Modul 5 STP

5.0 Úvod

Počítačové sítě jsou nezbytné pro práci téměř všech firem. Proto správci sítí konfigurují redundantní spoje mezi routery a switchi pro případ nějakého přerušení. To ale vede ke vzniku cyklů, které je třeba průběžně řešit. Případné cykly (deaktivaci linek) a výpadky spojů (aktivaci linek) pomáhá řešit STP (Spanning-tree protocol)

Tato kapitola popisuje:

- význam redundance v zapojení sítě
- funkci STP při eliminaci cyklů na úrovni L2
- postup konvergence STP
- implementaci PVST+ v LAN

5.1 Redundantní L2 topologie

5.1.1 Redundance

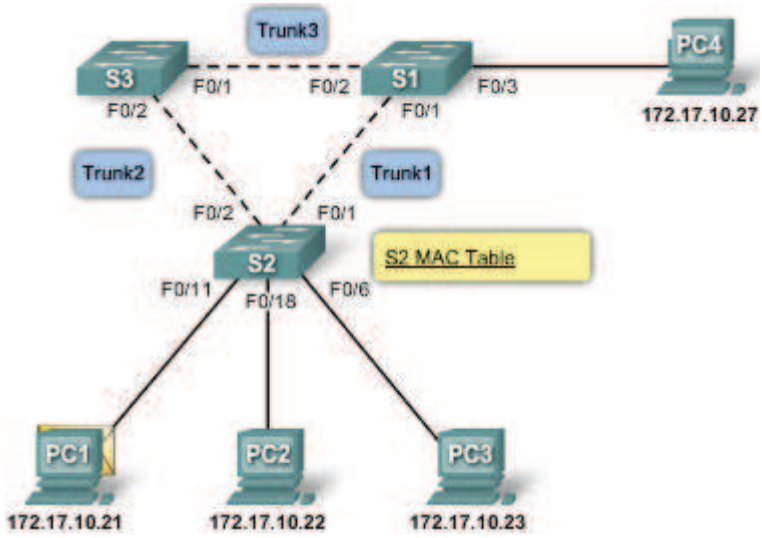
Redundance v hierarchické síti umožňuje zachovat funkčnost sítě i v případě výpadku některých linek. Pokud například switch detekuje výpadek linky, použije se pro předání zprávy jiná cesta. Pokud výpadek pomine, aktivuje se cesta původní.

Pokud máme dostatečné redundance na všech vrstvách hierarchického návrhu sítě, neměl by vadit výpadek jedné linky nebo jednoho switche (v kterékoliv vrstvě – přístupové, distribuční nebo páteřní) – záleží na způsobu propojení – aby existovala jiná cesta.

5.1.2 Problémy s redundancí

L2 cykly

Jakmile nakonfigurujeme redundantní spoje, vzniknou v síti cykly na úrovni druhé vrstvy (L2). Protože rámce nemají nic jako TTL u IP paketů, může se stát, že bude rámec v síti kolovat donekonečna, resp. dokud se síť nezahltí nebo nepřeruší – typicky broadcastové zprávy. Dalším důsledkem mohou být nesmyslné MAC tabulky. Příklad:



Pokud PC1 pošle broadcastovou zprávu, S2 ji přepošle na všechny ostatní porty (takže i na S1 a S3); S1 a S3 ji opět pošlou na ostatní porty, takže S1→S3,PC4 a S3→ S1; opět ji oba pošlou na ostatní porty – S1→S2,PC4 a S3→S2; S2 obě příchozí zprávy opět pošle na všechny ostatní por-

ty – S2→S1 a S2→S3 (samozřejmě také stanicím, ale ty teď nejsou důležité). Tím se uzavírá kruh, ve kterém bude tato zpráva posílána neustále.

Pokud postupně do takové sítě přijde více broadcastových zpráv, síť se časem zahltí – to nazýváme broadcastová bouře (broadcast storm). Ve výsledku to ale může ovlivnit i koncové stanice, kterým na síťové karty bude přicházet velké množství zpráv na zpracování a může tuto stanici zpomalit nebo také zahltit.

Ve stejné topologii také může dojít k duplikaci odeslaného rámce – například, pokud PC1 pošle zprávu pro PC4 a S2 ještě nemá o PC4 záznam v MAC tabulce (zatímco S1 a S3 už ano). V tom případě přepošle zprávu jak na S1, tak na S3, které oba pošlou zprávu směrem k PC4 a té přijde stejný rámec dvakrát.

5.1.3 Problémy s redundancí – reálně

Redundantní spoje jsou nutné (viz výše), ale vzniklé cykly je nutné řešit. Typickým problémem jsou cykly vzniklé omylem – např. špatným zapojením kabelu. Cyklus může vzniknout dvojitým spojením dvou switchů nebo propojením více switchů. Další příčinou může být připojení malých hubů/switchů do sítě (rozšíření připojení v rámci kanceláře) a následné propojení těchto malých hubů/switchů mezi sebou.

PKA 5.1.3.3 – ukázka redundantního zapojení a funkce STP

5.2 Úvod do STP

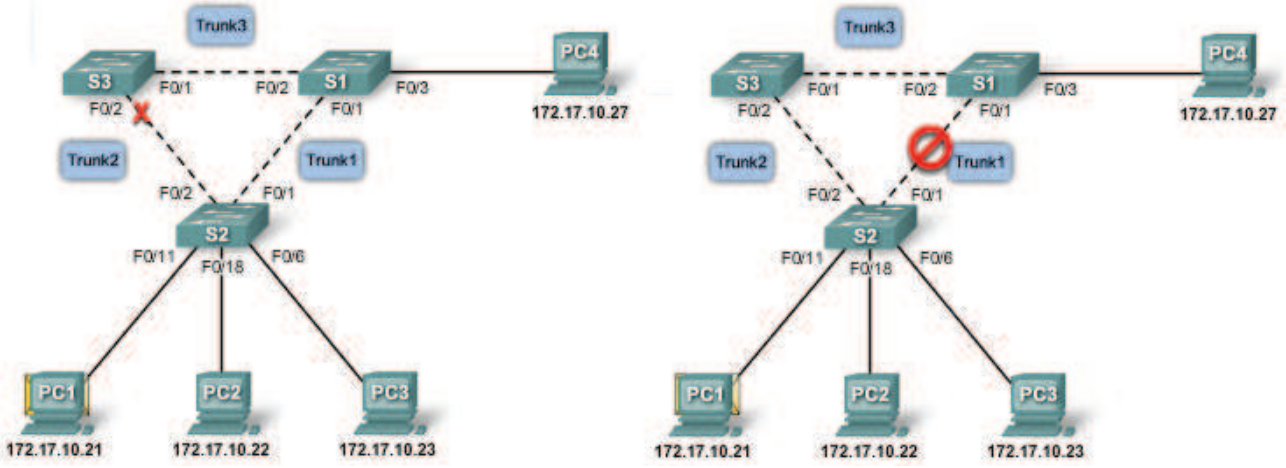
5.2.1 Spanning Tree algoritmus

STP topologie

Při instalaci redundantních linek nebo switchů vznikají cykly, takže některé linky je nutné přerušit, ale tak, aby v případě výpadku jiného zařízení byly tyto linky podle potřeby opět automaticky aktivovány. Toto řeší Spanning Tree Protocol – STP.

STP zajišťuje blokování portů těch linek, které mají být „vypnuté“. Takto blokovaný port nepřenáší data, ale přenáší BPDU=zprávy STP. Takže fyzicky redundantní spoje existují, ale jsou STP protokolem „vypnuty“ (=porty jsou blokovány pro datový přenos). Při výpadku přepočítá STP existující cesty a zapne potřebné porty, aby byl výpadek nahrazen.

