

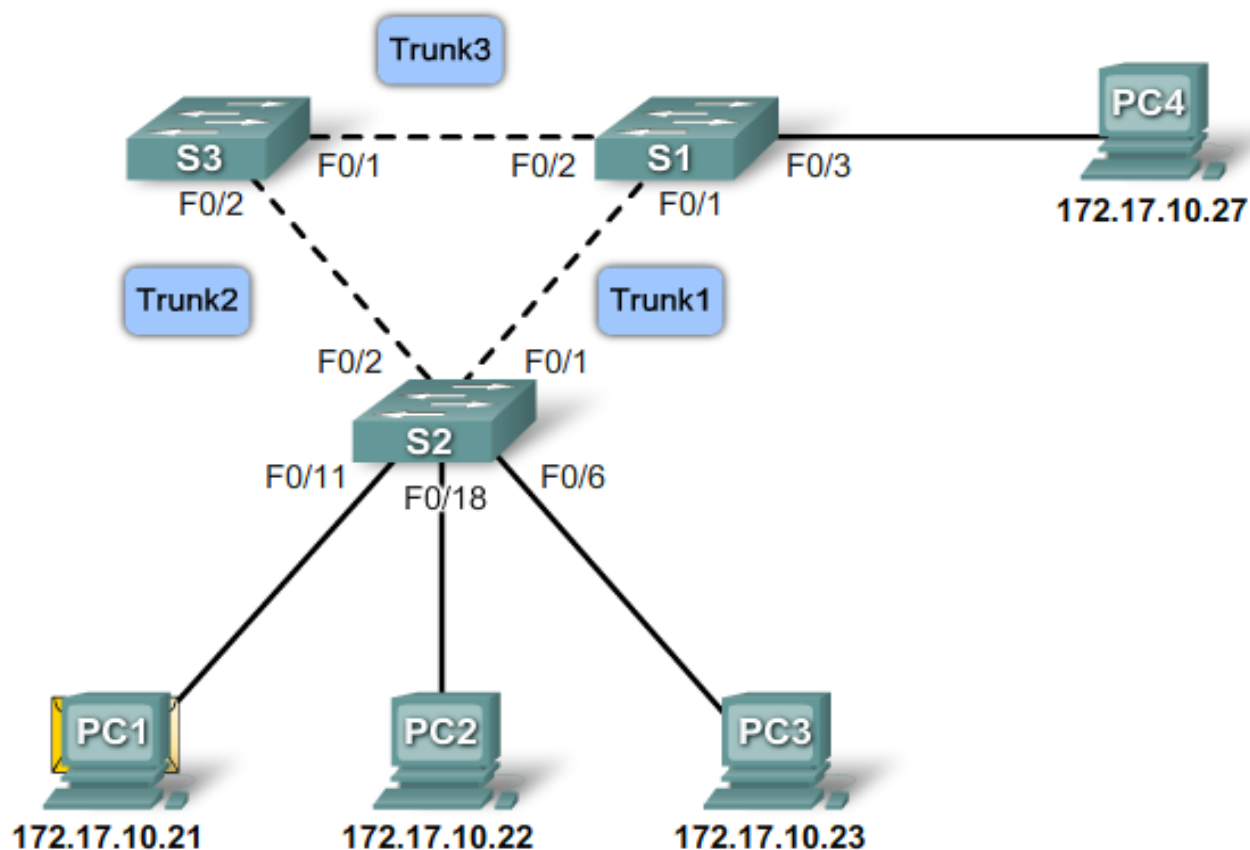
Des protocoles de Spanning Tree

Des protocoles de Spanning Tree

La redondance ?

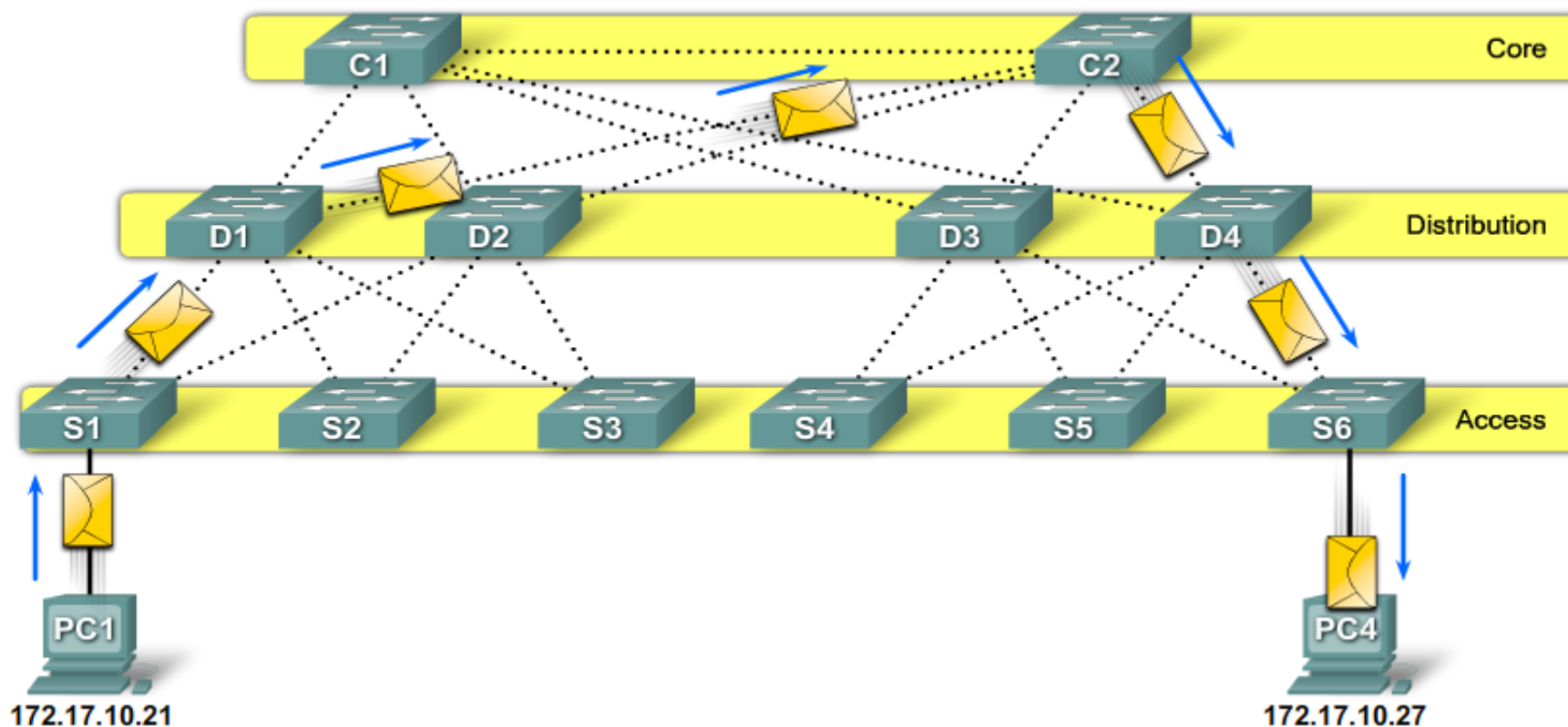
La redondance sur les liens

- ▶ Permet d'avoir plusieurs chemins en cas d'une défaillance d'un lien
- ▶ Problème : un broadcast provenant de PC1 ne pourra jamais être arrêté !



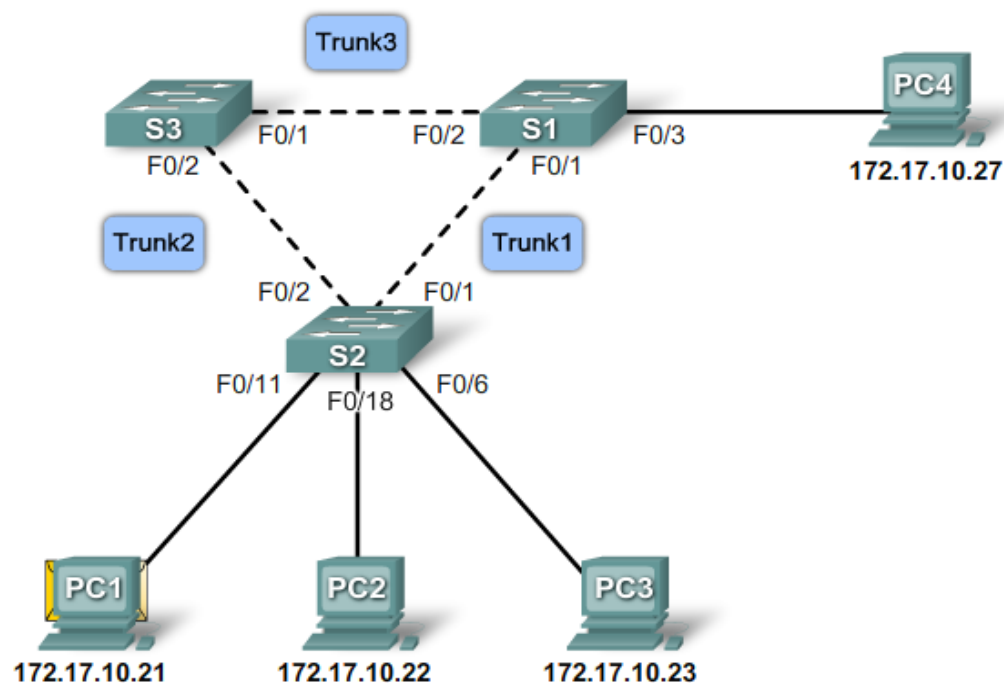
Autre type de redondance

- ▶ Permet de supporter la perte d'un lien ou d'un équipement
- ▶ Toujours le problème du broadcast



