

Institut de la Francophonie pour l'Informatique

Réseaux Avancés

Ngo Hong Son

Dept. Computer Networks and Communications

SoICT-HUST

Interconnection de la couche 2: Switch, Pont et STP

Contenu

- Pont transparent (commutateur)
- Protocole de l'arbre recouvrant (STP)
- Rapid Spanning Tree Protocol (RSTP)
- Réseaux locaux virtuels (VLANs)

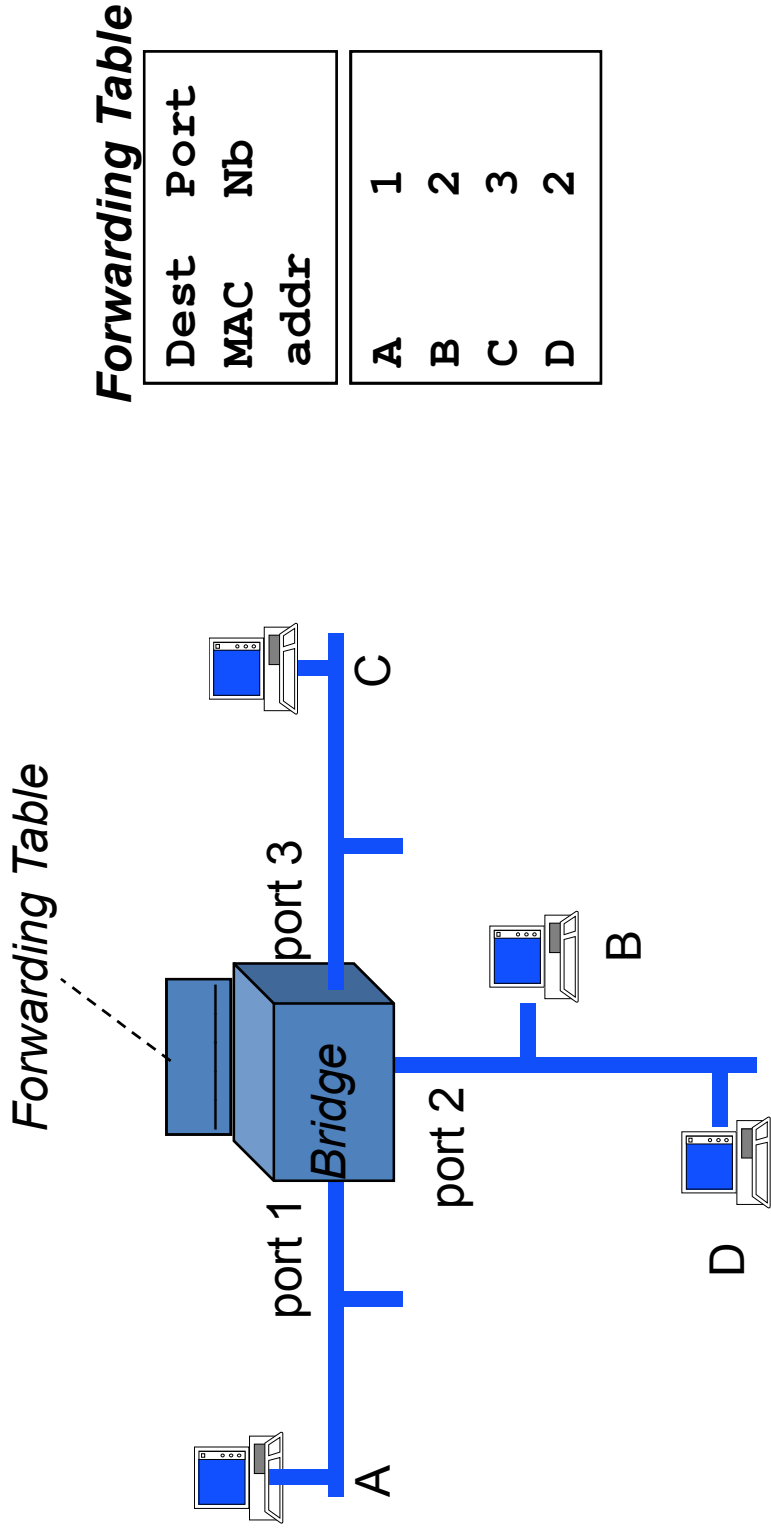
Pont Transparent

- Systèmes intermédiaires pour transférer les frames en basant sur l'adresse MAC
- Couche 2 seulement
 - Utilise les adresses MAC pour identifier les systèmes finaux
- PCs ne connaissent pas les ponts
 - Pont est transparent
 - Frames MAC ne sont pas changés par des ponts

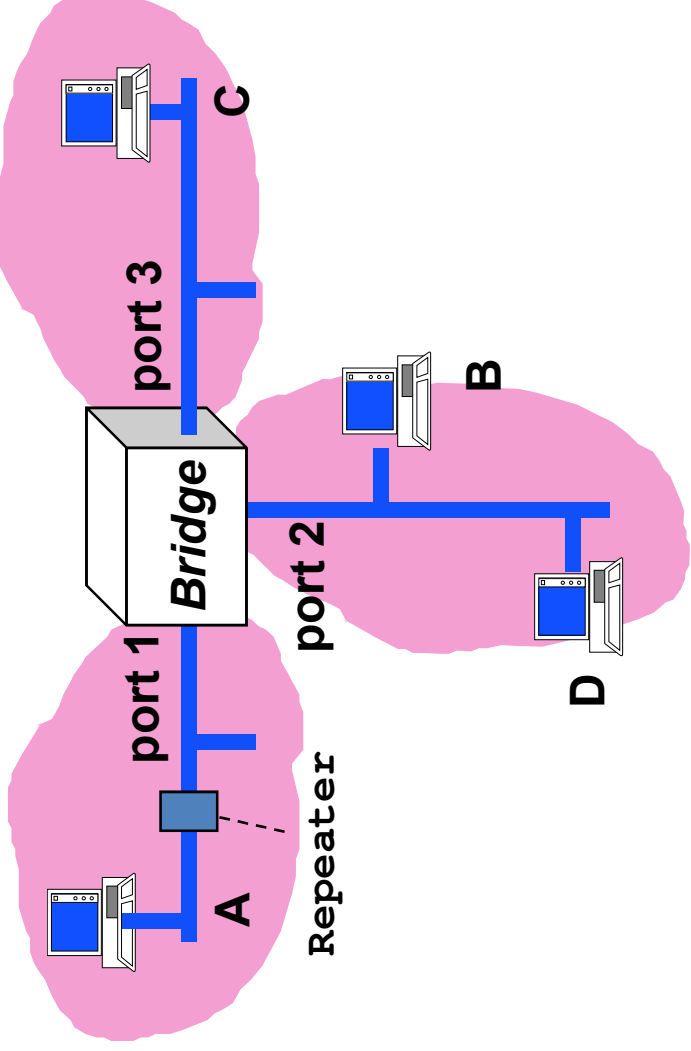


Pont Transparent

- Operate par “table de transfert”
- Opération
 - Sans-connexion, adresse non-structure