Ukázka funkce STP – při funkční síti vypne STP port F0/2 na S3 (obrázek vlevo), čímž přeruší cyklus. Při výpadku linky „Trunk1“ je automaticky tento port aktivován, takže je obnovena funkčnost sítě (obrázek vpravo):

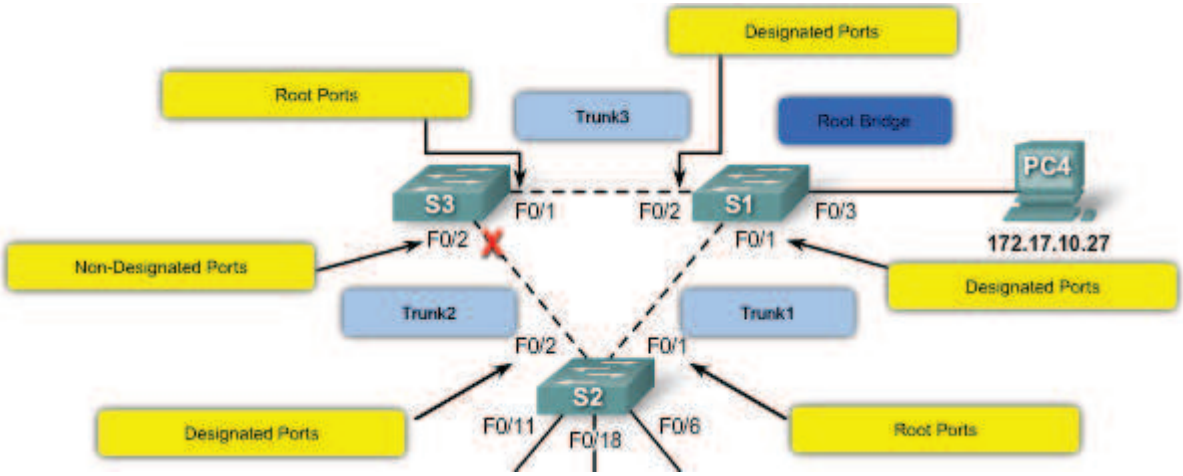


STP Algoritmus

K výpočtům, které porty mají být „blokovány“, aby se zabránilo cyklům, STP používá STA (Spanning Tree Algorithm). STA určí (vybere) jeden ze switchů jako tzv. „root bridge“ a použije jej jako výchozí bod při výpočtech cest v síti. V naší topologii je tímto root bridgem zvolen S1. Při volbě si všechny switche vyměňují BPDU zprávy, ve kterých je obsaženo Bridge ID (BID). Switch s nejmenším BID je zvolen jako root bridge.

Po volbě root bridge vypočítá STA od každého switche nejkratší cestu k root bridgi. Výpočet je prováděn pro každý switch (v broadcastové doméně) a v průběhu výpočtu je blokován klasický síťový provoz. Hodnota cesty se počítá jako součet hodnot pro jednotlivé porty v cestě (tyto hodnoty jsou určovány podle rychlosti portu). Pokud existuje více cest k cíli, vybere STA tu nejkratší (s nejnižším součtem hodnot). Jakmile jsou všechny cesty vypočítány, přidělí STA všem portům jejich roli (viz obrázek):

- root port – port switche, který je nejblíže k root bridgi
- designated port – ne-rootové porty, kterým je povolen přenos dat (jsou částí některé z vypočítaných nejkratších cest)
- non-designated port – port, který je nastaven pro blokování sítového provozu (s výjimkou BPDU), aby se zabránilo cyklům (nejsou částí žádné nejkratší cesty)



Volba root bridge

Každý switch si vytváří své BID z hodnot „priority“, „extended system ID“ a „MAC adresa switche“. Po startu začne switch rozesílat BPDU, obsahující jeho BID a root ID – zpočátku jsou stejné, protože každý switch považuje sebe sama za root bridge. Jakmile switch dostane od souseda jeho BPDU, porovná přijaté a svoje root ID a pokud je přijaté root ID nižší, aktualizuje si tuto informaci a sousední switch tímto bude považovat za root bridge (i když to ve skutečnosti nemusí být on). Na konci mají všechny switche jako root ID zvoleno BID switche s nejnižší hodnotou.

Nejkratší cesta k root bridgi

Délka cesty mezi switchi je definována jako součet hodnot jednotlivých portů v cestě mezi nimi. Defaultní hodnoty portů jsou v tabulce – kvůli vývoji nových (rychlejších) technologií bylo nutno původní hodnoty (sloupec vpravo) aktualizovat:

Rychlost linky	Hodnota portu (nová)	Hodnota portu (původní)
10 Gb/s	2	1
1 Gb/s	4	1
100 Mb/s	19	10
10 Mb/s	100	100

Pokud bychom chtěli výchozí hodnotu portu změnit (např. na 25), můžeme použít příkaz **`S2(config-if)#spanning-tree cost 25`**. Na původní hodnotu se vrátíme příkazem **`S2(config-if)#no spanning-tree cost`**.

V naší topologii můžeme jít z S2 na S1 buď přímo (Cesta1 = 1 x 19 = 19) nebo přes S3 (Cesta2 = 2 x 19 = 38), takže Cesta1 bude brána jako lepší a druhá cesta bude považována za redundantní a nastavena pro blokování provozu.

Ověření aktuálních hodnot jednotlivých portů a cest k root bridgi můžeme provést příkazem **`S2#show spanning-tree`**, případně **`S2#show spanning-tree detail`**.

5.2.2 STP zprávy (BPDU)

Zprávy STP protokolu (BPDU) obsahují celkem 12 hodnot:

- identifikátor protokolu, verzi, typ zprávy, označení stavu
- root ID, délku cesty, bridge ID, port ID
- časové údaje

K rozesílání BPDU používají switche speciální multicastovou cílovou MAC adresu 01:80:C2:00:00:00, díky které všechna ostatní zařízení správu zahodí.

Výpočet root bridge

Po startu začnou všechny switche pravidelně (co 2 vteřiny = hello interval) rozesílat BPDU. Zpočátku označují jako root bridge samy sebe, tj. v odchozích BPDU je root ID = BID switche. Při přijetí BPDU switch porovná přijeté root ID s vlastní informací o root ID. Pokud je přijatá hodnota nižší, aktualizuje svoji hodnotu a v dalších BPDU bude již rozesílat nové root ID. Pokud by přijatá hodnota byla vyšší, je zpráva zahozena. Při přeposílání BPDU je přepočítávána cesta – je přičtena vždy hodnota rozhraní (viz tabulka výše – např. 19 pro Fast Ethernet). Díky tomu je současně známá také délka cesty k root bridgi, takže je možné vybrat tu nekratší a porty s redundantními cestami vypnout.

5.2.3 Bridge ID

Bridge ID je hodnota rozesílaná v BPDU, pomocí které se volí root bridge. Skládá se ze tří hodnot – priorita bridge, extended system ID a MAC adresy.

- priorita – hodnota (od 1 do 65536), kterou můžeme nastavit, abychom ovlivnili volbu – čím nižší hodnota, tím vyšší priorita (tj. root bridgem se stane switch s nejnižší prioritou)
- extended system ID – před začátkem používání VLAN nebylo potřeba; pokud je to potřeba, zahrne se tato hodnota do priority (zabere část bitů) – viz PVST dále
- MAC adresa – pokud jsou předchozí hodnoty stejné (např. výchozí), je MAC adresa rozhodující hodnotou určující volbu root bridge, což ale není ideální, protože to není konfigurovatelná (předvídatelná) hodnota; také to může způsobit chaos v síti v případě, že přidáme nový switch a není vyvolána nová volba root bridge

Konfigurace a ověření hodnoty BID

Pokud má být některý switch root bridgem, musíme zajistit, aby měl nižší BID, než všechny ostatní switche. Máme dvě možnosti.

Příkaz `Switch1(config)#`**`spanning-tree vlan`** `vlan-id` **`root primary`** nastaví prioritu buď 24576, nebo nejbližší nižší násobek 4096 vzhledem k nejnižší detekované prioritě v síti. Pokud chceme určit záložní root bridge, zadáme na něm příkaz `Switch2(config)#`**`spanning-tree vlan`** `vlan-id` **`root secondary`** – ten nastaví prioritu na 28672, což je mezi 24576 a 32768 (výchozí hodnota).

Druhá možnost – nastavovat přímo hodnotu priority číselně – příkazem `Switch1(config)#spanning-tree vlan vlan-id priority hodnota` (hodnota je násobkem 4096). To umožňuje přesnější a jasnější kontrolu volby root bridge.

