	to SW1	outer Eth + TRILL
PC3	directly connected	local port X	none
PC4	directly connected	local port Y	none

Tabuľka 3.2: Smerovacie tabuľky jednotlivých prepínačov prislúchajúce príkladu z obrázku 3.5

Pre SW2 bude prepínacia tabuľka prázdna, pretože sa nachádza v jadre TRILL siete a nie sú k nemu pripojené žiadne koncové stanice (SW2 slúži iba ako tranzitný prepínač). Vystačí si preto iba so smerovacou tabuľkou.

Teraz si podrobnejšie rozoberieme vyššie uvedený príklad zasielania TRILL rámcu sieťou. Pre jednoduchšie pochopenie neuvádzam pri obsahoch jednotlivých rámcov všetky položky.

Obsah rámcu, ktorý na začiatku vyšle PC2 je znázornený na obrázku 3.6

Ethernet			
Destination MAC	Source MAC	Payload	FCS
PC3	PC2	DATA	CRC

Obr. 3.6: Obsah Ethernetového rámcu vyslaného z PC2 smerom k PC3

Tento rámec je následne prijatý prepínačom SW1. Pozrie sa na cieľovú MAC adresu a z prepínacej tabuľky zistí, že ho musí odoslať smerom k SW3. Zo smerovacej tabuľky zistí, že *next hop* pre SW3 je prepínač SW2. SW1 je vstupný RBridge (*Ingress RBridge*) do TRILL siete. Preto je nutné rámec zapúzdriť.

Určenie inicializačnej hodnoty hop countu nie je deterministicky definované. Hodnota však musí byť taká, aby v prípade chyby rámec necyklil po sieti príliš dlho a bol zahodený, aby sa zbytočne nezahlcovala sieť. Zároveň však musí rámec doraziť do cieľovej destinácie. V tomto prípade hop count nainicializujem na hodnotu 8.

Pôvodný Ethernetový rámec ostane nezmenený. Pridá sa k nemu TRILL hlavička + vonkajšia Ethernetová hlavička. Obsah rámca opúšťajúceho SW1 je popísaný v obrázku

3.7.

Outer Eth		TRILL			Inner Eth	Outer Eth
Destination MAC	Source MAC	Hop count	Egress RBridge	Ingress RBridge	Original Eth frame	FCS
SW2	SW1	8	SW3	SW1	header+payload	CRC

Obr. 3.7: Obsah rámcu vyslaného z SW1 smerom k SW2

Takto enkapsulovaný rámec prijme SW2. Na základe cieľovej MAC adresy zistí, že je rámec určený pre neho. *Ethertype* v hlavičke je nastavený na hodnotu `0x22F3` čo indikuje, že sa jedná o TRILL rámec. Preto odbalí vonkajšiu Ethernetovú hlavičku. Následne prečíta TRILL hlavičku, z ktorej zistí, že rámec má byť doručený prepínaču SW3. Na základe smerovacej tabuľky zistí, že SW3 ma pripojený na lokálnom porte. Hodnotu hop count zníži o 1, nahradí vonkajšiu Ethernetovú hlavičku novou a takto upravený rámec pošle ďalej (obrázok 3.8).

Outer Eth		TRILL			Inner Eth	Outer Eth
Destination MAC	Source MAC	Hop count	Egress RBridge	Ingress RBridge	Original Eth frame	FCS
SW3	SW2	7	SW3	SW1	header+payload	CRC

Obr. 3.8: Obsah rámcu vyslaného z SW2 smerom k SW3

SW3 prijme rámec. Podľa vonkajšej Ethernetovej hlavičky zistí, že je určený pre neho a že je to TRILL rámec. Z TRILL hlavičky zistí, že cieľový RBridge je on sám, takže rámec došiel na koniec TRILL siete. V prepínacej tabuľke vyhľadá cieľovú MAC adresu z vnútornej Ethernetovej hlavičky a správnym portom následne rámec doručí. Na konci cesty pri PC3 rámec vyzerá rovnako ako na začiatku keď ho vyslal PC2.

3.2.8 Existujúce implementácie

Napriek tomu že vývoja protokolu TRILL sa zúčastnili aj veľké firmy ako Huawei, Cisco Systems alebo Brocade, produktov podporujúcich TRILL sme sa zatiaľ stále nedočkali. Firmy vyvíjajú predovšetkým proprietárne riešenia pre dátové centrá ako náhradu za STP. Výnimkou je Cisco, ktoré vyvinulo protokol *FabricPath*, ktorý je kompatibilný s protokolom TRILL. Viac o proprietárnych protokoloch nájdete v kapitole 3.4.

Existuje však softwarová implementácia, ktorá je zahrnutá v operačnom systéme Oracle Solaris 2010.11 [18].

3.3 SPB

Alternatívny protokol nazvaný **Shortest Path Bridging (SPB)** vznikol v rámci výskumnej skupiny **IEEE 802.1aq**. [7] Zo spoločných vlastností protokolov TRILL a SPB by som zmienil najmä ich snahu kompletne nahradiť STP. Na výpočet najoptimálnejšej cesty používajú oba spomenuté protokoly smerovací protokol IS-IS, konkrétne jeho SPF (Shortest Path First) algoritmus, ktorý zároveň rieši problematiku výskytu slučiek v sieti. Protokol SPB sa budem snažiť predstaviť predovšetkým z kontrastného pohľadu voči protokolu TRILL a preto budem opisovať hlavné rozdiely medzi nimi.

TRILL bol vyvíjaný v prostredí IETF s použitím mechanizmov využívaných v L3 smerovaní. Aj napriek tomu, že vylepšuje škálovateľnosť L2 domén, nebolo to jeho primárnym cieľom.

SPB vznikol v prostredí IEEE ako riadiaci protokol pre siete MAC-in-MAC nahradzujúci ručnú konfiguráciu. Nových mechanizmov sa snaží využívať čo najmenej, aby ho bolo možné jednoducho implementovať. [17]

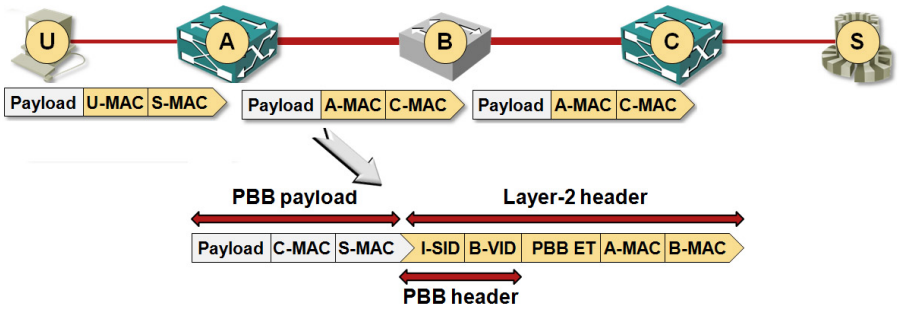
SPB využíva nasledujúce 2 existujúce protokoly:

- **IEEE 802.1ad** Provider Bridges - známy tiež ako **Q-in-Q**, špecifikuje označenie Ethernetového rámcu druhým tagom podľa štandardu 802.1Q. Využíva sa napríklad v L2 sietiach poskytovateľov služieb (service provider). Pridaním ďalšieho tagu pri vstupe do backbone siete sa predíde konfliktom rovnakých VLAN identifikátorov viacerých zákazníkov. Automatické použitie Q-in-Q v SPB sa označuje ako princíp zvaný SPBV. [17] [6]
- **IEEE 802.1ah** Provider Backbone Bridges - známy tiež ako **MAC-in-MAC**, špecifikuje pridanie druhej Ethernetovej hlavičky rámcu. Mechanizmus je analogický k Q-in-Q. Namiesto VLAN tagu sa však pridáva druhý pár MAC adries, ktoré oddeľujú vnútorné MAC adresy zákazníkov. Automatické použitie MAC-in-MAC v SPB je označované ako princíp SPBM. [17] [6]

Je nutné podotknúť, že Q-in-Q alebo MAC-in-MAC sa zvyčajne konfigurujú manuálne na vstupných zariadeniach do backbone siete. V prípade SPB sa tagovanie/pridávanie druhej Ethernetovej hlavičky deje automaticky.

Formát rámcu 802.1ah poskytuje identifikátor služby zvaný **I-SID**, ktorý kompletne separuje backbone MAC adresy a VLAN identifikátory, čo umožňuje jednoduchšiu virtualizáciu dátových centier (DC). Cieľom je plne separovať poskytované služby od fyzickej infraštruktúry siete pomocou odstránenia všetkých závislostí medzi protokolmi a fyzickou sieťou. Abstrakcia služieb siete je realizovaná tým, že mapuje jednu alebo viac VLAN-ov na jeden I-SID. [6] [2]

I-SID zároveň poskytuje mechanizmus pre granulárnu správu sieťovej prevádzky. Mapovaním služieb (aplikácií) na špecifický I-SID je možné jednoducho vytvoriť mission-specific end-to-end siete a kontrolovať prístup k týmto službám. [2]



Obr. 3.9: Paradigma prechodu rámcu sieťou pri použití protokolu SPB, zdroj [19]

Ako bolo bližšie popísané v kapitole 3.2, položky TRILL rámcu (MAC adresa, TTL, FCS) sú modifikované na každom uzle vnútri TRILL siete, čo pridáva sieti na komplexicite a zťažuje sa tým prípadné riešenie problémov. Pretože neexistuje jednoduchý spôsob

ako zistiť vybranú cestu pre konkrétny dátový tok, riešenie problémov musí byť vykonané hop-by-hop metódou na každom uzle siete. SPB implementácia eliminuje komplexnosť používaním jednoduchého vyhľadávania v prepínacej tabuľke a priradením najkratšej cesty sieťovej prevádzke smerom ku krajnému bodu. Táto skutočnosť zjednodušuje eventuálne riešenie problémov, pretože konkrétny dátovy tok môže byť identifikovaný na základe zdrojovej a cieľovej MAC adresy. [2]

	TRILL	**SPB**
Štandardizačné teleso	IETF	IEEE
Náhrada za STP?	Áno	Áno
Použitý protokol pre hľadanie najvýhodnejších ciest	IS-IS	IS-IS
Mechanizmus zaisťujúci bezslučkovosť	TTL a RPFC	RPFC
Zapúzdrenie Ethernetových rámcov	Formát TRILL + vonkajšia Ethernetová hlavička	Q-in-Q (pre SPBV), MAC-in-MAC (pre SPBM)
Úprava hlavičky na každom prvku	Áno, na hraničných aj tranzitných uzloch	Nie, iba na hraničných uzloch
TTL	Áno	Nie
Potreba nového HW	Áno, TRILL definuje nový spôsob zapúzdrenia	Nie, pokiaľ prepínač zvláda MAC-in-Mac alebo Q-in-Q
Cesta pre unicast	Najkratšia cesta na základe IS-IS výpočtov	Najkratšia cesta na základe IS-IS výpočtov
Cesta pre broadcast/multicast	Záleží na zvolenom Rbridge; cesty pre unicast a broadcast/multicast môžu byť rozdielne	Rovnaká ako v prípade unicastu
Možnosť doručenia rámcov v inom poradí	Áno, v prípade že na začiatku nie je známa MAC adresa príjemcu	Nie
Učenie zákazníckych MAC adries	Učenie na krajných prístupových portoch + ESADI protokol	Učenie na okraji SPB siete
Agregácia služieb	Nie	Áno, viac VLANov môže byť zlúčených pomocou I-SID
Zložitosť riešenia problémov	Vysoká zložitosť, treba analyzovať prevádzku hop-by-hop metódou na zistenie celej cesty	Menšia zložitosť, je viditeľná cesta cez celú sieť

Tabuľka 3.3: Porovnanie vlastností protokolov SPB a TRILL, vypracované z [17] [2]

Hlavným rozdielom medzi protokolmi SPB a TRILL v spôsobe prepínania rámcov. Potom čo IS-IS zostaví topológiu siete, SPB vyberie najkratšiu cestu na základe metriky linky

a následne nasmeruje sieťovú prevádzku touto vybranou cestu. Je veľmi jednoduché predikovať celú cestu toku dát, keďže je vypočítaná iba raz. Pri použití 802.1aq môžeme za pomoci sieťového analyzátoru určiť celú cestu, ktorou daná sieťová prevádzka pôjde, na základe zdrojovej MAC adresy, cieľovej MAC adresy a VLAN ID.

TRILL oproti tomu využíva dva rôzne mechanizmy na prepínanie paketov založené na type sieťovej prevádzky. V prípade unicastu, kedy je známy vonkajší RBridge, TRILL používa IS-IS link-state databázu na priradenie optimálnej cesty (mechanizmus podobný ako u SPB). Avšak v prípade multicastu a broadcastu TRILL využíva distribučné stromy a RBridge ako koreň pre prepínanie rámcov. V mnohých prípadoch tieto cesty nie sú zhodné a je možné že rámce nedorazia do cieľa v rovnakom poradí. Preto je v niektorých prípadoch obtiažne poznať presnú cestu doručenia rámcov. [2]

Pri reálnom nasadzovaní má SPB určitú výhodu v tom, že neprináša žiadny úplne nový typ zapúzdrenia rámcov. Využíva MAC-in-MAC a Q-in-Q zapúzdrenia, ktoré existujú už nejakú dobu. Nasadenie do existujúcich prepínačov, ktoré ovládajú tieto 2 typy zapúzdrení, by mala byť len otázkou upgradu firmwaru. Nie je teda potrebný nový hardware.

TRILL však definuje novú vlastnú hlavičku rámcu, na ktorú nie sú existujúce prepínače pripravené. Použitý hardware toto zapúzdrenie nezvládne a preto je nutná jeho výmena. [17]

Existuje množstvo ďalších rozdielov medzi týmito 2 protokolmi, ktoré sú zhrnuté v tabuľke 3.3.