Capacité d'apprendre les adresses MAC



Forwarding Table

Dest MAC addr	Port Nb
---------------------	------------

--

Forwarding Table

Dest MAC addr	Port Nb
---------------------	------------

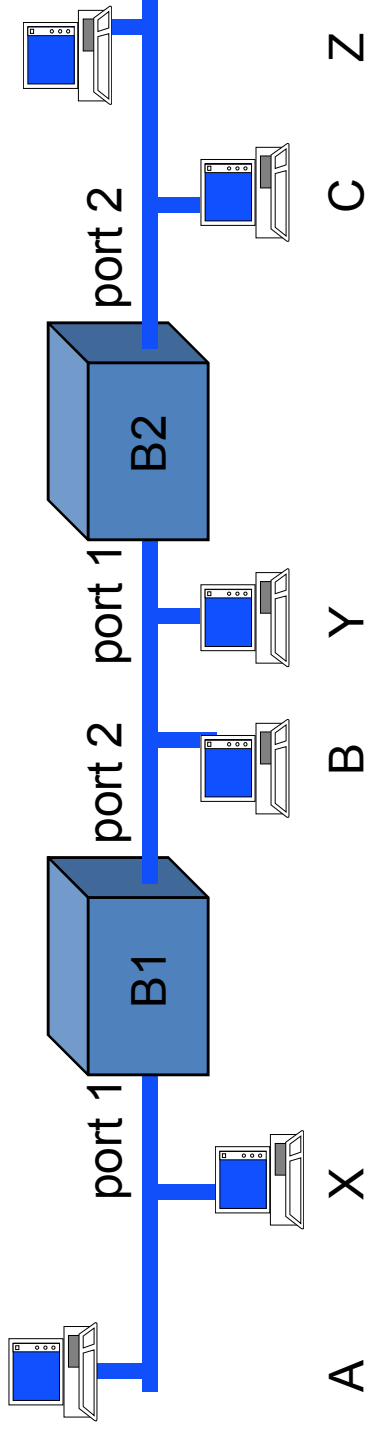
A	1
B	2
C	3
D	2

- Table de transfert est construite par “l'apprentissage”

- plug et play, pas de configuration (non IP)
- par le champ SA dans la trame MAC
- Broadcast si la destination est inconnue

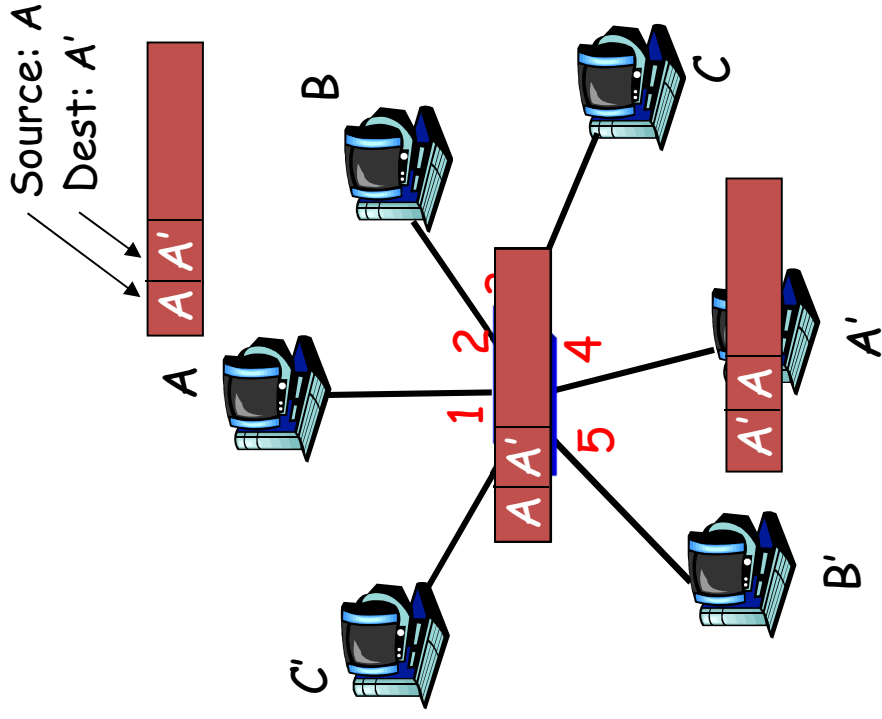
Reseau des ponts

- Le process est similaire?
- Comment B2 voit le reste du reseau?



Ex.

- A est inconnue:
Broadcast
- A est reconnait:
Transfert direct



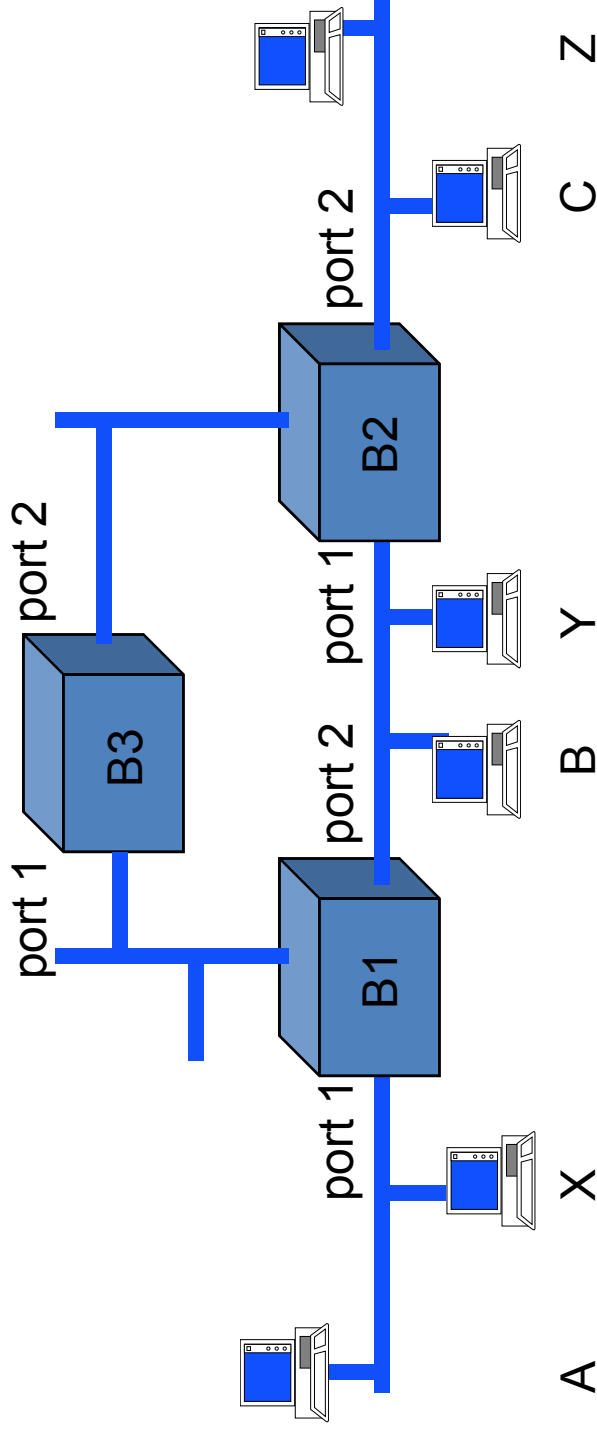
MAC addr	interface	TTL
A	1	60
A'	4	60

Table de transfert

STP

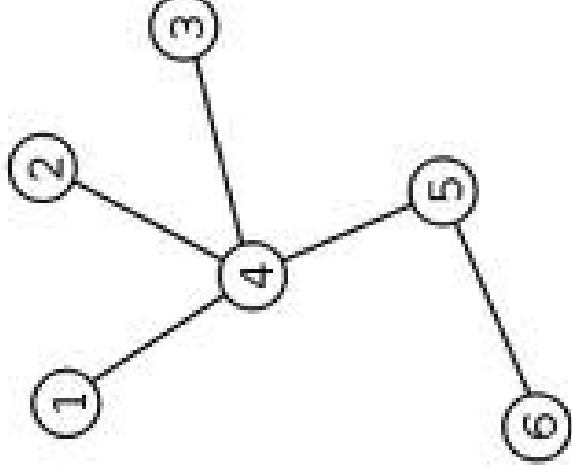
Boucle

- Si A envoie une trame à B?
 - Si “a table de transfert est vite -> **des tempêtes de broadcast**