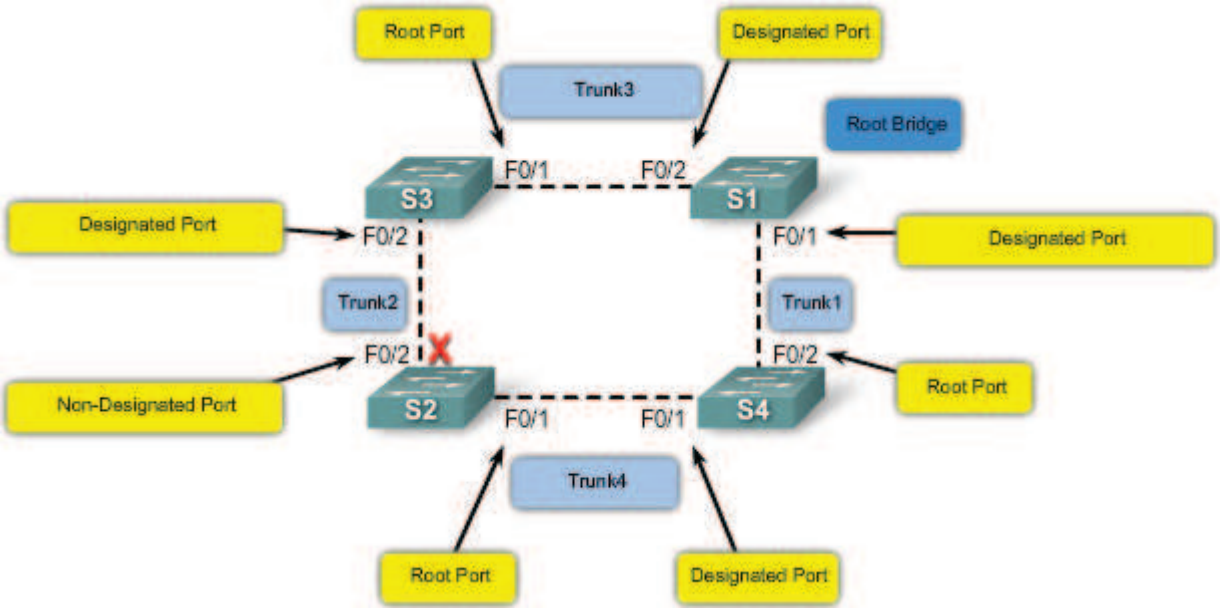
Ověření stavu – příkaz **`show spanning-tree`**.

5.2.4 Role portu

Při volbě root bridge se počítá také nejkratší cesta, pomocí které také určíme roli portu. Existují 4 odlišné role portu:

- root port – pouze na switchích, které nejsou root bridgem; je to port, přes který vede nejkratší cesta k root bridgi, může být na switchi pouze jeden; tento port předává zprávy směrem k root bridgi; adresy odesilatelů zpráv přicházejících na tento port se přidávají do MAC tabulky
- designated port – na root bridgi jsou všechny ostatní porty určeny designated porty; na ostatních switchích jsou to porty, které přijímají a předávají zprávy a adresy odesilatele z těchto zpráv se mohou přidávat do MAC tabulky; v každém segmentu může být pouze jeden takový port (druhý musí být buď root port nebo non-designated port); pokud je na jednom segmentu více switchů (např. linka S2 – S3 na obrázku na str. 42), je jeden switch vybrán jako designated (a tím i jeho port) a druhý jako non-designated
- non-designated port (alternate port) – port, který není ani root, ani designated; tento port je blokován pro síťový provoz, takže ani nepřidává adresy odesilatelů do MAC tabulky; slouží k přerušení cyklů;
- vypnutý (disabled) port – port, který je administrativně vypnutý, takže se neúčastní ani STP

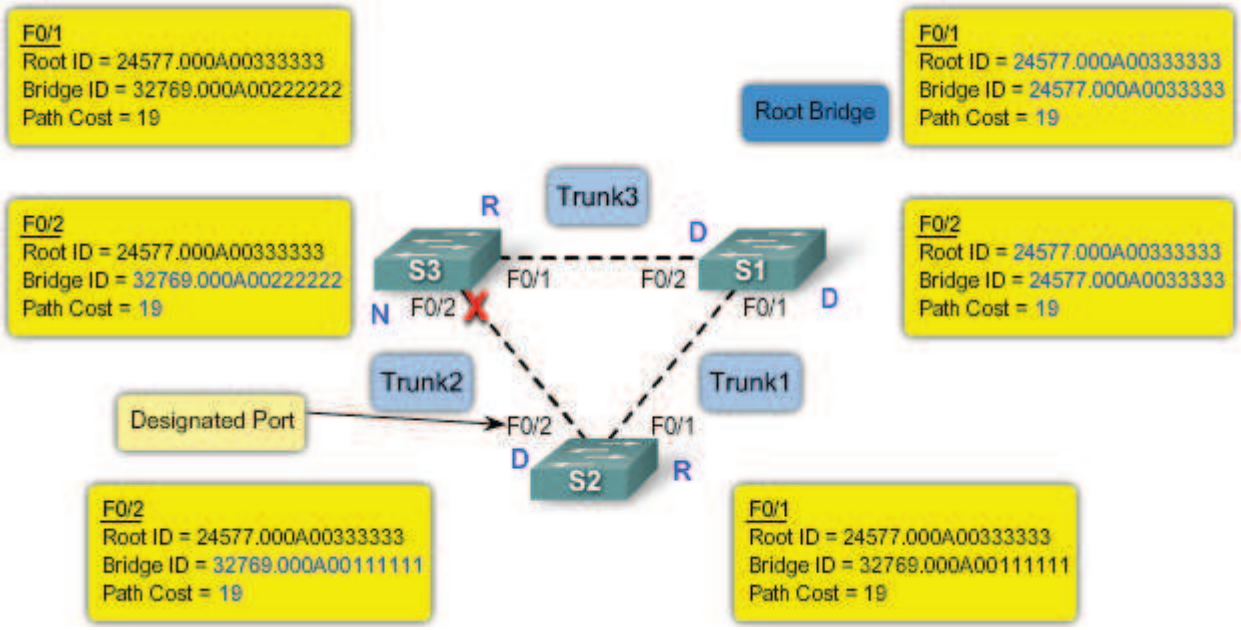
Role portů – příklad



Při určování root portu na switchi se zvolí port s nejnižší hodnotou cesty k root bridgi (tj. F0/1 na S3 a F0/2 na S4). Pokud jsou na switchi dva porty se stejnou (nejnižší) hodnotou cesty, použije se priorita portu. Výchozí hodnota je 128 (lze ji měnit) a za ni se přidává port ID (označení rozhraní). Takže na tomto obrázku mají porty na S2 priority 128.1 (F0/1) a 128.2 (F0/2). Vybrán je port s nižší hodnotou (F0/1). Prioritu portu nastavujeme příkazem `S2(config-if)#spanning-tree port-primary hodnota`.

Pokud jsou dva switche na jednom segmentu a oba chtějí určit tyto porty jako designated, musí být jeden z těchto portů blokován (non-designated). Jako designated je určen port s nižší hodnotou cesty k root bridgi. Pokud mají tyto hodnoty stejné, vymění si switche svá BID a ten, který

má nižší, nastaví svůj port jako designated, druhý switch nastaví svůj port jako non-designated. Příklad (už výsledek):



Ověření role portu – opět příkaz **`show spanning-tree`**.

5.2.5 STP – stavy portů, BPDU časovače

Při zapnutí switche/portu by mohlo dojít k vytvoření cyklu. Proto existuje 5 stavů portu, kterými musí port projít před plnou funkčností a 3 BPDU časovače.

- blokuje (blocking) – non-designated port, nepředává uživatelské zprávy, ale přijímá a vysílá BPDU, aby mohl určit root bridge
- naslouchá (listening) – port chce být aktivní (předávat zprávy), takže přijímá i vysílá BPDU
- učí se (learning) – připravuje se na předávání uživatelských zpráv a už je schopen z příchozích zpráv dodávat adresy do MAC tabulky
- předává zprávy (forwarding) – port je kompletně funkční, přijímá i vysílá jak uživatelské zprávy, tak BPDU
- vypnutý (disabled) – „administratively down“, port je úplně vypnut

Souhrn vlastností jednotlivých stavů je v tabulce:

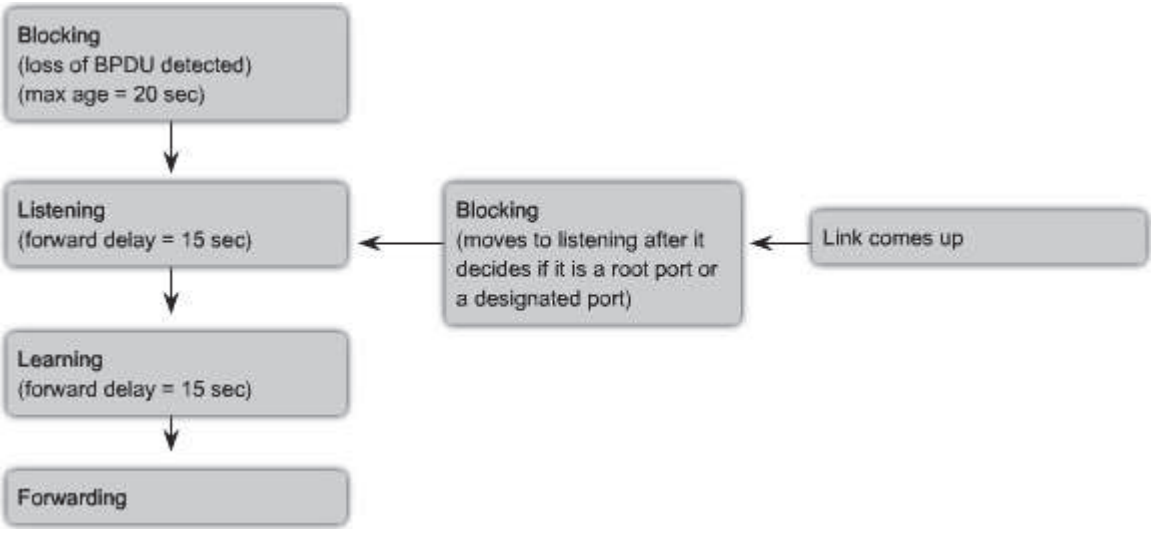
	Zpracovává BPDU	Předává uživatelské zprávy	Učí se MAC adresy (ukládá je do MAC tabulky)
Blocking	Ano	Ne	Ne
Listening	Ano	Ne	Ne
Learning	Ano	Ne	Ano
Forwarding	Ano	Ano	Ano
Disabled	Ne	Ne	Ne

Časovače

Každý port switche musí před rozhodnutím o finálním stavu projít jednotlivými stavy jako na obrázku. Nicméně mezi jednotlivými stavy jsou stanoveny minimální doby, jak dlouho se má čekat:

- hello time – jak často se odesílají BPDU (výchozí = 2 s, rozsah = 1 až 10 s)

- forward delay – jak dlouho má port zůstat ve stavu „listening“ a „learning“ (výchozí = 15 s, rozsah = 4 až 30 s)
- maximum age – maximální stáří informací z BPDU zpráv, které port switche uchovává (výchozí = 20 s, rozsah = 6 až 40 s)



Tyto časovače umožňují dosáhnout konvergence i v sítích s průměrem sedm switchů (to je maximum pro STP právě kvůli zajištění konvergence).

Nakonec je stav portu určen buď „forwarding“ nebo „blocking“. Pokud je detekována změna topologie, je port nastaven do stavu „listening“.

Poznámka – pokud chceme ovlivnit čas konvergence (např. v menší síti), je vhodnější to udělat nastavením průměru sítě (pouze na root bridgi!): `Switch(config)#`**`spanning-tree vlan`** `vlan-id` **`root primary`** `průměr`.

PortFast nastavení

Pokud má access port nastavenu možnost PortFast, přejde ze stavu „blocking“ do stavu „forwarding“ okamžitě. To se využívá u portů, na kterých je připojeno jediné PC. Pokud by ale na takový port dostal switch BPDU, je možné, že by jej přepnul do stavu „blocking“ (díky technice BPDU guard).

PortFast technologie je vhodná zejména u portů stanic, které získávají IP adresu z DHCP. Tam by kvůli časové prodlevě mohlo dojít k nepřidělení IP adresy.

Nastavení se provádí příkazem `Switch(config-if)#`**`spanning-tree portfast`**. Vypnutí opačným příkazem (**`no ...`**) a ověření pomocí výpisu aktuální konfigurace.

PKA 5.2.5.4 – volba root bridge

5.3 STP – konvergence

5.3.1 Konvergence STP

Konvergence je proces, v průběhu kterého je zvolen root bridge a všechny porty znají svoji roli v STP a dosáhnou finálního stavu (forwarding/blocking), díky čemuž se přeruší všechny případné cykly v síti. Někdy se konvergencí míní také čas, který je nutný na dosažení popsaného stavu.

Základními kroky procesu jsou:

- volba root bridge
- určení root portů
- určení designated a non-designated portů

5.3.2 Volba root bridge

Volba root bridge je vyvolána buď dokončením startu switche, nebo výpadkem některé cesty v síti. Na začátku jsou všechny porty blokovány (defaultně 20 s), aby se zabránilo vzniku cyklů ještě před dokončením konvergence. Teoreticky by měl proces volby root bridge trvat maximálně 14 sekund = průměr sítě (7) krát interval mezi BPDU (2 s).

BPDU jsou rozesílány pravidelně i nadále – pomocí toho lze určit výpadek linky, což se projeví několika po sobě jdoucími neobdrženými BPDU. Interval, jak dlouho má switch na BPDU čekat (a uchovávat původní BPDU informace), je „maximum age“ (defaultně 20 s).

Ověření volby – příkazem `switch(config)#`**`show spanning-tree`**. Na root bridgi se ve výpisu objeví řádek „`This bridge is the root`“.

5.3.3 Určení root portů

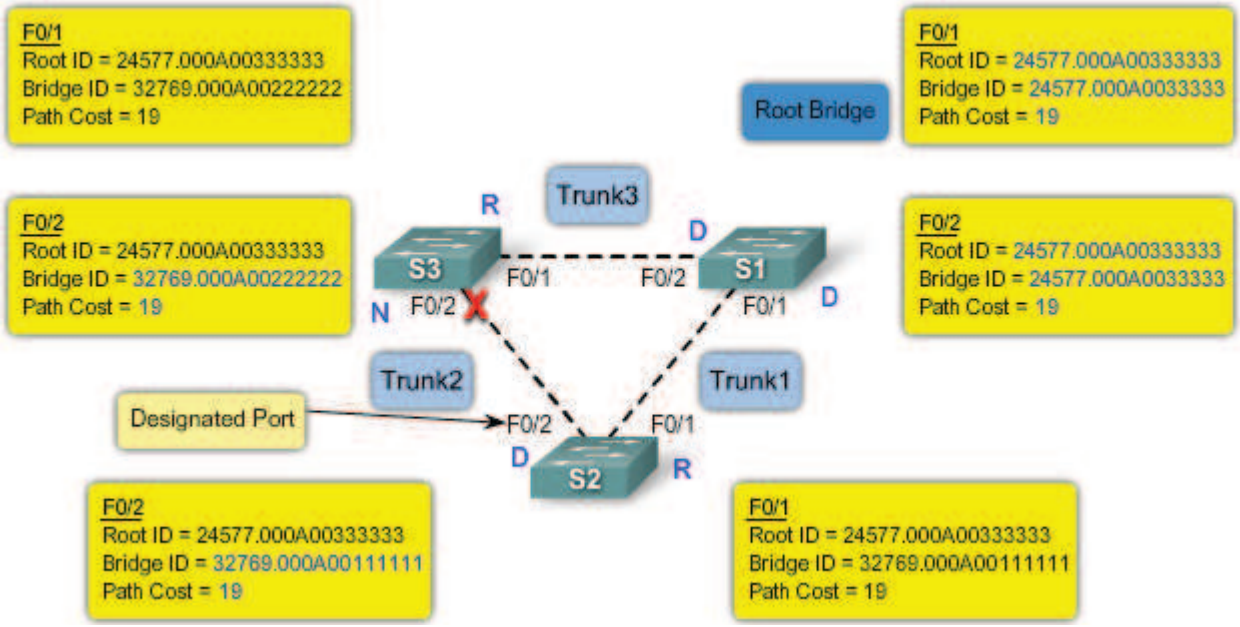
Root port switche je určen nejnižší hodnotou cesty k root bridgi. Pokud je takových portů více (např. při vícenásobném propojení switchů bez použití technologie EtherChannel), rozhoduje mezi nimi port ID (případně priorita portu). Vítězný port je označen jako root port, druhý port jako non-designated port (aby se vyhnulo cyklům).

Toto určování role portu probíhá současně s volbou root bridge – vždy, když se změní na switchi root bridge, jsou portům přepočítány jejich role. Takže na konci volby mají porty již určeny své finální role (v průběhu volby se mohly několikrát měnit).

Aktuální role a stavy portů jsou vidět v tabulce výpisu **`show spanning-tree`**.