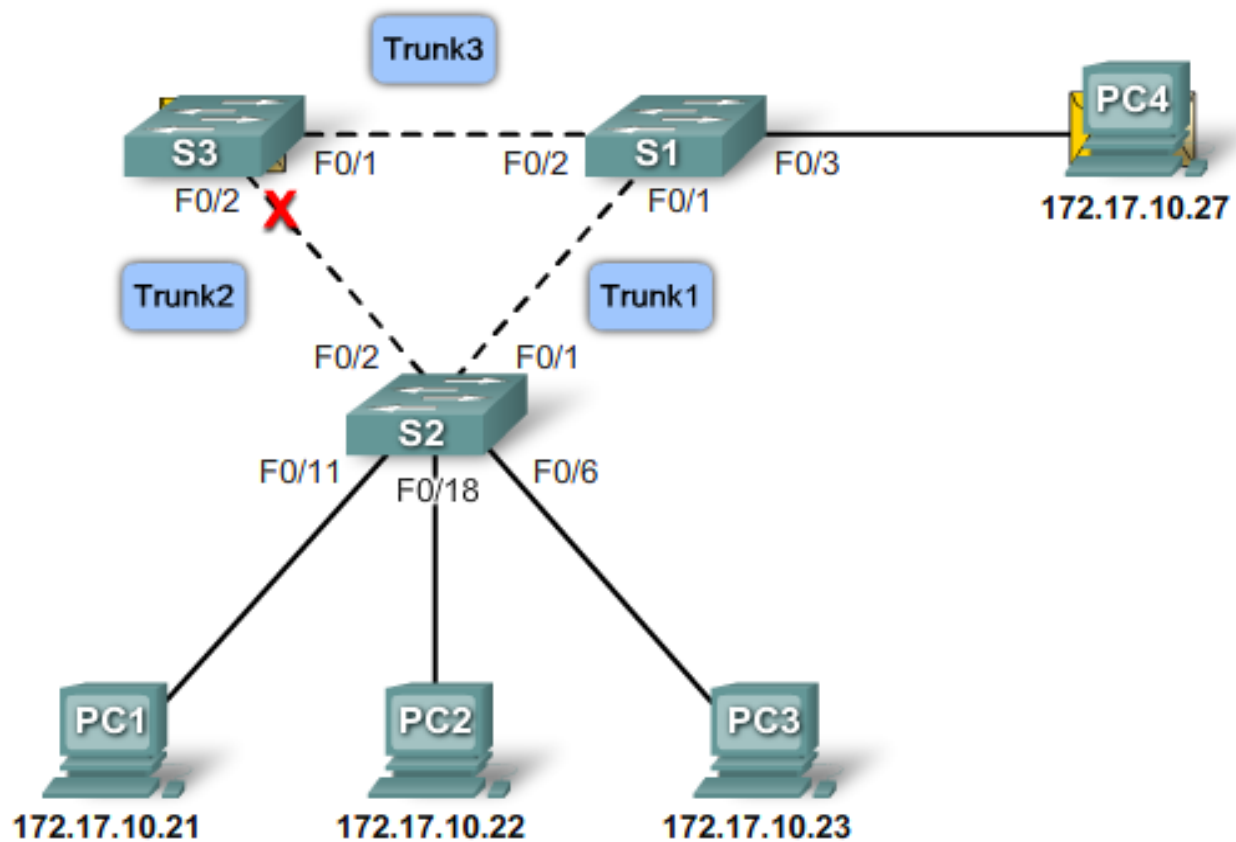
Les problèmes

- ▶ Le broadcast
 - ▶ Généré par exemple par les requêtes ARP
 - ▶ Comme il n'existe aucun moyen de supprimer les trames Ethernet, comme sur IP avec le TTL, cela encombre les liens et génère des broadcast storm
- ▶ Duplication de trames unicast
 - ▶ A cause des boucles, un PC peut recevoir plusieurs fois une même trame !



Une solution : Spanning Tree Protocol

- ▶ Les ports qui génèrent des boucles sont automatiquement « désactivés »



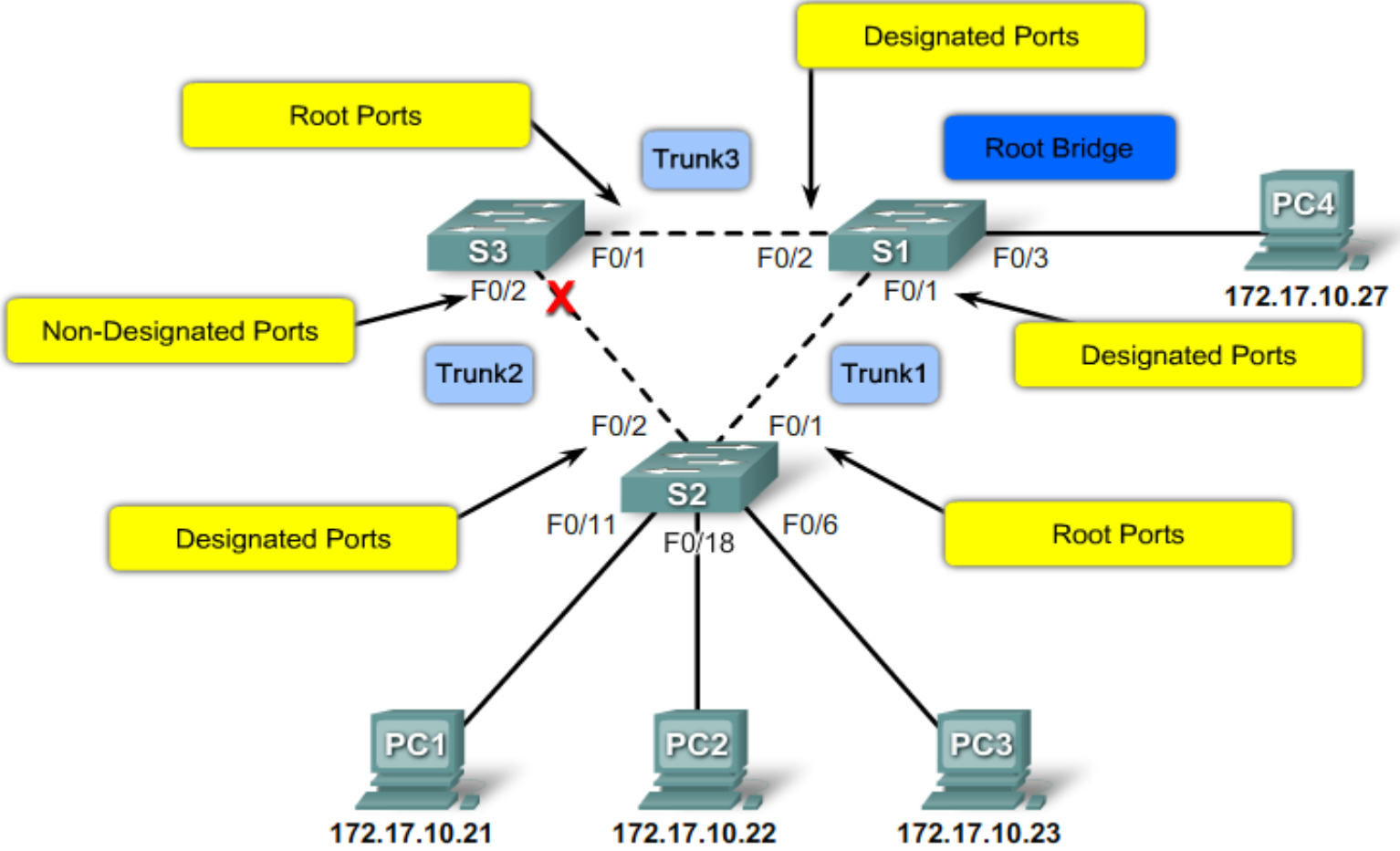
Des protocoles de Spanning Tree

Le plus simple : STP ou IEEE 802.1D

Le Spanning Tree Algorithm 802.1D

- ▶ Le switch qui possède la plus petite Bridged Identity (BID) est élu Root
 - ▶ Le BID est la concaténation
 - ▶ d'une priorité comprise entre 1 et 65536, par défaut 32768
 - ▶ de l'adresse MAC du switch
- ▶ Toutes les 2 secondes, chaque switch broadcast son BID ainsi que le BID du Root Bridge dont il a la connaissance
- ▶ Les ports des switchs sont appelés
 - ▶ Root Ports s'ils sont connectés au Root Bridge (ou Root Switch)
 - ▶ Designated Ports s'ils ne sont pas Root Port et s'ils autorisent le trafic à circuler
 - ▶ Non-designated Ports s'ils bloquent le trafic
- ▶ Pendant l'exécution de cet algorithme, les messages échangés entre les switchs sont appelés des BPDU