3.4 Proprietárne protokoly

Takmer každý veľký výrobca prepínačov má dnes vlastný proprietárny protokol určený ako náhrada za STP. Niekedy sú založené na otvorených štandradoch, inokedy majú vlastný základ. Spravidla platí, že podpora pre takéto protokoly existuje iba pre najvyšše rady prepínačov určených pre dátové centrá a poskytovateľov služieb. Ako už názov napovedá, prepínače rôznych výrobcov potom nie sú medzi sebou kompatibilné čo v reále znamená, že jadro sieťe musia tvoriť prepínače jedného výrobcu ak chceme vyťažiť potenciál siete na maximum.

3.4.1 Cisco

Cisco vyvinulo protokol zvaný **FabricPath**, ktorého základom je práve TRILL. V základnej konfigurácii nie je s protokolom TRILL kompatibilný kvôli prioprietárnej nadstavbe, ktorá určité vlastnosti TRILLu vylepšuje. Cisco uvádza, že FabricPath je možného prepnúť do režimu, kedy bude s TRILLom plne kompatibilný. Tieto dva módy sa ale navzájom vylučujú a pri zapnutej kompatibilite s protokolom TRILL nie je možné využívať výhody FabricPathu. FabricPath okrem iných vymožeností pridáva možnosť zlučovať viac VLAN do skupín, na ktoré je následne možné aplikovať rôzne politiky. Tento mechanizmus vykonáva podobnú funkciu ako *I-SID* v protokole SPB. TRILL v základe túto funkciu nemá. FabricPath je momentálne dostupný na prepínačoch Nexus rady 7000 a 5500. [21]

Na vybraných radách prepínačov Cisco, ktoré z HW dôvodov nebudú vedieť TRILL, je možné nasadiť proprietárny protokol *L2MP*(Layer 2 Multipathing). L2MP je podobný protokolu TRILL s tým rozdielom, že nezapúzdruje rámce a je v niektorých funkciách obmedzený (môže byť nasadený iba na dvojbodových linkách). Na druhú stranu L2MP oproti TRILLu obsahuje funkcie naviac (*MAC Conversational Learning* a spoluprácu s protokolom *vPC*).[17]

3.4.2 Ostatní výrobcovia

O konkrétnych princípoch fungovania proprietárnych protokolov týkajúcich sa problematiky bezslučkovosti L2 sietí sa vie všeobecne veľmi málo. Vždy sa však jedná o nahradenie STP a uvedenie nového spôsobu virtualizácie L2 sietí s cieľom maximálneho využitia danej siete (predovšetkým z hľadiska komplexnosti a škálovateľnosti). Výrobcovia sa sústredia predovšetkým na vývoj protokolov pre dátové centrá s ohľadom na rapídny nárast cloud computingu. Juniper predstavil proprietárny protokol *QFabric*, Brocade pomenoval svoj *VCS* (Virtual Cluster Switching).

Kapitola 4

IS-IS

Obsah kapitoly je zameraný na popis základných princípov IS-IS (Intermediate System-to-Intermediate System) smerovacieho protokolu. Predmetom popisu je špecifická sieťová vrstva, formát jednotlivých typov správ protokolu, nadväzovanie susedstva a šírenie smerovacích informácií. Hlavné informácie z tejto kapitoly sú čerpané zo zdrojov [12, 4, 13, 16]. Doplňujúce informácie pochádzajú zo zdrojov [9, 10, 8]

V predchádzajúcej kapitole bolo spomenuté, že protokoly TRILL a SPB využívajú upravenú verziu smerovacieho protokolu IS-IS na výpočet optimálnej cesty rámcov pre prenos po sieti. IS-IS má preto v rámci tejto práce nezanedbateľný význam a je mu venovaná samostatná kapitola. Popisovať však budem základnú verziu L3 IS-IS protokolu, z ktorej vychádzajú verzie modifikované pre TRILL a SPB. Princípy rôznych verzií sú rovnaké, líšia sa predovšetkým v type prenášaných informácií a upravenými typmi správ.

Protokol IS-IS bol štandardizovaný v roku 1992 a jeho najnovšia verzia je podrobne popísaná štandardom **ISO 10589:2002**. IS-IS je používaný v kombinácii s point-to-point linkami (sériové linky) a broadcast network (Ethernet LAN) a pre oba typy sa funkcionalita líši. Ja sa primárne zameriam na popis broadcastovej verzie. V reále je viac používaná verzia IS-IS rozšírená štandardom RFC 1195, ktorý popisuje použitie IS-IS v duálnom prostredí. To znamená, že okrem CLNS sa prenášajú smerovacie informácie aj pre IP. Ja sa však zameriam na čisto ISO CLNS verziu, ktorá je pochopenie základných princípov dostačujúca.

4.1 CLNS

V terminológii IS-IS protokolu sú všetky podporované zariadenia na sieti systémami. Rozlišujeme 2 druhy systémov:

- **End System (ES)** - Jedná sa o koncové zariadenia (hosts).
- **Intermediate System (IS)** - Smerovač podporujúci IS-IS.

IS-IS využíva službu OSI sieťovej vrstvy zvanú **CLNS** (Connectionless Network Service). CLNS nepotrebuje uzavrený okruh na doručovanie správ a jednotlivé správy sú doručované nezávisle na ostatných správach. CLNS nie je v dnešnej dobe ne Internete príliš rožšírené, pretože jeho úlohu zastáva majoritnejší IP protokol (ktorý využíva väčšina smerovacích protokolov ako RIP, OSPF, EIGRP a pod.). CLNS sa využíva predovšetkým v telekomunikačných sietiach po celom svete. Pakety, ktorými IS-IS komunikuje sú typu CLNS PDU (protocol data unit), takže všetky smerovače musia podporovať ISO CLNS

protokol. Preto nezáleží na IP adresách jednotlivých rozhraní (môžu byť z iných podsietí) a susedstvo sa aj tak naviaže.

Medzi sieťové protokoly využívajúce CLNS patria:

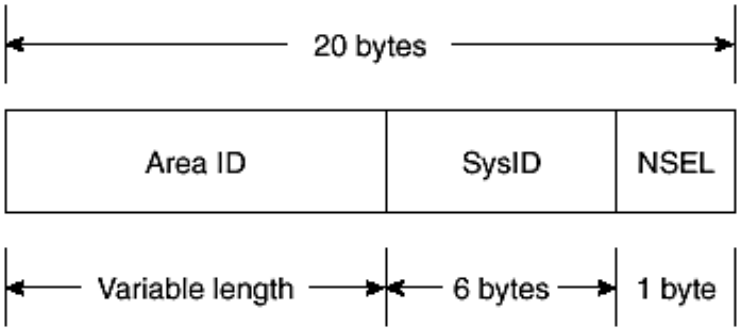
- **ISO 8473** - Connectionless Network Protocol (CLNP) je protokol poskytujúci komunikáciu s využitím CLNS.
- **ISO 9542** - End System-to-Intermediate System (ES-IS) je protokol slúžiaci na výmenu hello paketov a nadviazania susedstva mezi IS a ES.
- **ISO 10589** - Intermediate System-to-Intermediate System (IS-IS) je protokol určený na smerovanie v CLNS sieti.

4.2 NSAP

V počítačových sietiach je potrebné mať na každej vrstve konkrétny spôsob štandardizovaného adresovania. V IP sietiach sú to 32-bitové IP adresy (IPv4). CLNP adresy používané protokolom IS-IS sú takzvané **Network Service Access Points** (NSAP) adresy. Narozdiel od IP, NSAP adresa je asociovaná so sieťovým zariadením a nie sieťovým rozhraním. Každé zariadenie môže mať priradených viac NSAP adries (minimálne jedna je však potrebná).

NSAP adresa sa skladá z viacerých komponent, z ktorých najdôležitejšou časťou je *System Identifier*, ktorý musí byť unikátny v rámci CLNS siete. ISO terminológia označuje L2 adresy (MAC adresy, Frame Relay DLCI a pod.) ako **Subnetwork Point of Attachments** (SNPA). Pretože sieťové zariadenia môžu byť pripojené viacerými linkami, môžu mať viac SNPA adries ale potrebujú len 1 NSAP adresu, aby mohli bez problémov fungovať. Zvykom v IP sietiach je, že koncové stanice nemusia ovládať konkrétny dynamický smerovací protokol aby mohli komunikovať s ostatnými zariadeniami. Namiesto toho využívajú služby IP vrstvy ako ARP, DHCP alebo statické predvolené cesty. V čisto OSI sietiach sa koncové stanice (ES) spoliehajú čisto na ES-IS protokol, ktorého hlavnou úlohou je vykonávať mapovanie NSAP adries na SNAP adresy.

NSAP adresy nemajú fixne danú dĺžku a môžu byť až 160 bitov dlhé.



Obr. 4.1: Formát NSAP adresy [16]

Na obrázku 4.1 je znázornený zjednodušený formát NSAP adresy. V skutočnosti je adresa viac granulovaná a obsahuje viac menších častí, ktoré sú v tomto prípade zlúčené do väčších celkov. V rámci adresovania v IS-IS sieti je však toto delenie adekvátne a podrobnejší popis je nad rámec tejto práce.

NSAP adresu je preto možné rozdeliť na nasledujúce časti:

- **Area ID** - ID oblasti, do ktorej daný systém patrí. Jedná sa o pole s variabilnou dĺžkou. Bližší význam je popísaný v kapitole 4.3.
- **SysID** - ID systému, ktoré musí byť unikátne v rámci jednej oblasti. Dĺžka sa podľa štandardu líši od 1 do 8 bytov. Cisco však vo svojich implementáciách používa fixne danú dĺžku 6 bytov a tejto dĺžky sa budem držať aj naďalej.
- **NSEL** - NSAP Selector Value. Hodnota špecifikujúca užívateľa (network service user) služby sieťovej vrstvy. Môže mať nasledujúce hodnoty:

NSEL Value	Network Service User
`0x00`	Routing Layer
`0x21`	DECNet Phase IV Transport Layer
`0x22`	OSI Transport Layer TP4

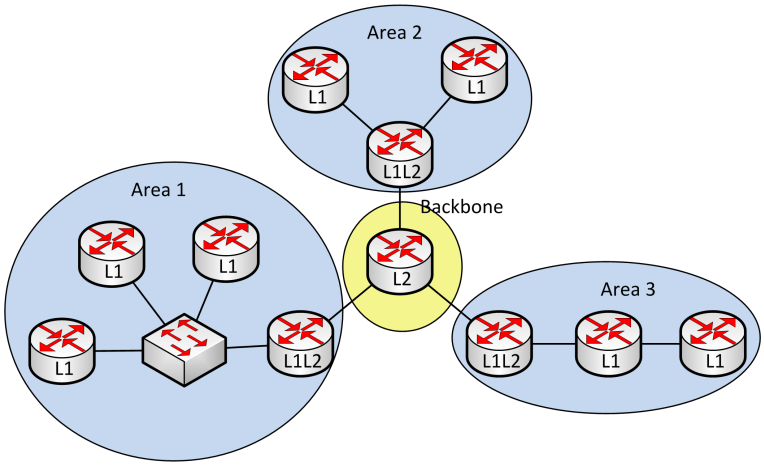
Tabuľka 4.1: NSEL hodnoty

Keďže IS-IS spadá do *Routing Layer*, NSEL bude mať vždy hodnotu 0. NSAP adresa s NSEL hodnotou rovnou 0 sa nazýva **Network Entity Title** (NET).

NSAP adresa môže vyzerať napríklad takto: `49.0001.1921.6800.1001.00` , kde `49.0001` je 3B Area ID, `1921.6800.1001` je 6B SysID a `00` reprezentuje NSEL.

4.3 Hierarchia systémov

Area ID v NSAP adrese smerovača identifikuje oblasť, do ktorej smerovač patrí. Každý IS-IS smerovač musí mať priradenú aspoň jednu NSAP adresu. V prípade, že ich má priradených viac, *SysID* ostáva rovnaké, ale *Area ID* môže byť rôzne a tým sa smerovač účastní výmeny smerovacích informácií vo všetkých zadefinovaných oblastiach (*multihoming*).



Obr. 4.2: Hierarchia systémov IS-IS [16]

Smerovače rozdeľujeme do 3 skupín:

- **Level 1** - Smerovače označené ako L1 sú si vedomé iba lokálnej topológie, ktorá zahrňuje všetky zariadenia (IS aj ES) v oblasti. To znamená že všetky takéto za-

riadenia musia mať priradenú rovnakú hodnotu *Area ID*. Proces smerovania v L1 je označovaný ako *intra-area routing*.

- **Level 2** - L2 smerovače sú používané v backbone sietiach na šírenie informácií medzi jednotlivými smerovacími doménami (oblasťami). L2 backbone zostavená z L2 smerovačov vytvára virtuálnu IS-IS oblasť. L2 oblasť musí byť súvislá a všetky smerovače v nej musia byť prepojené. Proces smerovania s využitím L2 je označovaný ako *inter-area routing*.
- **Level 1-2** - Hraničné smerovače sú označované ako L1L2 a predstavujú prepojenie medzi L1 a L2 oblasťou. Takýto smerovač obsahuje nezávislé link-state databázy pre oba typy oblastí. Smerovaču je potrebné určiť typ oblasti pre každé aktívne sieťové rozhranie.

Napriek tomu, že IS-IS podporuje súčastnú prítomnosť viacerých NET adries, je zvykom, že každý smerovač patrí práve do jednej oblasti. Hranice oblastí preto prechádzajú medzi smerovačmi. Príkladným protikladom je protokol OSPF, kde sa IP adresy zadávajú pre každé rozhranie a je preto bežné, že smerovače patrí do viacerých oblastí naraz.