5.3.4 Určení designated a non-designated portů

Porty switche, které nejsou root porty, jsou buď designated nebo non-designated. Určení probíhá mezi switch, které jsou danou linkou propojeny. Prvním kriteriem je délka (hodnota) cesty portů, a pokud je stejná, tak BID těchto switchů. Vítězí port s nižší hodnotou cesty, případně BID – ten se stává designated portem. Tento proces opět funguje již v průběhu volby root bridge. Pokud je portu nastavena role non-designated, ukáže se to ve výpisu příkazu **`show spanning-tree`** jako „Altn“ (zkratka z „Alternate“).



Na obrázku je vidět výsledek – volba role portů mezi D a N proběhla mezi F0/2 na S3 a F0/2 na S2. Protože S2 má nižší BID (díky MAC adrese), je jeho port zvolen designated (D).

5.3.5 STP – změna topologie

Switch detekuje změnu topologie například tím, že funkční port se vypne. V tom případě switch pošle upozornění root bridgi a ten informuje celou síť.

V normálním (konvergovaném) stavu STP switch na svém root portu BPDU jen přijímá, ale nevysílá. Proto existuje speciální zpráva TCN („Topology Change Notification“), která informuje o změně topologie právě root bridge. Switch ji odešle na svůj root port – na switch, který je vzhledem k němu „designated“. Ten mu odpoví pomocí standardní BPDU zprávy s vlastností TCA (potvrzení TCN) a následně pošle TCN směrem na svůj root port. Takto TCN postupně projde až k root bridgi. Root bridge po příjmu TCN zprávy začne vysílat konfigurační BPDU s vlastností TC („Topology Change“), kterou switche předávají dál. Díky tomu se všechny switche v síti „dozví“ o změně topologie.

5.4 PVST+, RSTP, Rapid-PVST+

5.4.1 Varianty STP

Existuje mnoho variant STP protokolu – některé jsou proprietární Cisco protokoly, jiné jsou definovány jako standardy IEEE (zpravidla se vyvinuly z Cisco verzí). Každopádně je nutné mít stručný přehled o tom, co která varianta nabízí.

Cisco:

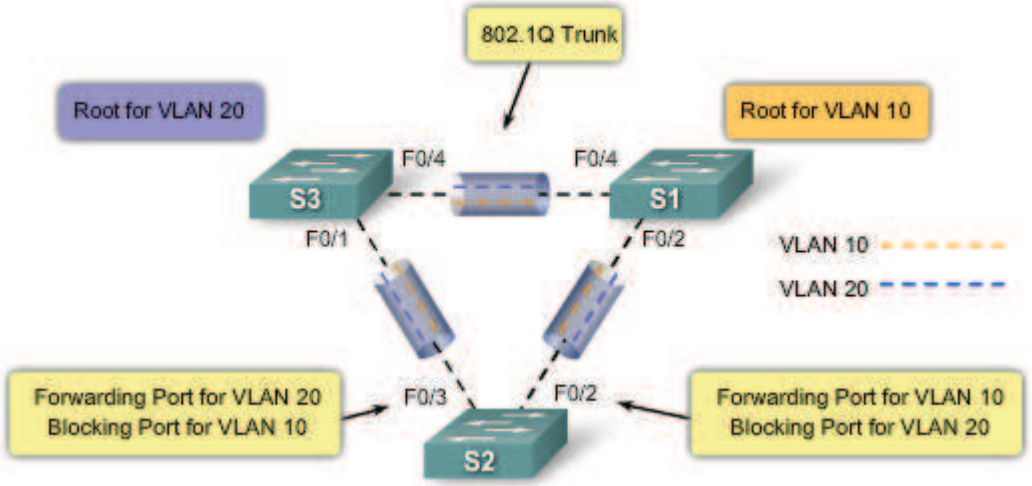
- PVST (Per-VLAN STP)
    - používá proprietární Cisco ISL trunking protokol
    - každá VLAN má vlastní STP, díky čemuž umožňuje load balancing už na linkové vrstvě
    - podporuje rozšíření BackboneFast, UplinkFast a PortFast
- PVST+
    - oproti PVST přidává podporu IEEE 802.1Q trunking protokolu a podporu dalších rozšíření – BPDU Guard a Root guard
- Rapid-PVST+
    - založen na standardu IEEE 802.1w
    - má rychlejší konvergenci než 802.1D

IEEE:

- RSTP
    - rychlejší konvergence než 802.1D
    - podporuje základní Cisco rozšíření
    - zapracován do 802.1D a označen jako specifikace 802.1D-2004
- MSTP
    - podpora více VLAN v jednom stromě STP
    - inspirován Cisco MISTP (Multiple Instances STP)
    - obsažen v 802.1Q-2003

5.4.2 PVST+

PVST+ umožňuje mít každé VLAN v síti vlastní STP. Musíme si ale uvědomit, že to znamená také větší zátěž sítě kvůli přenášení BPDU pro každou VLAN zvlášť. Na druhou stranu je možné poté optimalizovat síť například tak, aby jedna linka přenášela provoz jedné poloviny VLAN a druhá linka provoz té druhé poloviny VLAN – viz obrázek:



S1 je nakonfigurován, aby byl root bridge pro VLAN 10 a S2 root bridge pro VLAN 20. Díky tomu půjde provoz z S2 pro VLAN 10 na S1 (viz informace o portu F0/2) a provoz pro VLAN 20 na S3 (přes F0/3). Poznámka – v klasickém STP by to dopadlo tak, že jedna linka z S2 by byla kompletně vypnuta, takže VLAN 10 a VLAN 20 by se o tu druhou linku musely dělit.

PVST+ Bridge ID

V klasickém STP nebylo potřeba rozlišovat BID pro jednotlivé VLAN – proto bylo nutné toto BID upravit. Původní BID bylo priorita (2B=16b) + MAC adresa (6B=48b). Nově byla priorita rozdělena – část je vyhrazena pro určení VLAN: BID = priorita (4b) + „extended system ID“ (=VLAN ID, 12b) + MAC adresa (6B=48b).

Příklad: priorita + VLAN ID + MAC adresa = BID

32768 + 10 + 000A00333333 = 32778.000A00333333 pro VLAN 10

32768 + 20 + 000A00333333 = 32788.000A00333333 pro VLAN 20

Poznámka – pokud by měly switche výchozí nastavení, byly by jejich priority stejné, takže o BID (a root bridgi) by rozhodovaly MAC adresy. Proto je vhodnější nastavit prioritu pro jednotlivé VLAN podle potřeby (také proto, že starší MAC jsou zpravidla nižší a tím pádem také voleny).

Ukázka konfigurace (topologie viz obrázek výše).

Nastavení, aby S1 byl root bridgem pro VLAN 10 a záložním root bridgem pro VLAN 20 (nastavení S3 by bylo analogické):

```
S1(config)#spanning-tree vlan 10 root primary
S1(config)#spanning-tree vlan 20 root secondary
```

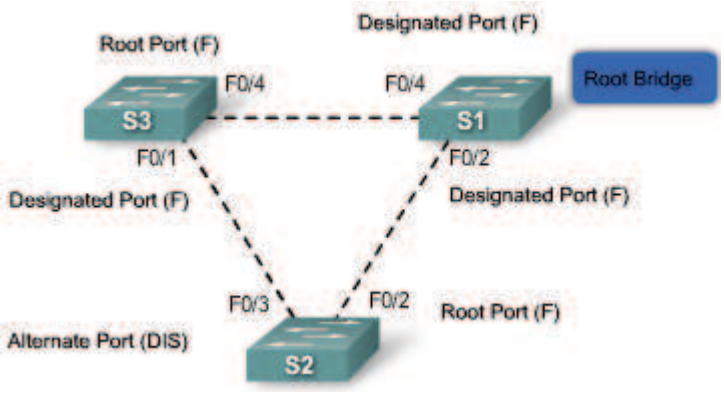
Jiná možnost řešení – pomocí nastavení přímo hodnoty priority:

```
S1(config)#spanning-tree vlan 10 priority 4096
```

Ověření – příkaz **`show spanning-tree active`** (ukáže aktuální stav switche v rámci STP pro jednotlivé VLAN), **`show running-config`** (zobrazí provedenou konfiguraci – jestli jsme něco nezapomněli nebo nezadali chybně).

5.4.3 RSTP

RSTP se vyvinul ze standardu 802.1D, princip a terminologie se proti 802.1w téměř nezměnily. RSTP definuje nový typ portu, nepodporuje stav portu „blocking“ – definuje stavy „discarding“, „learning“ a „forwarding“.