Topologie sans boucle

- Réseau de ponts travail bien dans un topology sans boucle
 - Graphe: nœud = pont, arête = connection par LAN
 - Graphe sans boucle = arbre
 - Étoile
 - Un chemin de A à B



“Spanning Tree” Protocol

- Topology -> Arbre
 - Algorithm “Spanning Tree” dans tous les ponts
 - Quelques portes sont bloqués
 - <> “forwarding ports”
- Interconnection of bridges
 - Chemins parallèles pour la fiabilité
 - Algorithm “Spanning Tree” sélectionne un chemin par une fois

Methode de transfert

Transfert de
PDU

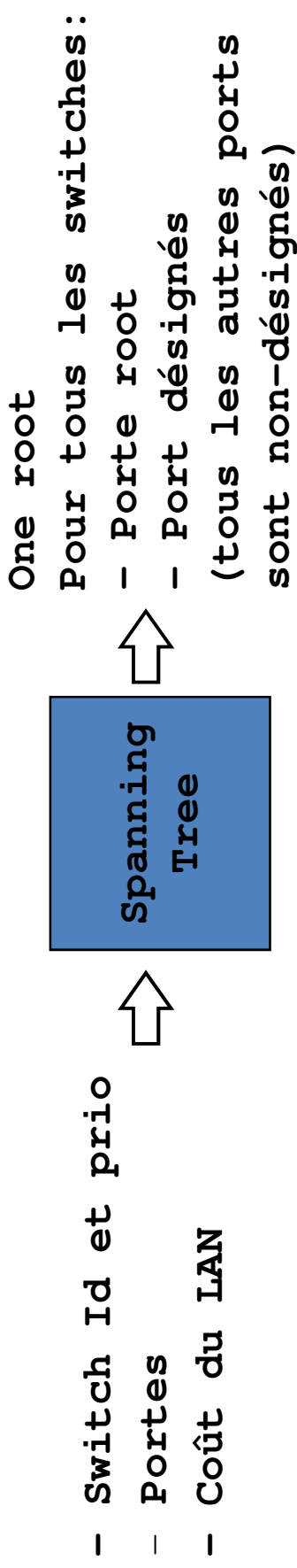
+

Controle

```
Trame de porte i ->
/* porte i est forwarding */
Si DA est unicast (vers porte j),
    copier vers porte
else broadcast vers portes ≠ i
Metre a jour la table avec (i, SA)
```

```
Maintenir le spanning tree et les etats
des portes
Apprendre des adresses
```

Specification



- Switch Root sera le point central de l'arbre STP
- Switch qui aura l'ID la plus faible sera celui qui sera élu Root
- Spanning Tree = les chemins les plus courts vers root
 - Le plus bande passante, le moins coût
 - Bellman-Ford principle:
$$\text{cost_to_root} = \text{best_announced_cost} + \text{local_cost}$$

Le cout d'un lien

Vitesse du lien	Coût	Plage de coût recommandée
4Mbps	250	100 to 1000
10Mbps	100	50 to 600
16Mbps	62	40 to 400
100Mbps	19	10 to 60
1Gbps	4	3 to 10
10Gbps	2	1 to 5

Liaisons entre switch, port et etat

Port	<->	Port	Etats	<->	Etats	Switch	<->	Switch
Root	<->	Désigné	Forwarding	<->	Forwarding	Non-root	<->	Root
Désigné	<->	Root	Forwarding	<->	Forwarding	Non-root	<->	Non-root
Désigné	<->	Non-désigné	Forwarding	<->	Blocking	Non-root	<->	Non-root

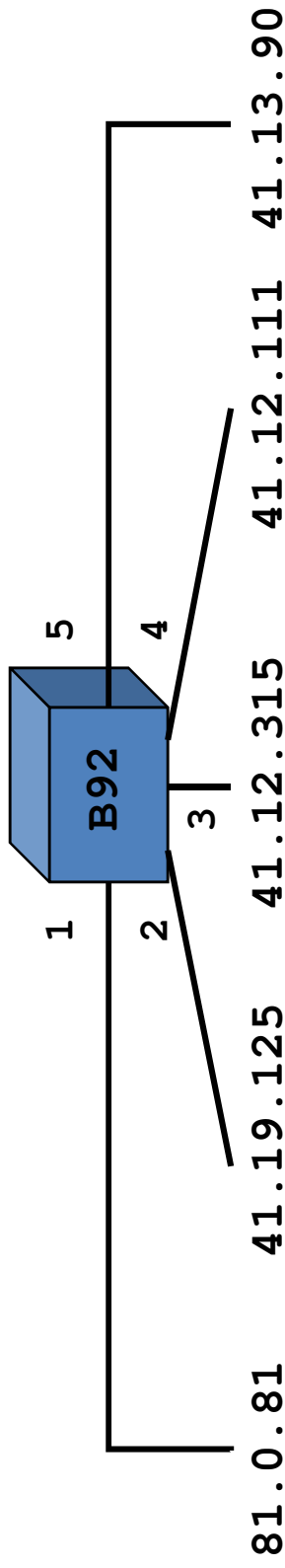
Messages de configuration:

