1.1 Moderní technologie přepínaných lokálních sítí

1.1.1 Přepínání a přepínače

Většina současných instalací lokálních sítích je založena na technologii Ethernet s použitím přepínačů (angl. switch). Přepínače propojují stanice do hvězdicové topologie s tím, že jednotlivé hvězdy mohou být dále propojeny do stromu a za předpokladu použití algoritmu Spanning Tree pro dynamické zablokování smyček i do složitějších topologií s redundandními spoji.

1.1.1.1 Funkce přepínače

Oproti rozbočovačům (angl. hub) přinášejí přepínače tu výhodu, že zavedením vyrovnávacích pamětí eliminují kolize v síti a dovolují plně duplexní přenos na dvoubodových spojích mezi stanicí a přepínačem a přepínači navzájem. Přepínače také šetří přenosové pásmo tím, že podobně jako dříve mosty analyzují hlavičky procházejících rámců a posílají rámce selektivně na správně cílové porty, za kterými leží stanice s MAC adresou příjemce. V případě potřeby současného přenosu více rámců na tentýž výstupní port přepínače dočasně uloží rámce ve vyrovnávací paměti (bufferu). Na rozdíl od klasických mostů (bridge), což jsou více či méně specializované počítače, jsou přepínače realizovány hardwarově a dovolují současný průchod více rámců přepínačem, pokud procházejí mezi vzájemně různými dvojicemi portů.

1.1.1.2 Automatické učení přepínací tabulky

Přepínač je zařízení, které není třeba nijak konfigurovat. Přepínací tabulky, které udržují informaci, za kterým portem se nachází která MAC adresa, si totiž budují zcela automaticky z procházejících rámců. Kdykoli do přepínače vstoupí rámec, přepínač se podívá na MAC adresu zdroje a zaeviduje si, ze kterého portu rámec s touto adresou vstoupil. Jelikož aktivní topologie nesmí obsahovat smyčky, bude rámec od jisté pevně připojené stanice vstupovat vždy jediným konkrétním portem. Zaznamenanou informaci o portu, za kterým leží jistá MAC adresa, přepínač využije v okamžiku, kdy bude mít za úkol na správný port přeposlat rámec právě s touto cílovou MAC adresou. Jestliže na přepínač dojde rámec s cílovou MAC adresou, který v přepínací tabulce doposud není, přepínač jej rozešle na všechny porty (což principiálně odpovídá chování rozbočovače). Tomuto rozeslání říkáme flooding (zaplavování). Přepínač bude posílat rámec s neznámou MAC adresou na všechny porty do té doby, než cílová stanice sama vyšle nějaký rámec, čímž přepínači prozradí, za kterým portem se ukrývá. V souvislosti s automatickým učením přepínací tabulky často mluvíme o samoučících nebo také transparentních mostech (přepínačích).

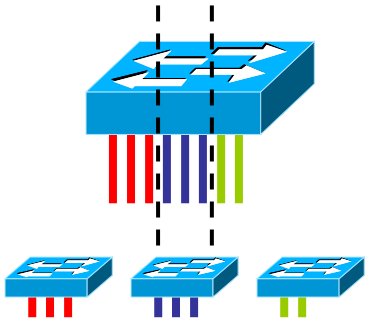
Aby si přepínač nemusel pamatovat MAC adresy všech stanic, které se v síti jedenkrát objevily a pak byly odpojeny (např. notebooky zákazníků apod.), udržuje se každá MAC adresa v přepínací tabulce jen po určitou dobu (typicky 3 minuty od okamžiku, kdy z dané MAC adresy přišel poslední rámec). Pokud měl uživatel stanici zapojenu v některém portu a následně ji přemístí do jiného, budou se rámce pro danou stanici posílat na původní port až do té doby, než stanice vyšle nový rámec se svou zdrojovou MAC adresou z nového portu. V tomto okamžiku se v přepínací tabulce ihned přepíše port, za kterým je daná stanice právě připojena a další provoz se již k této stanici bude posílat správně.

1.1.2 Virtuální lokální sítě

Řada výrobců v posledních letech do přepínačů implementujte podporu tzv. virtuálních sítí (virtual LAN, VLAN). Tento mechanismus dovolují rozdělit porty přepínače do skupin - virtuálních sítí (obr. 1.1a) s tím, že provoz může procházet vždy jen mezi porty téže skupiny (na obr. 1.1a zobrazené toutéž barvou).

Pro každou virtuální síť si přepínač buduje a používá samostatnou přepínací tabulku. Přepínač je tak logicky rozdělen na více nezávislých logických přepínačů a to čistě na základě konfigurace přepínače, bez potřeby jakéhokoli fyzického přepojování (obr. 1.1b). Správce sítě tak získává možnost seskupování stanic do vzájemně nezávislých (virtuálních) sítí vzdáleně, zpravidla pomocí WWW rozhraní nebo Telnetu. Logická struktura sítě se tím stává nezávislou na fyzické topologii.