Na obrázku je role portu F0/ 3 na S2 určena jako „alternate“ se stavem „discarding“.

Charakteristika RSTP

- RSTP urychluje přepočet ST při změně topologie – pokud je port „alternate“ nebo „backup“, může být okamžitě přepnut do stavu „forwarding“ bez nutnosti čekat na konvergenci sítě
- RSTP je v současnosti upřednostňovaný protokol pro řešení L2 cyklů
- nepodporuje předchozí vlastnosti – UplinkFast a BackboneFast
- nahrazuje STP (802. 1D) a zachovává zpětnou kompatibilitu (např. totožný výpočet root bridge)
- má stejný formát BPDU jako STP
- nepotřebuje časovače (timery), na které se má čekat při přepínání do „forwarding“ stavu

RSTP BPDU

Protože RSTP má formát BPDU (typ 2, verze 2) kompatibilní s STP, mohou spolu bez problémů komunikovat na jedné lince. Zprávy jsou ale odesílány trochu jinak:

- jestliže vyprší časovač „max-age“ nebo se ztratí 3 po sobě jdoucí příchozí „hello“ zprávy (tj. ve výchozí konfiguraci 6 sekund), jsou informace okamžitě ztraceny
- tři ztracené po sobě jdoucí BPDU znamenají ztrátu konektivity (výpadek linky) s root nebo designated bridgem – to umožňuje rychlou detekci výpadků

Na obrázku je znázorněna struktura BPDU a využití bytu označeného „Flags“:

- bity 0 a 7 označují, zda došlo ke změně topologie a její potvrzení
- bity 1 a 6 označují proces návrhu řešení a souhlasu (pro urychlení konvergence)
- bity 2-5 označují roli a stav portu odesílajícího BPDU, z toho
- bity 4 a 5 obsahují zakódovanou roli portu – viz obrázek

RSTP Version 2 BPDU

Field	Byte Length
Protocol ID=0x0000	2
Protocol Version ID= 0x02	1
BPDU Type= 0x02	1
Flags	1
Root ID	8
Root Path Cost	4
Bridge ID	8
Port ID	2
Message Age	2
Max Age	2
Hello Time	2
Forward Delay	2

Flag Field

Field Bit	Bit
Topology Change	0
Proposal	1
Port Role Unknown Port Alternate or Backup Port Root Port Designated Port	2-3 00 01 10 11
Learning	4
Forwarding	5
Agreement	6
Topology Change Acknowledgement	7

5.4.4 Hraniční porty (edge ports)

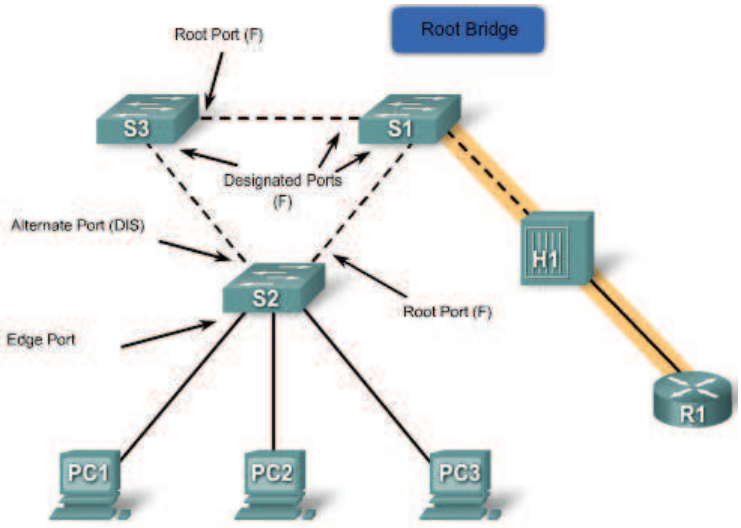
V terminologii RSTP je „edge port“ takový port, ke kterému nebude nikdy připojen jiný switch, takže po zapnutí se automaticky přepne do „forwarding“ stavu (nečeká na STP). Je to podobné předcházející vlastnosti PortFast, rozdíl je v tom, že edge port v případě obdržení BPDU ztrácí status edge portu a stává se standardním portem zapojeným do spanning-tree.

Nicméně nastavení edge portu zůstává stejné – příkazem **`spanning-tree portfast`**.

5.4.5 Typy linek (spojů)

Typ linky určuje, zda může být port za daných okolností okamžitě převeden do „forwarding“ stavu. Podmínky, za kterých toto může být provedeno, jsou různé pro „edge“ porty a ostatní porty. Typ linky je automaticky detekován, ale může být překonfigurován ručně.

Typy linek jsou buď bod-bod (point-to-point) nebo sdílená linka. Na obrázku je zvýrazněna sdílená linka (protože je tam hub), vše ostatní jsou spoje bod-bod:



Přehled možností přepnutí do „forwarding“ stavu:

- root porty parametr „typ linky“ nevyužívají – po inicializaci jsou automaticky schopny okamžitého přepnutí

- alternate a backup porty zpravidla tento parametr také nevyužívají
- nejvíce tento parametr využívají designated porty – pokud je linka připojená k danému portu typu point-to-point, přepne se port okamžitě do „forwarding“ stavu

5.4.6 RSTP – stavy a role portů

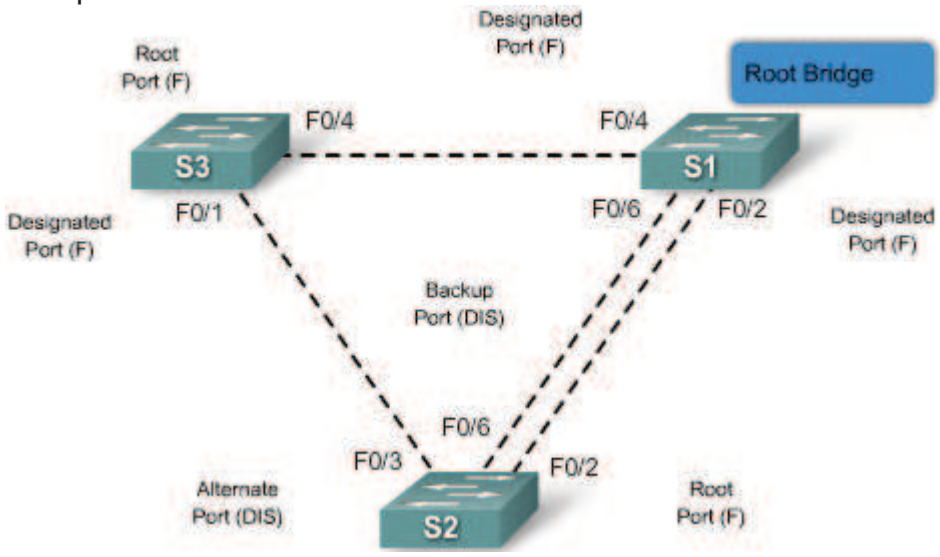
RSTP poskytuje rychlou konvergenci sítě po výpadku linky nebo náhradě zařízení. V RSTP je role portů nezávislá na stavu portu. Možné stavy portů jsou:

- discarding – existuje jak ve stabilní síti, tak v průběhu synchronizace změn topologie; data obdržená na tomto portu jsou zahazována, díky čemuž jsou přerušeny případné L2 cykly
- learning - existuje jak ve stabilní síti, tak v průběhu synchronizace změn topologie; port akceptuje zprávy pouze pro účely doplňování MAC tabulky a tím omezení propouštění unicastových zpráv pro neznámé MAC adresy
- forwarding – existuje pouze v konvergované síti; port detekuje a přijímá změny topologie; pokud v průběhu změny topologie přijme datový rámec, musí před předáním proběhnout proces schválení („proposal and agreement process“).