Exemple pour le STP

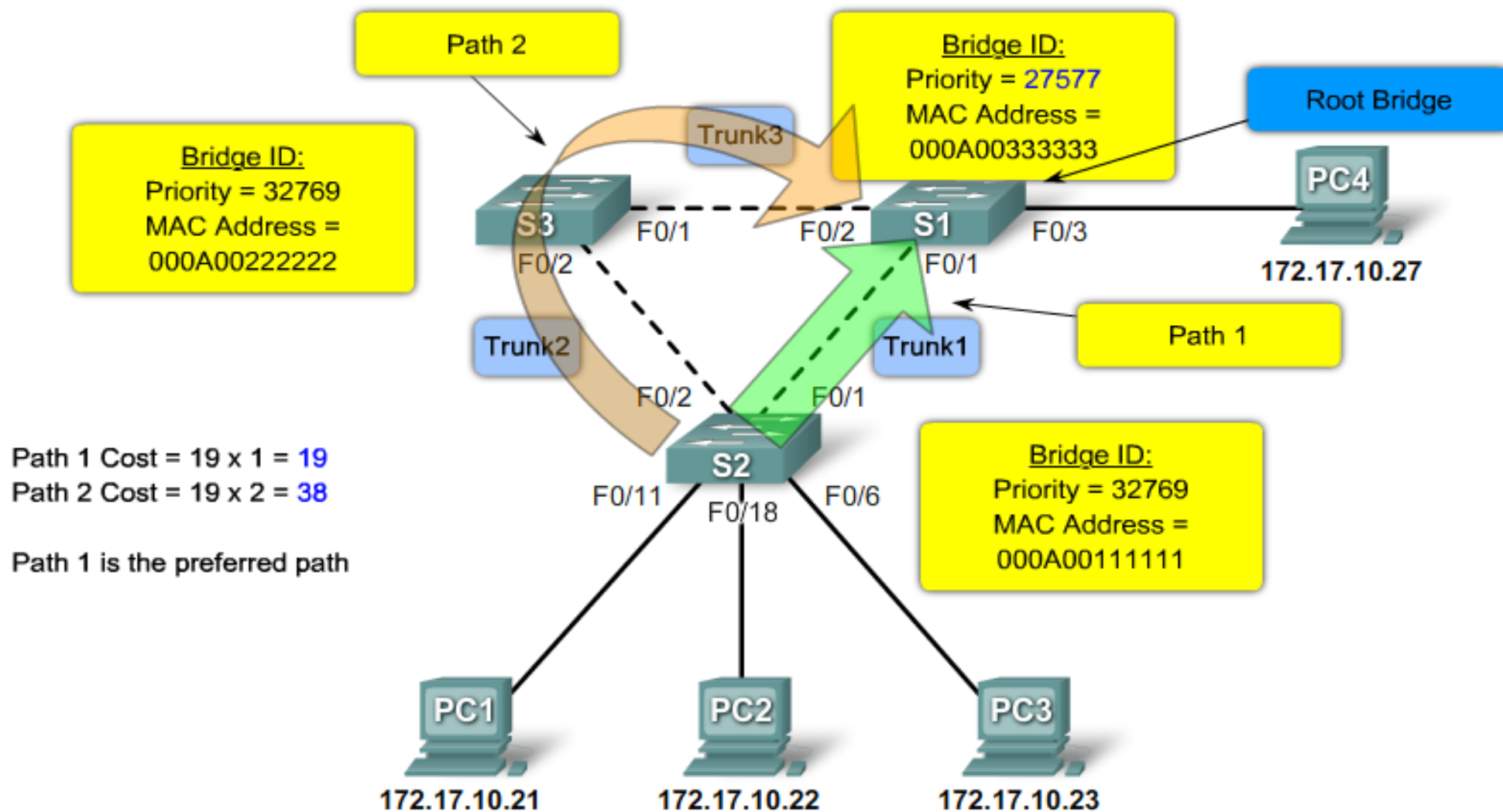


Sélection des meilleurs chemins

- ▶ Sur chaque port du switch, le coût de chaque chemin vers la racine va être évalué

Link Speed	Cost (Revised IEEE Specification)	Cost (Previous IEEE Specification)
10 Gb/s	2	1
1 Gb/s	4	1
100 Mb/s	19	10
10 Mb/s	100	100

Calcul des couts des chemins



En cas d'égalité ?

- ▶ En cas de liens parallèles entre 2 switchs ?
- ▶ Quel est le chemin le plus court ?
- ▶ En cas d'égalité (coût du chemin et BID), le port ID est utilisé. Le port ID le plus petit « forward » les trames et les autres ports les bloqueront

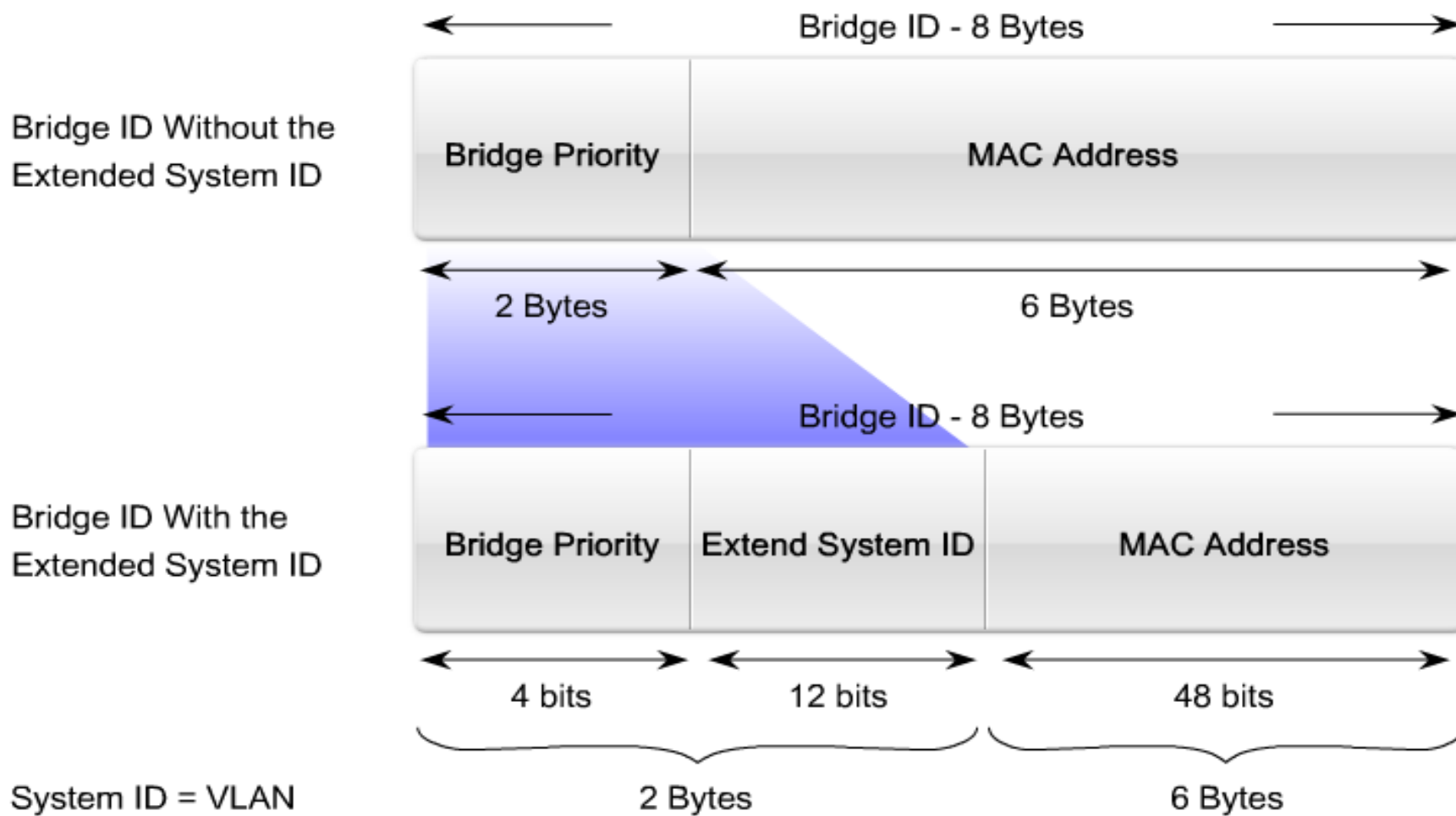
Configuration

- ▶ Changer le cout d'un port
 - ▶ **Switch(config-if)#spanning-tree cost** *number*
- ▶ Revenir à cout par défaut
 - ▶ *Switch(config-if)#no spanning-tree cost*
- ▶ Vérifier l'état du spanning-tree sur un switch
 - ▶ Switch#show spanning-tree
 - ▶ Permet de connaître le Root ID, le BID du switch et le cout de chaque port vers le Root Bridge

Les instances STP

- ▶ Une instance de spanning tree est définie de la façon suivante :
 - ▶ Quand tous les switches participant aux échanges de BPDU et à la négociation du spanning tree sont associés à une unique racine
- ▶ Si nous avons qu'une instance pour tous les VLAN, on parle de **Common Spanning Tree (CST)**, sinon, quand c'est effectué par VLAN, on parle de **Per VLAN Spanning Tree (PVST)**
 - ▶ Avec **PVST**, nous avons donc **une racine par VLAN**
- ▶ Par VLAN, cela nécessite d'avoir l'information de l'identifiant du VLAN concerné (le VID) dans la trame !
 - ▶ Le BID est donc étendu pour pouvoir contenir le VID