Obr. 1.1a – Přiřazení portů přepínače do virtuálních sítí

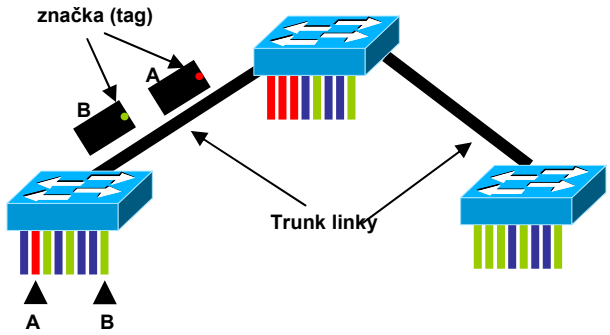


Obr. 1.1b – Rozdělení přepínače na oddělené logické přepínače pro jednotlivé VLAN

1.1.2.1 Trunk spoje a standard IEEE 802.1q

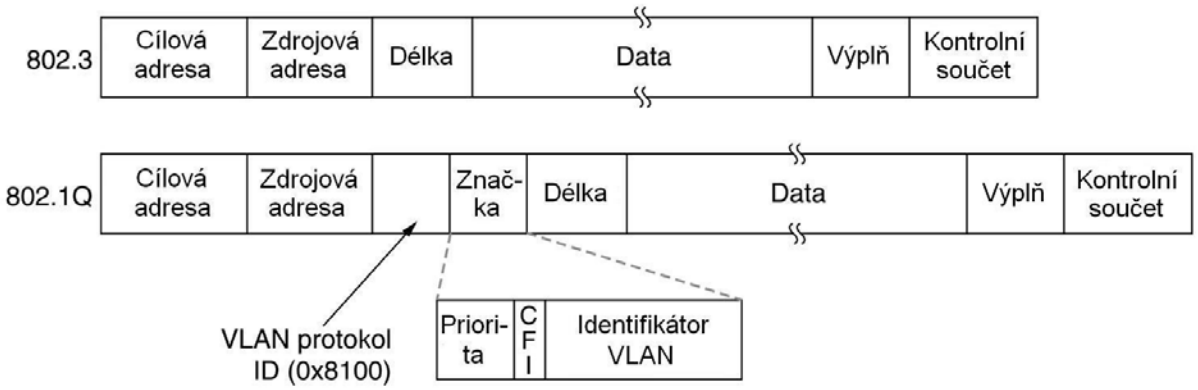
Pro praktické použití je často užitečné seskupit do společné virtuální sítě nejen porty jediného přepínače, ale porty různých, vzájemně propojených přepínačů. Spoje mezi přepínači však v tomto případě musí nést současně provoz z více virtuálních sítí. Protože přepínač, na který došel rámec musí být schopen rozlišit, do jaké virtuální sítě příchozí rámec patří, je nutné na spojích mezi přepínači identifikovat příslušnost jednotlivých rámců k virtuální síti, ze které byl vyslány (obr. 1.2). K tomu se ke každému rámci připojuje zvláštní hlavička – tzv. značka (angl. tag). Spoje mezi přepínači, na kterých jsou rámce označovány značkou, jsou často nazývány „trunk“ linkami a musí být do tohoto režimu zpravidla explicitně konfigurovány. Jen některé modely přepínačů se (s

použitím proprietárních protokolů) dokáží samy dohodnout, zda linka bude pracovat v režimu trunk nebo standardní.



Obr. 1.2 – Indikace příslušnosti rámce k virtuální síti na trunk linkách

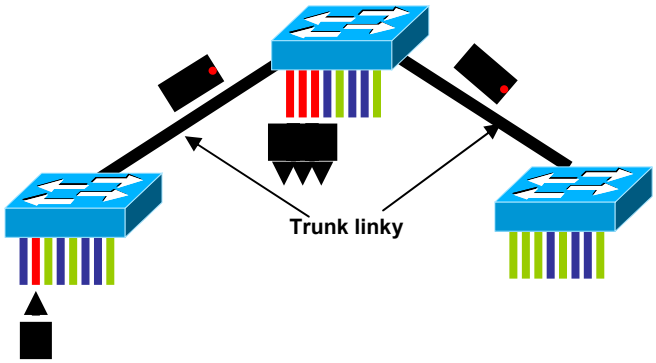
Je zřejmé, že na trunk spojích musí být rámce Ethernetu poněkud modifikovány, aby bylo kam uložit značku identifikující VLAN. V počátcích implementace virtuálních sítí do přepínačů jednotliví výrobci zaváděli svůj vlastní formát rámce rozšířeného o značku a tudíž nebylo možné trunk linkami propojovat zařízení různých výrobců. Později došlo ke standardizaci a dnes prakticky všichni výrobci podporují standard IEEE 802.1q, který používá standardního formátu rámce Ethernetu a přítomnost hlavičky obsahující číslo VLAN v rozsahu 0-4095 vložené na začátek datového pole rámce indikuje zvláštní hodnotou pole EtherType (obr. 1.3). Mimo čísla VLAN hlavička podle normy 802.1q může nést i prioritu rámce. Původní hodnota pole EtherType pak následuje za hlavičkou 802.1q.



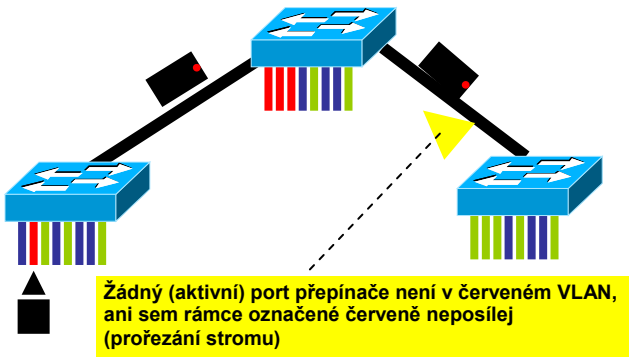
Obr. 1.3 – Rámec Ethernetu s hlavičkou 802.1q nesoucí identifikátor VLAN

Záhlaví podle IEEE 802.1q prodlužuje rámec o 2B. Hardware síťových karet, resp. portů přepínačů tak musí být schopny zpracovávat rámce s maximální délkou (MTU) o 2B vyšší, než definuje klasická norma Ethernetu.