Srovnání stavu portů v STP a RSTP – v podstatě jsou 3 stavy STP shrnuty v RSTP do jednoho stavu:

Operativní stav portu	Stav portu v STP	Stav portu v RSTP
vypnutý (disabled)	blocking	discarding
vypnutý (disabled)	listening	discarding
vypnutý (disabled)	learning	learning
vypnutý (disabled)	forwarding	forwarding
zapnutý (enabled)	disabled	discarding

Role portů



V RSTP může mít port jednu ze čtyř rolí:

- root port – pouze na switchích, které nejsou root bridgem; je to port, přes který vede nejkratší cesta k root bridgi, může být na switchi pouze jeden; ve stabilní síti je ve stavu „forwarding“
- designated port – v každém segmentu je právě jeden designated port (a příslušný switch je určen pro tento segment jako designated switch); všechny switche na tomto segmentu na-

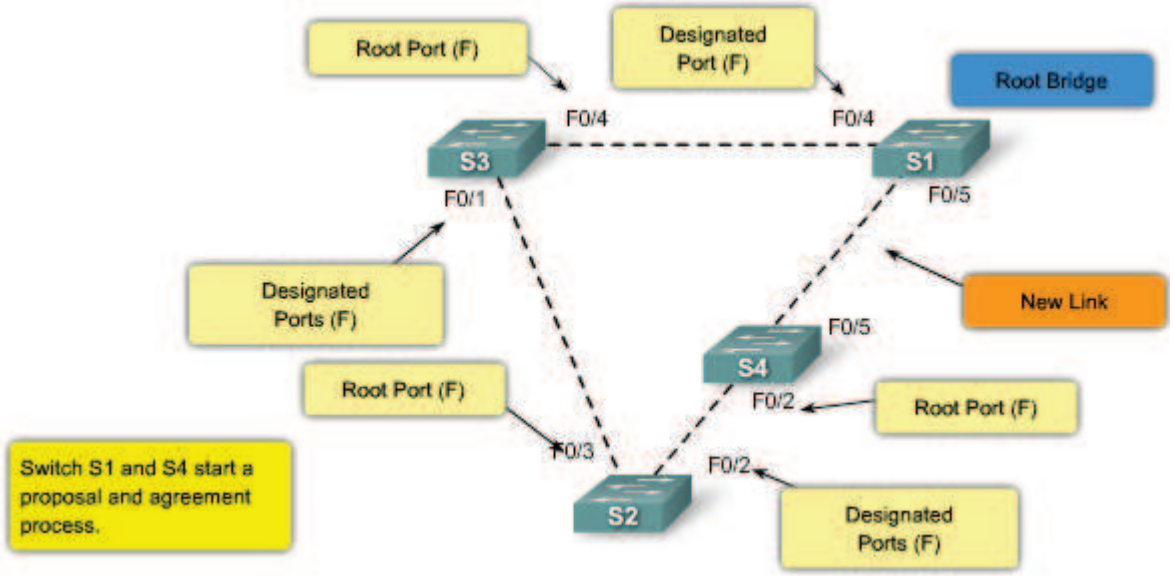
slouchají BPDU a vybírají právě designated switch; ve stabilní síti je designated port ve „forwarding“ stavu a přijímá zprávy, které následně směřují k root bridgi

- alternate port – je to port switche, který poskytuje alternativní cestu k root bridgi, která ale zatím není potřeba; ve stabilní síti je tento port ve stavu „discarding“; „alternate“ port se může vyskytnout pouze na switchích, které nejsou designated switchi a v případě výpadku hlavní cesty se změní na designated port
- backup port – je záložní port u redundantní linky, kde je switch určen jako designated; tento port má vyšší port ID, než odpovídající designated port; ve stabilní síti je ve stavu „discarding“

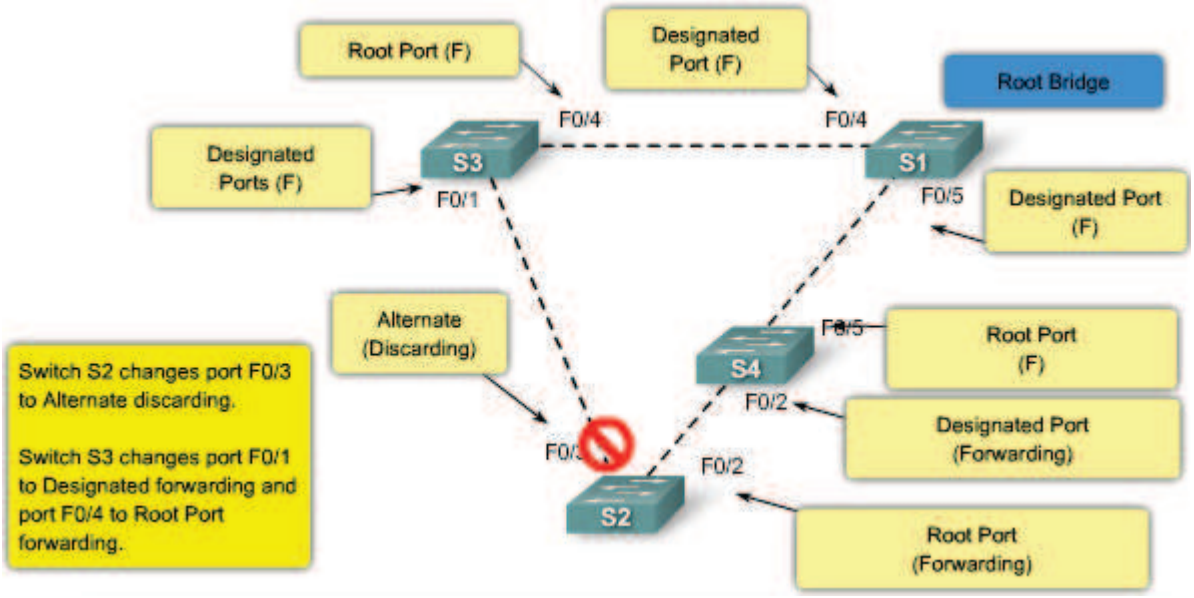
RSTP proces „Proposal and Agreement“

V STP musí nově zvolený designated port čekat dvojnásobek času „forward delay“, než je stav portu nastaven na „forwarding“. Protože RSTP řeší konvergenci pro každou linku nezávisle, nejsou tyto časovače potřeba, takže některé porty je možné přepnout do „forwarding“ stavu okamžitě – splňují to edge porty na point-to-point linkách, což jsou v důsledku designated porty ve stavu „discarding“.

Na obrázku je v topologii naznačeno zapojení nové linky (mezi S1 a S4).



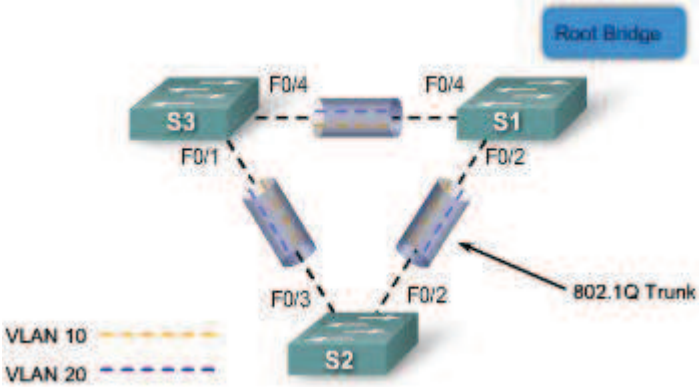
Nejprve se oba nově aktivované porty nastaví jako „discarding“. Poté S1 pošle k S4 „proposal BDU“ (začne proces). S4 zjistí, že toto je kratší cesta k root bridgi, takže po dobu synchronizace (procesu) zablokuje všechny ST porty (ne edge porty – tady F0/2). Poté odpoví „agreement BPDU“, čímž se dohodnou, že F0/5 na S4 bude „root port“ a F0/5 na S1 bude „designated port“ – oba ve stavu „forwarding“. Port F0/2 na S4 zůstává blokovaný, protože nyní S4 vyvolá synchronizační proces s S2 – výsledek bude podobný (S2 zablokuje F0/3). Následně se S2 synchronizuje s S3 – výsledek je, že port F0/3 na S2 se stává „alternate“ portem ve stavu „discarding“ (cesta přes S4 je vyhodnocena jako lepší) – viz obrázek:



Při další synchronizaci S3 s S1 už se na výsledku nic nezmění.

5.4.7 Konfigurace Rapid-PVSTP+

Rapid- PVSTP+ je Cisco implementace RSTP. Podporuje ST pro každou VLAN a RSTP v sítích s Cisco zařízeními. V topologii jsou dvě VLAN – 10 a 20, které vytvoříme a nakonec nakonfigurujeme Rapid-PVST+ na switchi S1, který je root bridgem a STP serverem.



Instance Rapid-PVST+ stromu pro VLAN je vytvořena při přidání prvního rozhraní do dané VLAN a odstraněna při přesunutí posledního rozhraní do jiné VLAN.