`rootId.cost_to_root.senderId.port (41.13.92.3)`

Simplifie: `rootId.cost_to_root.senderId`

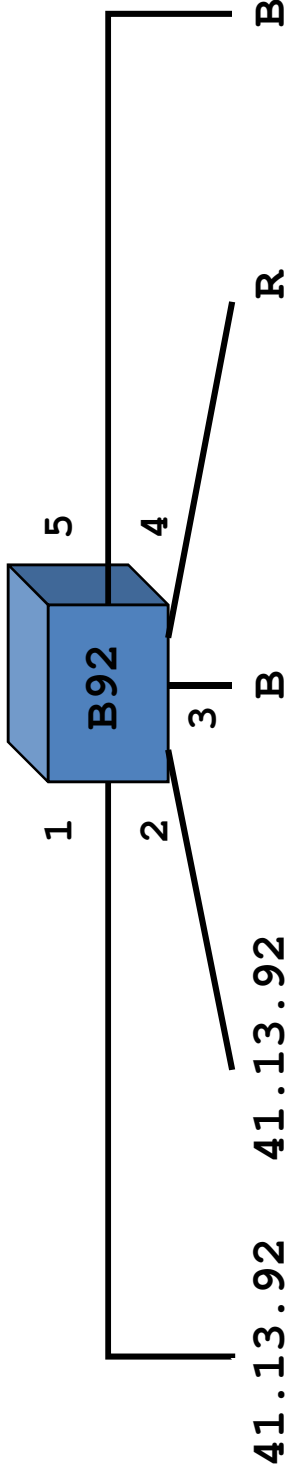
Example de STP

Spanning Tree example



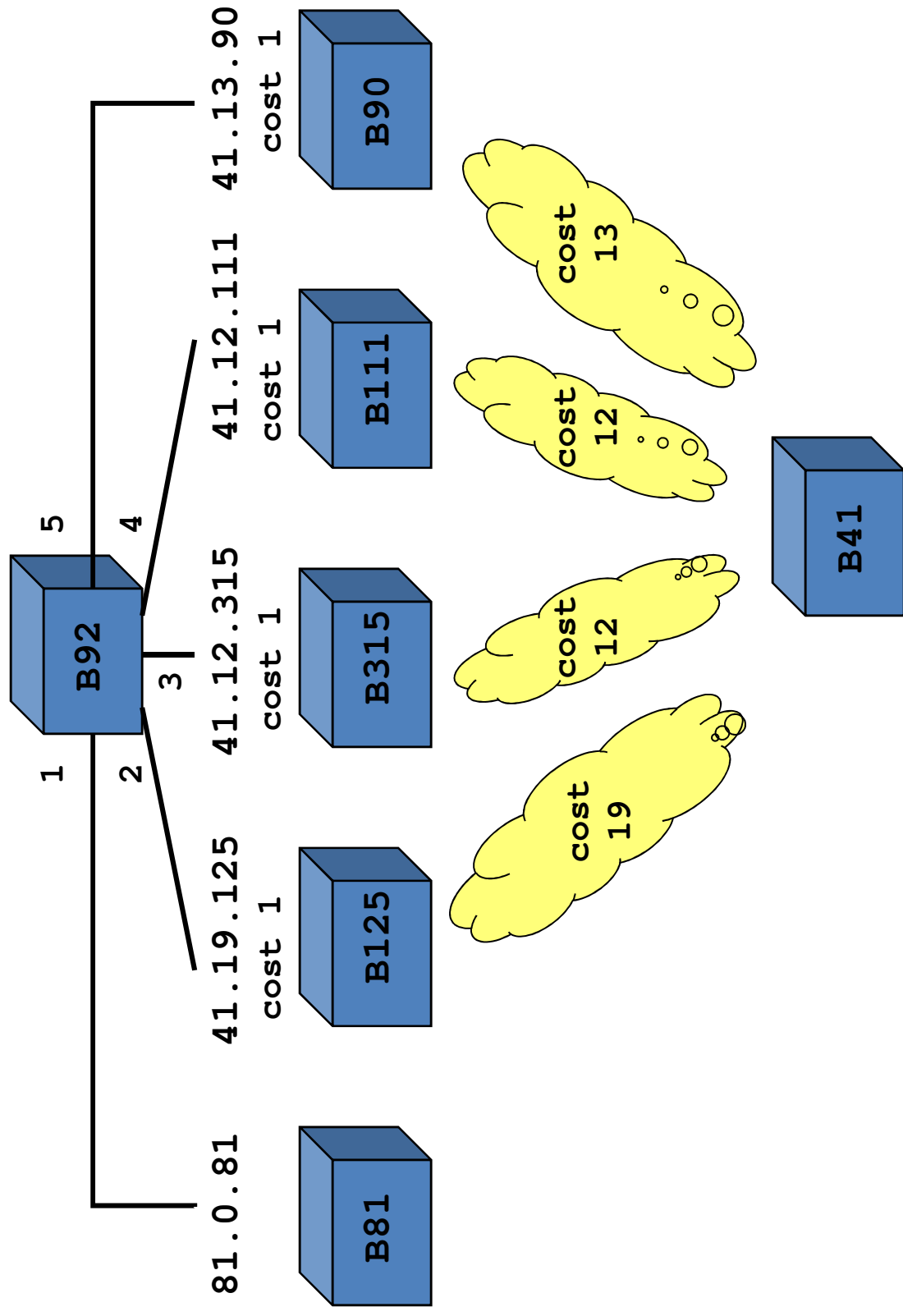
- Mieux root: **41**
- Mieux cout: $12 + 1 = 13$, a porte 3 ou 4 (cout=1)
- **Root port: 4**, car $111 < 315$
- Nouveau message: **41.13.92**
- Portes 1 et 2 sont **designated**: **41.13.92** est meilleur que 81.0.81 et 41.19.125
- Portes 3 et 5 sont **blocked**: **41.13.92** n' est pas meilleur que **41.12.315**, ni 41.13.90

Spanning Tree example

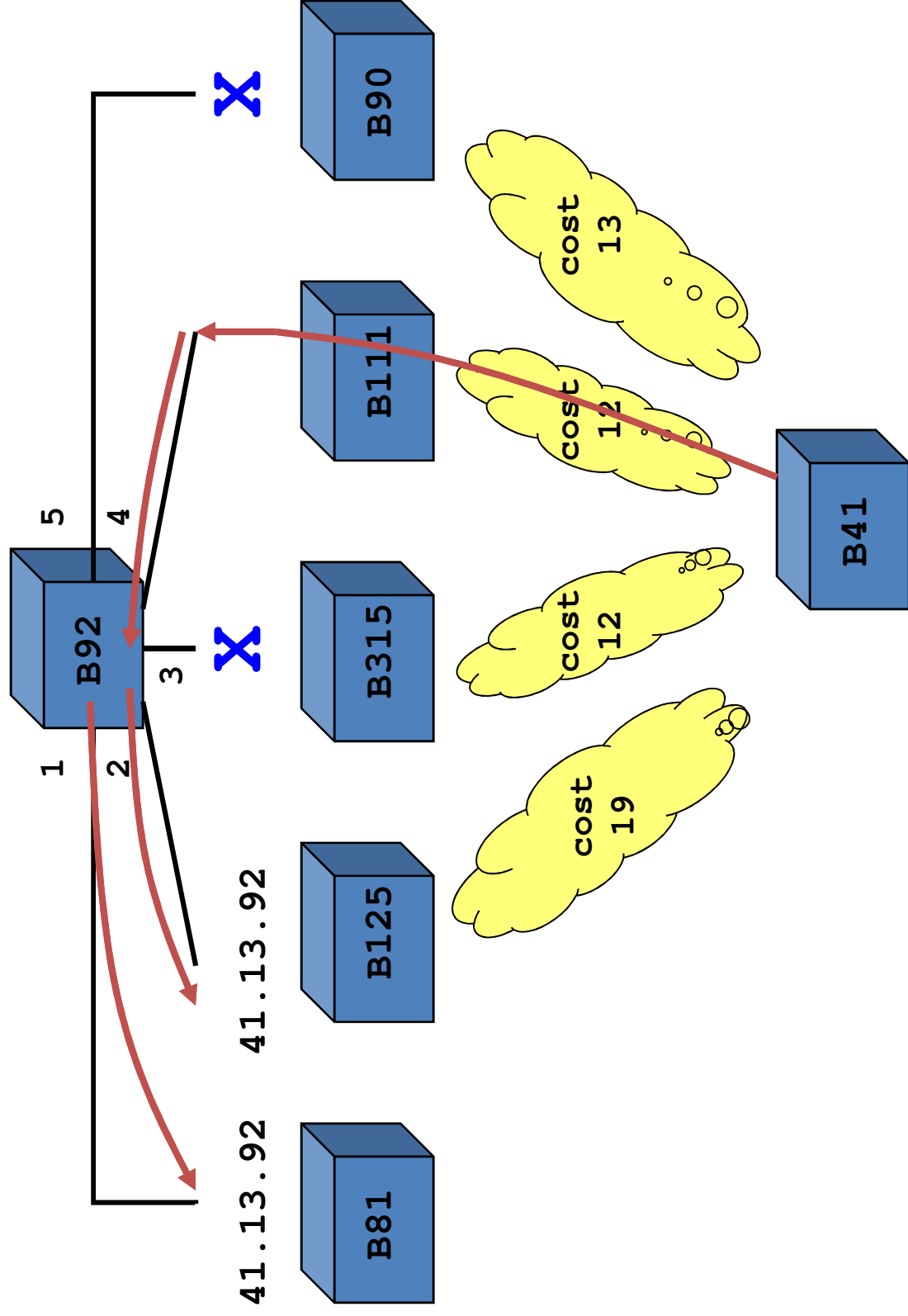


- Message **41.13.92** est envoyé périodiquement vers portes **1** et **2**
- Portes **1**, **2**, **4** participent au “forwarding” (II sont dans le Spanning Tree)