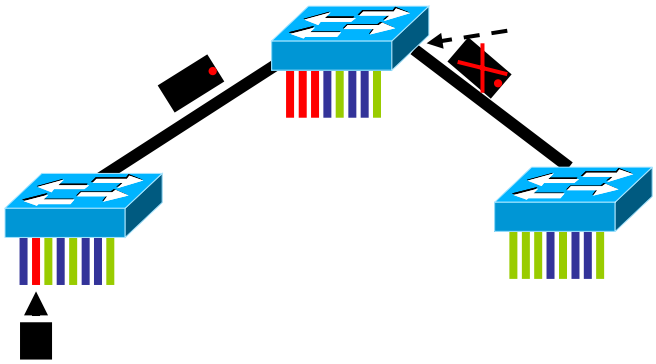
Rámce vyslané na broadcast adresu nebo rámce, u kterých není přepínači znám port příjemce, jsou zaslány na všechny porty přepínače v téže VLAN a dále označkované na všechny trunk porty (obr. 1.4). Zde však dochází k situaci, kdy přes trunk linky k cílovému přepínači rámce putují rámce i z těch VLAN, do nichž není na přijímajícím přepínači zařazen žádný port. Proto přepínače některých výrobců implementují protokol pro prořezávání topologie jednotlivých VLAN, jímž se jednotlivé přímo propojené přepínače informují o číslech VLAN, do nichž mají připojení aktivní porty (obr. 1.5). Jedná se zpravidla o protokoly proprietární (např. Cisco VTP), bylo však již definováno i standardní řešení GVRP (GARP VLAN Registration Protocol) na bázi IEEE protokolu Generic Attribute Registration Protocol (GARP) . V situacích, kdy jsou kombinována zařízení různých výrobců, která nepodporují shodný protokol dynamického prořezávání, je často možné filtrovat VLAN, která mají být propouštěny jednotlivými trunk linkami s použitím statické konfigurace (obr. 1.6).



Obr. 1.4 – Rozeslání rámce po virtuální síti



Obr. 1.5 – Pořezávání topologie jednotlivých VLAN



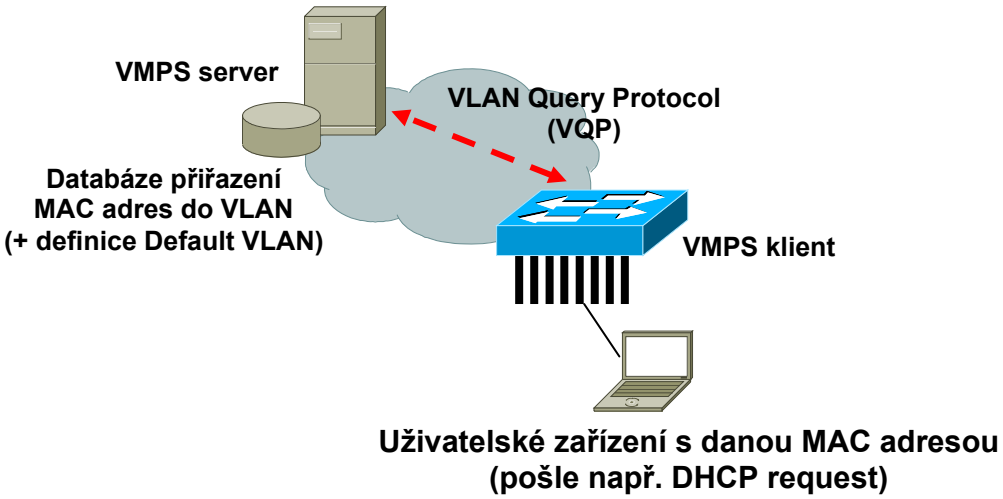
Obr. 1.6 – Statické omezení topologie VLAN

1.1.2.2 Přiřazování stanic do virtuální sítě

Stanice mohou být do virtuálních sítí přiřazovány buďto staticky nebo dynamicky. Statické přiřazení je založeno na statickém přiřazení jednotlivých portů přepínačů do VLAN a stanice je pak přiřazena do VLAN na základě toho, přes který port se do přepínače připojí. U dynamického přiřazování naopak nezáleží na tom, do kterého portu přepínače je stanice připojena a její přiřazení se děje na základě údajů z rámce, který do portu vstupuje – typicky podle zdrojové MAC adresy, ale lze (alespoň teoreticky) uvažovat i o přiřazení na základě IP adresy, protokolu nebo i údajů ze 4. vrstvy OSI RM, určujících aplikaci.

Z implementačních důvodů dynamické přiřazování stanic do VLAN typicky předpokládá, že na portu je připojena jen jediná stanice nebo je stanic i více, ale všechny mají být přiřazeny do téže VLAN. Pak stačí přepínači zjistit přiřazení stanice do VLAN jen při příchodu prvního rámce na port (obvykle jde o DHCP request), svázat tento port se zjištěnou VLAN a dále již pracovat stejným způsobem jako u statického přiřazení portu do VLAN. Přiřazení stanic do VLAN na základě MAC adresy se realizuje za použití databáze, která požadované mapování udržuje. V současné době bohužel neexistuje protokol, který by formát dotazu a odpovědi na přiřazení do VLAN standardizoval.