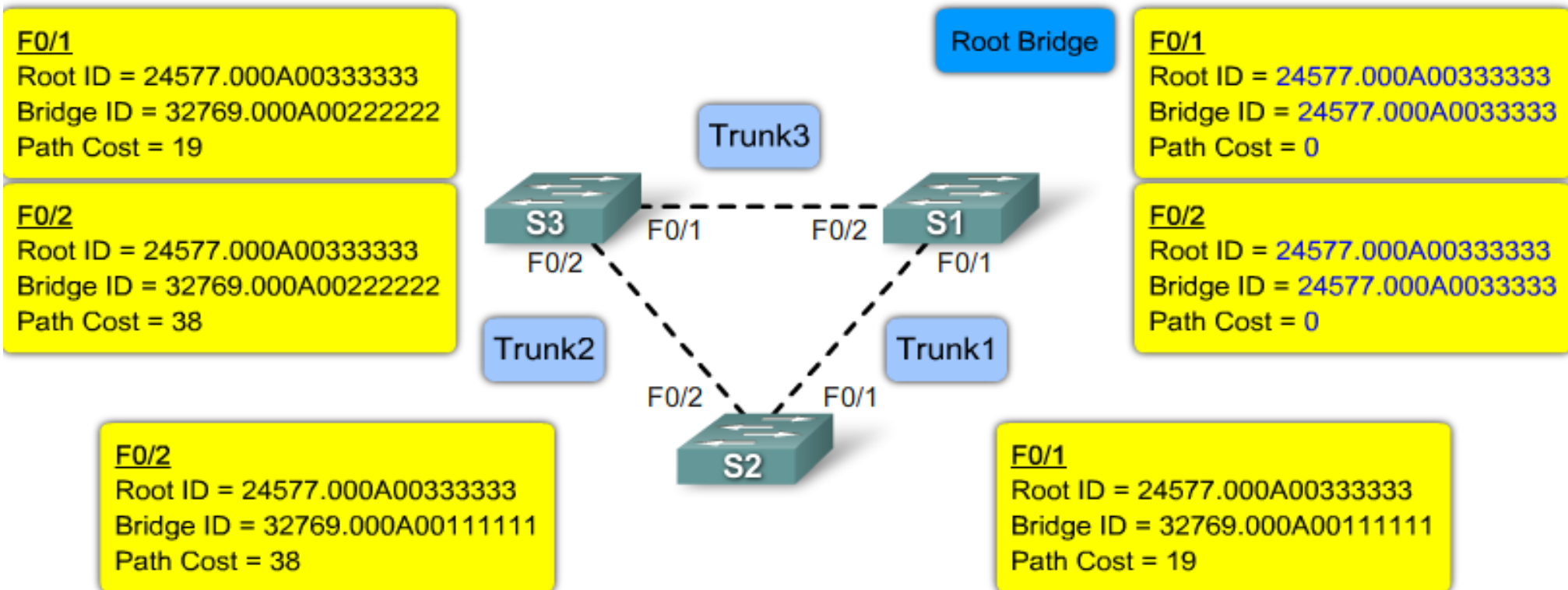
Le BID



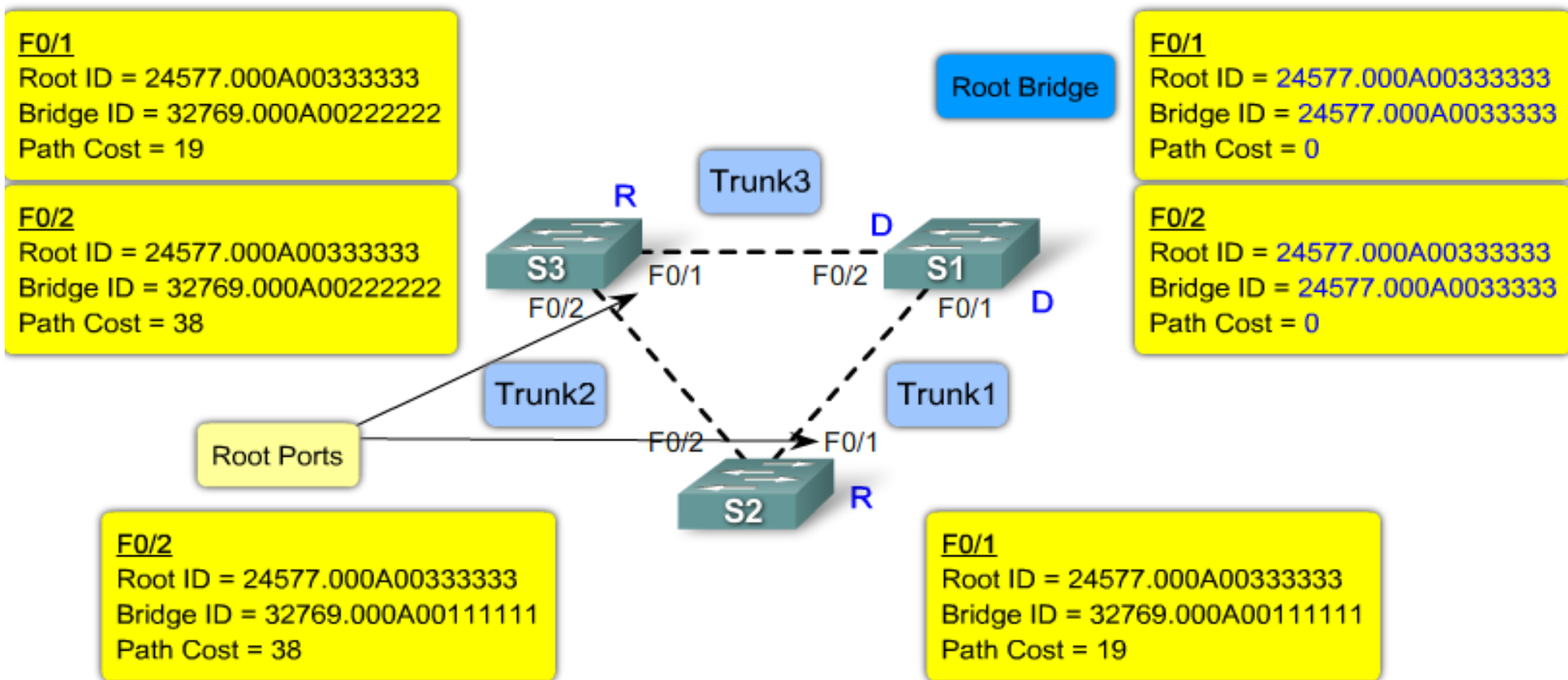
Le BID

- ▶ Avant l'apparition des VLAN, la priorité était codé sur 16 bits
- ▶ Depuis l'existence des VLAN, la priorité est codé sur 4 bits auquel on ajoute 12 bits pour l'identifiant du VLAN sur lequel se construit le Spanning Tree
 - ▶ La priorité ne peut donc prendre que des valeurs multiples de 4096 (2^{12})
- ▶ Pour configurer cette priorité 2 solutions
 - ▶ **Switch(config)#spanning-tree vlan 1 root primary**
 - ▶ Fixe la priorité à la valeur de 24576
 - ▶ puis éventuellement
 - ▶ **Switch(config)#spanning-tree vlan 1 root secondary**
 - ▶ Fixe la priorité à la valeur de 28672
 - ▶ **Switch(config)#spanning-tree vlan 1 priority *number***

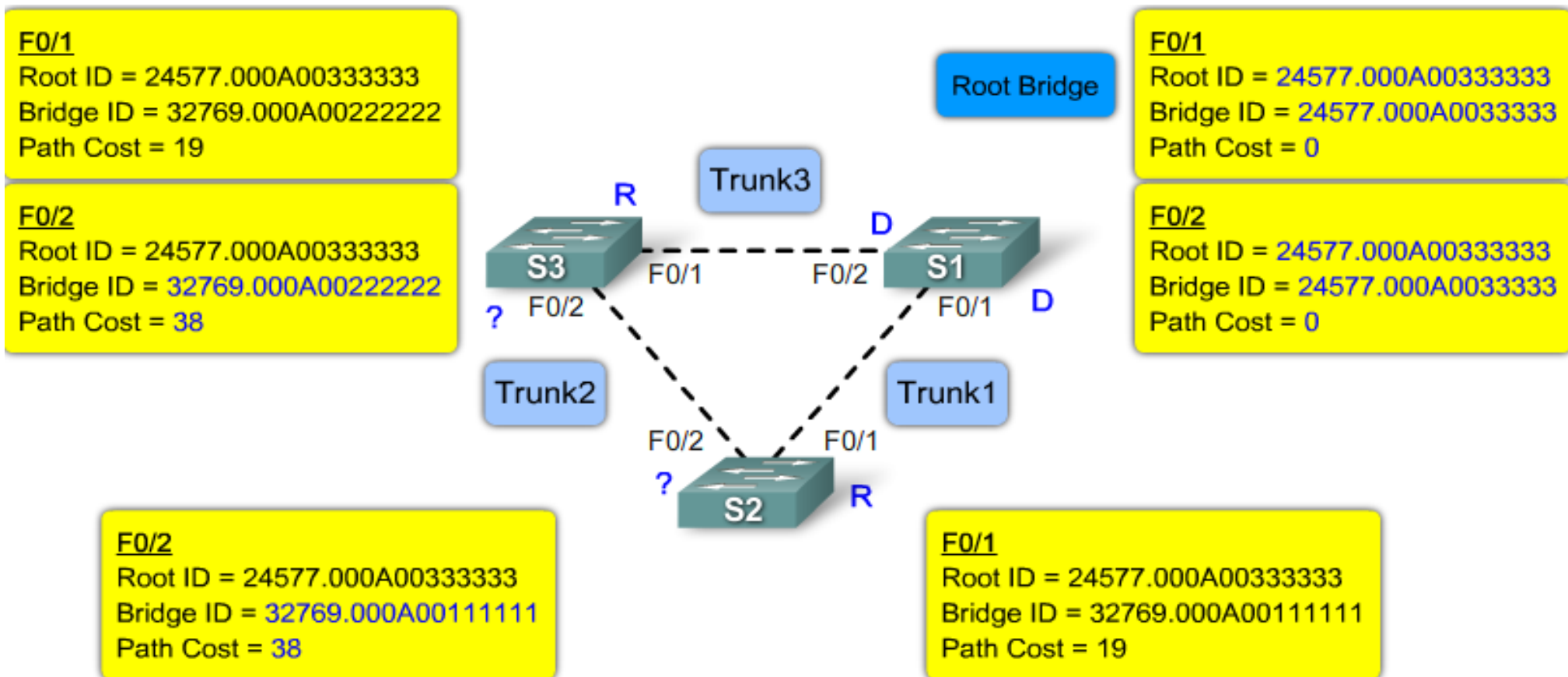
Exemple



Exemple

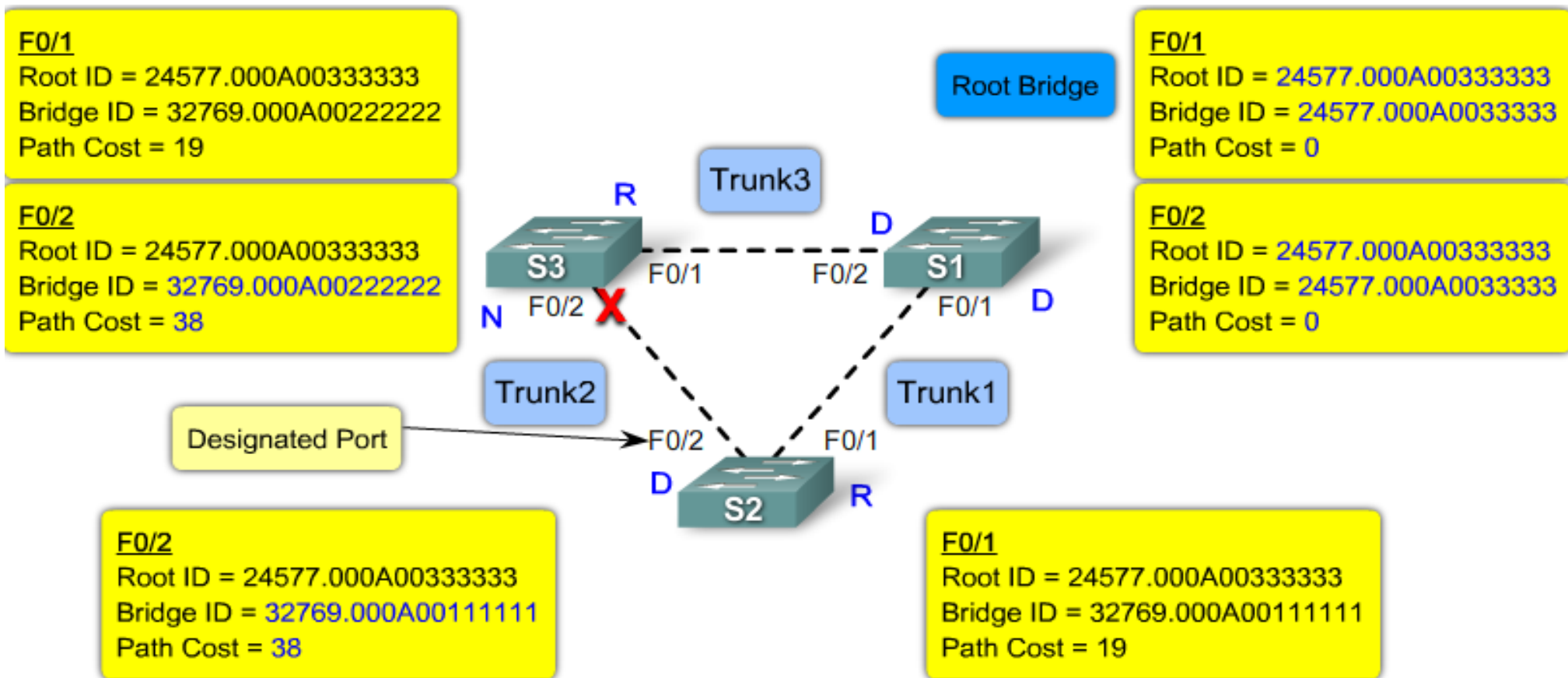


Exemple



En cas de cout égaux vers le root, le plus petit BID gagne.
F0/2 de S2 sera donc designated et F0/2 de S3 Non designated