Spanning Tree example



Spanning Tree example



Resume: Procedure de STP (IEEE 802.1d)

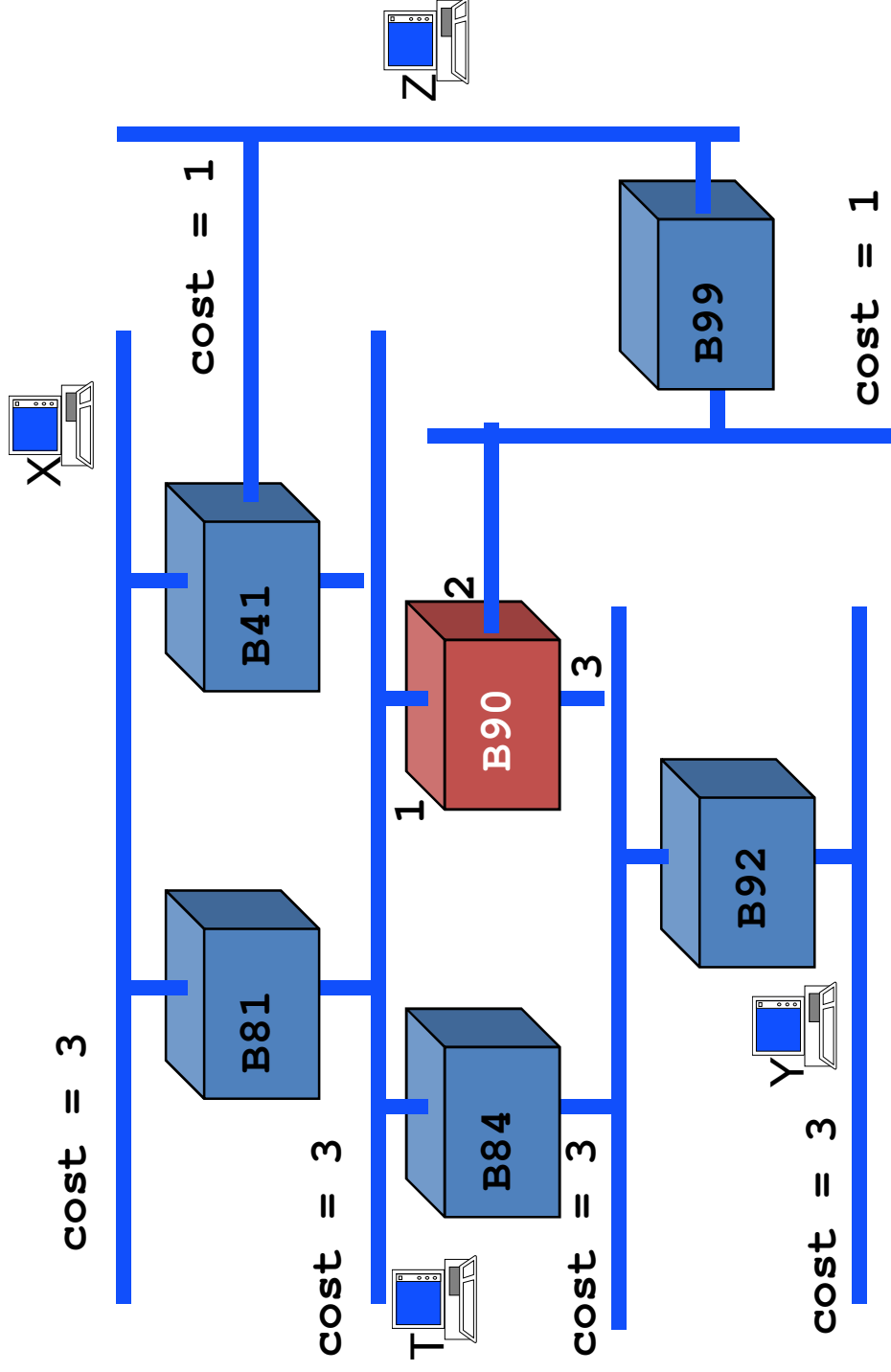
```
config BPDU received on any port or port enabled ->

compute new root;
compute new cost to root;  /* Bellman Ford */
build_new_config_BPDU;
for all ports i do
    if new_config_BPDU better than stored_config[i]
        then store and send on port i;
    end

compute root port /* smaller distance to root */
designated ports = ports where config BPDU was sent
blocked ports = other ports
```

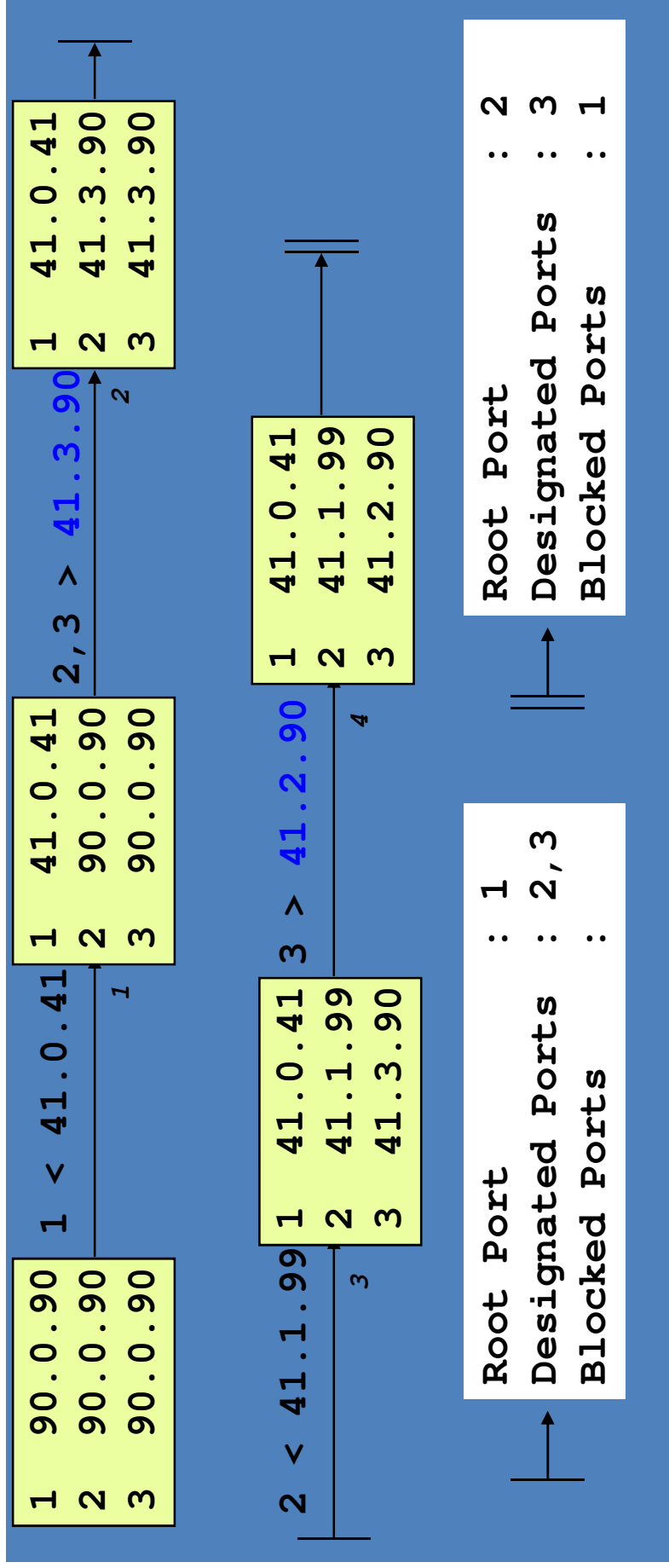
```
r.c.s better than r'.c'.s' iff
    (r<r') or (r=r' and c<c') or (r=r' and c=c' and s<s')
```