Obecně lze říci, že dynamické přiřazování do VLAN v dnešní době není běžné. Známá je hlavně implementace na přepínačích Cisco, které ke komunikaci se strojem udržujícím přiřazení stanic do VLAN používají proprietární protokol VLAN Query Protocol (VQP). Databázi mapování MAC adres do VLAN zde udržuje tzv. VLAN Membership Server (VMPS Server), který prostřednictvím VQP odpovídá na dotazy přepínačů -VMPS klientů- dynamicky zařazujících stanice do VLAN (obr. 1.7). Stanice, které nejdou v databázi explicitně uvedeny, je možné nechat přiřazovat do implicitní (default) VLAN. Poněkud nepříjemný je fakt, že zatímco funkci VMPS klienta mohou zastávat i nejnižší modely přepínačů Cisco, funkce VMPS serveru je implementována pouze ve vyšších a dražších modelech. To dosti omezuje použitelnost dynamického přiřazování stanic do VLAN v menších sítích, ve kterých vyšší modely Cisco přepínačů schopné provozovat VMPS server nejsou obvykle k dispozici. Toto však lze řešit např. použitím OpenVMPS – implementace VMPS serveru pro OS Linux [1]

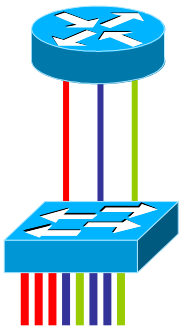


Obr. 1.7 – Přiřazování stanic do VLAN pomocí VMPS serveru

1.1.2.3 Další použití trunk spojů

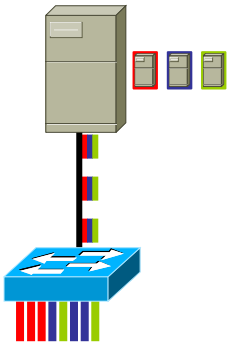
Trunk linky nesoucí značkované rámce z různých VLAN nemusí vždy vést pouze mezi přepínači. Často však bývá užitečné trunk linku přivést i ke směrovači a umožnit směrování mezi VLAN s použitím směrovače s jediným rozhraním (obr. 1.8) – tzv. „router on the stick“, „router na tyči“. Výhodou tohoto přístupu je ušetření počtu rozhraní směrovače a lepší škálovatelnost, kdy počet požadovaných rozhraní směrovače neroste s počtem VLAN, mezi kterými má směrovač směrovat (viz „klasické“ řešení na obr. 1.9). Nevýhodou je naopak nižší propustnost, protože při směrování paketu mezi VLAN, musí paket vždy projít trunk linkou k routeru a pak (s jinou značkou VLAN) zpět k přepínači.

Obr. 1.8 – Směrování mezi VLAN pomocí trunk portu směrovače („router on the stick“)



Obr.1.9 – Klasické řešení směrování mezi VLAN

Další možností užitečnou možností použití je přivést trunk linku k serveru, který je vybaven kartou s podporou 802.1q (obr. 1.10). Operačnímu systému se pak tato síťová karta jeví jako více nezávislých logických síťových rozhraní, z nichž každé přísluší k jedné virtuální síti. Na jednom počítači tak může běžet více procesů svázaných s jednotlivými virtuálními sítěmi, což může být výhodné pro různé poskytovatele služeb.



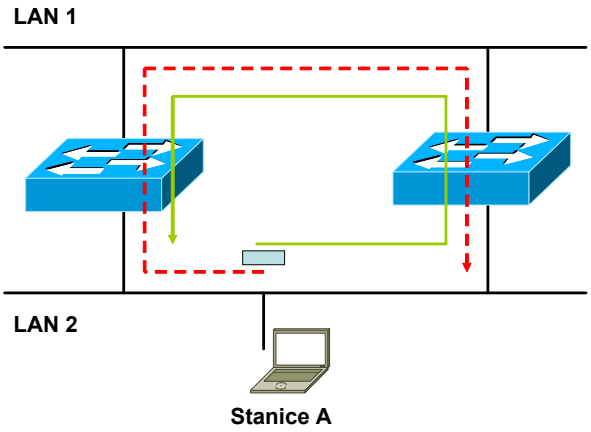
Obr. 1.10 - Přivedení trunk linky k serveru se síťovou kartou s podporou 802.1q

1.1.3 Protokol Spanning Tree

Princip samoučících se mostů předpokládá stromovou topologii. V praxi však často potřebujeme zapojit i redundandní linky, které v topologii sítě vytvoří smyčky. Na těchto smyčkách by však docházelo k nekonečné cirkulaci rámců, pročež musíme hledat způsoby, jak smyčky v topologii sítě eliminovat.

1.1.3.1 Problém smyček v přepínané síti

Problém cirkulace rámců v topologii se smyčkami ukazuje názorně obr. 1.11. Uvažujme dvě lokální sítě pro názornost ve sběrnicové topologii (místo sběrnic si můžeme představit i rozbočovač) propojené více než jedním přepínačem, čímž vzniká smyčka. Představme si nyní, že stanice A vyšle broadcast (ve skutečnosti by mohlo jít o rámec s libovolnou adresou, která není v přepínací tabulkách přepínačů). Vyslaný rámec každý přepínač kopíruje na všechny porty mimo příchozího. Oba přepínače tedy přenesou rámec na LAN2, na které se tak postupně objeví dvě kopie původního rámce od stanice A. Obě tyto kopie převezmou všechny přepínače na LAN2 mimo toho, který každou kopii vyslal (zde to bude druhý z přepánačů) a zašlou jej zpět na LAN1. Tím vznikne situace, kdy v síti budou donekonečna protisměrně cirkulovat dva broadcast rámce.