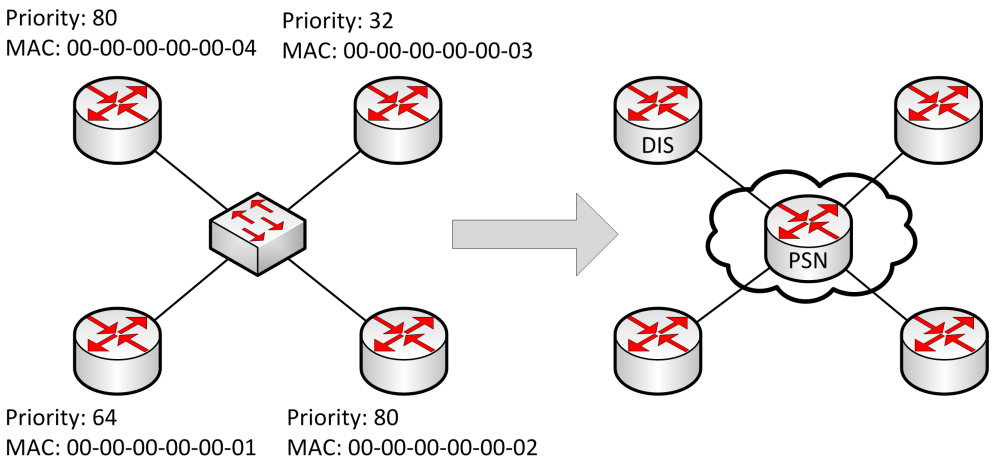
Na Cisco zariadeniach je implicitne predvolený L1L2 typ zariadenia, pokiaľ nie je manuálne nakonfigurovaný ináč.

Pri použití IS-IS s protokolom TRILL predstavuje celá TRILL doména jednu celistvú L1 oblasť, ktorá má *Area ID* nastavenú na hodnotu 0.

4.4 Designated Intermediate System

Veľkú úlohu protokolu IS-IS zohráva tzv. *Designated Intermediate System* (DIS). Jedná sa o IS, ktorý je volený na každej LAN. DIS je asociovaný len s multiaccess rozhraniami akým je Ethernet. V prípade point-to-point žiadny DIS neexistuje. DIS zastáva podobnú úlohu ako Designated Router v protokole OSPF.

DIS je volený na každej LAN na základe obsahu IIH (hello) paketov (kapitola 4.5.1), ktoré obsahujú v hlavičke pole *Priority*. Táto hodnota určuje prioritu IS, kde vyššia hodnota znamená vyššiu prioritu. Prípustné hodnoty sú od 0 do 127 (7 bitov). Implicitná hodnota je 64. V prípade, že IS majú nastavené rovnaké priority, vyhráva IS s vyššou hodnotou MAC adresy rozhrania, z ktorého bol IIH paket odoslaný.



Obr. 4.3: Voľba DIS v LAN sieti [16]

Majme scenár zobrazený na obrázku 4.3. Štyri IS sú navzájom prepojené Ethernetovými linkami za pomoci prepínača. IS-IS modeluje multiaccess linky ako uzly zvané *Pseudonodes*. Ako názov napovedá, jedná sa o virtuálny IS, ktorého rolu zastáva práve zvolený DIS. Smerovače si postupne prepošlú hello pakety a z obsiahnutých priorít sa určí, kto bude DIS. Keďže dva zo štyroch IS majú prioritu 80, čo je najvyššia hodnota na sieti, rozhoduje vyššia MAC adresa. IS, ktorý sa stane bude zvolený ako DIS sa pre všetky IS v LAN (vrátane seba) tvári ako ďalší virtuálny smerovač.

DIS má za úlohu minimalizovať veľkosť zasielaných link-state paketov po LAN a takisto udržovať link-state databázy jednotlivých IS na LAN konzistentné pomocou sequence number paketov. DIS je volený separátne pre L1 a pre L2. Narozdiel od OSPF, kde nájdeme *Backup Designated Router*, v IS-IS nič podobné neexistuje. V prípade, že aktuálny DIS stratí konektivitu, čo najskôr sa vyberie IS s druhou najvyššou prioritou. Voľba DIS je preemptívna a ak sa do sieti pripojí nový IS s najvyššou prioritou, okamžite preberá rolu DIS a nečaká sa kým doterajší DIS nebude schopný túto funkciu vykonávať.

Či je potrebné DIS voliť alebo nie, záleží na type linky. Pri multiaccess linkách, akým je Ethernet, môže byť prítomnosť DIS na LAN zbytočná v prípade, že sú takouto linkou priamo prepojené práve dva smerovače. Štandard ISO10589 takýto prípad nerozlišuje od akéhokoľvek iného zapojenia a DIS je preto aj v tomto prípade volený. Týmto problémom sa zaoberá až draft RFC 5309 (Point-to-Point Operation over LAN in Link State Routing Protocols).

Podrobnejšie informácie o funkciách DIS sa dozviete v nasledujúcich kapitolách. DIS v kombinácii s protokolom TRILL zastáva rolu DRB a má mierne odlišné povinnosti, ktoré boli popísané v kapitole 3.2.4.

4.5 Typy paketov

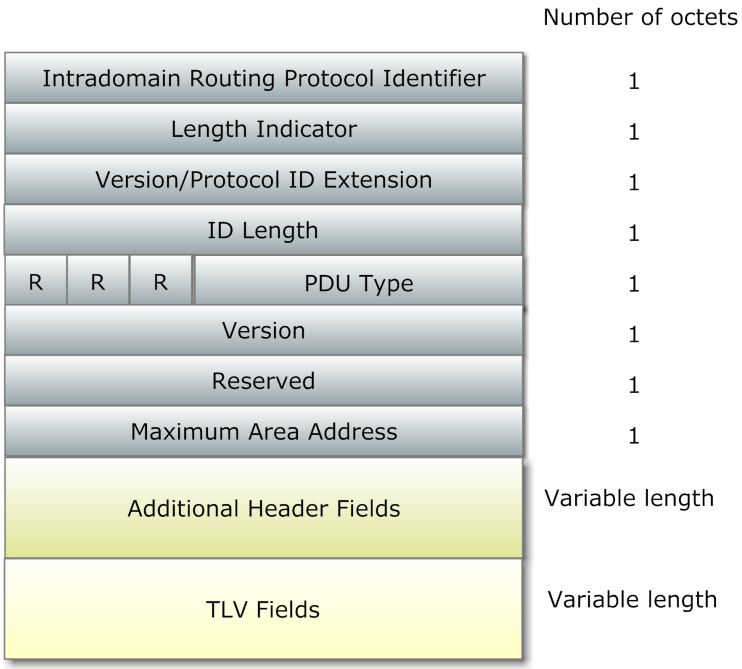
Keďže IS-IS nevyužíva k prenosu paketov IP, vyskytujú sa tu špecifické PDU. V IS-IS nájdeme 3 hlavné skupiny paketov (niekedy uvádzané 4), ktorých význam bude postupne ozrejmený v nasledujúcich podkapitolách. Na prenos smerovacích informácií sú využívané **Type-Length-Value (TLV)** polia, ktoré môžu niesť všeobecne akýkoľvek druh informácií a IS-IS je vďaka tejto skutočnosti jednoducho rožšíriteľný. Formát TLV polí je nasledovný:

- **Type** - Hodnota určujúca typ informácie obsiahnutej v poli *Value*. Do úvahy beriem iba typy popísané štandardom ISO 10589. Existujú rožšírenia akými sú napríklad typy obsiahnuté v štandarde RFC 1195 (Použitie IS-IS v duálnom prostredí) alebo RFC 6326 (Použitie IS-IS v protokole TRILL). Pole má veľkosť 1 byte.
- **Length** - Dĺžka poľa *Value* v bytoch. Jedná sa o 1 bytovú hodnota a celková veľkosť poľa *Value* môže byť preto maximálne 255 bytov. V TLV sa prenáša napríklad obsah link-state databáz, ktoré môžu byť značne obsiahle a v prípade značných veľkostí sa tieto informácie fragmentujú.
- **Value** - Hodnota s významom určeným v poli *Type* s veľkosťou určenou poľom *Length*.

Všeobecná IS-IS hlavička spoločná pre všetky typy paketov je znázornená na obrázku 4.4. Význam jednotlivých polí je nasledovný:

- **Intradomain Routing Protocol Discriminator** - Identifikátor sieťovej vrstvy priradený IS-IS špecifikovaný v štandarde ISO 9577. Hodnota pre IS-IS je `0x83` .

- **Length Indicator** - Veľkosť hlavičky paketu v bytoch.
- **Version/Protocol ID Extension** - Aktuálna verzia má hodnotu 1.
- **ID Length** - Veľkosť poľa source ID (SysID). Hodnota 0 znamená veľkosť 6bytov, hodnota 255 veľkosť 0 bytov. Iné prístupné hodnoty sú od 1 do 8.
- **PDU Type** - Pole určujúce typ paketu. Možné hodnoty sú zobrazené nižšie.
- **Version** - Aktuálna verzia má hodnotu 1.
- **Reserved** - Nepoužité bity určené pre eventuálne budúce použitie. Nesú hodnotu 0.
- **Maximum Area Addresses** - Najväčší možný počet priradených Area ID adries. Čísla 1 až 254 priamo reprezentujú hodnoty. 0 predstavuje maximálne 3 adresy.
- **Additional header fields** - Doplňujúce polia hlavičky špecifické pre každý druh paketu. Dĺžka je rôzna.
- **TLV Fields** - TLV polia špecifické pre každý druh paketu nesúce rôzne typy informácií.



Obr. 4.4: Všeobecná IS-IS hlavička [12]

IS-IS využíva nasledujúce typy paketov:

- LAN Level 1 hello packets (PDU Type 15)
- LAN Level 2 hello packets (PDU Type 16)
- Point-to-point hello packets (PDU Type 17)

- Level 1 link-state packets (PDU Type 18)
- Level 2 link-state packets (PDU Type 20)
- Level 1 complete sequence number packets (PDU Type 24)
- Level 2 complete sequence number packets (PDU Type 25)
- Level 1 partial sequence number packets (PDU Type 26)
- Level 2 partial sequence number packets (PDU Type 27)

Význam jednotlivých typov paketov bude bližšie popísaný v nasledujúcich kapitolách. L1 a L2 pakety rovnakého druhu sa líšia v poli *PDU Type*, niekedy v povolených typoch *TLV* polí a obsahom *TLV* polí.

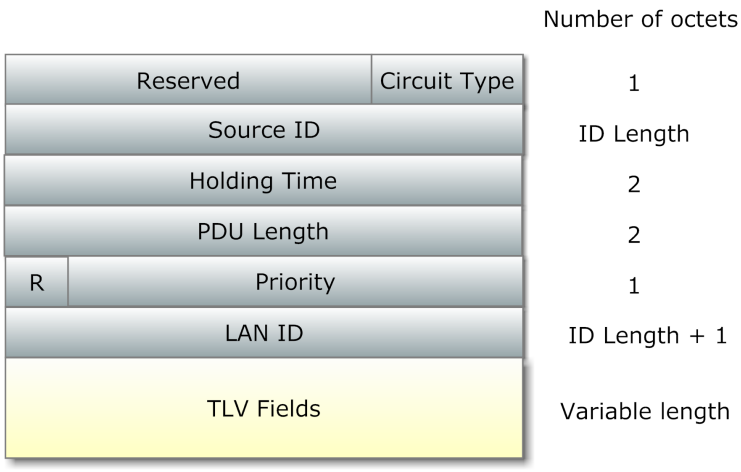
IS-IS využíva pri šírení všetkých druhov paketov systém floodingu lokálnych portov. Pri posielaní správ sa v hlavičke Ethernetového rámcu nastaví cieľová MAC adresa na vopred známu multicastovú adresu, na ktorej príjemci očakavájú príchod dát. K dispozícii sú nasledujúce MAC adresy:

Target Systems	**MAC address**
All L1 IS	`01-80-C2-00-00-14`
All L2 IS	`01-80-C2-00-00-15`
All IS	`09-00-2B-00-00-05`
All ES	`09-00-2B-00-00-04`

Tabuľka 4.2: Multicastové MAC adresy používané protokolom IS-IS

4.5.1 Hello packets

Hello pakety slúžia na nadviazanie susedstva prepojených systémov (ES aj IS). Sú označované aj ako IIH pakety (IS-IS Hellos). Rozlišujeme tri druhy IIH: Point-to-point IIH (PDU Type 17, Level 1 LAN IIH (PDU Type 15) a Level 1 LAN IIH (PDU Type 16). Venovať sa budem budem LAN hello paketom, ktoré majú rovnaký formát pre oba typy oblastí (L1 aj L2). Štruktúru LAN IIH paketu bližšie popisuje obrázok 4.5.



Obr. 4.5: IS-IS LAN hello paket [12]

Význam jednotlivých polí je nasledovný:

- **Reserved** - 6 bitová hodnota rezervovaná pre budúce použitie (hodnota 0).
- **Circuit Type** - 2 bitová hodnota určujúca typ smerovača vzhľadom na oblasti, v ktorých operuje:
  - 0 - rezervovaná hodnota (celé PDU by malo byť ignorované)
  - 1 - Level 1
  - 2 - Level 2
  - 3 - Level 1-2
- **Source ID** - *System ID* smerovača, ktorý hello paket vyslal.
- **Holding Time** - Čas v sekundách, za ktorý majú ostatné systémy považovať tento smerovač za mŕtvy pokiaľ sa dovtedy neozve pomocou IIH. Po uplynutí tejto doby je smerovač vyhodený z tabuľky susedov (maximálna hodnota 65535).
- **PDU Length** - Dĺžka celého PDU vrátane hlavičky a TLV.
- **Priority** - Bit 8 rezervovaný (hodnota 0) a ignorovaný pri prijatí. Zvyšných 7 bitov určuje prioritu smerovača pri volení Designated IS. Vyššia hodnota znamená vyššiu prioritu.
- **LAN ID** - Hodnota skladajúca sa zo *System ID* + 1 bytovej hodnoty určenej Designated IS identifikujúca konkrétnu LAN. Každé Ethernetové rozhranie v separátnej LAN má priradené unikátne LAN ID.
- **TLV Fields** - Špecifické TLV polia popísané nižšie.

Formát point-to-point IIH sa mierne líši, ale popis tohoto druhu linky nie je primárnym cieľom práce.