Complex example



Initialization de Spanning Tree

- Pont B90 -> BPDU 90.0.0.90 envoie a toutes les portes; Les tables configuration de B90 :

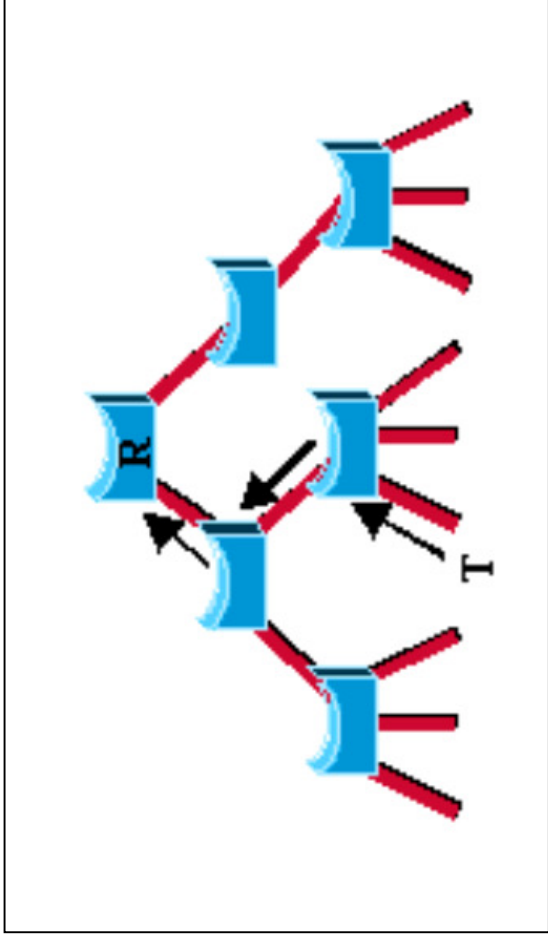


message received on port 1: 1 < 41.0.0.41 message format: rootId.cost_to_root.senderId

Commentaires

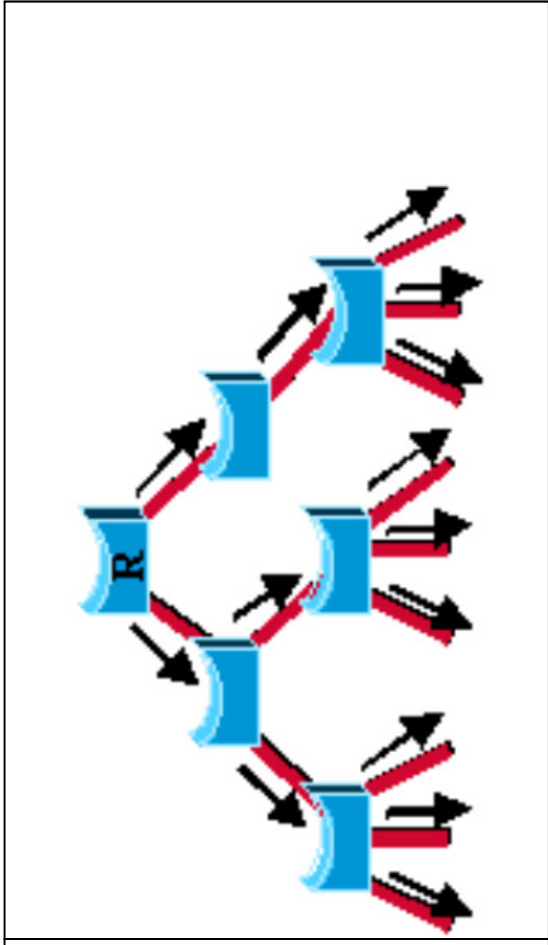
- Lors de la réception d'un message, on compare le coût (le coût local inclus), mais nous conservons le message reçu (sans le coût)
- Lors de la réception de **41.0.41** au porte 1:
 - 41.3.41 < 90.0.90? yes -> 1 devient root
 - Nouveau config msg = 41.3.90
 - 41.3.90 < 90.0.90? yes -> 2 devient designated
 - 41.3.90 < 90.0.90? yes -> 3 devient designated
- Lors de la réception de **41.1.99** au porte 2:
 - 41.2.99 < 41.3.41? yes -> 2 devient root
 - Nouveau config msg = 41.2.90
 - 41.2.90 < 41.3.90? yes -> 3 devient designated
 - 41.2.90 < 41.0.41? no -> 1 devient blocked

Problem de STP: Changement de topologie



Nouveau lien ajouté au switch

Topologie mise à jour envoyée à root



Topologie mise à jour envoyée par root
sur ST dans

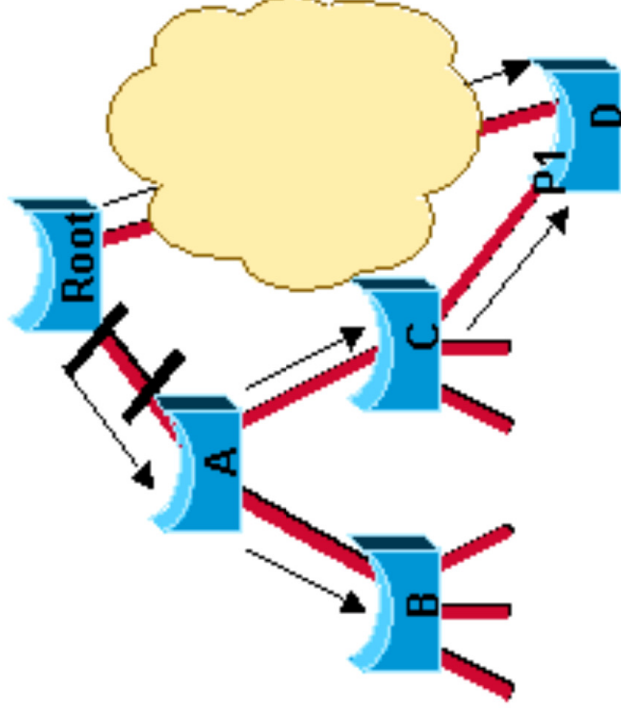
MaxAge + ForwardDelay

Tous les ponts recalcul le ST

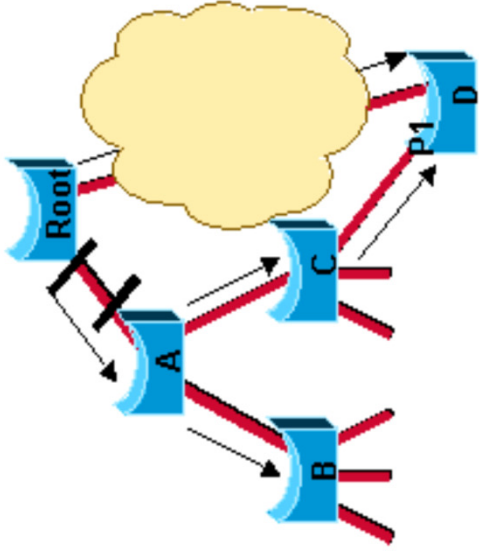
Source: CISCO RSTP White Paper

Example

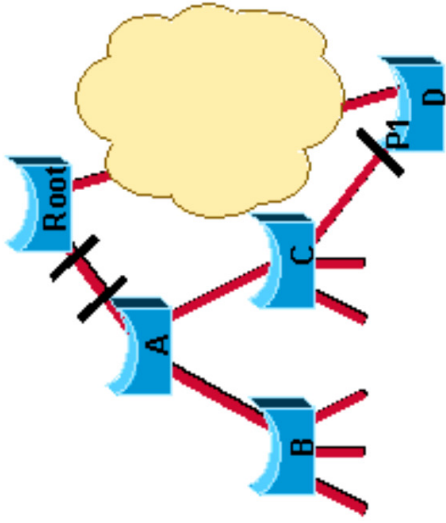
- Un pont nouvellement connecté au root. Qu'est-ce qui se passe?