Obr. 1.11 – Cirkulace rámců v topologii se smyčkami

V případě, že by obě LAN byly propojeny pomocí třech paralelních přepínačů, počet broadcast rámců v síti by dokonce neustále exponenciálně narůstal. Broadcast vyslaný ze stanice A by se totiž dostal na LAN2 každým ze třech přepínačů (pozdržený v bufferech, aby nedošlo ke kolizi), takže se na LAN2 objeví 3 kopie. Každá z těchto kopií vyslaná jedním ze tří přepínačů dorazí k ostatním dvěma, které ji přepošlou opět na LAN1, čímž z každé kopie na LAN2 vzniknou 2 kopie na LAN1, tedy celkově 6 rámců. V následných krocích vznikne 12 rámců, 24 rámců a tak dále až do úplného zahlcení přepínačů rámci, které budou donekonečna cirkulovat sítí.

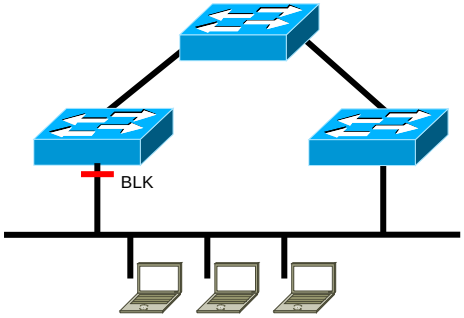
Všimněme si, že přítomnost smyčky nejen vede k cirkulaci nebo generování kopií rámců, ale zcela nabourává i samotný princip automatického učení přepínací tabulky. Cirkulující rámce se zdrojovou adresou stanice A jednou přicházejí do přepínačů z LAN1 a jindy z LAN2, čímž si přepínače neustále střídavě a v polovině případů nesprávně podle portu příchozího rámce aktualizují informaci, za kterým portem je vlastně stanice A připojena.

Z uvedeného příkladu je vidět, že si smyčku v topologii nesmíme dovolit. Při správě sítě však smyčka může omylem při zapojování nebo i zlým úmyslem uživatele velmi snadno vzniknout. Také z praktických důvodů není obvykle vhodné konstruovat topologii sítě jako strom, ale poskytnout i redundantní linky pro případ výpadku a izolace některé větve. Proto musíme mít k dispozici mechanismus, který v topologii obsahující smyčky zablokuje některé porty přepínačů tak, aby

výsledná topologie byla stromová. A právě takovýto mechanismus poskytuje protokol Spanning Tree, standardizovaný v normě IEEE 802.1d a podporovaný většinou přepínačů alespoň střední třídy.

1.1.3.2 Základní funkce protokolu Spanning Tree

Algoritmus (resp. protokol) Spanning Tree má za úkol zkonstruovat a neustále udržovat nad danou topologii přepínačů strom, který pokryje celou topologii – odtud také jeho anglický název. Je důležité si uvědomit, že linky tvořící smyčky nejsou blokovány z obou stran, ale vždy jen na jednom konci. Je tomu tak proto, že linka může být ve skutečnosti segmentem sítě, kde jsou připojeny stanice, které samozřejmě musí být do stromu také napojeny (viz obr. 1.12).



Obr. 1.12 – Způsob blokování smyček v topologii

Je důležité, že algoritmus pracuje neustále - v případě výpadku linky nebo portu přepínače se strom automaticky přepočte (odblokuje se některý doposud zablokovaný port).

Konstrukce stromu se děje ve dvou krocích, které však probíhají neustále a současně:

1. volba kořenu stromu (root bridge)
2. vytvoření stromu nejkratších (nejlevnějších) cest z každého přepínače ke kořeni

Kořen stromu se volí podle nakonfigurovaných priorit, v případě shody (např. pokud byla ponechána tovární nastavení) podle Bridge ID pevně přiděleného každému přepínači (má tvar MAC adresy). Priority přepínačů je obvykle vhodné konfigurovat tak, aby kořenem byl přepínač blízko směrovače do WAN nebo k serveru. Volba kořenového přepínače funguje tak, že každý přepínač ze začátku prohlásí za kořen sám sebe a vysílá o tom do svém okolí tzv. BPDU (Bridge Protocol Data Unit), které mj. obsahují nastavenou prioritu přepínače. Pokud přepínač přijme BPDU, ve kterém se za kořen prohlašuje přepínač s prioritou nižší, než má on sám, BPDU zahodí. V opačném případě BPDU rozešle do svého okolí a sám přestane BPDU generovat (podřídí se kořeni s lepší prioritou). Tímto způsobem po chvíli na síti zbyde jediný přepínač generující BPDU – vybraný kořen stromu.