Hello pakety obsahujú TLV polia popísané v tabuľke 4.3

Type name	**Type value**	**Popis**
Area Addresses	1	Area ID nakonfigurované na prepínači
IS Neighbors	6	MAC adresy (SNPA) L1/L2 susedov od ktorých smerovač dostal aspoň jeden hello paket
Padding IS	8	IIH pakety sa posielajú zaplnené na hodnotu max MTU size, ktorú daný smerovač zvláda. Pole obsahuje samé nuly a plní funkcie výplne zvyšného miesta v pakete.
Authentication	10	Nepovinné pole obsahujúce heslo pre prípadnú autentizáciu.

Tabuľka 4.3: TLV polia IIH paketov. Vypracované z [12]

Každý IS periodicky vysiela každých 10 sekúnd IIH pakety všetkými lokálnymi portami, na ktorých je IS-IS povolené. L1 portami vysiela samozrejme L1 hello pakety, ktoré v TLV *IS Neighbors* majú obsiahnutých L1 susedov, od ktorých prijali aspoň jeden hello paket. Analogicky k tomu fungujú L2 hello pakety. Smerovače si držia separátne L1 a L2 tabuľky susedov.

Susedia v tabuľke sa môžu nachádzať v 2 stavoch: *Init* a *Up*. *Init* znamená iba jednosmerne ustanovené susedstvo a *Up* indikuje, že obaja susedia o sebe vedia a susedstvo je plne funkčné.

Pre nadviazanie susedstva sa využíva mechanizmus 3-way handshake. Pri štarte nemá smerovač v tabuľke žiadnych susedov. Preto prvý hello paket neobsahuje TLV *IS Neighbors*. Pripojený smerovač prijme hello paket a pokúsi sa podľa hodnoty *System ID* daný smerovač vyhľadať v tabuľke susedov. Tam ho však nenájde a preto vytvorí nový záznam obsahujúci System ID odosielateľa, sieťové rozhranie, ktorým paket dorazil a stav susedstva, ktorý označí ako *Init*. Tento smerovač rozpošle rovnakým spôsobom hello pakety, ktoré už ale obsahujú v TLV *IS Neighbors* prvý smerovač (ako aj všetky ostatné IS, ktoré sú v stave *Up* alebo *Init*). Ten po prijatí paketu zistí, že dostal hello paket od suseda, ktorého vo svojej tabuľke nemá, a že o ňom vie na základe informácií obsiahnutých v TLV. Preto pridá do tabuľky rovnakým spôsobom záznam, ale stav susedstva označí rovno ako *Up*, pretože našiel v cudzom TLV svoju MAC adresu. Prvý smerovač vyšle ďalšiu sadu hello paketov s obsiahnutou MAC adresou suseda, ktorý keď to zistí, uvedie susedstvo taktiež do stavu *Up*. Po tejto výmene IIH je funkčné obojsmerné susedstvo smerovačov.

Až po ustanovení susedstva, ktoré je v stave *Up* sa môže daný IS zúčastniť voľby Designated IS a šírenia link-state databáz. Pokiaľ daný IS plní funkciu DIS pre danú LAN, hello pakety vysiela daným rozhraním v tretinových intervaloch tj. 3,33 sekundy. Pri použití Autentizácie pomocou TLV sa samozrejme obsiahnuté heslá musia zhodovať, ináč je paket ignorovaný.

Pokiaľ IS od nejakého suseda nedostane žiadny hello paket po dobu určenú hodnotou *Holding Time*, daný IS je zmazaný z tabuľky susedov a takisto záznamy v link-state databáze pochádzjúce od neho sú zneplatnené. Typická hodnota pre *Holding Time* je 30 až 40 sekúnd.

Príklad výpisu IS-IS susedov na Cisco smerovači:

```
Router# show isis neighbors
System Id      Type Interface IP Address      State Holdtime Circuit Id
0000.0000.0002 L1   Et0/0     192.168.128.2   UP    21       R5.02
0000.0000.0002 L2   Et0/0     192.168.128.2   UP    28       R5.02
```

IIH pakety sú rozosielané vždy všetkými sieťovými rozhraniami (s aktivovaným IS-IS) s cieľovou multicastovou MAC adresou `01-80-C2-00-00-15` pre L2 susedov a `01-80-C2-00-00-14` pre L1 susedov. Pri použití IS-IS v kombinácii s protokolom TRILL sú multicastové MAC adresy odlišné. Ďalším rozdielom v TRILL IS-IS je, že hello pakety sa zbytočne nevypĺňujú pomocou TLV *Padding* na maximálnu veľkosť paketu danú hodnotou *max MTU size*.

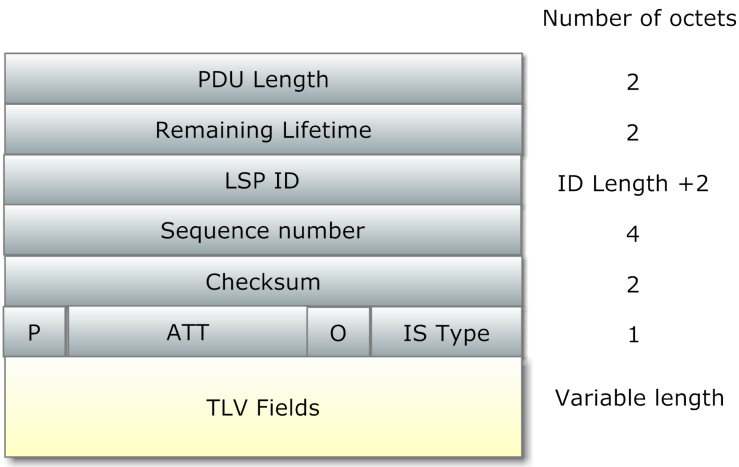
4.5.2 Link-state packets

Link-state pakety slúžia na synchronizáciu link-state databáz medzi jednotlivými IS. S tým je spojený proces tvorenia lokálnych link-state databáz smerovačov. Keďže IS-IS je link-state protokol, každý IS v oblasti si musí byť vedomý celej topológie, z čoho vyplýva, že všetky IS majú rovnaký obsah databáz. Databázy sú separátne pre L1 a L2 oblasti.

Formát IS-IS LSP paketu je znázornený na obrázku 4.6. Význam jednotlivých polí LSP paketu je nasledovný:

- **PDU Length** - Veľkosť PDU vrátane hlavičky v bytoch.

- **Remaining Lifetime** - Počet sekúnd, za ktoré má byť LSP záznam považovaný za neplatný, v prípade, že nedôjde k jeho obnoveniu.
- **LSP ID** - Identifikátor LSP záznamu skladajúceho sa z *Source ID* zdrojového IS (variabliná dĺžka), *Pseudonode ID* (1B) a *LSP Number* (1B). *Pseudonode ID* rovné 0 indikuje, že LSP pochádza od fyzického smerovača. Nenulové hodnoty identifikujú *Pseudonode* ako pôvodcu LSP paketu. *LSP Number* jednoznačne určuje fragment LSP paketu. Pokiaľ je LSP záznam príliš veľký a nezmestí sa do jedného paketu, *LSP Number* sa inkrementuje o 1 pre každý ďalší fragment (prvý fragment má hodnotu 0).
- **Sequence number** - Prvá instancia LSP záznamu má hodnotu 1. S každou ďalšou zmenou sa hodnota inkrementuje o 1 a porovnaním verzií u susedov sa overuje aktuálnosť záznamov.
- **Checksum** - Kontrolný súčet vypočítaný od poľa LSP ID do konca paketu.
- **Partition** - 8. bit indikujúci, či zdrojový IS podporuje funkciu *Partition repair*.
- **Attached** - 4 bity určujúce typy metrík, ktoré zdrojový IS používa
  - 4. bit - Default metric
  - 5. bit - Delay metric
  - 6. bit - Expense metric
  - 7. bit - Error metric
- **Overload** - 3. bit indikujúci preťaženie smerovača (CPU, pamať). LSP s týmto flagom nebude použité pre výpočet cesty skrz zdroj tohoto LSP paketu.
- **IS Type** - 2 bitová hodnota určujúca typ smerovača. Pre L1 je nastavený 1. bit, pre L2 sú nastavené oba bity, iné hodnoty nie sú definované.
- **TLV Fields** - Špecifické TLV polia popísané nižšie.



Obr. 4.6: Formát IS-IS LSP paketu [12]

Proces plnenia link-state databáze začína v momente, keď má daný IS aspoň s jedným susedom ustanovené susedstvo v stave *Up*. Každý IS a DIS má svoj separátny záznam v link-state databáze (záznam je označovaný ako LSP). LSP záznam je identifikovaný jednoznačne pomocou LSP ID. Má priradené aj ďalšie parametre ako sekvenčné číslo, kontrolný súčet a remaining lifetime

Príklad výpisu IS-IS link-state databázy na Cisco smerovači:

```
Router# show isis database
IS-IS Level-1 Link State Database
LSPID                 LSP Seq Num  LSP Checksum LSP Holdtime ATT/P/OL
0000.0000.0005.00-00  0x000007EF   0xDD14       667          0/0/0
0000.0000.0006.00-00  0x000007E7   0x2ECA       1126         0/0/0
0000.0000.0007.00-00* 0x000007FB   0x6FCB       960          1/0/0
0000.0000.0007.01-00* 0x000007E3   0xA91D       782          0/0/0
```

Samotný LSP záznam v sebe nesie *System ID* pripojených susedov daného IS (susedia v stave *Up*) a metriky k nim.

ISO 10589 definuje nasledovné typy metrík:

- **Default metric** - Hodnota nepriamoúmerná šírke pásma danej linky, čo znamená, že menšia hodnota znamená vyššiu šírku pásma linky a tým pádom lepšiu metriku. Tento druh metriky musí byť ako jediný podporovaný všetkými IS v sieti. Cisco vo svojich prepínačoch podporuje ako jedinú práve túto metriku. Implicitná metrika linky má hodnotu 10.
- **Delay metric** - Voliteľný druh metriky reprezentujúci oneskorenie linky.
- **Expense metric** - Voliteľná metrika udávajúca finančne závislú cenu linky.
- **Error metric** - Voliteľný druh metriky, ktorý zodpovedá miere chybovosti linky.

Záleží na konkrétnej implementácii, ktorý druh metriky sa použije pri výpočte optimálnej cesty.

IS-IS LSP pakety využívajú TVL zhrnuté v tabuľke 4.4

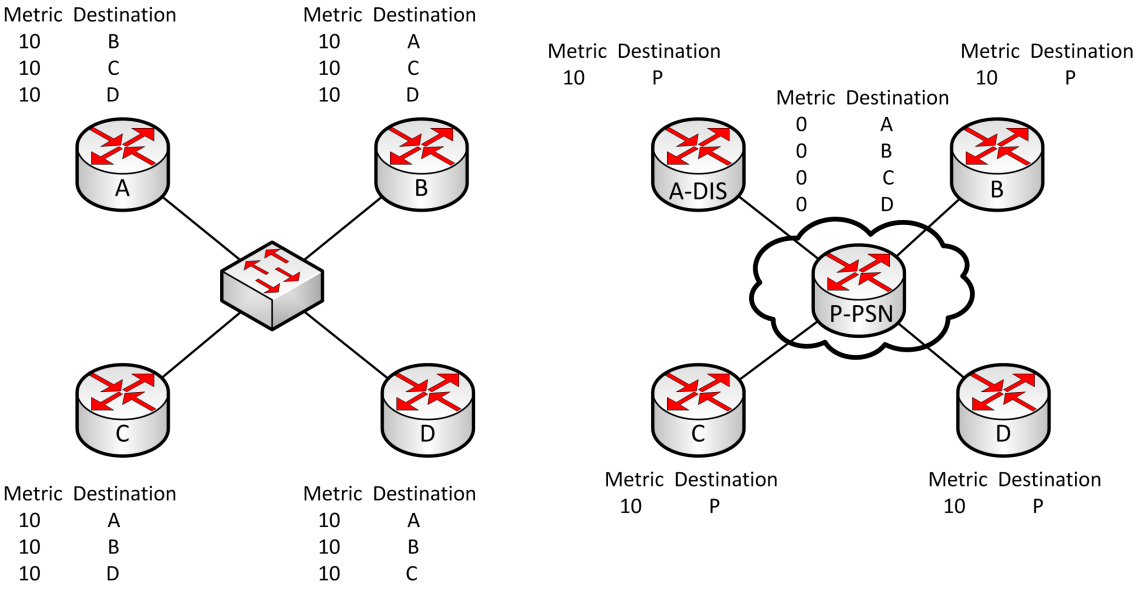
Po vytvorení lokálnej link-state databázy daný IS vyšle LSP pakety všetkými pripojenými rozhraniami s aktivovaným IS-IS na známe MAC multicastové adresy popísané v kapitole 4.5. Keď susedia prijmú LSP paket, porovnajú obsiahnuté LSP záznamy (predovšetkým hodnoty LSP ID a sequence number) so svojou databázou. Pokiaľ ich link-state databáza daný LSP záznam neobsahuje, vytvorí sa a následne sa paket ďalej prepošle na všetky aktívne IS-IS porty okrem toho, z ktorého paket prišiel. Ak daný záznam databáza obsahuje a lokálna hodnota sequence number je menšia ako má prijatý LSP paket, záznam sa aktualizuje a rovnako sa paket prepošle zvyšným susedom. Ak daný záznam lokálna link-state databáza obsahuje, ale hodnota sequence number je väčšia ako v prijatom pakete, vytvorí sa nový LSP paket s najaktuálnejšou verziou LSP záznamu a pošle sa všetkým priamo pripojeným susedom.

Kompletná link-state databáza sa z IS flooduje každých 900 sekúnd (15 minút) alebo v momente, keď sa nejakým spôsobom zmení topológia siete. Implicitná hodnota remaining lifetime je 1200 sekúnd (20 minút). Podotýkam, že na periodické kontroly verzií jednotlivých LSP záznamov slúžia práve *Sequence Number* pakety, ktoré sú posielané v omnoho kratších

Type name	Type value	IS Type	Popis
Area Address	1	L1,L2	Area ID nakonfigurované na prepínači.
IS Neighbors	2	L1,L2	System ID + Pseudonode ID identifikujúce dané IS asociované spolu s ich metrikou.
ES Neighbors	3	L1	System ID identifikujúce dané ES asociované spolu s ich metrikou.
Partition-Designated L2 IS	4	L2	TLV umožňujúce prepojiť prerušenú L1 oblasť cez virtuálnu cestu vedúcu cez L2 backbone. V súčasnosti nie je táto funkcia na Cisco smerovačoch podporovaná
Prefix Neighbors	5	L2	TLV združujúce NSAP prefixy s rovnakou metrikou.
Authentication	10	L1,L2	Nepovinné pole obsahujúce heslo pre prípadnú autentizáciu.

