Algoritmi de evitare a buclelor la nivel data link

Studenti :

Nastase Alexandru **441A**

Sandru Traian **441A**

Cuprins

I. Introducere

In retelele LAN modern legaturile dintre echipamentele de retea sunt realizate prin cabluri fizice dar se pot realiza si prin legaturi wireless. Aceste conexiuni odata facute dau posibilitatea componentelor retelei sa primeasca si sa transmita date.

Echipamentele care folosesc legaturi fizice transmit datele respectand standardele stabilite de Ethernet. Standardele Ethernet stabilesc detaliile legate de cablare dar si regulile ce guverneaza nivelul legatura de date, ce se refera la incapsularea pachetelor in cadre(frame-uri) si la adresarea MAC.

In general principalele componente ale retelelor de tip LAN(Local Area Network)sunt urmatoarele:

- Dispozitive dotate cu placi de retea;
- Switch-uri;
- Legaturi intre dispositive(cabluri UTP,FTP, fibra optica);

Protocolul STP(Spanning Tree Protocol) previne modul de transfer al frame-urilor de catre switch cu scopul limitarii buclelor in retea. Se poate intampla ca in reteaua LAN switch-urile sa transmita frame-urile in asa fel incat acestea sa fie transmise in bucla, ducand la traficul lor infinit prin retea, lucru ce favorizeaza congestia.

Pentru a demonstra cum pot sa apara frame-uri trebuie mai intai explicat modul si logica dupa care switch-urile transmit frame-uri. [1]

Modul de transmitere al frame-urilor atunci cand se ignora functionalitatea STP-ului:

1. Se determina VLAN-ul din care face parte frame-ul si unde acesta trebuie trimis:
   a. Daca frame-ul vine pe o interfata in mod acces se va folosi ca referinta VLAN-ul interfetei pe care a venit.
   b. Daca frame-ul vine pe o interfata trunk, se va folosi VLAN-ul afisat in incapsularea trunk.
2. Se adauga adresa MAC sursa in tabela de adrese MAC.

3. Se verifica adresa MAC destinatie in tabela de adrese MAC :
    a. Daca se gaseste se trimite pe interfata care e asociata cu adresa MAC respectiva.
    b. Nu se gaseste: Se trimite frame-ul pe toate interfetele din acelasi VLAN cu interfata pe care a venit sau pe interfetele trunk.

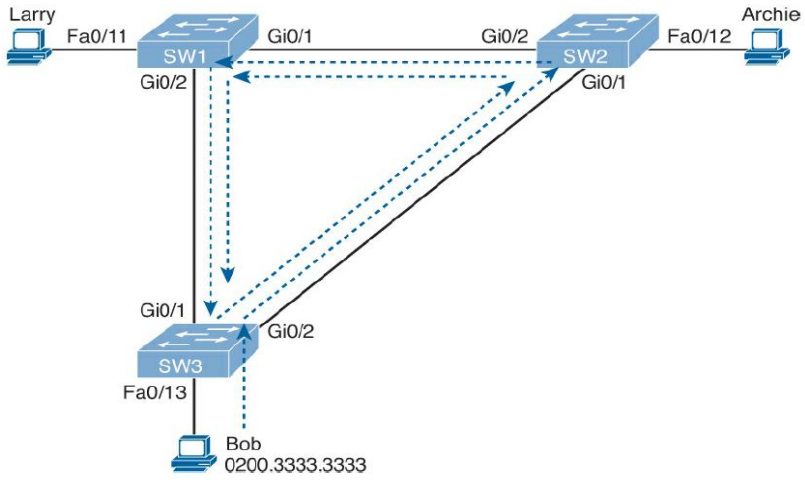
Intr-o retea cu legaturi redundante pot aparea situatii in care frame-urile se invart in retea la infinit. Protocolul STP modifica statutul porturilor switch-ului din modul de trimitere(forwarding) in modul blocare(blocking) cu scopul evitarii buclelor. Protocolul blocheaza in mod inteligent porturile astfel incat sa nu se intrerupa legatura intre doua dispozitive din retea.

In continuare sunt enumerate principalele probleme care pot sa apara in lipsa utilizarii STP in retea:

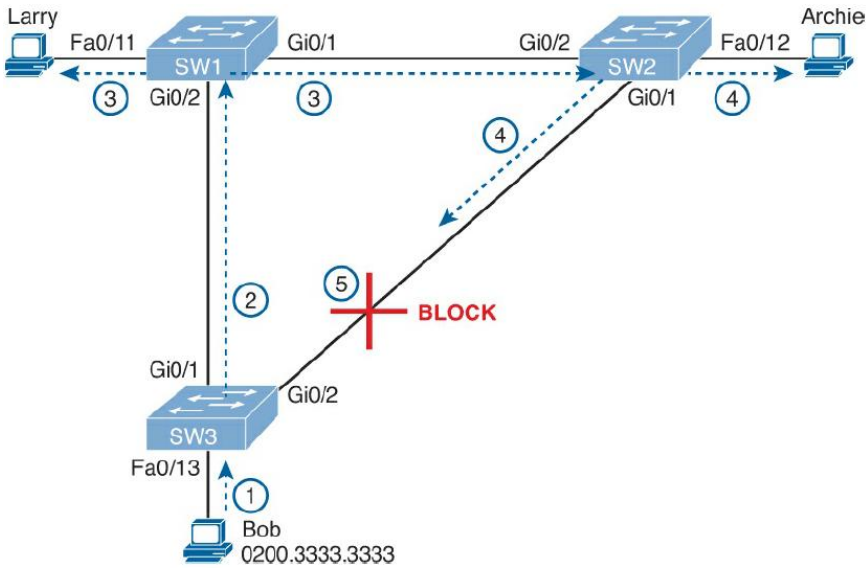
- Broadcast storms = trimiterea unor frame-uri pe aceleasi porturi in retea in mod repetat ceea ce duce incarcarea buffere-lor porturilor.
- Instabilitate in cadrul tabelei de MAC-uri = se traduce prin umplerea tabelei cu adrese MAC incorecte ce duc la trimiterea frame-urilor la adrese gresite.
- Transmisia de frame-uri multiple = un dezavantaj al buclarii in retea este ca multiple copii ale aceluiasi frame ajung la gazdele destinatie, acestea putand fi induse in eroare.

Protocolul STP incearca sa evite situatiile descrise mai sus punand porturile in stare de forwarding sau stare de blocking.
Starea de forwarding reprezinta starea in care portul se comporta in mod normal in sensul ca trimite frame-urile la destinatii indiferent de VLAN-ul din care fac parte sau daca sunt frame-uri de tip unicast(o singura destinatie), multicast(un grup de destinatii) sau broadcast(toate destinatiile din retea).

In continuare e prezentata o topologie in care nu functioneaza Spaning Tree Protocol-ul cu problemele ce pot sa apara:



Bob trimite un frame broadcast. Acesta e transmis pe toate porturile de catre SW3 mai departe SW1 trimite pe toate porturile mai putin pe cel care primeste. Larry primeste frame-ul. SW2 primeste frame-ul si trimite din nou pe toate porturile exceptand Gi0/2. Acest lucru e facut in prima faza si de SW2 cand primeste frame-ul de la SW3. Astfel apar bucle in care frame-urile circula la infinit, statiile Larry, Bob si Archie primesc in mod continuu copii ale frame-urilor si astfel de aici porneste nevoia folosirii unui protocol de evitare a acestor probleme.
Problema e rezolvata in topologia urmatoare in care protocolul STP e implementat:



Acelasi lucru se intampla si aici: Bob trimite un frame broadcast catre SW3. Algoritmul ruleaza pe cele 3 switch-uri si ia decizia sa blocheze portul Gi0/2. Numerele din imagine descriu modul de propagare al frame-ului:

1. Bob transmite.
2. SW3 transmite doar pe Gi0/1 care e in modul propagare.
3. SW1 transmite la Larry si la SW2.
4. SW2 trimite catre Archie si catre SW3 si nu catre SW1 deoarece pe acolo a venit.
5. Portul Gi0/2 e in modul blocking(blocare) si nu mai primeste frame-ul.

Se observa astfel din pasii de mai sus ca problemele aparute in topologia de retea in care algoritmul nu era implementat aici dispar.[1] [2][3]

II. Algoritmul Spanning Tree Protocol (Nastase Alexandru – 441A)

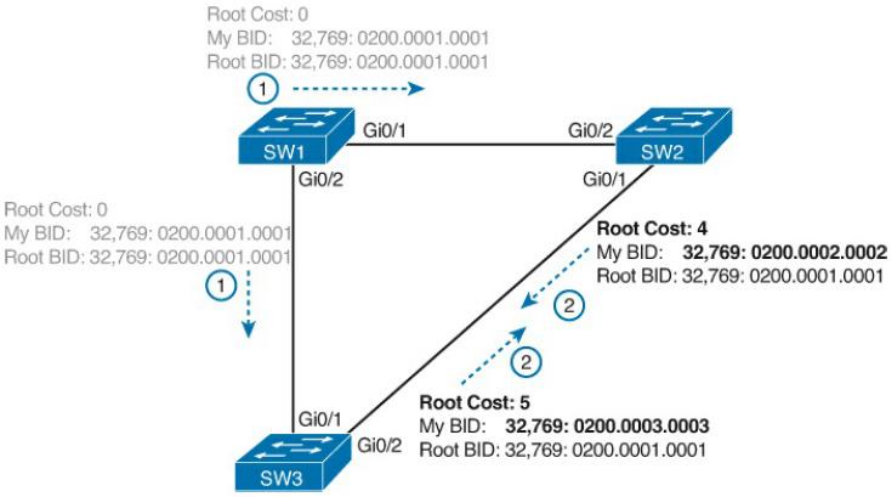
2.1. Descrierea algorimului

Algoritmul creeaza un arbore al nodurilor din retea care transmit frame-uri. Astfel se incearca sa se gaseasca o singura cale in arbore, intre doua noduri care transmit. Protocolul STP denumit si de multe ori STA(Spaning Tree Algorithm) practic decide care porturi sa fie blocate si care porturi nu. Porturile blocate nu primesc frame-uri si nu transmit adica nu se afla in stari de ascultare sau invatare.
Pasii principali in functionarea algoritmului sunt urmatorii:

- Se alege un switch care va fi considerat radacina(root). Toate porturile acestui switch se vor pune in modul forwarding;
- Fiecare switch care nu este root isi calculeaza costul administrativ catre root, aceasta numindu-se costul de radacina (root cost) iar portul cu cost administrativ mai mic este desemnat port radacina(root port).
- Legatura dintre 2 switch-uri se numeste segment. Pe un segment se calculeaza costul catre root iar portul cu cost mai mic e pus in modul desemnat (designated).

Toate celelalte porturi care nu au fost catalogate cu una din caracteristicile root sau designated vor fi puse in mod blocking.

Primul pas al algoritmului este alegerea switch-ului root. Pentru a face asta switch-urile schimba intre ele mesaje de hello la inceputul algoritmului pentru a putea desemna root-ul. Aceste mesaje de hello se numesc mesaje BPDU (bridge protocol data units). Ele sunt pe 8 octeti si contin un camp de prioritate pe 2 octeti si un camp system ID pe restul de 6 octeti. Acest system ID este reprezentat de catre adresa de MAC lucru care da unicitate acestui camp deoarece nu se vor gasi doua dispozitive cu acelasi MAC in retea. Mesajele de tip BPDU contin informatii despre bridge ID-ul root-ului, bridge ID-ul sursei, costul root al sursei si o serie de contori(timeri).



In topologia de mai sus se va alege switch-ul root. La inceput toate switch-urile se considera root. SW1 va trimite mesaje BPDU catre vecini adica SW2 si SW1. La inceput cum toate switch-urile cred ca sunt root au costul BID-ul egal cu BID-ul root-ului. Cand SW2 primeste BPDU-ul de la SW1 compara BID-ul sau cu BID-ul trimis de vecin. In functie de aceasta valoare(BID SW2 > BID SW1) SW2 stabileste ca root-ul e SW1. Se vor compara mai intai campul prioritate care de obicei are valori multiplu de 4096 in mod implicit. Daca acestea sunt egale se trece la compararea campurilor system ID care fiind reprezentate de adresele MAC nu pot fi egale. Acest lucru se intampla si in cazul switch-ului SW3. Primeste un mesaj BPDU de la SW1 iar dupa compararea BID-urilor SW3 considera ca SW1 e root petru ca are system id-ul mai mic. Astfel putem rezuma: SW1 se considera

root, SW2 il considera pe SW1 root, SW3 il considera pe SW1 root. Astfel a fost ales switch-ul root.

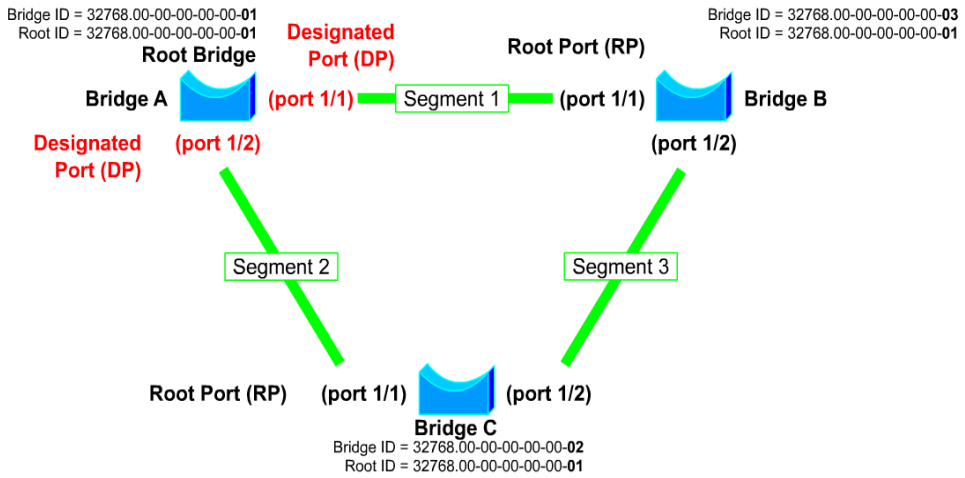
Dupa alegerea radacinii aceasta va continua sa trimita mesaje de hello BPDU la anumite interval de timp ce vor fi discutate mai tarziu.