Při vytváření stromu nejkratších (nejlevnějších) cest z každého přepínače ke kořeni se bere v úvahu nakonfigurovaná preference (cena) jednotlivých linek. Pokud cena nebyla nastavena, implicitně se zpravidla počítá jako nepřímo úměrná přenosové rychlosti linky. Cena linky se konfiguruje

zpravidla stejně na všech portech přepínačů připojených k lince (nebo segmentu sítě). Pokud by ceny portů na témže segmentu byly různé, počítal by se segment s různou cenou podle toho, na které straně segmentu by byl zvolený kořenový přepínač.
Porty, které jsou součástí vytvořeného stromu, budou funkční (forwarding), ostatní budou blokovány (blocking).

Technicky algoritmus funguje tak, že kořenový přepínač generuje co 2 sekundy zprávu BPDU, která se postupně šíří po topologii a při tom do sebe „sbírá“ sumární cenu linek, kterými už prošla. Když přepínač slyší BPDU z různých portů, vybere si tu z nich, ve které je menší sumární cena cesty ke kořeni a připojí se do stromu právě tímto portem (tzv. root port). Pokud by BPDU z root portu přestaly přicházet, znamená to výpadek v topologii a jako root port se zvolí ten, ze kterého přicházejí BPDU s nejnižší hodnotou sumární ceny cesty ke kořeni.

Z důvodu zabránění smyčkám během konvergence (přechodu na jiný strom) algoritmus definuje přechodné stavy portů Learning a Listening. Port ve stavu Listening nepřeposílá rámce, ale pouze po dobu 15 sekund sleduje okolní provoz, aby mohl rozhodnout, zda se přepne do stavu Forwarding nebo Blocking. Před přechodem do stavu Forwarding se navíc 15 sekund ve stavu Learning bez přeposílání rámců pouze učí MAC adresy okolních stanic, aby se omezilo procento rámců, jejichž příjemce není v přepínací tabulce a musí být tudíž rozeslány na všechny porty.

S ohledem na skutečnost, že přepínač prohlásí původní cestu za nefunkční až ve chvíli, kdy z portu nepřicházejí po dobu 20 sekund BPDU, může po výpadku přechod na nový strom trvat až 50 sekund. To bylo naprosto dostačující v době před několika desetiletími, kdy standard 802.1d vznikal, avšak pro dnešní požadavky sítí přenášející multimédia již tato hodnota příliš uspokojivá není (a to ani ve srovnání s rychlostí konvergence sítí založených na směrovačích a moderních směrovacích protokolech).

1.1.3.3 Rapid Spanning Tree (IEEE 802.1w)

Pro vyřešení neuspokojivé konvergence klasického protokolu Spanning Tree byl nově navržen a postupně je implementován protokol Rapid Spanning Tree, jehož specifikaci lze najít v normě IEEE 802.1w. Od tohoto protokolu je očekáváno, že poskytne konvergenci v řádu jednotek sekund. To je zajištěno kompletní změnou filosofie oproti klasickému Spanning Tree (802.1d) založenému na časovačích. Rapid Spanning Tree je založen na reakci na oznamované události a na aktivních dotazech mezi sousedními přepínači, které vedou k rychlejšímu přechodu portů do stavu Forwarding. Další změnou je i to, že BPDU převzaly funkci dozoru nad funkčností linek (keepalives) a negeneruje je pouze kořenový, ale nezávisle každý přepínač. Také informace o změně topologie se již neposílá nejprve na kořenový přepínač a následně dolů po stromu, ale šíří se přímo z přepínače, který změnu topologie detekoval. Podle toho, zda je linka halfduplexní nebo fullduplexní rozlišuje Rapid Spanning Tree dva typy linek – dvoubodové a sdílené. Funkce algoritmu na dvoubodových linkách je jednodušší a tedy konvergence rychlejší.

Mimo uvedené změny principu zavádí Rapid Spanning Tree řadu vylepšení, které byly již jako rozšíření k dispozici i v některých implementacích standardního Spanning Tree, avšak nebyly obsaženy ve standardu. Jedná se zejména o možnost explicitního označení portu za hraniční (tzv.

edge port), za kterým již není žádný přepínač, ale pouze koncová stanice a tak může být ihned po hardwarovém náběhu odblokován bez dalších testů. Jsou rovněž k dispozici mechanismy, které dovolují přepínači připojenému dvěmi redundantními linkami k přepínačům vyšší hierarchické úrovně bez čekání ve stavech Listening a Learning odblokovat redundandní linku ve velmi krátké době poté, co detekuje výpadek dříve aktivní linky,

1.1.3.4 Spanning Tree v prostředí s virtuálními sítěmi

Zajímavým problémem je způsob aplikace protokolu Spanning Tree v prostředí používajícím virtuální sítě. Jelikož virtuální sítě jsou na sobě vzájemně nezávislé a jsou logicky zcela odděleny, provozují se obvykle na jednotlivých VLAN nezávislé instance protokolu Spanning Tree. To dává možnost pro jednotlivé VLAN nastavit různé stromy a tím využít některé linky pro provoz určitých VLAN a jiné linky pro provoz jiných VLAN a tím zefektivnit využití infrastruktury. Je si však třeba uvědomit, že při velkém množství VLAN toto představuje značnou zátěž pro procesory jednotlivých přepínačů. Např. pro 500 VLAN a standardní časování vysílání BPDU co 2 sekundy v každé VLAN musí CPU každého přepínače za sekundu zpracovat 250 BPDU.