Exemple



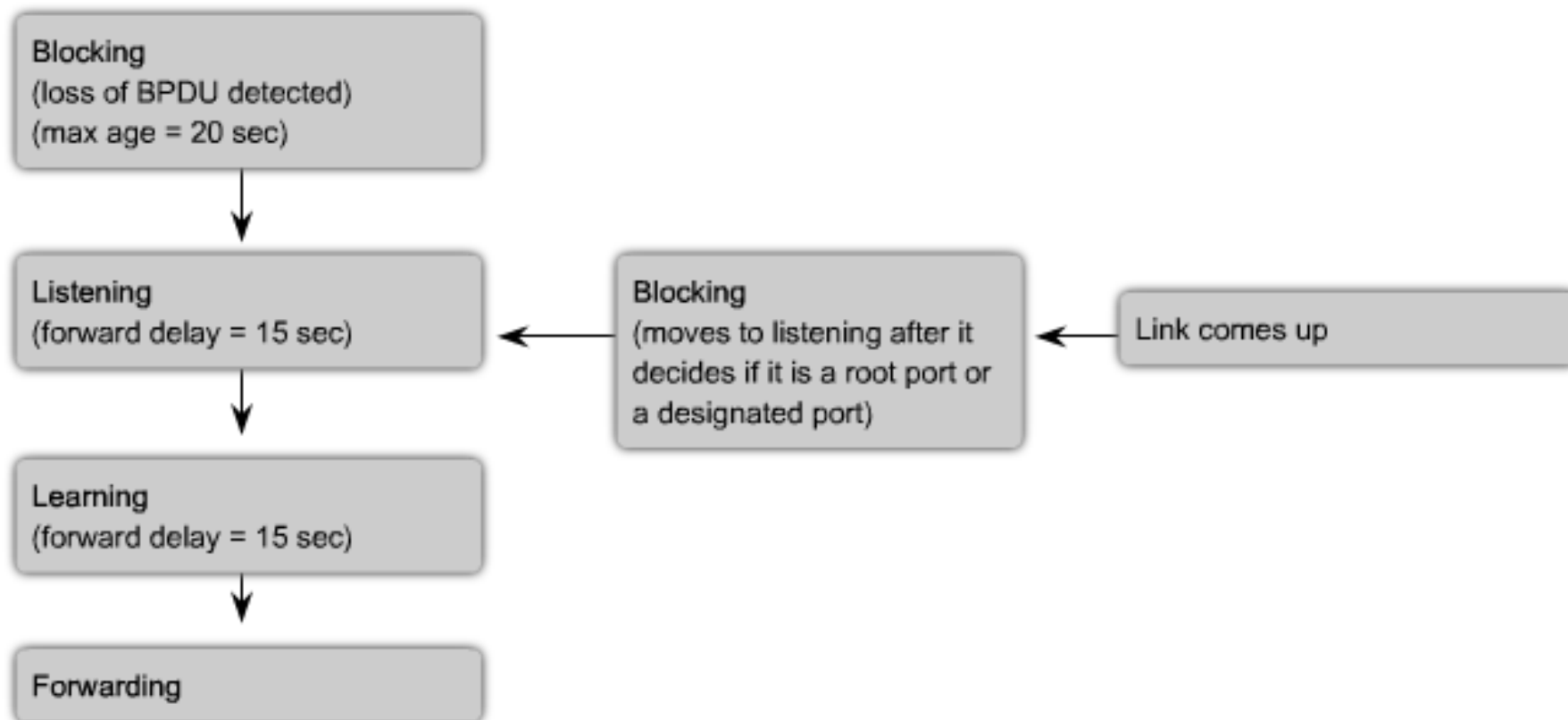
L'état des ports sur STP

- ▶ Un port de switch peut prendre 5 états différents pendant le processus de fabrication de l'arbre convrant via le protocole STP
 - ▶ Blocking : port non-designated. Ne fait juste qu'envoyer des BPDU et les analyser
 - ▶ Reste dans cet état pendant 20 secondes max. (maximum age)
 - ▶ Listening : reçoit les BPDU et transmet les siens également
 - ▶ Reste dans cet état pendant 15 secondes max. (forward delay)
 - ▶ Learning : le port ne forward toujours pas les trames mais apprend les adresses MAC sources contenues dans celles-ci
 - ▶ Reste dans cet état pendant 15 secondes max. (forward delay)
 - ▶ Forwarding : toutes les trames sont transmises
 - ▶ Disable : port administrativement désactivé (shutdown)
- ▶ Chaque switch envoie toutes les 2 secondes ses BPDU
 - ▶ Configurable entre 1 et 10 secondes
- ▶ Toutes les valeurs de timer ont été optimisées pour un réseau de diamètre 7
- ▶ Existence de 2 types de BPDU : configuration ou Topologie Change Notification (TCN)

L'état des ports sur STP

Processes	Blocking	Listening	Learning	Forwarding	Disable
Receives and process BPDUs	YES	YES	YES	YES	NO
Forward data frames received on interface	NO	NO	NO	YES	NO
Forward data frames switched from another interface	NO	NO	NO	YES	NO
Learn MAC addresses	NO	NO	YES	YES	NO

Temps dans chaque état avec STP



Des protocoles de Spanning Tree

Les modifications topologiques et les TCN

Topology Change Notification

- ▶ Un changement topologique est
 - ▶ Soit un port en forwarding qui passe down
 - ▶ Soit un port qui passe en forward
- ▶ Le switch qui constate la modification envoie un TCN BPDU au Root Bridge
- ▶ Le Root Bridge répond par un TCN BPDU portant le bit TCAck à 1
- ▶ En connaissance d'un changement topologique, le Root Bridge diffuse ses BPDU en mettant le bit TC (Topology Change) à 1.
 - ▶ Il laisse ce bit TC à 1 pendant $\text{Max Age} + \text{Forwarding Age}$, soit $20+15=35$ s par défaut

Le Bridge Protocol Data Unit (BPDU)

Bytes	Field
2	Protocol ID
1	Version
1	Message type
1	Flags
8	Root ID
4	Cost of path
8	Bridge ID
2	Port ID
2	Message age
2	Max age
2	Hello time
2	Forward delay

- ▶ Message Type : Configuration ou TCN
- ▶ Root ID : le BID le plus petit
- ▶ Cost of Path : coût de tous les liens du switch expéditeur vers le root bridge
- ▶ BID : BID de l'expéditeur
- ▶ Port ID : ID du port qui expédie le BPDU
- ▶ STP timers values : Max age, hello time et forward delay
- ▶ Flags : dépend du protocole de spanning tree (au moins TC et TCAck bits)

Des protocoles de Spanning Tree

Améliorations et variantes de STP

Le Cisco PortFast

- ▶ C'est une solution propriétaire Cisco qui fait passer un port instantanément de Blocking à Forwarding
- ▶ A faire sur les switches d'access layer sur lesquels sont connectés directement des workstations ou serveur
 - ▶ Ceci est à faire sur un réseau avec support du DHCP. Sans cela, une station va envoyer sa requête et en raison du temps de convergence du STP, n'obtiendra pas d'adresse IP
- ▶ Pour l'activer, sur l'interface correspondante :
 - ▶ **spanning-tree portfast**
- ▶ Pour l'activer sur toutes les interfaces non trunk
 - ▶ **spanning-tree portfast default**
- ▶ Vérification par
 - ▶ **show running-config interface fastethernet ...**