Dupa ce a fost ales switch-ul root toate porturile acestuia se vor pune in mod designated. Acestea trimit si primesc frame-uri de orice tip.

Celelalte switch-uri trebuie acum sa isi aleaga portul cu statut de root. Acesta va fi ales in functie de costul catre switch-ul root. Se iau toate caile posibile catre root iar calea cu costul cel mai mic este prioritara, in sensul ca portul acela va fi desemnat port root.

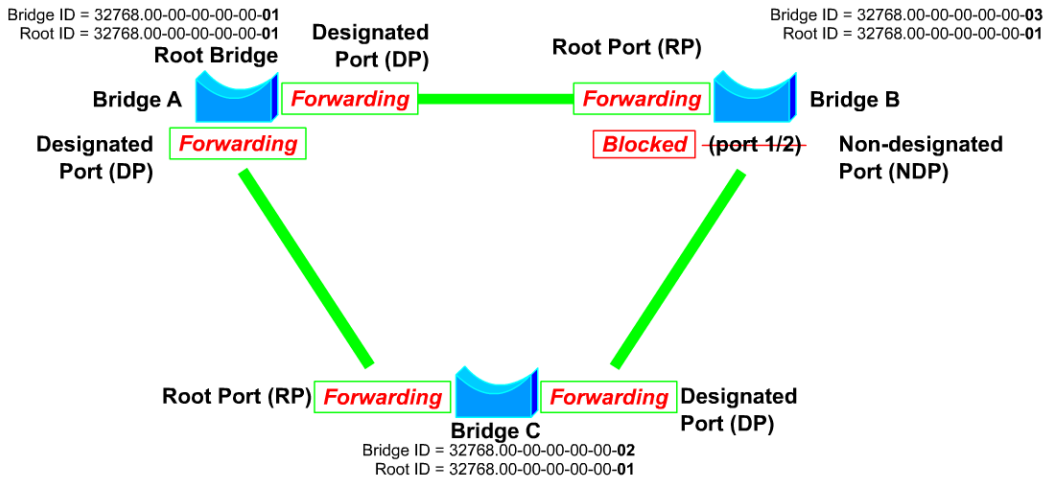
Dupa ce au fost alese toate porturile de root toate celelalte porturi trebuie sa fie in stare designated sau blocking.

Pe segmentele de retea se stabilesc aceste tipuri de porturi.



Pe topologia de mai sus se vor alege porturile designated asa cum se vede. Root-ul alege toate porturile sale asa cum am spus. In cazul segmentului 3 se compara costul pe segment intre port ½ al switch-ului c si port ½ al switch-ului b. In functie de valoarea mai mica se alege care dintre ele e port designated si care va fi port blocking. In cazul in care costul e egal si nu ne putem decide vom folosi valoarea bridge ID mai mica. [1][2]

La final topologia are urmatoarea forma:



Costurile se aleg in functie de vitezele Ethernet in felul urmator(stabilit de IEEE):

10Mbps -> cost=100
100Mbps -> cost=19
1Gbps -> cost=4
10Gbps -> cost=2

2.2. Metode de implementare

In mod implicit STP-ul porneste pe switch-uri in mod automat fara sa fie nevoie sa fie activat. Isi desfasoara algoritmul in mod normal si blocheaza porturile pe care le considera. Atunci cand functioneaza si au loc modificari in topologie el reactioneaza in acest sens. Exista 3 timpi cu care protocolul lucreaza:

- Hello este intervalul de timp intre mesajele de hello trimise de switch-uri (e de 2 secunde).
- Timpul MaxAge reprezinta cat trebuie sa astepte switch-ul dupa ce a incetat sa mai primeasca mesaje de hello si va trebui sa modifice configuratia porturilor. (Are valoarea de 20 secunde).
- Timpul Forward Delay care afecteaza procesul in care un port trece din starea de blocking in starea de forwarding. Portul intra intr-o stare de ascultare iar mai apoi in starea de invatare.

Protocolul intra in functiune in mod implicit dar se pot face modificari ce duc la alegerea modului in care se desemneaza root-ul astfel incat reteaua sa functioneze intr-un anumit fel.

Exemplu in care configuratia se aplica pentru a modifica switch-ul root in functie de preferinta:

```
SW2# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW2(config)# spanning-tree vlan 10 root primary
SW2(config)# ^Z
```

```
! Next, SW1 is configured to back-up SW1
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# spanning-tree vlan 10 root secondary

SW1(config)# ^Z
SW1#
```

In cazul de fata SW2 a primit prioritate de a fi ales root iar SW1 va fi urmatorul candidat.