Opačný přístup, kdy bude pro všechny VLAN použito téhož stromu (tzv. Common Spanning Tree), navrhuje norma pro virtuální sítě IEEE 802.1q. V tomto případě běží pouze jedna instance Spanning Tree a to ve VLAN 1, jejíž výpočet se použije pro všechny VLAN. Zatížení procesorů přepínačů se tak minimalizuje, ale za cenu dosti neefektivního využití infrastruktury.

Kompromisní řešení definuje norma IEEE 802.1s (Multiple Spanning Tree nebo též Multi-Instance Spanning Tree), která dovoluje provozovat na síti několik nezávislých instancí (Rapid) Spanning Tree a pro jednotlivé VLAN definovat, od které instance převezmou strom určující zablokované a funkční porty.

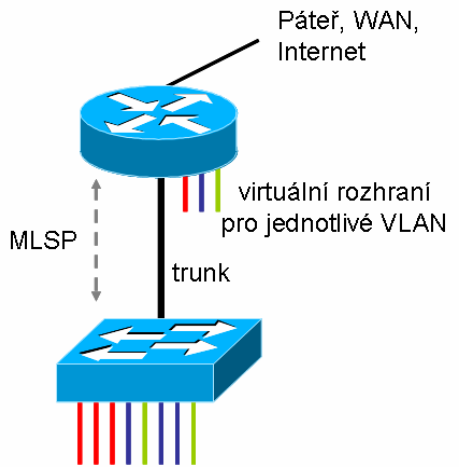
1.1.4 Přepínání na více vrstvách

V poslední době se často hovoří o přepínaní na třetí vrstvě či obecně o přepínání na více vrstvách (multilayer switching, MLS). Principiálně se jedná o různé technologie, které mají za cíl urychlit hardwarovými prostředky směrování a přepínat pakety, resp. jednotlivé (jednosměrné) datové toky definované i parametry z vyšších vrstev (např. porty TCP/UDP) bez potřeby zpracování každého paketu procesorem a časově náročného procházení celé směrovací tabulky. Cílem je dosáhnout rychlosti přepínaných sítí, avšak využít při tom všech výhod škálovatelnosti sítí směrovaných. Technicky může být multilayer switching realizován různými způsoby, buďto formou specializovaného zařízení (tzv. L3 přepínače) integrujícího celou funkci nebo speciální vazby mezi přepínačem a směrovačem, dovolující přepínači na základě informace ze směrovače hardwarově přímo přepínat rámce i mezi porty různých segmentů (virtuálních sítí), mezi nimiž by normálně provoz musel procházet přes směrovač.

1.1.4.1 Princip přepínání na více vrstvách

Prvotní implementace MLS byly realizovány jako vylepšení funkcionality přepínačů a definice speciálního protokolu mezi přepínačem a směrovačem (obr. 1.13). Směrovač může být realizován

buďto jako externí (s dodatečnou funkcionalitou pro MLS), nebo jako směrovací modul do modulárního přepínače. Protokol mezi přepínačem a směrovačem není bohužel standardizován, my jej zde (v souladu s implementací firmy Cisco) budeme označovat jako MLSP – Multilayer Switching Protocol.



Obr. 1.13 – Způsob implementace MLS

Princip fungování MLS je následující. Směrovač nejprve pomocí speciální zprávy MLSP oznámí formou skupinového vysílání přepínači (resp. všem přepínačům, je-li jich vzájemně propojeno více) MAC adresy všech svých virtuálních rozhraní, jimiž jsou ke směrovači přivedeny jednotlivé VLAN. Na základě této informace může přepínač rozlišit, které rámce jsou připojenými zařízeními posílány na směrovač (typicky na výchozí bránu nakonfigurovanou na stanicích umístěných na jednotlivých VLAN) nebo naopak přicházejí (po přesměrování mezi VLAN) ze směrovače. Dále se u MLS předpokládá, že na určitou cílovou adresu (resp. i port TCP/UDP) typicky nebude zaslán pouze jeden paket, ale že první paket tvoří pouze začátek déle trvajícího datového toku. Přepínač se tak snaží ke každému prvnímu paketu toku jdoucímu na směrovač (tzv. candidate packet) najít odpovídající paket (tzv. enable packet) v některé jiné VLAN, do které byl candidate packet přesměrován. Candidate packet bude mít jako cílovou adresu MAC adresu virtuálního rozhraní směrovače pro některou VLAN, enable packet naopak ponese MAC adresu některého virtuálního rozhraní směrovače jako zdrojovou. Spárování candidate a enable paketu se samozřejmě nepodaří v případě, že odchozí candidate paket směrovač přesměroval mimo přepínanou doménu ven do páteře, resp. Internetu.

Při příchodu candidate paketu si přepínač vytvoří zárodek položky pro budoucí přímé přepínání dalších paketů toku. Uvede v něm MAC adresu zdroje, cílovou MAC adresu (adresu rozhraní směrovače) a kompletní hlavičku třetí a čtvrté vrstvy, podle kterých později nalezne enable paket. Při příchodu každého rámce z některé MAC adresy virtuálního rozhraní směrovače přepínač testuje, zda se nejedná o enable paket. Pokud se obsah hlaviček 3. a 4. vrstvy shoduje s informací uloženou v některém zárodku položky tabulky pro přímé přepínání, příslušná položka je doplněna o zdrojovou a cílovou MAC adresu z enable paketu a číslo VLAN, ze které enable paket přišel. Pokud zárodek položky tabulky pro přímé přepínání není do určité doby dokončen, odešel patrně příslušný candidate paket pryč do páteře a zárodek je odstraněn.