Tabuľka 4.4: TLV polia LSP paketov. Vypracované z [12]

intervaloch. Po uplynutí 20 minút bez aktualizácie záznamu je daný záznam označený ako neplatný. IS, ktorému časovač vyprší ako prvému vyšle LSP paket s LSP ID daného záznamu a sequence number nastaví na hodnotu 0, čím indikuje ostatným IS, že tento záznam je potrebné vymazať. Takýto LSP paket sa postupne šíri celou sieťou až kým nemajú všetky IS zosynchronizované link-state databázy. IS-IS má definovanú minimálnu dobu medzi poslaním 2 LSP paketov rovnú 33ms.



Obr. 4.7: Význam DIS v súvislosti s LSP

Majme zapojenú topológiu znázornenú na obrázku 4.7. Predpokladáme, že sieť je skonvergovaná, čo sa susedstva týka. V ľavej časti obrázku je zobrazený scenár, keby na sieti neexistoval DIS. Každý IS by mal vo svojom lokálnom LSP zázname adresy susedov a ich metriky k nim (všetky metriky rovné 10 v tomto prípade). Keď však smerovač `A` preberie úlohu DIS, všetky ostatné IS budú mať vo svojom lokálnom LSP zázname iba cestu k *Pse-*

udonodu (ktorý predstavuje smerovač `A`) s danou metrikou, čím sa zmenší obsah databáze a pri synchronizácii LSP záznamov sa tak nezahltí sieť ako pri absencii DIS. *Pseudonode* má svoj vlastný LSP záznam obsahujúci všetky fyzické smerovače v LAN sieti a metriky k nim majú vždy hodnotu 0.

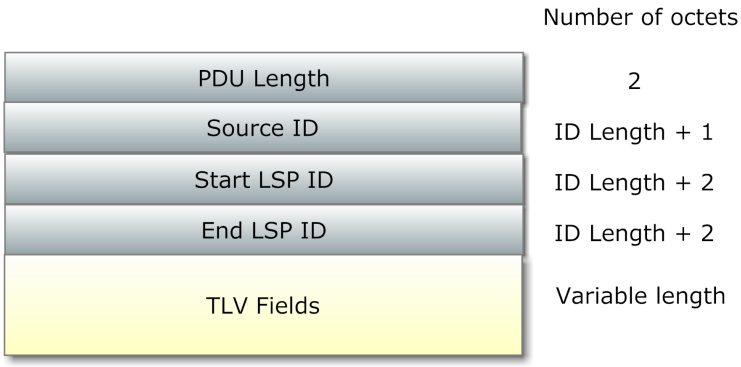
4.5.3 Sequence number packets

SNP pakety poskytujú efektívny spôsob synchronizácie link-state databáz medzi IS bez potreby prenášať kompletné link-state databázy.

Rozlišujeme 2 druhy SNP paketov:

- **Complete Sequence Number Packets (CSNP)** - Obrázok 4.8.
- **Partial Sequence Number Packets (PSNP)** - Obrázok 4.9.

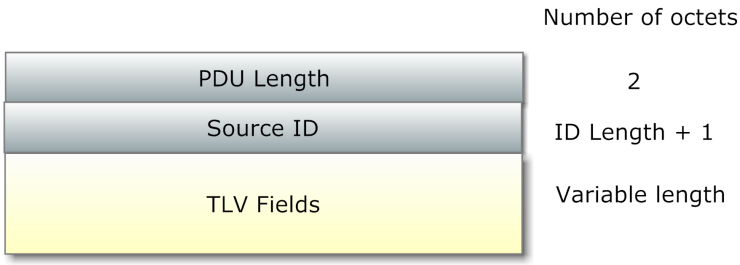
Oba typy paketov sa ďalej delia na L1 a L2, kde jediný rozdiel je, že pomocou L1 CNSP/PSNP sa prenášajú údaje z L1 link-state databáze a analogicky k tomu fungujú L2 CNSP/PSNP.



Obr. 4.8: Formát CSNP paketu [12]

Význam jednotlivých polí CSNP je nasledovný:

- **PDU Length** - Celková dĺžka PDU vrátane hlavičky v bytoch.
- **Source ID** - System ID smerovača, ktorý generoval NSP paket + 0 ako hodnota *Circuit ID* .
- **Start LSP ID** - Najnižšia hodnota LSP ID spomedzi všetkých prenášaných LSP záznamov v TLV. Cisco smerovače nastavujú hodnotu tohoto poľa na samé nuly.
- **End LSP ID** - Najvyššia hodnota LSP ID spomedzi všetkých prenášaných LSP záznamov v TLV. Cisco smerovače nastavujú hodnotu tohoto poľa na maximálnu možnú hodnotu. Polia *Start LSP ID* a *End LSP ID* sú teda fixne dané a dohromady pokrývajú celý rozsah možných LSP ID.
- **TLV Fields** - Špecifické TLV polia popísané nižšie.



Obr. 4.9: Formát PSNP paketu [12]

Význam jednotlivých polí vrátane TLV pre PSNP paket sú identické s CSNP. V CSNP aj PSNP paketoch sa vyskytujú TLV popísané tabuľkou 4.5

Type name	**Type value**	**Popis**
LSP Entries	9	Všetky LSP záznamy obsiahnuté v link-state databáze daného IS. Pole *Value* obsahuje hodnoty *LSP ID*, *Remaining lifetime*, *Sequence number* a *LSP checksum*.
Authentication	10	Nepovinné pole obsahujúce heslo pre prípadnú autentizáciu.

Tabuľka 4.5: TLV polia CSNP a PSNP paketov. Vypracované z [12]

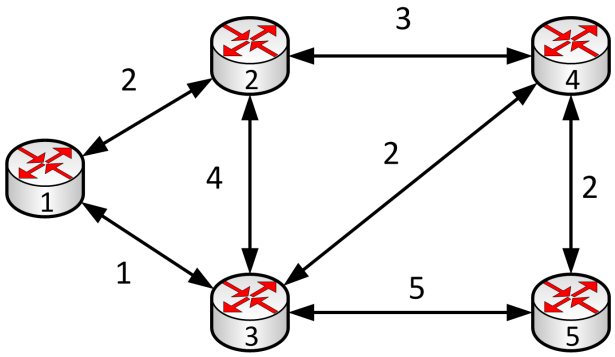
DIS flooduje CSNP pakety každých 10 sekúnd do celej LAN. Obsahom TLV *LSP Entries* je kompletný zoznam LSP záznamov daného IS. Neobsahuje však obsah záznamov (tj. zoznam susedných IS s metrikami) ako je tomu pri LSP paketoch a šetrí sa tým prenosové pásmo. Pri prijatí CSNP paketu sa porovná obsah link-state databázy s obsahom CSNP paketu. V prípade, že daný IS zistí, že niektorý záznam mu chýba alebo je neaktuálny, vyšle do siete PSNP paket, ktorý obsahuje v TLV LSP ID záznamu, ktorý chce nechať poslať a tým ho aktualizovať. Pôvodný odosielateľ CSNP paketu (DIS) odpovie tak, že do siete začne floodovať LSP paket so želaným obsahom LSP ID. Na point-to-point linkách slúžia PSNP pakety na potvrdzovanie doručenia a jedná sa teda o spoľahlivé doručovanie. V broadcastových sietiach však všetka komunikácia prebieha na základe mechanizmu floodingu a žiadne typy paketov nie sú potvrdzované.

4.6 SPF

IS-IS využíva **Shortest Path First (SPF)** algoritmus na výpočet optimálnej cesty ku všetkým ostatným systémom v sieti. Je známy aj pod menom Dijkstrov algoritmus, podľa Holandského vedca menom Edsger Dijkstra. Jedná sa o základný algoritmus z teórie grafov, ktorého primárnym využitím je hľadanie najkratšej cesty v hranovo ohodnotenom digrafe. Takýto graf pozostáva z množiny vrcholov N, množiny orientovaných hrán L a funkcie c, ktorá zobrazuje množinu hrán do množiny reálnych čísiel väčších ako 0.

Každý IS-IS smerovač v sieti pozná po skonvergovaní kompletnú topológiu siete v danej oblasti. Pri predstave, že každý IS predstavuje vrchol (*Vertex*) grafu, linky medzi nimi sú hrany (*Arcs*) grafu a metrika linky je výsledok ohodnocovacej funkcie dostaneme digraf, na ktorý je jednoducho aplikovateľný SPF algoritmus.

Majme topológiu (graf), ktorá je znázornená na obrázku 4.10. Uzly (vrcholy) grafu predstavujú smerovače označené 1 - 5. Keďže linky v počítačových sietiach bývajú obojsmerné



Obr. 4.10: Ukážková topológia pre výpočet optimálnych ciest pomocou SPF algoritmu

a hrany grafy musia byť orientované, je každá linka považovaná za dvojicu protichodných hrán. Každá z dvojíc hrán môže mať rôznu cenu (metriku) avšak v tomto prípade sú ceny v oboch smeroch rovnaké. Označenia používané pri popise algoritmu:

- N - Množina vrcholov.
- L - Množina hrán.
- $d(i, j)$ - Vzdialenosť od vrcholu i k vrcholu j.
- P - Množina vrcholov, ku ktorým bola už vypočítaná cesta od referenčného vrcholu s.
- $L(n)$ - Aktuálna cena (vzdialenosť/metrika) cesty od vrcholu s k vrcholu n.

Množiny pre uvedenú ukážkovú topológiu budú vyzerať nasledovne:

- $N = \{1, 2, 3, 4, 5\}$
- $L = \{(1, 2), (2, 1), (1, 3), (3, 1), (2, 3), (3, 2), (2, 4), (4, 2), (3, 4), (4, 3)(3, 5), (5, 3), (4, 5), (5, 4)\}$

Algoritmus sa skladá z 3 hlavných krokov: Inicializácia (1), výber nasledujúceho vrcholu (2) a aktualizácia najlepšej cesty (3) [16]. Formálnejší zápis vyzerá nasledovne:

1. Nastav $i = 0, P_0 = \{v_0 = s\}, L(s) = 0$
   $L(n) = d(s, n)$ ak n je priamo spojené s s. Ináč $L(n) = \infty$. Označkuj každý vrchol n ako $[L(n), s]$. Nastav $i = 1$.
2. Nájdi nasledujúci vrchol v_i taký, že $v_i \notin P$ a zároveň $L(v_i) = min\{L(n)\}$ pre všetky n. Presuň v_i do P. Nový označený vrchol je v_i.
3. Aktualizuj najkratšie cesty vrcholov, ktoré nie sú obsiahnuté v P:
   $L(n) = min\{L(n), L(v_i) + d(v_i, n)\}$
   Ak je $L(n)$ nahradené, aktualizuj označenie vrcholu n na $[L(n), NH(v_i)]$, kde $NH(v_i)$ je next hop k vrcholu v_i z s.
   Inkrementuj i o 1 $(i = i + 1)$.
   Ak $i == |N|$, zastav. Ináč pokračuj na krok 2.

Kapitola 5

Návrh a implementácia

Kapitola popisuje postup implementácie IS-IS modulu v prostredí OMNeT++ a hodnotí dosiahnutú funkcionalitu.

Po dohode s vedúcim práce som sa rozhodol pre implementáciu protokolu TRILL. Základným predpokladom pre funkčnosť TRILLu je prítomnosť IS-IS protokolu. Po dôkladnom naštudovaní som však zistil, že IS-IS je veľmi rozsiahly protokol. Výsledkom práce bolo, že k programovaniu TRILLu som sa ani nedostal a sústredil som sa predovšetkým na implementáciu protokolu IS-IS. Reálnym cieľom bolo preto vytvorenie IS-IS modulu v jazyku C++ do prostredia OMNeT++.

5.1 OMNeT++

OMNeT++ je simulačné prostredie s kalendárom diskrétnych udalostí, slúžiace primárne na simuláciu komunikácie v počítačových sietiach [22]. Základom je hierarchický systém modulov s presne definovanou funkciou, ktoré sú vytvárané v jazyku C++. Prepojenia medzi modulmi sú popísané pomocou jazyka NED. Jednotlivé moduly komunikujú medzi sebou pomocou takzvaných brán (*gates*). OMNeT++ je možné vďaka flexibilnej architektúre použiť aj v iných odvetviach akými je napríklad simulácia hardwarových architektúr alebo agentných systémov.

OMNeT++ je skôr platforma pre simuláciu sieťovej komunikácie ako samotný simulátor. V praxi sa často využíva v kombinácii s rôznymi frameworkami, ktoré doplňujú funkcionalitu formou rôznych modulov. Framework, ktorý budem využívať nesie názov INET. Obsahuje množstvo modulov pre podporu v oblasti TCP/IP, ktoré uľahčujú simulácie počítačových sietí.