Pentru a verifica modul in care sunt facute configuratiile pentru functionarea protocolului STP folosim comenzile urmatoare.[4]

```
SW2# show spanning-tree root

                                        Root    Hello Max Fwd
Vlan                Root ID             Cost    Time  Age Dly  Root Port
---------------- ------------------- --------- ----- --- ---  ------------
VLAN0001         32769 1833.9d5d.c900        23    2   20  15  Gi0/1
VLAN0010         32778 1833.9d7b.0e80         4    2   20  15  Gi0/2
VLAN0020         32788 1833.9d7b.0e80         4    2   20  15  Gi0/2
VLAN0030         32798 1833.9d7b.0e80         4    2   20  15  Gi0/2
VLAN0040         32808 1833.9d7b.0e80         4    2   20  15  Gi0/2

SW2# show spanning-tree vlan 10 bridge

                                                Hello  Max  Fwd
Vlan                    Bridge ID                Time  Age  Dly  Protocol
---------------- --------------------------------- -----  ---  ---  --------
VLAN0010         32778 (32768,  10) 1833.9d7b.1380    2    20   15  ieee
```

Aici se observa root ID-ul pentru fiecare VLAN in parte precum si costurile si timpii folositi. Se prezinta si portul root.

```
SW3# show spanning-tree vlan 10

VLAN0010
  Spanning tree enabled protocol ieee
  Root ID     Priority    32778
              Address     1833.9d7b.0e80
              Cost        8
              Port        26 (GigabitEthernet0/2)
              Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID   Priority    32778  (priority 32768 sys-id-ext 10)
              Address     f47f.35cb.d780
              Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
              Aging Time  300 sec

Interface           Role Sts Cost      Prio.Nbr Type
------------------- ---- --- --------- -------- --------------------------
Fa0/23              Desg FWD 19        128.23   P2p
Gi0/1               Altn BLK 30        128.25   P2p
Gi0/2               Root FWD 4         128.26   P2p
```

Pentru VLAN-ul 10 se afisaza toate informatiile: prioritati, cost precum si starile in care se afla porturile.

2.3. Avantaje si dezavantaje asupra modului de functionare al retelei

STP ofera redundanta pentru toate dispozitivele din retea. Termenul de redundanta se refera la faptul ca toate rutele au o cale de back-up care poate sa fie utilizata in caz de probleme cu legatura primara. In timp ce fiecare dispozitiv are mai multe cai pentru a transmite datele, doar o singura cale va fi activa pentru transmiterea datelor. In cazul in care acea cale pica se va calcula si gasi o alta cale pentru transmiterea de date. Acest lucru permite tot timpul sa existe legatura intre dispozitive chiar daca exista instabilitate in retea.[5]

Cum am spus si in capitolul de introducere avantajele aduse de protocol pentru retea sunt:

- Mentinerea redundantei retelei;
- Prevenirea buclelor;
- Gasirea unor cai optime de transmitere;
- Convergenta rapida a retelei in urma unei probleme;

III. Variante imbunatatite ale Spanning Tree Protocol si protocoale ajutatoare (Sandu Dorin Traian – 441A)

3.1. Rapid Spanning Tree Protocol - descriere si diferente fata de Spanning Tree Protocol

Protocolul STP a fost adoptat intr-o perioada in care era ideal cazul in care o legatura in retea era refacuta in mai putin de un minut, in cazul in care aceasta pica.

Rapid Spanning Tree protocol (RSTP) poate fi vazut ca o imbunatatire adusa STP-ului. Terminologia ramane in principal aceeasi la fel ca si termenii folositi atunci cand vorbim de STP.

STP sau asa cum e cunoscut dupa standard 802.1D definea urmatoarele stari pentru porturi:

- Disabled(inchis);
- Listening(ascultare);
- Learning(invatare);
- Blocking(blocare);
- Forwarding(trimitere);

In cadrul RSTP exista doar trei stari ale porturilor ce sunt echivalente cu starile porturilor din STP dupa cum urmeaza:

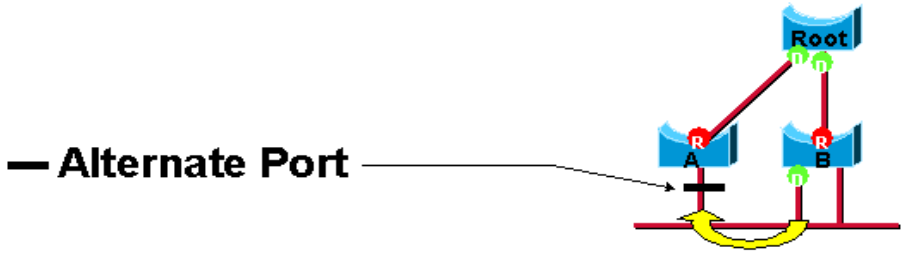
- Disabled(STP) →Discarding(RSTP)
- Listening(STP) →Discarding(RSTP)
- Learning(STP) →Discarding(RSTP)
- Blocking(STP) →Learning (RSTP)
- Forwarding(STP) →Forwarding(RSTP)

Rolulile de port root si port designated raman si in cadrul RSTP dar starea de blocking este impartita in 2 stari: starea back-up si starea alternate.