Source: CISCO RSTP White Paper



1. A et root fait fonctionner la ST procedure dans les nouvelles portes.
2. Nouveaux BPDUs sont envoyes vers B et C
3. D compute la porte p1 comme nouveau root



1. P1 de D est mise en "listening" etats dans 15 s
2. P1 de D est mise en "listening" etats dans 15 s

⇒ Changement de topologie est rapide mais le nouvel etat n'est par mis a jour immediatement

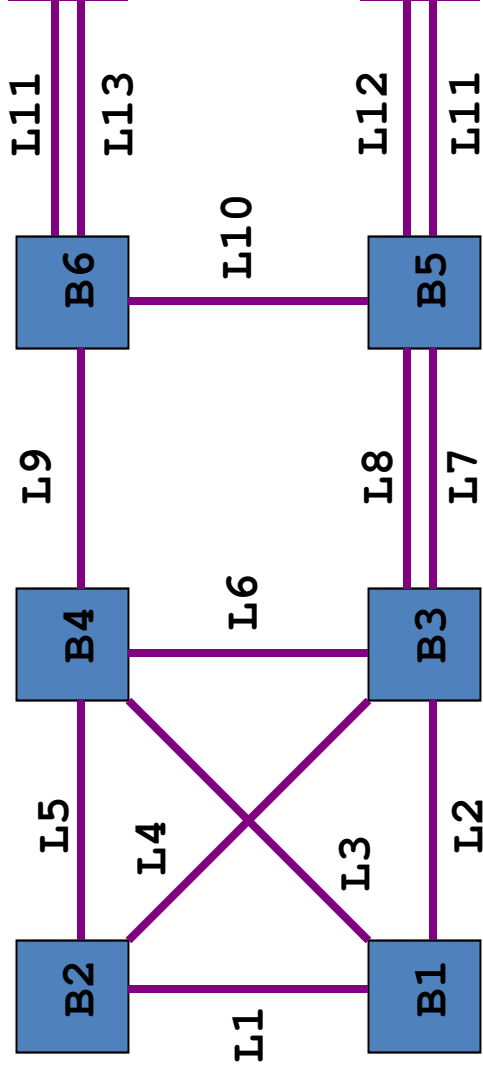
RSTP - Rapid ST protocol

- IEEE 802.1W
- Evolution de STP
- But: Rapide reconfiguration
- Principale amelioration:
 - Rapide reconfiguration: alternative chemins vers root (ou backup path vers un LAN)
 - Rapide reconfiguration vers “forwarding” etats utilisant le protocole de negociation (<> timers)

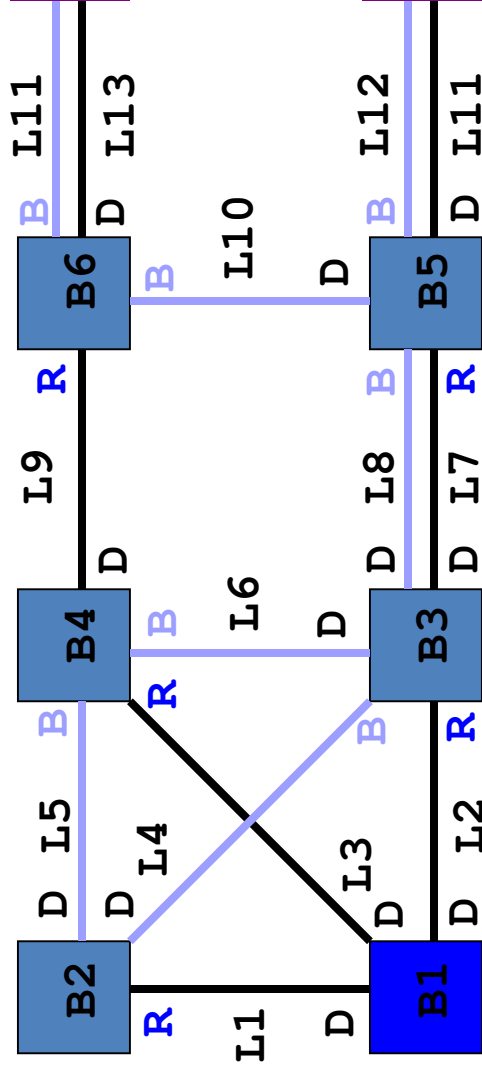
Etats des portes dans RSTP

- Une port soit : root, designated, alternate, backup, blocked
- **root** port = port vers root (même de STP)
- **designated** port = connecte LAN au spanning tree (même de STP)
- Une porte qui est ni root, ni designated
 - **alternate**: connecte le pont vers root après le changement de topologie
 - **backup**: connecte LAN au ST après le changement de topologie
 - **blocked**: est désactivé dans le spanning tree