5.2 Projekt ANSA

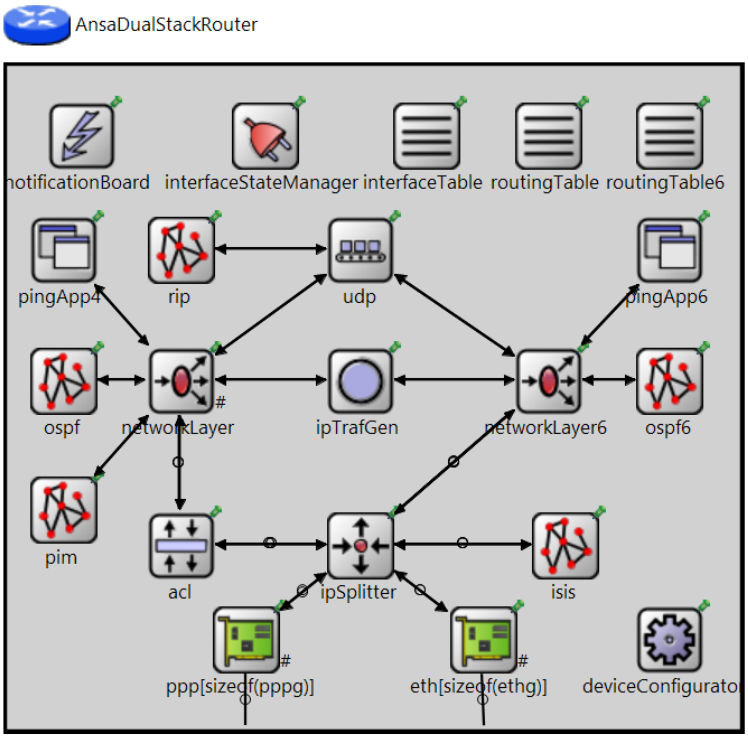
Projekt ANSA (Automated Network Simulation and Analysis) na FIT VUT v Brne [1] sa zaoberá skúmaním a rozširovaním balíku INET. Cieľom je poskytnúť komplexnú sadu modulov na simulácie reálnych počítačových sietí. Topológia siete a základná konfigurácia zariadení je načítaná z XML konfiguračného súboru.

Súčasne sa vývoj sústredí na rozširovanie komplexných modulov sieťových zariadení o nové protokoly. Jedná sa predovšetkým o moduly dual-stack smerovača a prepínača. Bakalárske a diplomové práce v posledných rokoch posunuli vývoj značne dopredu. Za zmienku

určite stojí podpora protokolov RIP, OSPF alebo STP. Došlo k úpravám smerovacej tabuľky a tabuľky sieťových rozhraní a bola pridaná podpora pre ACL.

5.3 Architektúra smerovača

Po zvážení kritérií som sa rozhodol implementovať IS-IS modul, ktorý bude ekvivalentom L3 IS-IS protokolu. Relatívne jednoduchými úpravami paketov a doplnením TLV polí je následne možné implementovať rozšírenie pre protokol TRILL alebo IP. Výsledný modul bude zakomponovaný do existujúceho modelu smerovača s názvom `ANSADualStackRouter`. Aktuálna architektúra tohoto modelu je znázornená na obrázku 5.1.



Obr. 5.1: Architektúra ANSA dual-stack smerovača

Súčasne sú podporované IPv4 a IPv6 protokoly, ktoré sú implementované ako samostatné sieťové moduly (`networkLayer` a `networkLayer6` na obrázku 5.1). CLNS však implementované nie je. CLNS sa v praxi v podstate využíva jedine v kombinácii s protokolom IS-IS a preto som CLNS neimplementoval ako samostaný modul, ale zahrnul som ho priamo do modulu `isis` spolu s rovnomenným smerovacím protokolom.

Všetky PDU vstupujúce do smerovača prechádzajú modulom `ipSplitter`, ktorý ich rozdeľuje na IPv4, IPv6 a po poslednom rozšírení už aj na CLNS pakety smerujúce do modulu `isis`.

Implementácia protokolu ES-IS ako samotného modulu použiteľného napríklad v modeli `ANSADualStackHost` nebola cieľom práce. Modul `isis` je preto možné použiť v smerovači, kde bude zastávať funkciu IGP (Interior Gateway Protocol).

5.4 Implementačné obmedzenia

Napriek tomu, že som sa snažil vytvoriť modul protokolu IS-IS podľa štandardu ISO10589, niektoré princípy sa mi nepodarilo vôbec implementovať alebo som ich funkčnosť pozmenil.

Pri implementácii som sa sústredil na Ethernet ako L2 nosný protokol. Spomínaný štandard definuje aj použitie IS-IS s point-to-point linkami, ale rozdiely vo funkčnosti sú oproti Ethernetu značné. Nepovažoval som to však za jednu z kľúčových funkcií a preto som ju vynechal.

Ďalšie obmedzenie sa týka hierarchií IS-IS smerovacích domén. Primárne som implementoval funkcionalitu v L1 oblasti. Funkčné je posielanie a prijímanie IIH (hello) paketov a s tým súvisiace nadväzovanie L1 aj L2 susedstiev, voľba DIS (len L1) a obsluha LSP paketov a vytváranie L1 link-state databázy (L2 zatiaľ nefunkčné).

Implementácia zatiaľ chýba pre obsluhu SNP paketov a pre vytváranie smerovacej tabuľky nad link-state databázou. Kľúčové funkcie protokolu sú implementované a pre zvyšok sú vytvorené kostry funkcií, ktoré by mali udávať formu budúceho vývoja.

Area ID a *SysID* polia NET adresy (NSAP) používané v IS-IS môžu mať podľa štandardu variabilnú dĺžku. Implementácie rôznych výrobcov sieťových zariadení sa však líšia a častým javom je obmedzovanie dĺžky podľa vlastných pravidiel (napr. Cisco používa fixne danú dĺžku SysID rovnú 6B). Ja som zvolil pevne danú dĺžku 10B pre celú NET adresu, z čoho 3B sú vyhradené pre *Area ID*, 6B *SysID* a 1B pre *NSEL*. Príklad takejto adresy je uvedený v kapitole 4.2. Pri aktivovanom IS-IS je povinné mať nakonfigurovanú práve jednu NET adresu (v súčasnej verzii nie je viac adries podporovaných).

Za zmienku stojí zmena týkajúca sa cieľovej MAC adresy pri posielaní. Pri simuláciách v prostredí OMNeT++ neboli rozpoznávané MAC adresy `01-80-C2-00-00-14` a `01-80-C2-00-00-15` ako multicastové a pri prijatí sa takéto rámce rovno zahadzovali. Preto som hodnotu cieľovej MAC adresy nastavil vždy na broadcast tj. `FF-FF-FF-FF-FF-FF`.

5.5 IS-IS modul

Modul je implementovaný ako samotná C++ trieda s názvom `ISIS`. Obsahuje všetky potrebné metódy pre funkčnosť protokolu. Priamo v triede sú zakomponované aj entity ako tabuľka susedov, link-state databáza alebo lokálna tabuľka sieťových rozhraní.

5.5.1 Adjacency table

Tabuľka susedov (Adjacency table) zohráva dôležitú úlohu v procese objavovania ostatných IS na sieti ako aj pri vytváraní link-state databáze. Tabuľka je reprezentovaná ako vektor všetkých susedov `std::vector<ISISadj>`. Existujú dve separátne tabuľky pre L1 a L2 susedov pomenované ako `adjL1Table` a `adjL2Table`. Štruktúra `ISISadj` obsahuje NET adresu suseda, jeho MAC adresu, port ku ktorému je pripojený, stav (*Up/Init*) a časovač udávajúci životnosť záznamu.

5.5.2 Link-state database

Link-state databáza je samostatná pre L1 a L2 podobne ako je tomu u tabuľky susedov. Implementovaná je ako vektor LSP záznamov `std::vector<LSPrecord>`. Štruktúra `LSPrecord` predstavujúca samotný záznam pozostáva z LSP ID, jeho sekvenčného čísla, časovača udávajúceho platnosť záznamu a zoznamu susedov, ktorý sa vzťahuje na IS, od

ktorého daný LSP záznam pochádza. Zoznam susedov daného IS obsahuje LAN ID suseda (*SysID + Pseudonode ID*) a metrika k nemu. Práve na základe zoznamu susedov a metriky k nim je možné zostaviť graf siete reprezentujúci danú topológiu. Po skonvergovaní majú všetky IS v oblasti rovnakú link-state databázu. Ďalším krokom po naplnení link-state databázy je zostavenie smerovacej tabuľky. Tento krok však v súčasnej verzii nie je implementovaný.

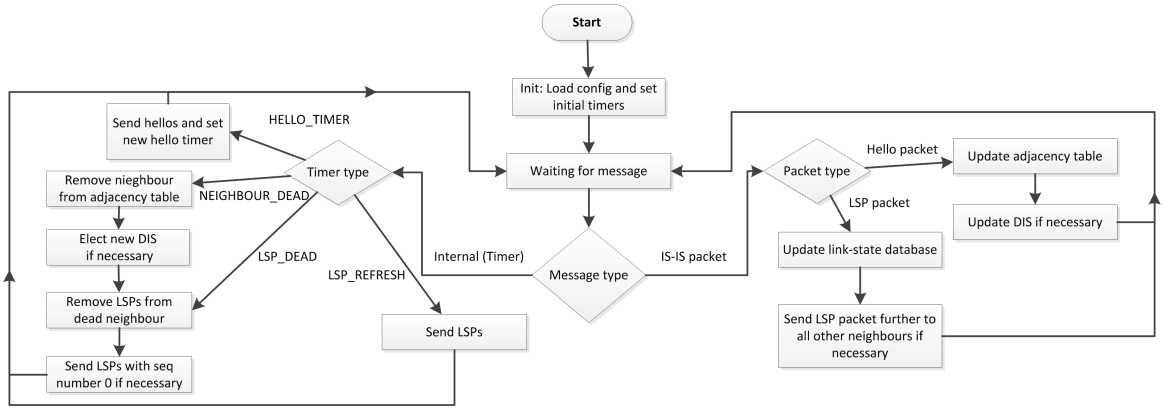
5.5.3 Interface table

Modul obsahuje vlastnú tabuľku sieťových zohraní. Základné informácie o existujúcich rozhraniach čerpá z modulu `InterfaceTable`. Pridávajú sa však informácie potrebné pre chod IS-IS protokolu. Jedná sa predovšetkým o parametre ako priorita rozhrania a metrika. Nesmieme zabudnúť ani na prioritu a LAN ID (*SysID + Psudonode ID*) aktuálneho DIS, ktorý sa nachádza na pripojenej LAN k danému rozhraniu.

Po zapnutí IS-IS na Cisco zariadeniach sa implicitne sieťové rozhrania neúčastnia žiadnej výmeny IS-IS informácií. Každé rozhranie sa preto musí explicitne aktivovať. V tom sa moja implementácia líši a po globálnej aktivácii IS-IS protokolu sú všetky sieťové rozhrania aktívne.

5.5.4 Aktivita modulu

Prvým krokom simulácie je inicializácia. Primárnou úlohou je načítať konfiguračný XML súbor a na základe získaných informácií nastaviť parametre IS-IS na smerovači. Potrebné je určiť predovšetkým NET adresu smerovača, priority rozhraní a metriky. V prípade, že niektoré parametre v konfiguračnom súbore chýbajú, použijú sa implicitné hodnoty. Viac o formáte a načítaní XML súboru sa dozviete v kapitole 5.6.



Obr. 5.2: Diagram aktivity IS-IS modulu

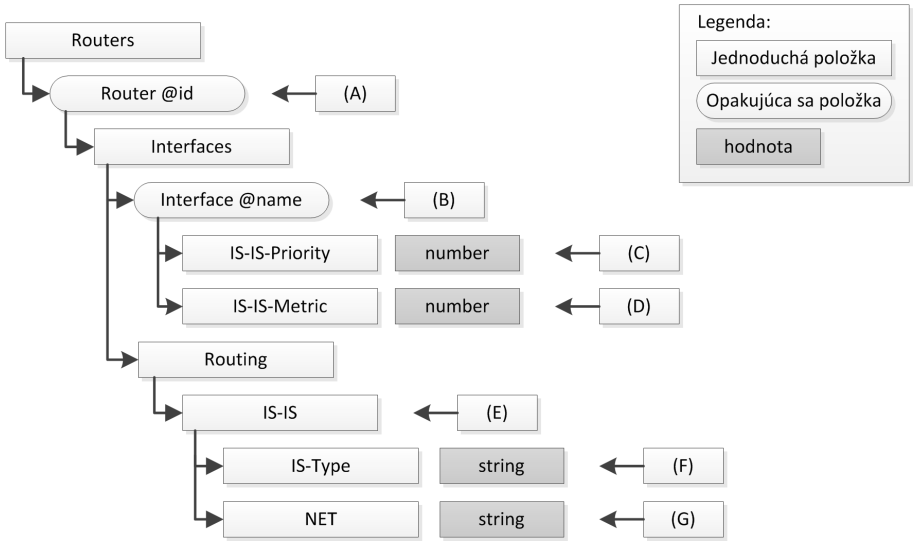
Súčasťou inicializácie je nastavenie dvoch časovačov. Prvý časovač sa spustí po uplynutí 1 sekundy prvé rozoslanie IIH (hello) paketov. Druhý časovač odštartuje po 15 sekundách rozoslanie LSP paketov. LSP pakety by mali byť vyslané pri každej zmene aktívnej topológie siete alebo každých 900 sekúnd, v závislosti na tom čo nastane skôr. Pri spustení smerovača sa v prvých sekundách známa topológia zo strany smerovača neustále mení z dôvodu prijímania nových hello paketov od susedov. Preto je vhodné počkať uvedenú dobu, kedy by už každý IS mal vedieť o svojich priamo pripojených susedoch a sieť by mala byť skonvergovaná z hľadiska nadviazaných susedstiev. Súčasťou inicializácie je v neposlednej

rade nastavenie parametrov sieťových rozhraní. Obrázok 5.2 popisuje zvyšnú činnosť IS-IS modulu.

Smerovač nasledovne periodicky vysiela hello a LSP pakety. Prijíma interné správy, ktoré reprezentujú svoje vlastné časovače po uplynutí stanovenej doby. Externé správy predstavujú pakety od ostatných IS na sieti. Uvedený diagram znázorňuje reálne implementovaný proces rozhodovania a nezobrazuje obsluhu SNP paketov, ktorých implementácia nie je zatiaľ prítomná.

5.6 Konfigurácia

Úspešné načítanie a rozparsovanie XML konfiguračného súboru je základom pre simuláciu IS-IS protokolu. Cieľom konfigurácie je nastaviť parametre zariadenia. Využil som existujúci formát konfiguračného súboru a rozšíril ho o nové tagy (obrázok 5.3).



Obr. 5.3: Štruktúra konfigurácie IS-IS v XML

(a) Definuje sekciu pre daný smerovač, `id` je určujúci reťazec ktorý sekciu identifikuje.