Algoritmul de alegere al switch-ului root este acelasi si anume prin trimiterea de mesaje BPDU din care se calculeaza si se stabileste root-ul. [2]

Portul de root are aceeasi semnificatie ca si in cadrul STP. Este portul cu cel mai mic cost atunci cand vine vorba de traseul catre switch-ul root.

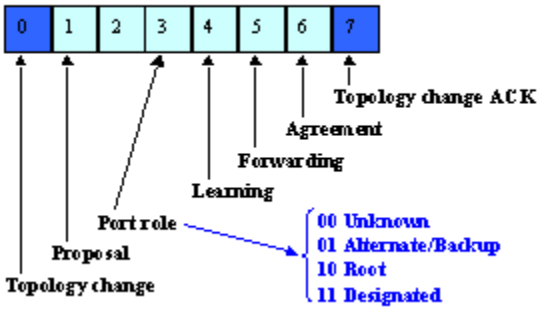
Porturile designated sunt cele de pe switch-ul root si porturile de pe celelalte switch-uri din retea care nu sunt root si in cadrul segmentului de retea au costul cel mai mic catre root.



Portul alternate primeste mai multe BPDU de la un port vecin dar acesta este in modul de blocare.

Portul de back-up primeste BPDU-uri de la portul de pe acelasi switch la care e conectat dar este tot in modul blocking . [2]

Noul format al frame-urilor de tip BPDU:



Formatul e schimbat foarte putin. Flag-urile TC si TCA sunt definite de standardul 802.1D. RSTP foloseste toti cei 6 biti din flag pentru:

- Codarea rolului si statutului in care se afla portul ce transmite un BPDUs
- Sa se ocupe de mecanismul de propunere/acceptare

In cadrul RSTP exista si conceptul de port de margine (edge port). Acest tip de port se presupune ca nu va fi niciodata conectat la alt switch in retea. El se va conecta doar la host-uri.

RSTP-ul asigura convergenta foarte rapida a retelei in cazul in care un port a cazut. Exista o diferenta intre rolul portului si starea portului. De exemplu un port in mod

designated poate sa se gaseasca in starea discarding(arunca frame-urile) temporar, chiar daca starea sa finala va fi starea forwarding.

In STP cand un port a fost desemnat port designated trebuie sa astepte de 2 ori timpul denumit “forward delay” pana sa ajunga in mod forwarding. In cazul RSTP acest timp e scurtat deoarece convergenta e refacuta pe principiul nod la nod si nu se mai bazeaza pe timerii folositi de STP. [3]

Tranzitia rapida la modul forwarding poate fi obtinuta doar de catre porturile terminale.

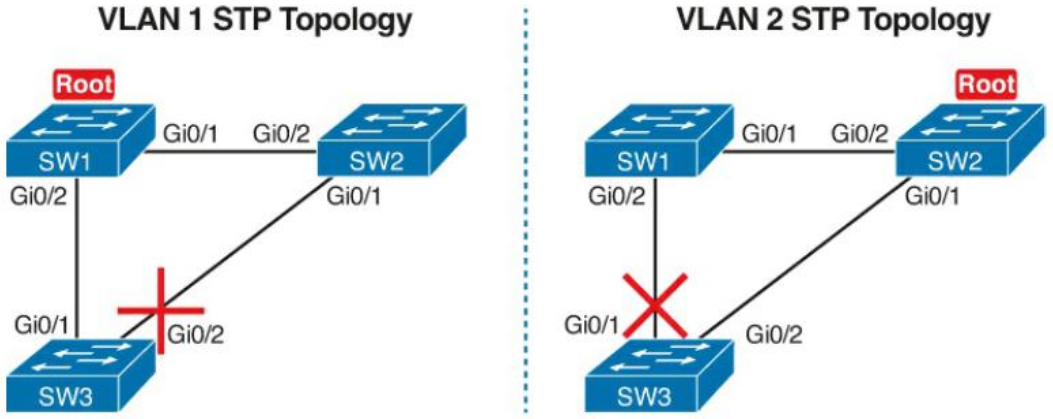
3.2. PerVLAN Spanning Tree Protocol – descriere si diferente fata de Spanning Tree Protocol

Acest tip de STP se refera la faptul ca se ruleaza o instanta a STP-ului pentru fiecare VLAN in parte. Acest lucru ajuta administratorii retelei in sensul ca ei pot sa configureze STP-ul in mod diferit pentru fiecare VLAN.