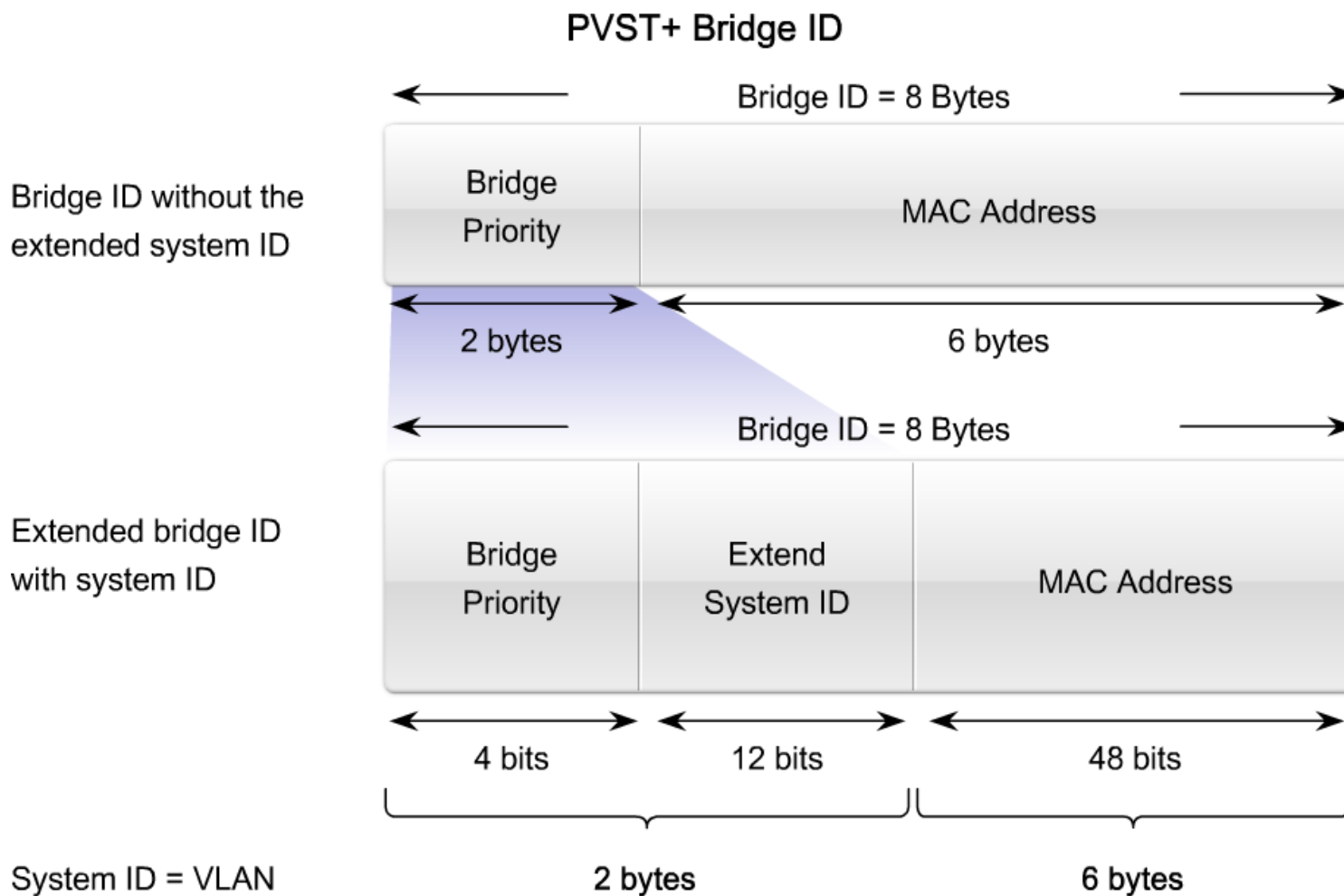
Les variantes

Cisco Proprietary	PVST <ul style="list-style-type: none"> • Uses the Cisco proprietary ISL trunking protocol • Each VLAN has an instance of spanning tree • Ability to load balance traffic at layer-2 • Includes extensions BackboneFast, UplinkFast, and PortFast
	PVST+ <ul style="list-style-type: none"> • Supports ISL and IEEE 802.1Q trunking • Supports Cisco proprietary STP extensions • Adds BPDU guard and Root guard enhancements
	rapid-PVST+ <ul style="list-style-type: none"> • Based on IEEE802.1w standard • Has faster convergence than 802.1D
IEEE Standard	RSTP <ul style="list-style-type: none"> • Introduced in 1982 provides faster convergence than 802.1D • Implements generic versions of the Cisco proprietary STP extensions • IEEE has incorporated RSTP into 802.1D, identifying the specification as IEEE 802.1D-2004
	MSTP <ul style="list-style-type: none"> • Multiple VLANs can be mapped to the same spanning-tree instance • Inspired by the Cisco Multiple Instances Spanning Tree Protocol (MISTP), • IEEE 802.1Q-2003 now includes MSTP

PVST+

- ▶ Protocole propriétaire Cisco
- ▶ Il est possible d'avoir un process STP pour chaque VLAN
 - ▶ Un port peut donc être en forwarding pour un VLAN et blocking pour un autre
 - ▶ Un switch peut même être root pour un VLAN et un autre switch pour un autre VLAN
- ▶ L'identification des VLAN se fait grâce une modification de l'interprétation du champ Bridge ID

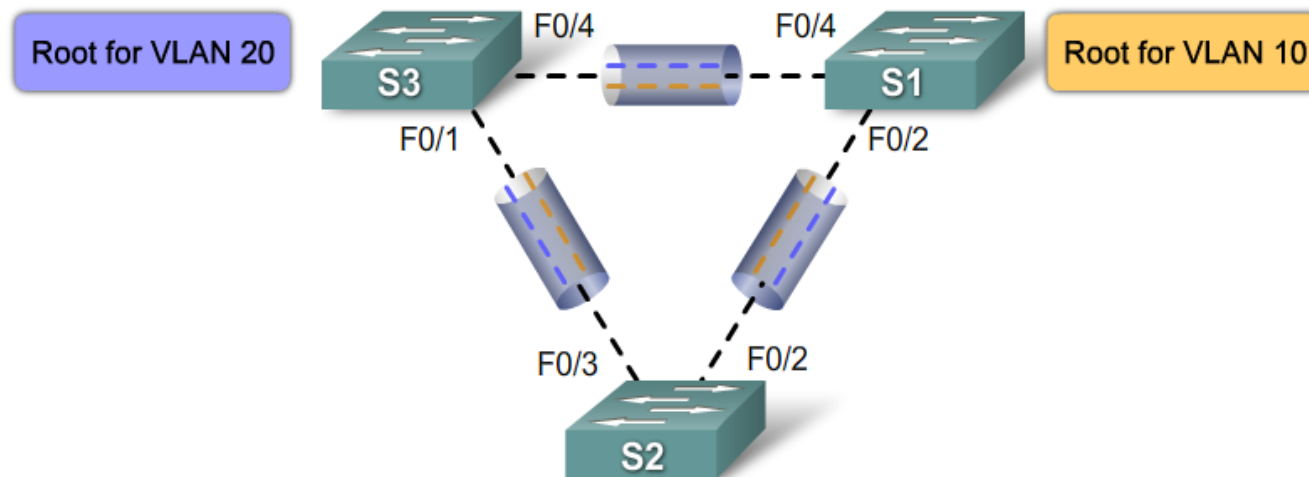
Le Bridge ID sur PVST+



Exemple sur PVST+

PVST+ Bridge ID

Priority	+	VLAN ID + MAC Address	=	BID
32768	+	10 + 000A00333333	=	32778.000A00333333
32768	+	20 + 000A00333333	=	32788.000A00333333



VLAN 10 - - - - -

VLAN 20 - - - - -

Configuration par défaut

- ▶ Par défaut, sur un switch Cisco Catalyst 2960

Feature	Default Setting
Enable state	Enabled on VLAN 1
Spanning-tree mode	PVST+ (Rapid PVST+ and MSTP are disabled.)
Switch priority	32768
Spanning-tree port priority (configurable on a per-interface basis)	128
Spanning-tree port cost (configurable on a per-interface basis)	1000 Mb/s: 4, 100 Mb/s: 19, 10 Mb/s: 100
Spanning-tree VLAN port priority (configurable on a per-VLAN basis)	128
Spanning-tree VLAN port cost (configurable on a per-VLAN basis)	1000 Mb/s: 4, 100 Mb/s: 19, 10 Mb/s: 100
Spanning-tree timers	Hello time: 2 seconds Forward-delay time: 15 seconds Maximum-aging time: 20 seconds Transmit hold count: 6 BPDUs

RSTP

- ▶ Rapid STP, normalisé par IEEE sous le numéro 802.1w
- ▶ Évolution du 802.1D
- ▶ Convergence plus rapide
- ▶ Notion de Edge port qui reprend la notion de PortFast de Cisco
 - ▶ Utilisé uniquement sur des ports qui ne seront jamais en liaison avec d'autres switchs
 - ▶ Port qui passe immédiatement à l'état forwarding
 - ▶ **Ne génère pas de TCN**, comme les PortFast

Les états des ports

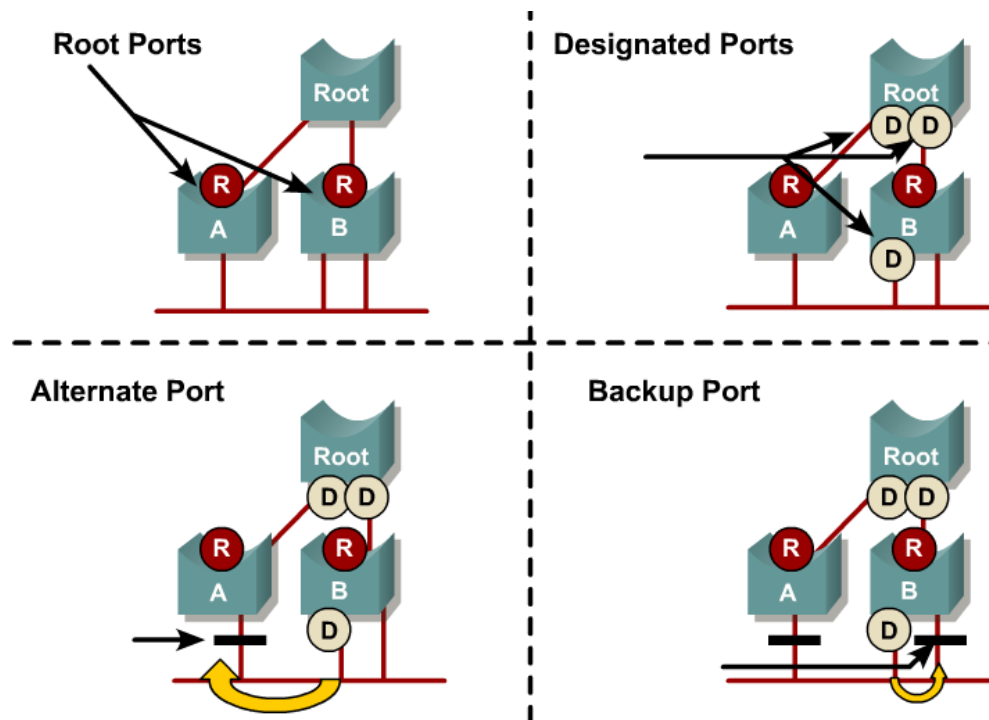
- ▶ 3 états uniquement pour les ports dans RSTP
 - ▶ **Discarding**
 - ▶ Aucune communication de paquet
 - ▶ **Learning**
 - ▶ Apprentissage des adresses MAC provenant des trames reçues
 - ▶ **Forwarding**
 - ▶ Toutes les trames sont transmises

Comparaison STP / RSTP

Operational Port State	STP Port State	RSTP Port State
Enabled	Blocking	Discarding
Enabled	Listening	Discarding
Enabled	Learning	Learning
Enabled	Forwarding	Forwarding
Disabled	Disabled	Discarding