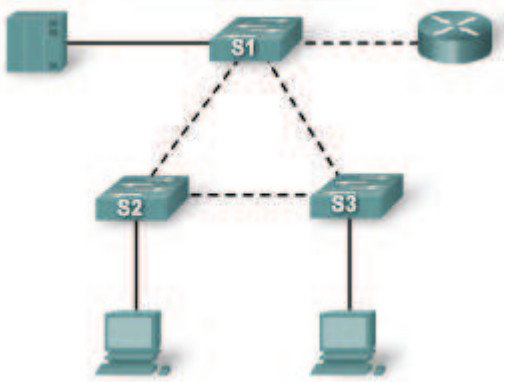
Příkazy použitelné pro konfiguraci Rapid-PVST+ :

- `(config)#`**`spanning-tree mode rapid-pvst`** – nastaví režim STP na Rapid-PVST+
- `(config-if)#`**`spanning-tree link-type point-to-point`** – nakonfiguruje typ linky na daném rozhraní na point-to-point (druhá možnost – shared) – umožní za vhodných podmínek rychlé přepnutí do „forwarding“ stavu (viz výše)
- `#`**`clear spanning-tree detected-protocols`** – odstraní veškeré detekované STP (následně se vytvoří ty „správné“)
- `#`**`show spanning-tree vlan`** `10` – ověření aktuálního stavu STP pro danou VLAN - informace o root ID, bridge ID, rolích portů, typech linek
- `#`**`show running-config`** – ověření aktuální konfigurace (zadané příkazy) STP

5.4.8 Návrh STP – předcházení problémům

Určení root bridge

Není příliš vhodné ponechat na STP volbu root bridge. Při známé topologii je vhodné nastavit switche tak, aby STP „vybral“ námi určený switch (např. pomocí priorit). Vhodné je určit výkonný switch někde „uprostřed“ topologie.

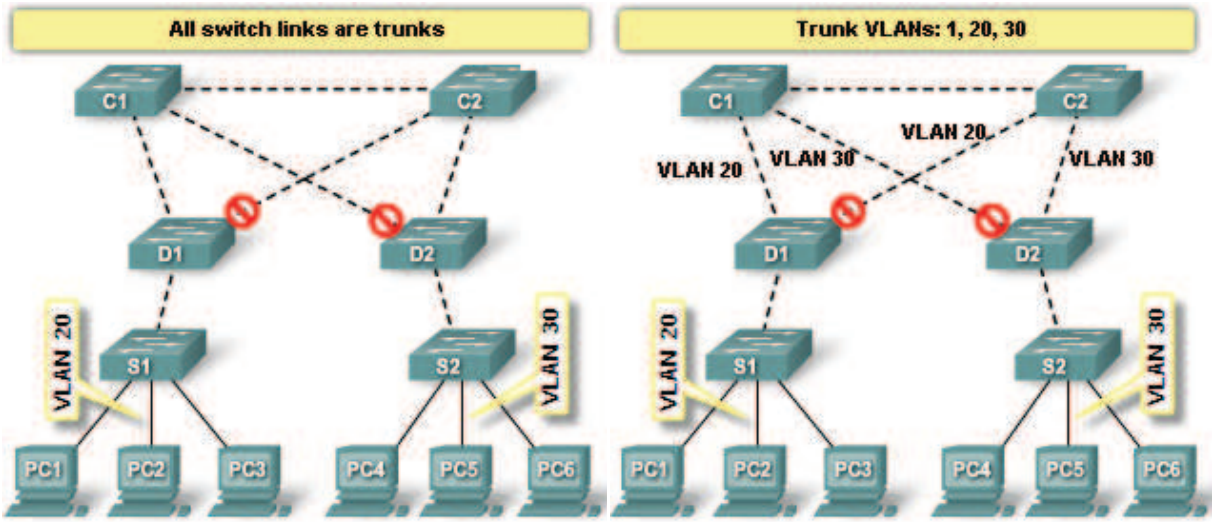


Například pro tuto topologii je vhodné nastavit jako root bridge S1 – výsledkem je, že stanice připojené k oběma switchům (S2 a S3) mají stejně dlouhou cestu jak k serverům, tak k routeru (internetu), protože deaktivována bude právě linka mezi S2 a S3. V každém jiném případě by byla deaktivována jiná linka, což by způsobilo jinak dlouhé cesty k serverům pro stanice (a když se tomu můžeme vyhnout).

Toto nastavení je nutné provést pro každou VLAN (root bridge a backup root bridge).

Plánování redundantních linek a blokovaných portů, pruning

Je vhodné již dopředu plánovat nastavení redundantních linek a předpovědět, které porty budou blokované, případně toto vyřešit „ručně“ – vhodnou konfigurací trunk spojů. Blokované porty jsou jedním z mála kritických míst v STP (nevhodné přepnutí do „forwarding“ režimu může mít nepříjemné důsledky pro velkou část sítě). Na obrázku jsou dvě možná řešení. Vlevo - automatické blokování pomocí STP – switch D1 musí blokovat jeden port pro VLAN 20 i pro VLAN 30, ačkoliv přes něj žádná komunikace pro VLAN 30 nemusí procházet, protože všechny jeho stanice jsou ve VLAN 20 (analogicky D2 pro VLAN 20). Vpravo – blokované porty jsou blokované pouze pro jednu VLAN.



Využití L3 přepínání

Router má dva hlavní úkoly – vytvářet si směrovací tabulky pomocí ostatních routerů a směrovacích protokolů a směrování zpráv mezi rozhraními podle cílové adresy a směrovacích tabulek. Toto směrování zvládají také L3 switche – s rychlostí blízkou L2 přepínání.

Důsledkem je také, že spoj mezi C1 a C2 je směrován (není to trunk), takže z pohledu redundance není potřeba vypínat žádné porty.

Závěr – doporučení

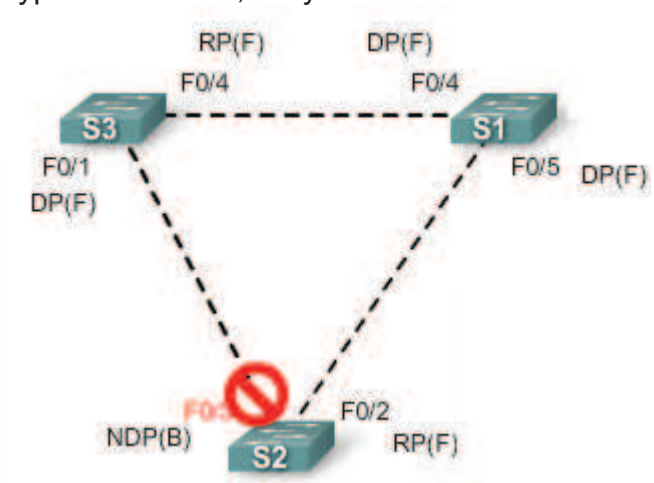
STP nevypínejte, ani když už by nemuselo být potřeba – není příliš náročné ani na procesor, ani na linky, ale může zabránit případnému pádu sítě (např. při zapojení dalšího switche nebo chybném zapojení kabelu).

Oddělte administrativní VLAN od reálného provozu – přílišná zátěž administrativní VLAN (např. broadcasty) by mohla ohrozit doručování BPDU.

Rozdělte síť na více domén pomocí L3 switchů (routerů) – přerušíte tím vznik možných L2 cyklů na VLAN 1 (výchozí VLAN, kde má standardně každý switch svoji adresu).

5.4.9 Řešení problémů s STP

Výpadek switche, linky



Pokud F0/3 je blokován, je vše v pořádku. Jestliže přestane (z nějakého důvodu) F0/3 uzlu S2 dostávat od F0/1 (S3) BPDU, přepne se port do aktivní role a může vysílat data. A jestli mezitím dojde k opětovnému propojení, je uzavřen L2 cyklus.

Výpadek v síti

Abychom správně mohli identifikovat výpadek v síti, je vhodné mít informace o topologii sítě, o vybraném root bridgi, blokovaných portech a redundantních linkách. Většinu problémů odhalíme příkazem **`show`**.

Chybná konfigurace PortFast

Typicky, když k portu s konfigurovanou vlastností PortFast, který dříve sloužil k připojení PC, připojíme switch. Tím dojde k uzavření cyklu. Ten se automaticky přeruší až ve chvíli, kdy blokovanému portu přijde BPDU.

Průměr sítě

Dodržujte max. průměr sítě (7). Pokud bude větší, nemusí nutně všechny switche slyšet BPDU ostatních switchů.