(b) Definícia sieťového rozhrania, reťazec `name` identifikuje konkrétne rozhranie (napr. `eth0`).

(c) Priradenie priority pre IS-IS rozhranie. Význam má pri voľbe DIS. Jedná sa o voliteľný tag, ktorý by mal obsahovať číselnú hodnotu v rozsahu v od 0 do 127. Ináč je nastavená implicitná hodnota 64.

(d) Voliteľný parameter udávajúci prioritu linky na danom rozhraní. Očakáva sa číselná hodnota v rozsahu od 0 do 63. Implicitná hodnota je 10.

(e) Povinný tag ak má byť aktivovaný IS-IS na smerovači.

(f) Voliteľný parameter udávajúci typ IS. Typ určuje reťazec s možnými hodnotami `level-1`, `level-2` a `level-1-2`. Implicitne je nastavená hodnota `level-1-2`.

(g) Definícia NET adresy IS. Očakáva sa reťazec v tvare, ktorý je popísaný v ukážke kapitoly 4.2. Zadanie práve jednej adresy je povinné pre funkčnosť IS-IS. Adresa musí byť

unikátna v rámci danej oblasti. Pri prijímaní hello paketov sa kontroluje eventuálna duplicita a v kladnom prípade sa vypíše varovanie informujúce o tomto probléme (rovnako sa v takomto prípade chovajú Cisco zariadenia). V prípade výskytu duplicitných adries nie je správanie definované.

Na načítanie XML súboru a parsovanie parametrov som využil existujúcu knižnicu `xmlParser`, ktorú som rozšíril o metódu `cXMLElement *GetIsisRouting(cXMLElement * device)`. Tá má za úlohu zistiť, či sa v súbore nachádza konfigurácia IS-IS a v kladnom prípade sa tento protokol aktivuje. Zvyšné parametre IS-IS majú len lokálny význam v rámci IS-IS modulu a preto sú načítané separátne bez použitia knižnice xmlParser. Na načítavanie konfiguračného súboru sa využíva modul `deviceConfigurator` ktorý obsahuje práve spomínanú knižnicu. Príklad XML konfiguračného súboru sa nachádza v prílohe B.

Kapitola 6

Simulácia

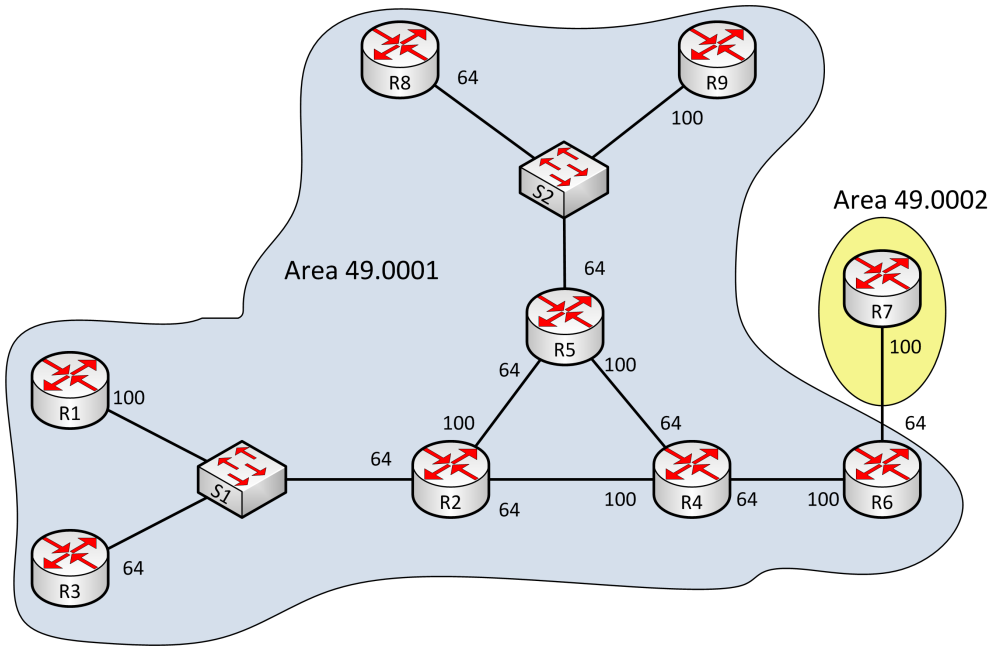
Kapitola popisuje proces simulácie vzorovej topológie, časovú analýzu jednotlivých udalostí a nakoniec validáciu samotnej implementácie.

V priebehu vývoja modulu som vykonal množstvo simulácií na rôznych typoch topológií. Pre ukážku si uvedieme jednu z nich, ktorá pokrýva väčšinu dosiahnutej funkcionality.

Validácia prebiehala porovnávaním komunikácie s identickou topológiou nakonfigurovanou v sieťovom simulátore GNS3 [3]. Ako referenčné zariadenia som použil smerovače Cisco 2691. Sledoval som časy odosielania jednotlivých paketov a ich obsah.

6.1 Topológia

Pre účel demonštrácie činnosti IS-IS som zvolil topológiu znázornenú na obrázku 6.1.



Obr. 6.1: Modelová topológia siete

V sieti sa nachádza 9 smerovačov označené ako R1-R9, ktoré sú priradené do 2 rôznych oblastí. Vzhľadom na to že medzi rôznymi oblasťami je zatiaľ funkčné iba nadväzovanie

susedstiev, nebolo v rámci demonštrácie potrebné vytvárať viac oblastí s väčším počtom smerovačov. Simulácia má predovšetkým ukázať proces nadväzovania susedstiev a vytváranie link-state databáze na L1 úrovni.

Aby bola voľba DIS na prvý pohľad jednoznačná, nastavil som priority rozhraní tak ako je uvedené na obrázku 6.1. Všetky linky majú priradenú implicitnú hodnotu 10. Smerovač R7 má hodnotu *IS-Type* nastavenú na `level-2`, hraničný smerovač R6 drží hodnotu `level-1-2` a zvyšné smerovače `level-1`.

Každý smerovač má NET adresu nastavenú podľa šablóny `AA.AAAA.1921.6801.200R.00`, kde `AA.AAAA` je adresa oblasti a `R` je číslo smerovača. Adresa smerovača R6 má teda hodnotu `49.0001.1921.6801.2006.00`.

6.2 Časová analýza

V rámci demonštrácie činnosti IS-IS je potrebné ukázať proces konvergencie na časovej osi. Podstatné udalosti nastanú v nasledujúcich časových okamihoch.

t=0s - Začína sa proces inicializácie zariadenia kedy sa načíta konfiguračný XML súbor, nastavia sa hodnoty parametrov IS-IS a sieťových rozhraní. Nakoniec sa nastavia časovače pre vyslanie prvých hello paketov (1 sekunda) a link-state paketov (15 sekúnd).

t=1-15s - Časovač štartuje vysielanie hello paketov všetkým priamo pripojeným susedom. Odosielanie sa na rozhraniach periodicky opakuje každé 3,33s v prípade, že smerovač zastáva úlohu DIS na pripojenej LAN. V opačnom prípade sú hello pakety vysielané v intervaloch 10s. Intervaly vysielania hello paketov na Cisco smerovači nájdete v obrázku 6.2. Komunikácia bola zachytená medzi prepínačom SW1 a smerovačom R2. Smerovače na tejto LAN vysielajú hello pakety v intervale 10s, až na smerovač R1, ktorý zastáva úlohu DIS a intervaly sú preto tretinové.

No.	Time	Source	Destination	Protocol	Length	Info
292	323.772000	c0:06:23:b4:00:00	ISIS-all-level-1-IS's	ISIS	1514	L1 HELLO, System-ID: 1921.6801.2001
293	324.333000	c0:04:19:90:00:00	ISIS-all-level-1-IS's	ISIS	1514	L1 HELLO, System-ID: 1921.6801.2003
294	325.243000	c0:03:19:90:00:00	ISIS-all-level-1-IS's	ISIS	80	L1 LSP, LSP-ID: 1921.6801.2002.00-00, Se
295	325.653000	c0:06:23:b4:00:00	ISIS-all-level-1-IS's	ISIS	80	L1 LSP, LSP-ID: 1921.6801.2001.00-00, Se
297	326.917000	c0:03:19:90:00:00	ISIS-all-level-1-IS's	ISIS	1514	L1 HELLO, System-ID: 1921.6801.2002
298	326.920000	c0:06:23:b4:00:00	ISIS-all-level-1-IS's	ISIS	1514	L1 HELLO, System-ID: 1921.6801.2001
299	329.581000	c0:06:23:b4:00:00	ISIS-all-level-1-IS's	ISIS	1514	L1 HELLO, System-ID: 1921.6801.2001
300	330.751000	c0:06:23:b4:00:00	ISIS-all-level-1-IS's	ISIS	116	L1 CSNP, Source-ID: 1921.6801.2001.00, S
302	332.403000	c0:04:19:90:00:00	ISIS-all-level-1-IS's	ISIS	1514	L1 HELLO, System-ID: 1921.6801.2003
303	332.413000	c0:06:23:b4:00:00	ISIS-all-level-1-IS's	ISIS	1514	L1 HELLO, System-ID: 1921.6801.2001
304	335.373000	c0:06:23:b4:00:00	ISIS-all-level-1-IS's	ISIS	1514	L1 HELLO, System-ID: 1921.6801.2001
306	336.543000	c0:03:19:90:00:00	ISIS-all-level-1-IS's	ISIS	1514	L1 HELLO, System-ID: 1921.6801.2002
307	338.143000	c0:06:23:b4:00:00	ISIS-all-level-1-IS's	ISIS	1514	L1 HELLO, System-ID: 1921.6801.2001
309	340.423000	c0:06:23:b4:00:00	ISIS-all-level-1-IS's	ISIS	116	L1 CSNP, Source-ID: 1921.6801.2001.00, S
310	341.433000	c0:06:23:b4:00:00	ISIS-all-level-1-IS's	ISIS	1514	L1 HELLO, System-ID: 1921.6801.2001
311	341.863000	c0:04:19:90:00:00	ISIS-all-level-1-IS's	ISIS	1514	L1 HELLO, System-ID: 1921.6801.2003
312	343.943000	c0:06:23:b4:00:00	ISIS-all-level-1-IS's	ISIS	1514	L1 HELLO, System-ID: 1921.6801.2001
315	346.173000	c0:03:19:90:00:00	ISIS-all-level-1-IS's	ISIS	1514	L1 HELLO, System-ID: 1921.6801.2002
316	346.753000	c0:06:23:b4:00:00	ISIS-all-level-1-IS's	ISIS	1514	L1 HELLO, System-ID: 1921.6801.2001
317	349.583000	c0:06:23:b4:00:00	ISIS-all-level-1-IS's	ISIS	116	L1 CSNP, Source-ID: 1921.6801.2001.00, S

Obr. 6.2: Úsek zachytenej komunikácie medzi SW1 a R2.

Na začiatku predstavuje každý smerovač Designated IS pre každú pripojenú LAN. Ako postupne sieť konverguje, smerovače si na základe uvedených priorít nastavujú adekvátny DIS.

V 15. sekunde simulácie máme zaručene skonvergovanú sieť z hľadiska susedstiev. Uveďme si výpis tabuľky susedov po skonvergovaní. Tabuľka susedov smerovačov R5 a R6 v prostredí OMNeT++:

```
L1 adjacency table of IS 49.0001.1921.6801.2005.00
No. of records in Table: 4
        1921.6801.2002        0a:aa:00:00:00:02        Up
        1921.6801.2004        0a:aa:00:00:00:0a        Up
        1921.6801.2008        0a:aa:00:00:00:15        Up
        1921.6801.2009        0a:aa:00:00:00:16        Up
------------------------------------------------------------------------
L2 adjacency table of IS 49.0001.1921.6801.2005.00
No. of records in Table: 0


L1 adjacency table of IS 49.0001.1921.6801.2006.00
No. of records in Table: 1
        1921.6801.2004        0a:aa:00:00:00:0b        Up
------------------------------------------------------------------------
L2 adjacency table of IS 49.0001.1921.6801.2006.00
No. of records in Table: 1
        49.0002.1921.6801.2007        0a:aa:00:00:00:11        Up
```

Pre porovnanie uvádzam výpis tabuľky susedov Cisco smerovačov R5 a R6:

```
R5#show isis neighbors
System Id      Type Interface  State Holdtime Circuit Id
R8             L1   Fa0/0      UP    29       R9.01
R9             L1   Fa0/0      UP    8        R9.01
R2             L1   Fa0/1      UP    8        R2.03
R4             L1   Fa1/0      UP    28       R5.01

R6#show isis neighbors
System Id      Type Interface  State Holdtime Circuit Id
R7             L2   Fa0/1      UP    7        R7.01
R4             L1   Fa0/0      UP    22       R6.01
```

Cisco na svojich smerovačoch pre sprehľadnenie výpisu zobrazuje samotný *hostname* namiesto *System ID*.

t=15s - Časovač odštartuje prvé vyslanie LSP paketov. Cisco vo svojich smerovačoch vysiela LSP pakety na začiatku pri každej zmene. Každý smerovač začne svojim susedom vysielať obsah svoje LSP záznamy, ktoré sú narozdiel od hello paketov preposielané ďalej až kým ich nepríjmu všetky smerovače v rovnakej oblasti. Po ukončenej konvergencii budú mať všetky smerovače v oblasti rovnaké link-state databázy. Pre uvedenú topológiu je link-state databáza príliš rozsiahla a preto je uvedená na konci tohto dokumentu. V prílohe C sa nachádza kompletný výpis link-state databázy v prostredí OMNeT++. Ekvivalentný výpis z Cisco smerovača nájdete v prílohe D. Ich porovnaním zistíme, že podstatné informácie sa zhodujú. Líšia sa v označovaní *Pseudonode ID*

časti LSP ID. Spôsob číslovania však nie je štandardizovaný a každá implementácia môže mať vlastný spôsob.

t>15s Periodicky sa opakuje rozosielanie hello paketov. LSP pakety sú vysielané každých 900 sekúnd alebo pri zmene aktívnej topológie siete.