RSTP port Roles

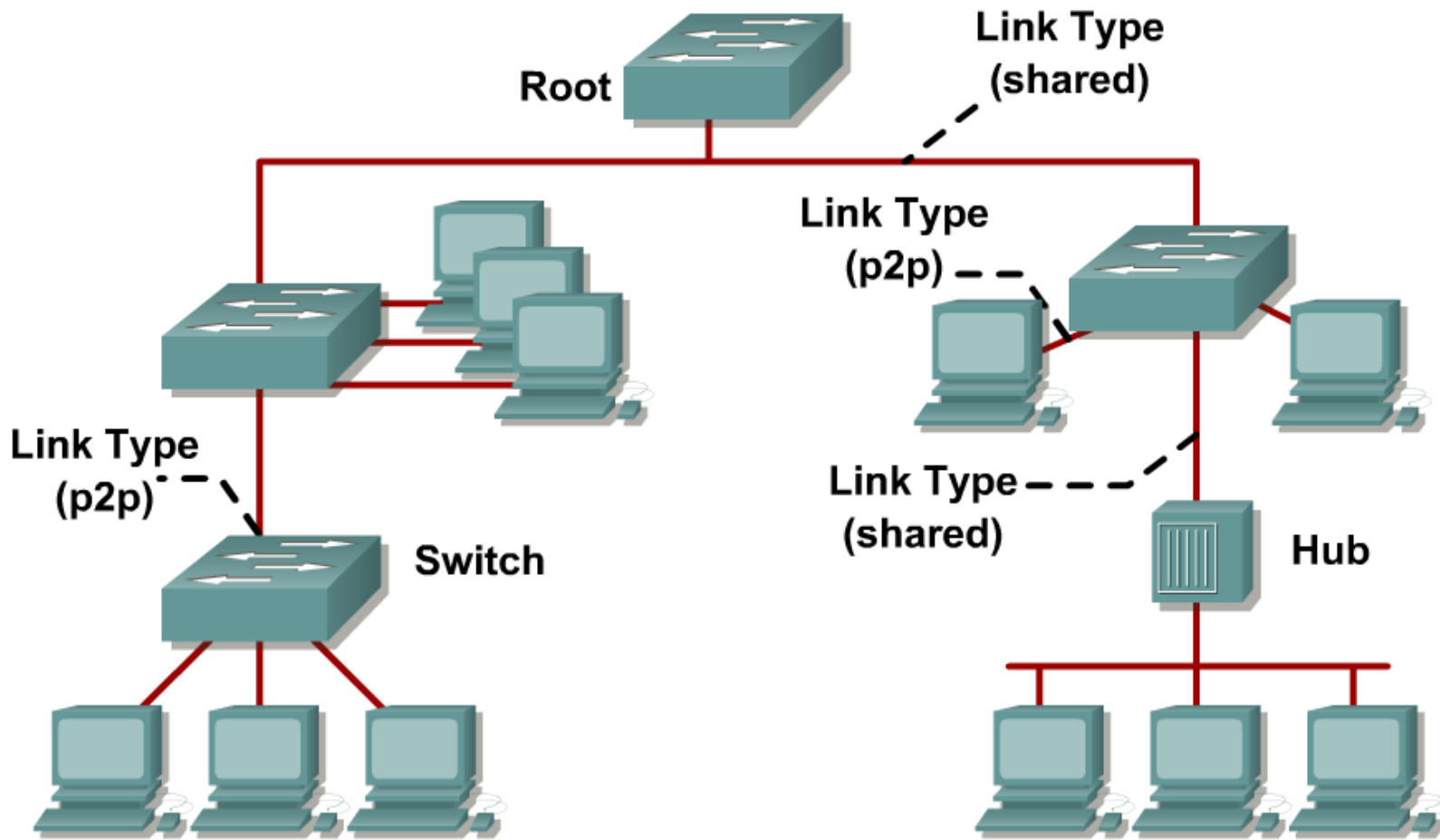
- ▶ **Alternate Port** : Port en discarding pendant la configuration stable et active. Permet d'obtenir un autre chemin sur un switch non designated si le chemin du designated switch est en échec
- ▶ **Backup Port** : Port en discarding pendant la configuration stable et active. Permet d'obtenir un autre port vers un même segment sur un designated switch. Port avec un ID plus élevé que le designated port du switch



Edge port et type de lien

- ▶ Edge ports : ports qui ne seront jamais connectés à un autre équipement actif. Ce sont des ports qui seront automatiquement dans l'état forwarding
 - ▶ Contrairement au PortFast, si un edge port reçoit un BPDU, il redevient automatiquement un port normal de spanning tree
- ▶ Type de lien : distinction entre les liens p2p et shared
 - ▶ Shared link : fonctionne en half-duplex, suppose que le média est partagé par plusieurs switches
 - ▶ Point-to-point (p2p) link : fonctionne en full-duplex : le port est supposé être connecté à un unique switch, à l'autre extrémité du lien
- ▶ Le type de lien est utilisé pour déterminer si un port peut ou pas passer immédiatement en forwarding port
 - ▶ Edge et le p2p passent instantanément Forwarding

Type de lien



Utilisation du type de lien

- ▶ Root ports ne s'en servent pas
- ▶ Alternate et Backup, dans la majorité des cas, non plus
- ▶ Designated port utilise le plus souvent ce paramètre : type de lien
- ▶ Transition rapide vers le forwarding state, pour un designated port, si le type est point-to-point

RSTP BPDU

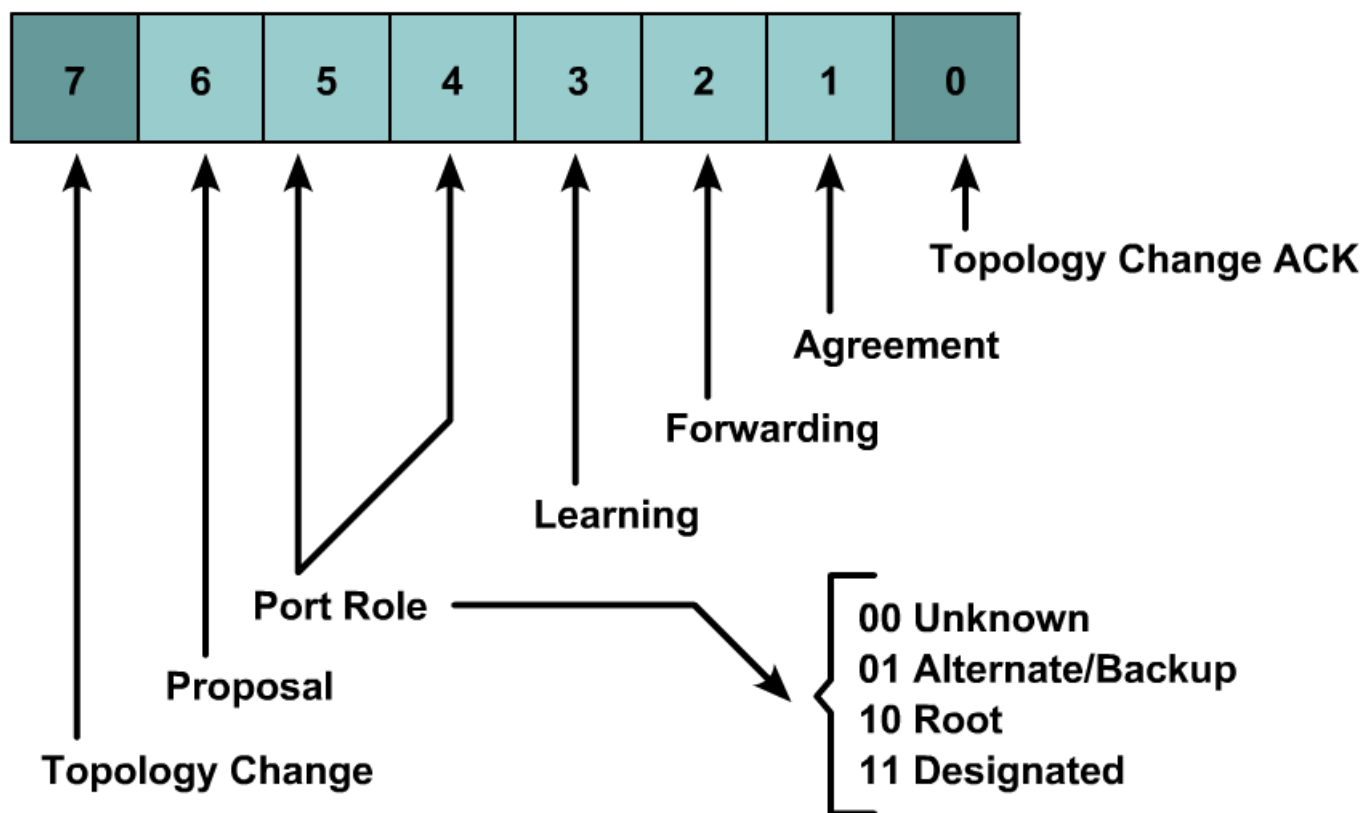
- ▶ RSTP (802.1w) utilise des BPDU de type 2, version 2
- ▶ Chaque switch BPDU envoie des BPDU à chaque intervalle hello (2 secondes par défaut). Même s'il n'a rien reçu du root bridge
 - ▶ Dans STP (802.1D), le BPDU n'est envoyé que par le root bridge
 - ▶ Dans RSTP (802.1w), chaque switch envoie des BPDU
- ▶ Les informations du protocole sont périmées si au bout de 3 hello times ou du age time, rien n'est reçu
 - ▶ Cela permet de faire une détection rapide de panne
 - ▶ On considère, qu'au bout de 3 hellos perdus, le root bridge, les switchs voisins et le designated bridge ne peuvent plus communiquer entre-eux

Rappel du Format d'un BPDU

Bytes	Field
2	Protocol ID
1	Version
1	Message type
1	Flags
8	Root ID
4	Cost of path
8	Bridge ID
2	Port ID
2	Message age
2	Max age
2	Hello time
2	Forward delay

- ▶ Flags
 - ▶ Seul 2 bits sont utilisés
 - ▶ Le 1er pour TC
 - ▶ Le 8ème pour TCAck
 - ▶ Les autres inutilisés