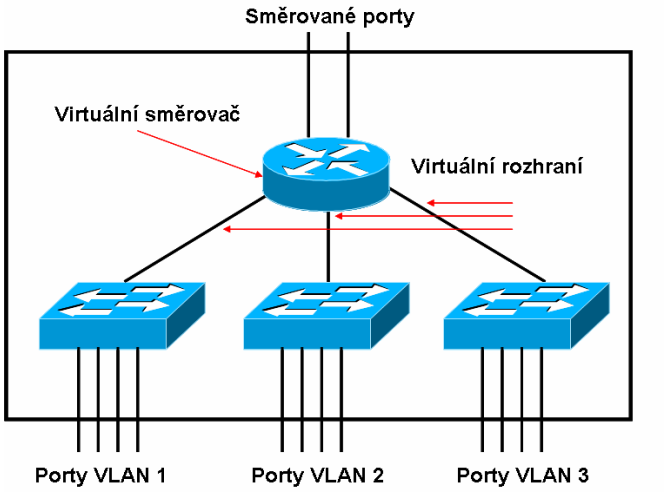
Samotný multilayer switching pak pracuje tak, že každý rámec jdoucí na MAC adresu některého virtuálního rozhraní směrovače je testován, zda nevyhovuje některé položce tabulky pro přímé přepínání. V kladném případě je z této položky vyzvednuta zdrojová MAC adresa a VLAN odpovídající rozhraní směrovače, ze kterého odešel enable paket a MAC adresa cílové stanice. Tyto adresy v rámci přepíší původní MAC adresy, zmenší se pole TTL v IP hlavičce, přepočte se kontrolní součet záhlaví a rámec se pošle do příslušné VLAN. Přijímající stanice pak rámec vnímá jako by normálně prošel směrovačem.

Při testování shody paketů s položkami tabulky pro přímé přepínání běžně postačí shoda cílové IP adresy. Pouze v případě, že by na směrovači byly aplikovány paketové filtry (Access Control Lists, ACL), je nezbytné testovat i další položky, aby enable paket vyhovující ACL neotevřel možnost přímého průchodu i pro jiné toky na tutéž cílovou adresu, ale nevyhovující ACL (např. jdoucí z nepovolené zdrojové IP adresy nebo směřující na nepovolený TCP/UDP port). Proto v případě aplikace ACL na směrovači musí být tato skutečnost pomocí protokolu MLSP oznámena přepínači, který pak musí příslušnost paketu k toku a shodu s položkou tabulky pro přímé přepínání testovat důkladněji, tj. v závislosti na obsahu daného ACL i podle zdrojové IP adresy, protokolu 4. vrstvy nebo zdrojového a cílového portu 4. vrstvy.

Nevýhodou uvedeného mechanismu je skutečnost, že při jakékoli změně směrovací tabulky nebo úpravě ACL na směrovači musí být tabulka pro přímé přepínání na přepínači (resp. všech přepínačích, je-li jich více) kompletně vymazána. Požadavek na vymazání je přepínačům distribuován pomocí protokolu MLSP.

Pro úplnost poznamenejme, že směrovač nutně nemusí být k přepínači připojen pouze trunk linkou, ale jeho jednotlivé VLAN mohou být připojeny k normálním rozhraním směrovače. V takovémto případě musí MLSP propagovat MAC adresy všech rozhraní, které slouží pro přesměrovávání mezi VLAN účastnícími se mechanismu multilayer switchingu.

Druhým typickým způsobem realizace přepínání na více vrstvách je L3 přepínač. Obecně jde o integrované řešení přepínače a směrovače s maximem přepínacích funkcí realizovaným hardwarovými prostředky a současně s plnou funkcionalitou směrovače provozujícího software implementující dynamické směrovací protokoly. Konkrétní technická realizace může být různá a v úplnosti téměř nikdy není výrobcem zveřejněna. A pohledu konfigurace se L3 přepínač jeví jako standardní přepínač podporující VLAN a směrování mezi nimi. Jehož jednotlivé porty mohou být navíc přepnuty do režimu směrovaných portů. Logicky se pak L3 přepínač může jevit jako virtuální topologie podobná té na obr. 1.14. Výhodou je nejen rychlost, ale i velká flexibilita konfigurace takového zařízení. Oproti standardním směrovačům však postrádá WAN rozhraní, všechna rozhraní jsou z důvodu realizovatelnosti maxima funkcí v hardware stejná (typicky Ethernet). Množství rozhraní zpravidla odpovídá přepínači, tj. je výrazně vyšší, než u klasického směrovače (typicky např. 24).



Obr. 1.14 – Logická struktura L3 přepínače

1.1.4.2 Modely použití VLAN v prostředí s přepínači na 3. vrstvě

S nástupem L3 přepínačů se poněkud mění filosofie používání VLAN v rozsáhlejších lokálních sítích. Ukázalo se, že rozsáhlé VLAN pokrývající celou topologii nejsou efektivní jedna z důvodu nižší stability protokolu Spanning Tree (ten je dimenzován na průměr grafu sítě max 7 přepínačů) a jednak proto, že není možné využít rozkládání zátěže (load balancing) mezi redundandní spoje v topologii. Proto je dnes tendence využívat VLAN spíše jen v okrajových částech lokální sítě a v částech blíže páteře raději aplikovat rychlé L3 přepínače, poskytující všechny výhody směrování za cenu stále více se blížící ceně kvalitních L2 přepínačů.