Kapitola 7

Záver

V rámci diplomovej práce som naštudoval protokoly zaisťujúce bezslučkovosť. Zameral som sa na nové protokoly nahradzujúce STP, ich výhody a nevýhody. Analyzoval som existujúce implementácie nových protokolov. Po dohode s vedúcim práce som sa rozhodol pre implementáciu protokolu TRILL. Po podrobnom naštudovaní som však zistil, že TRILL je pre kompletnú implementáciu v rámci diplomovej práce príliš rozsiahly protokol. Pre funkčnosť TRILLu je kritická prítomnosť smerovacieho protokolu IS-IS a preto som sa naň zameral. Samotnému návrhu a implementácii predchádzalo podrobné naštudovanie štandardov a z nich vyplývajúce mechanizmy fungovania.

Práca bola vytvorená v rámci výskumnej skupiny ANSA, kde som načerpal cenné skúsenosti. Bolo potrebné zoznámiť sa so simulačným prostredím OMNeT++ a predovšetkým s aktuálnym stavom vyvíjaných modelov sieťových zariadení. Zameral som sa na naštudovanie architektúry modelu dual-stack smerovača, do ktorého bolo mojim cieľom zakomponovať modul simulujúci IS-IS protokol.

Rozsah samotného IS-IS protokolu je obrovský a nepodarilo sa mi vytvoriť implementáciu, ktorá by obsahovala všetky funkcie protokolu. Elementárna funkcionalita však bola dosiahnutá. Funkčné sú procesy nadväzovania susedstva smerovačov, voľby Designated IS, tvorby a synchronizácie link-state databáze. Existujúca forma XML konfiguračného súboru bola rozšírená o nové tagy slúžiace na konfiguráciu IS-IS modulu.

Dosiahnutú funkcionalitu som porovnával s IS-IS komunikáciou nasadenou v reálnom prostredí. Jednalo sa o rôzne zapojenia Cisco zariadení a odchytávanie komunikácie medzi nimi. Pracoval som so zapojením fyzických zariadení v Cisco laboratóriu ale aj s virtuálnymi smerovačmi v simulátore GNS3. V miestach, kde štandard mätúcim spôsobom popisoval činnosť IS-IS protokolu, som za referenčné informácie bral práve zachytenú komunikáciu medzi Cisco smerovačmi. Tieto dáta zároveň slúžili na overenie správnosti fungovania implementovaného modulu.

Práca poskytuje potenciál pre budúci vývoj. Pre začiatok je potrebné doimplementovať obsluhu SNP paketov, upraviť existujúce mechanizmy pre fungovanie na L2 úrovni a vytvoriť smerovaciu tabuľku na základe obsahu link-state databázy. Eventuálne je možná implementácia zameraná na podporu point-to-point liniek. Ďalej je možné sa vybrať dvomi smermi: buď implementovať TLV pre podporu IP alebo mierne upraviť existujúce typy paketov a doplniť TLV pre podporu TRILLu.

Práca mi umožnila prehĺbiť si znalosti o fungovaní protokolov TRILL a IS-IS a dozvedieť sa viac o oblasti modelovania a simulácií počítačových sietí.

Literatúra

[1] ANSAWiki.
URL https://nes.fit.vutbr.cz/ansa/pmwiki.php

[2] Avaya: Compare and Contrast SPB and TRILL [online]. 2011.
URL http://www.avaya.com/uk/resource/assets/whitepapers/SPB-TRILL_Compare_Contrast-DN4634.pdf

[3] GNS3: Documentation.
URL http://www.gns3.net/documentation/

[4] Gredler, H.; Goralski, W.: *The Complete IS-IS Routing Protocol*. Springer, 2005, ISBN 1-85233-822-9.

[5] IEEE: IEEE Standard for Local and metropolitan area networks: Media Access Control (MAC) Bridges [online]. 2004-06-09.
URL http://standards.ieee.org/getieee802/download/802.1D-2004.pdf

[6] IEEE: 802.1ad - Provider Bridges [online]. 2006.
URL http://www.ieee802.org/1/pages/802.1ad.html

[7] IEEE: 802.1aq - Shortest Path Bridging [online]. 2011.
URL http://www.ieee802.org/1/pages/802.1aq.html

[8] IETF: Use of OSI IS-IS for Routing in TCP/IP and Dual Environments [online]. 1990.
URL http://tools.ietf.org/rfc/rfc1195.txt

[9] IETF: RFC 6325: Routing Bridges (RBridges): Base Protocol Specification [online]. 2011.
URL http://www.rfc-editor.org/rfc/rfc6325.txt

[10] IETF: RFC 6326: Transparent Interconnection of Lots of Links (TRILL) Use of IS-IS [online]. 2011.
URL http://www.rfc-editor.org/rfc/rfc6326.txt

[11] IETF: RFC 6327: Routing Bridges (RBridges): Adjacency [online]. 2011.
URL http://www.rfc-editor.org/rfc/rfc6327.txt

[12] ISO/IEC 10589:2002: *Information technology — Telecommunications and information exchange between systems — Intermediate System to Intermediate System intra-domain routeing information exchange protocol for use in conjunction with the protocol for providing the connectionless-mode network service (ISO 8473).*

ISO, Geneva, Switzerland, 2002.
URL http://standards.iso.org/ittf/PubliclyAvailableStandards/c030932_ISO_IEC_10589_2002(E).zip

[13] ISO/IEC 8473-1:1998: *Information technology — Protocol for providing the connectionless-mode network service: Protocol specification*. ISO, Geneva, Switzerland, 1998.

[14] Kozierok, C.: TCP/IP Guide: Circuit Switching and Packet Switching Networks [online]. 2005-09-20.
URL http://www.tcpipguide.com/free/t_CircuitSwitchingandPacketSwitchingNetworks.htm

[15] Kraus, Z.: *Modelování a analýza spolehlivosti počítačové sítě VUT*. FIT VUT v Brně, 2011, diplomová práce.

[16] Martey, A.; Sturgess, S.: *IS-IS Network Design Solutions*. Cisco Press, 2002, ISBN 1-57870-220-8.

[17] Miroslav Matuška: Seriál TRILL: Konečně náhrada za Spanning Tree? [online]. 2010.
URL http://www.lupa.cz/serialy/trill/

[18] Oracle: ORACLE SOLARIS 11 EXPRESS 2010.11: WHAT'S NEW [online]. 2010.
URL http://www.oracle.com/technetwork/server-storage/solaris11/documentation/solaris-express-whatsnew-201011-175308.pdf

[19] Pepelnjak, I.: TRILL and 802.1aq are like apples and oranges [online]. 2010-08-02.
URL http://blog.ioshints.info/2010/08/trill-and-8021aq-are-like-apples-and.html

[20] Perlman, R.: *Interconnections: Bridges, Routers, Switches and Internetworking Protocols*. Addison-Wesley 2nd editioní, 2001, ISBN 0-201-63448-1.

[21] Systems, C.: Scale Data Centers with Cisco FabricPath [online]. 2011.
URL http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9402/white_paper_c11-605488.pdf

[22] Varga, A.: OMNeT++ 4.2 Manual [online]. 2011.
URL http://www.omnetpp.org/doc/omnetpp/Manual.pdf

[23] Wikipedia.: Ethernet - Wikipedia, The Free Encyclopedia [online]. 2012.
URL http://en.wikipedia.org/wiki/Ethernet

Dodatok A

Obsah CD

/ansainet/ - revízia zdrojových súborov projektu ANSAINET z 22.5.2012

/ansainet/src/ansa/isis/ - zdrojové súbory IS-IS modulu

/ansainet/src/ansa/ipSplitter/ - upravené zdrojové súbory modulu ipSplitter

/ansainet/src/ansa/deviceConfigurator/ - upravené zdrojové súbory modulu deviceConfigurator

/ansainet/src/ansa/ANSADualStackRouter.ned - rozšírený model dual-stack smerovača

/ansainet/src/ansa/isis/ - zdrojové súbory IS-IS modulu

/ansainet/examples/ansaExamples/ISIS/ - ukážková simulácia činnosti IS-IS

/ansainet/examples/ansaExamples/ISIS2/ - simulácia činnosti protokolu IS-IS s odlišnou topológiou

/text/ - zdrojové súbory výsledného textu tejto práce

/projekt.pdf - vysádzaný text práce

/readme.txt - popis prekladu a spustenia projektu ANSAINET

Dodatok B

XML konfiguračný súbor pre IS-IS

```
<?xml version="1.0"?>
<Routers>
    <Router id="192.168.12.1">
        <Interfaces>
            <Interface name="eth0">
               <IPAddress>192.168.12.1</IPAddress>
               <Mask>255.255.255.0</Mask>
               <IS-IS-Priority>30</IS-IS-Priority>
               <IS-IS-Metric>2</IS-IS-Metric>
             </Interface>
        </Interfaces>
        <Routing>
             <IS-IS>
                <IS-Type>level-1</IS-Type>
                <NET>49.0001.1921.6801.2001.00</NET>
             </IS-IS>
        </Routing>
    </Router>
</Routers>
```

Dodatok C

Obsah link-state databázy v prostredí OMNeT++

```
L1 LSP database of IS 49.0001.1921.6801.2003.00
No. of records in database: 14
        1921.6801.2003.00-00        0x00000001
                1921.6801.2001.01       metric: 10
        1921.6801.2001.00-00        0x00000001
                1921.6801.2001.01       metric: 10
        1921.6801.2002.00-00        0x00000001
                1921.6801.2002.01       metric: 10
                1921.6801.2001.01       metric: 10
                1921.6801.2004.01       metric: 10
        1921.6801.2001.01-00        0x00000001
                1921.6801.2002.00       metric: 00
                1921.6801.2003.00       metric: 00
                1921.6801.2001.00       metric: 00
        1921.6801.2002.01-00        0x00000001
                1921.6801.2005.00       metric: 00
                1921.6801.2002.00       metric: 00
        1921.6801.2004.00-00        0x00000001
                1921.6801.2004.01       metric: 10
                1921.6801.2005.01       metric: 10
                1921.6801.2006.01       metric: 10
        1921.6801.2005.00-00        0x00000001
                1921.6801.2005.01       metric: 10
                1921.6801.2002.01       metric: 10
                1921.6801.2009.01       metric: 10
        1921.6801.2004.01-00        0x00000001
                1921.6801.2002.00       metric: 00
                1921.6801.2004.00       metric: 00
        1921.6801.2005.01-00        0x00000001
                1921.6801.2004.00       metric: 00
                1921.6801.2005.00       metric: 00
        1921.6801.2006.00-00        0x00000001
```

```
        1921.6801.2006.01       metric: 10
1921.6801.2008.00-00        0x00000001
        1921.6801.2009.01       metric: 10
1921.6801.2006.01-00        0x00000001
        1921.6801.2004.00       metric: 00
        1921.6801.2006.00       metric: 00
1921.6801.2009.00-00        0x00000001
        1921.6801.2009.01       metric: 10
1921.6801.2009.01-00        0x00000001
        1921.6801.2005.00       metric: 00
        1921.6801.2008.00       metric: 00
        1921.6801.2009.00       metric: 00
```

Dodatok D

Obsah link-state databázy na Cisco smerovači

```
R3#show isis database detail
IS-IS Level-1 Link State Database:
LSPID                 LSP Seq Num  LSP Checksum  LSP Holdtime      ATT/P/OL
R1.00-00              0x0000000E   0x54C3        1035              0/0/0
  Area Address: 49.0001
  Hostname: R1
  Metric: 10         IS R1.01
  Metric: 0          ES R1
R1.01-00              0x0000000B   0x5952        763               0/0/0
  Metric: 0          IS R1.00
  Metric: 0          IS R2.00
  Metric: 0          IS R3.00
R2.00-00              0x00000010   0xC291        1182              0/0/0
  Area Address: 49.0001
  Hostname: R2
  Metric: 10         IS R2.03
  Metric: 10         IS R4.01
  Metric: 10         IS R1.01
  Metric: 0          ES R2
R2.03-00              0x0000000A   0x5D9B        1094              0/0/0
  Metric: 0          IS R2.00
  Metric: 0          IS R5.00
R3.00-00            * 0x0000000E   0xA46D        967               0/0/0
  Area Address: 49.0001
  Hostname: R3
  Metric: 10         IS R1.01
  Metric: 0          ES R3
R4.00-00              0x0000000E   0x400A        821               0/0/0
  Area Address: 49.0001
  Hostname: R4
  Metric: 10         IS R4.01
  Metric: 10         IS R6.01
```

```
  Metric: 10         IS R5.01
  Metric: 0          ES R4
R4.01-00              0x0000000A   0x2CCD        580               0/0/0
  Metric: 0          IS R4.00
  Metric: 0          IS R2.00
R5.00-00              0x0000000E   0xD370        1118              0/0/0
  Area Address: 49.0001
  Hostname: R5
  Metric: 10         IS R5.01
  Metric: 10         IS R9.01
  Metric: 10         IS R2.03
  Metric: 0          ES R5
R5.01-00              0x0000000A   0x6491        794               0/0/0
  Metric: 0          IS R5.00
  Metric: 0          IS R4.00
R6.00-00              0x0000000B   0xA557        922               1/0/0
  Area Address: 49.0001
  Hostname: R6
  Metric: 10         IS R6.01
  Metric: 0          ES R6
R6.01-00              0x00000008   0x7083        634               0/0/0
  Metric: 0          IS R6.00
  Metric: 0          IS R4.00
R8.00-00              0x0000000B   0x34C9        553               0/0/0
  Area Address: 49.0001
  Hostname: R8
  Metric: 10         IS R9.01
  Metric: 0          ES R8
R9.00-00              0x0000000B   0x5C9E        871               0/0/0
  Area Address: 49.0001
  Hostname: R9
  Metric: 10         IS R9.01
  Metric: 0          ES R9
R9.01-00              0x0000000A   0x6A2A        1051              0/0/0
  Metric: 0          IS R9.00
  Metric: 0          IS R8.00
  Metric: 0          IS R5.00
```