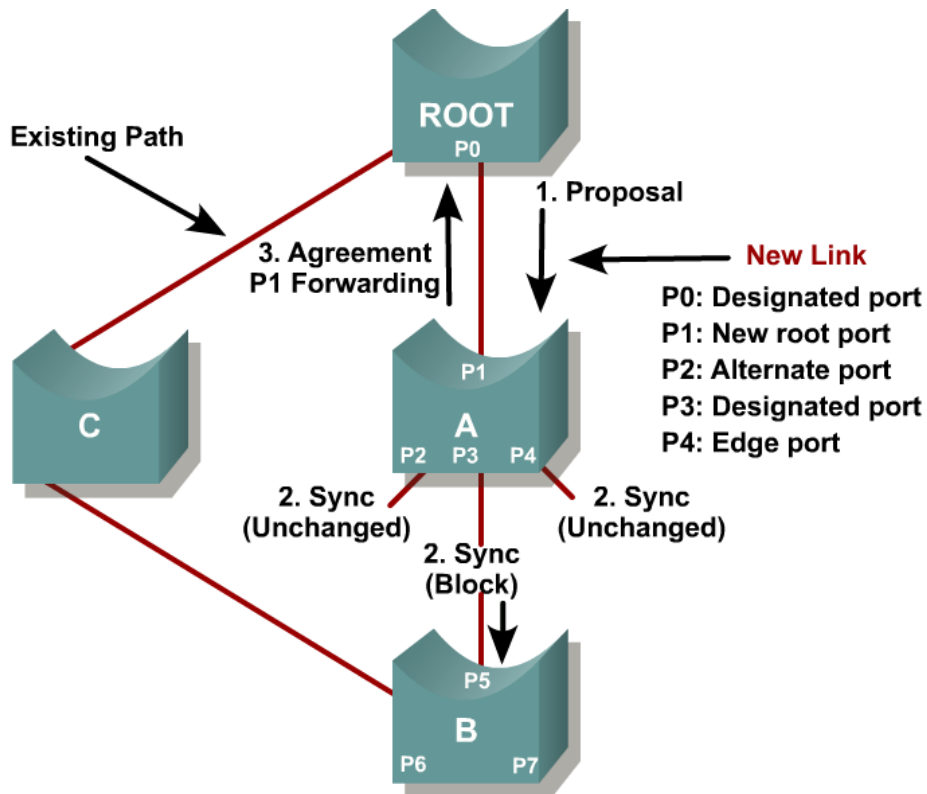
Flags d'un BPDU pour RSTP



Fonctionnement

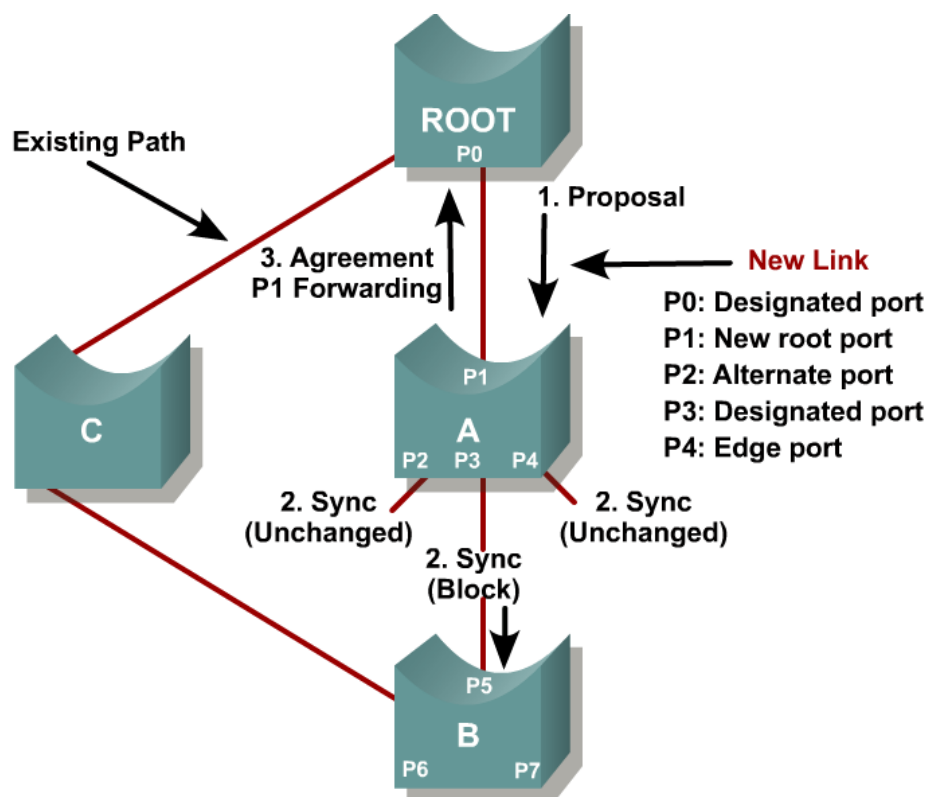
- ▶ Avec STP (802.1D), quand un port devient « designated port », il faut attendre 2 forward time pour être en forwarding port
- ▶ Avec RSTP, la transition ne se fait pas à expiration d'un compteur
- ▶ La transition rapide vers le forwarding port ne se fait que sur des Edge ports et sur des point-to-point links
- ▶ Quand un switch reçoit un BPDU supérieur d'un voisin, (donc avec une priorité plus petite, ou un chemin moins coûteux) (dit supérieur car les informations contenues donnent plus d'importance à ce BPDU), ce port passe en root port
 - ▶ Avec RSTP, les non-edge port passent en proposal-agreement process
 - ▶ Uniquement les non edge en discarding ou learning

RSTP Proposal et Agreement process (1/2)



- ▶ Switch A a un chemin vers Root par B et C
- ▶ Les 2 ports P0 et P1 sont en designated blocking state
- ▶ Quand un designated port est en discarding ou learning, il met le Proposal bit du BPDU à 1 et envoie ce BPDU
- ▶ A met alors tous les autres non-edge designated port en blocking : cela évite les boucles pendant le proposal-agreement process
 - ▶ Ce opération est dite être l'opération sync
- ▶ Switch A envoie un agreement à Root et met le P1 en Forwarding

RSTP Proposal et Agreement process (2/2)



- ▶ Ensuite A va envoyer un proposal à B qui va faire de même
 - ▶ Mettre ses non-edge designated port, autre que P5, en sync, envoyer un agreement à A et mettre P5 en forwarding
- ▶ Le processus continue avec les autres switches, « sous » B

Changement topologique

- ▶ En 802.1D, chaque changement génère un BPDU de type Topologie Change Notification (TCN)
- ▶ Quand un switch constate un changement topologique, il envoie au Root Bridge un TCN
- ▶ Le Root Bridge met le flag TC à 1 et diffuse à tous les switches l'information
 - ▶ A la réception d'un BPDU avec le flag TC à 1, toutes les adresses MAC associées aux non edge ports sont vidées afin de refaire un apprentissage

Mécanismes après un TCN

Action	Notes
The RSTP bridge starts the TC-while timer.	RTSP sets the TC-while timer with a value equal to twice the hello time for all its non-edge designated ports and the root port, if necessary.
The RSTP bridge flushes the MAC addresses associated with all these ports.	
The TC flag bit is set on all outbound BPDUs.	BPDUs are sent on the root port as long as the While timer is active.
The bridge receives a BPDU with the TC bit set from a neighbor and clears the MAC addresses on all ports.	The port that received the TC BPDU retains learned MAC addresses.
The bridge starts the TC-while timer and sends BPDUs with a TC bit set out of all its designated ports and root port.	RSTP does not use the specific TCN BPDU, unless a legacy bridge needs to be notified.

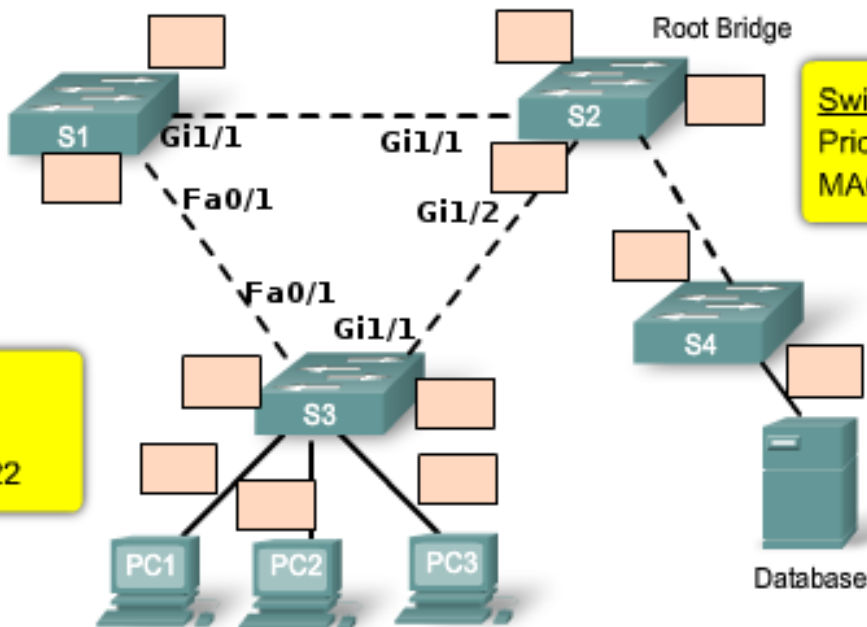
Trouver le type de chaque port ?

Switch S1:
Priority = 32769
MAC Address = 000A00111111

Switch S3:
Priority = 32769
MAC Address = 000A00222222

Switch S2:
Priority = 24577
MAC Address = 000A00333333

Switch S4:
Priority = 32769
MAC Address = 000A00444444



RP =Root Port

AP =Alternate Port

DP =Designated Port

EP =Edge Port

Des protocoles de Spanning Tree

La configuration de RSTP

La configuration

- ▶ Par défaut, les switchs Cisco implémente PVST+
 - ▶ Une instance STP 802.1D par VLAN
- ▶ Il est possible de mettre en place RSTP
 - ▶ Le mode est alors Rapid PVST+ (RPVST+)
 - ▶ On a alors un RSTP par VLAN

Les commandes

Command	Description
Switch(config)# spanning-tree mode rapid-pvst	Sets spanning tree mode to Rapid PVST+
Switch# show spanning-tree <i>vlan</i> <i>vlan-number</i> [detail]	Shows commands that are VLAN-based rather than instance-based
Switch# debug spanning-tree pvst+	Debugs PVST+ events
Switch# debug spanning-tree switch state	Debugs port state changes

Des protocoles de Spanning Tree

Questions ?