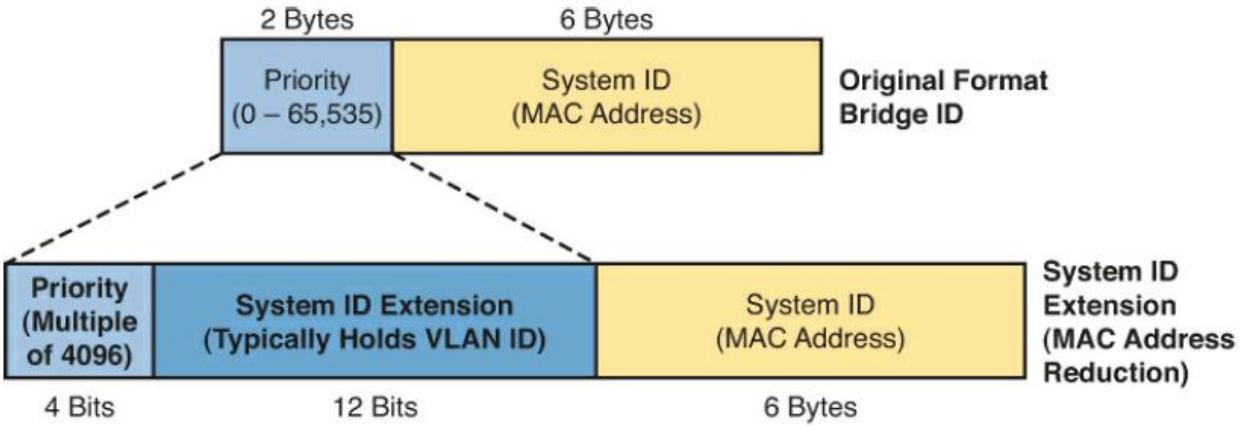
Conceptul de VLAN: Acest concept se refera la faptul ca un switch poate fi impartit din punct de vedere logic in mai multe switch-uri virtuale. Porturile nu vor mai face parte din acelasi domeniu toate. Un port sau mai multe porturi pot fi configurate pentru a face parte dintr-un VLAN astfel incat sa nu existe conectivitate intre ele. VLAN-urile se reprezinta prin numere cu valori cuprinse intre 1 si 4096.

Conceptul de VLAN este foarte utilizat deoarece imbunatateste securitate petru ca utilozatorii vor fi grupati in VLAN-uri si nu se vor mai vedea intre ei.

PVSTP ofera posibilitatea ca inginerii sa modifice configuratiile STP pentru fiecare VLAN in parte si astfel sa aleaga porturile root sau designated dupa nevoile retelei.



In figura e demonstrat cum pentru fiecare VLAN avem root-uri diferite si astfel putem sa balansam traficul asa cum avem nevoie.



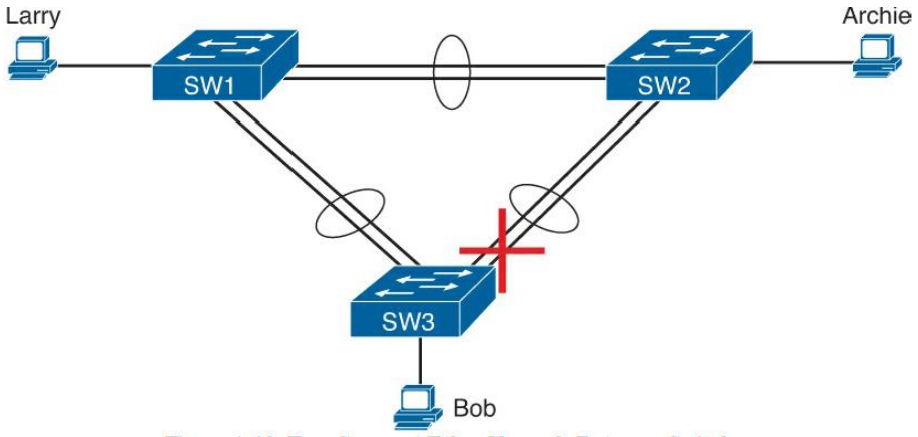
In mod original formatul bridge ID-ului era ca e format din 2 octeti de prioritate si 6 octeti cu system ID.

Pentru a putea lucra cu VLAN-uri acest lucru a fost modificat iar in campul priority acum se gaseste un camp pe 4 biti ce da prioritatea si un camp de 12 biti ce da id-ul VLAN-ului.[1]

3.3. Tehnologiile EtherChannel si PortFast – avantaje aduse

EtherChannel:
Cea mai buna metoda pentru a micsora timpul de convergenta este sa nu se permita convergenta cu totul. Tehnologia EtherChannel face ca atunci cand un singur port cade sa nu se mai incerce restabilirea convergentei.

Tehnologia functioneaza in felul urmator: mai multe porturi, pana la maxim 8 sunt configurate ca un singur link intre switch-uri. Ele trebuie sa respecte regula ca viteza lor sa fie aceeasi. In cazul in care unul din ele cade nu se reface convergenta deoarece STP-ul considera toate legaturile ca pe o singura legatura. Doar in cazul in care toate legaturile dispar se incepe recalcularea parametrilor si alegerea unei cai optime.
Figura de mai jos explica modul in care arata o topologie ce foloseste EtherChanel.



Legaturile dintre switch-uri sunt considerate ca o singura legatura. Trebuie ca ambele sa cada pentru ca STP-ul sa recalculeze caile.

Astfel reteaua este valabila altfel fata de modul clasic in care fiecare legatura era tratata de una singura.

De asemenea exista si un alt avantaj. Atunci cand switch-ul e configurat pentru EtherChannel el poate lua decizia sa trimita toate datele pe o singura legatura dintre legaturile paralele sau sa balanseze traficul pe toate legaturile paralale. Acest lucru face ca traficul sa fie impartit si trimis pe toate legaturile si nu se ocupa banda inutil.

Pentru a se configura EtherChannel trebuie ca porturile sa fie introduse in grupul corect.

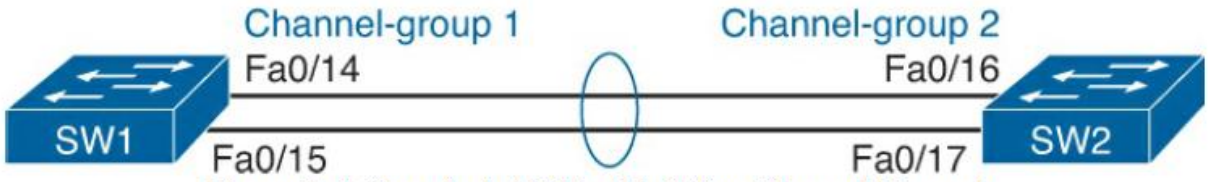


Figure 2-6. Sample LAN Used in EtherChannel Example

```
SW1# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
SW1(config)# interface fa 0/14
SW1(config-if)# channel-group 1 mode on

SW1(config)# int fa 0/15
SW1(config-if)# channel-group 1 mode on

SW1(config-if)# ^Z
```

Se observacum cele 2 porturi sunt introduse in channel-group 1.

PortFast:

Aceasta tehnologie permite unui port sa treaca direct din starea de blocare in starea forwarding sarind peste starile de listening si learning.
Pentru securitate singurele porturi pe care aceast protocol trebuie sa fie activat sunt porturile switch-ului care sunt direct legata la gazde, niciodata cele legate la alte switch-uri.
Legarea la alte dispozitive ce sunt capabile de rulare de STP duce ca folosirea protocolului PortFast sa produca bucle.

Atunci cand optiunea e activata pe legaturile cu utilizatorii, imediat dupa ce PC-ul boot-eaza si placa de retea devine activa, legatura se ridica. Nu se mai asteapta trecerea prin starile de ascultare si invatare.

STP-ul poate sa aduca cu sine o serie de vulnerabilitati cum ar fi:

- un atacator se poate lega cu un switch la retea cu o prioritate mai mica si astfel el va deveni root. Acest lucru duce la o functionare mai slaba a retelei decat inainte.
- atacatorul se poate conecta in multe porturi ale multor switch-uri si poate deveni root transmitand o parte din traficul prin retea. El poate captura date importante in cadrul retelei.
- utilizatorii pot face rau retelei daca se conecteaza cu un switch ce nu suporta STP si astfel se pot creea bucle.

Pentru a imbunatatii PortFast exista un protocol care dezactiveaza un port ce primeste BPDU-ui cand acesta nu ar trebui sa primeasca deoarece e conectat la un calculator al unui utilizator.[5]

IV. Concluzii (Sandu Dorin Traian– 441A)

Protocolul STP se bazeaza pe un algoritm in care se creeaza o harta a retelei sub forma de arbore si in care intre 2 legaturi exista doar o singura cale.
Scopul lui este de a creea o retea cu legaturi redundante in sensul ca in cazul in care se intrerup legaturi va exista totusi coectivitate.

Varianta RSTP ofera un plus de viteza cand vine vorba de timp de convergenta, timp care poate pune sub semnul intrebarii performantele retelei in cazul in care aceasta este foarte mare.

Varianta PerVLANSTP ofera posibilitatea configurarii aceluiasi protocol pentru mai multe retele lan virtuale. Acest lucru duce la balansarea traficului in mod eficeint pentru a se evita congestia.

Vulnerabilitatiile pe care le aduce STP-ul sunt combatute in oarecare masura de tehnologiile PortFast si EtherChannel care aduc un plus de viteza dar incearca sa limiteze posibilitatile prin care cineva incearca sa se conecteze la retea si sa strice configuratia acesteia.

V. Bibliografie

[1] – Cisco CCNA Routing and Switching ICND2 200-101

[2] – Cisco CCNA Routing and Switchinc ICND1 100-101

[3]- http://www.cisco.com/en/US/tech/

[4] – www.netacad.com

[5]- http://en.wikipedia.org/wiki/Spanning_Tree_Protocol