Example de Rapid STP

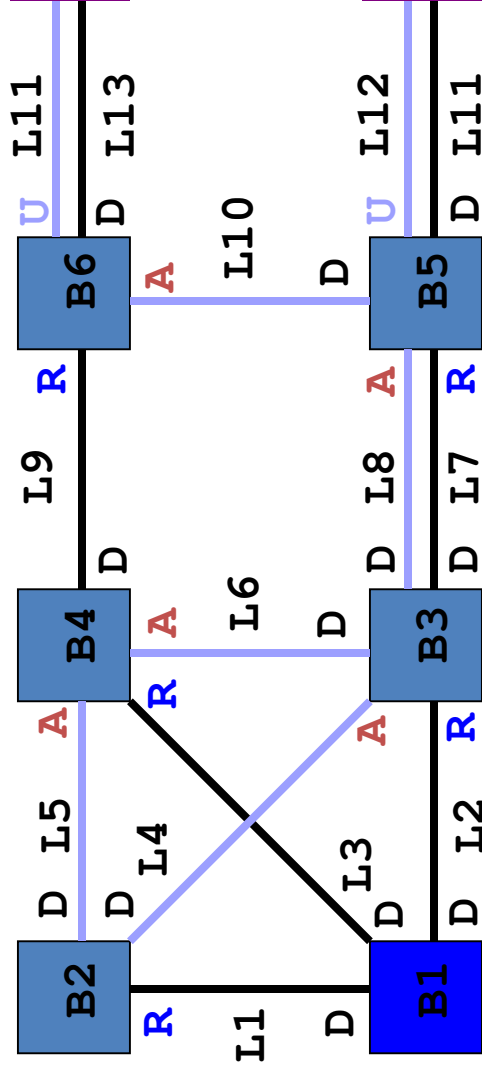


ST construit par STP



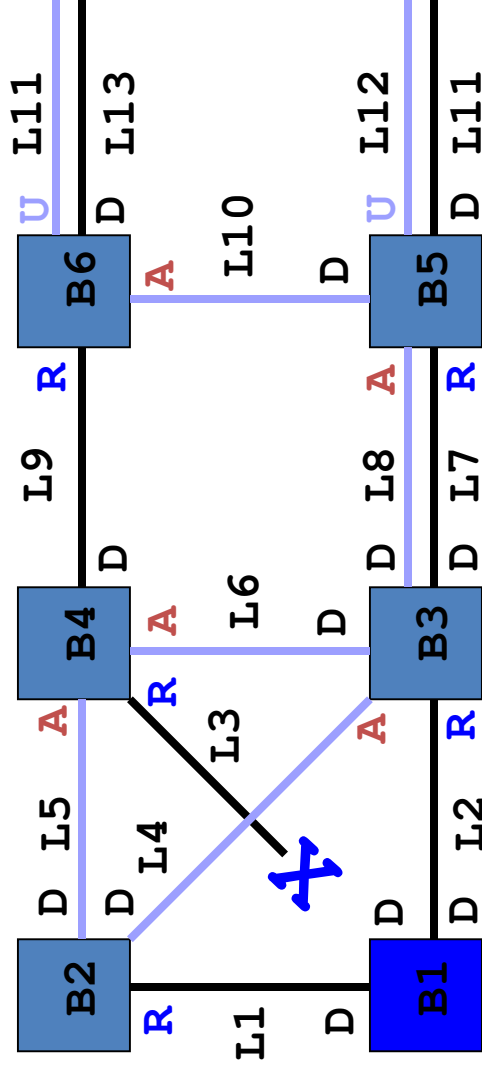
- B1 - root
- R - root ports, D - designated ports, B - blocked ports

ST construit par RSTP



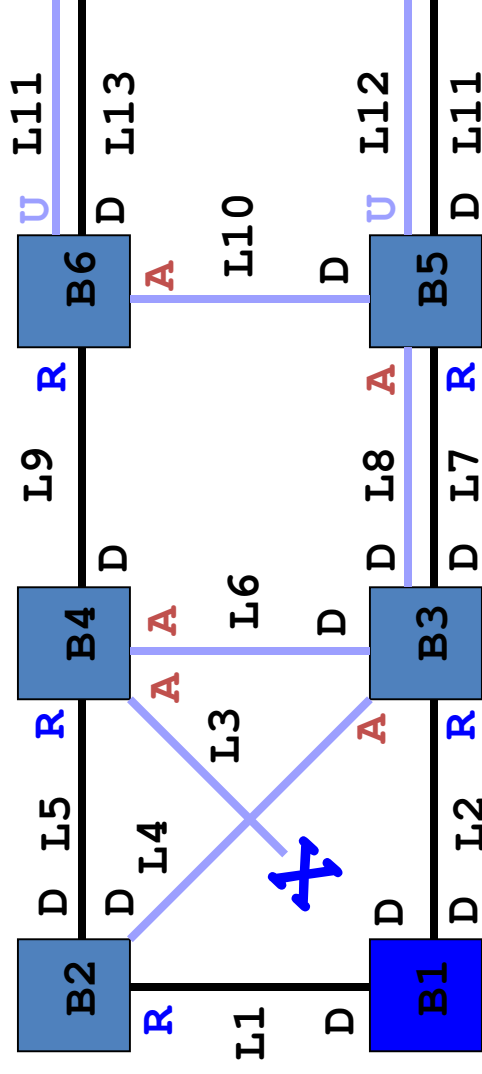
- B1 - root
- R - root ports, D - designated ports, B - blocked ports
- A - alternative ports, U - backup ports

L3 en failure



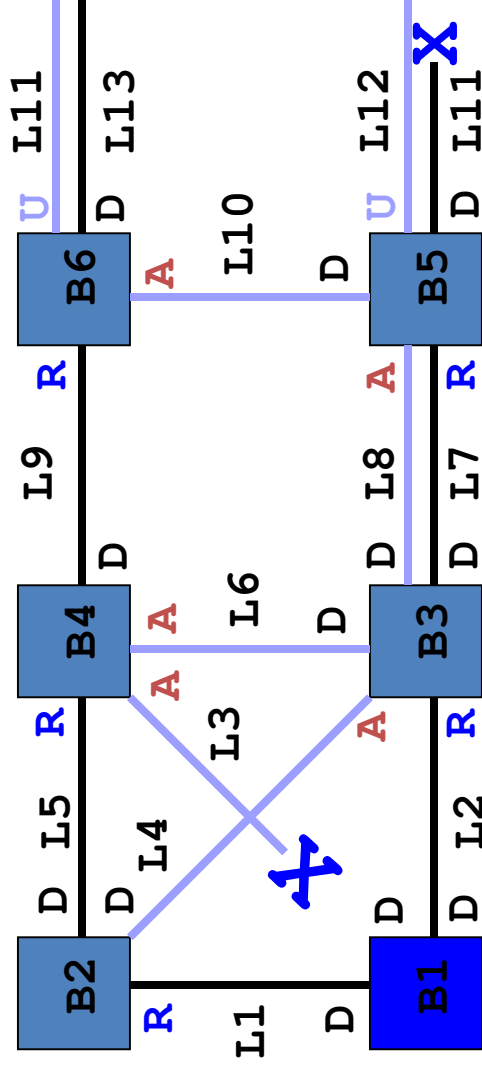
- B1 - root
- R - root ports, D - designated ports, B - blocked ports
- A - alternative ports, U - backup ports

L3 en failure



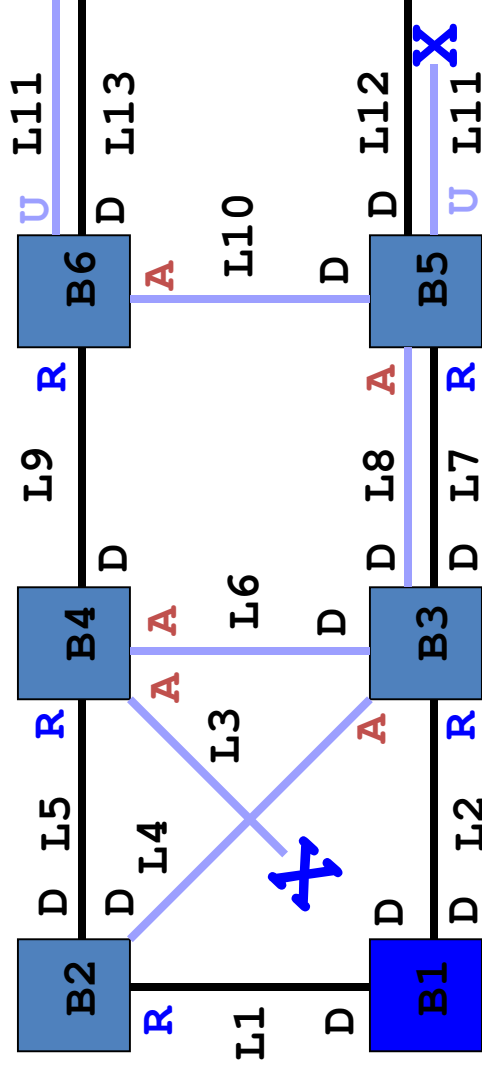
- Dans B4
 - porte sur L3 devient A, etat Discarding
 - porte sur L5 devient R, etat Forwarding

L3 en failure



- B1 - root
- R - root ports, D - designated ports, B - blocked ports
- A - alternative ports, U - backup ports

L3 en failure



- Dans B5
 - porte sur L11 devient U, etat Discarding
 - porte sur L12 devient D etat Forwarding

Conclusion

- Ethernet switches/ponts sont largement utilisés
- 100 Mb/s switches pour PCs, 1 Gb/s dans le backbone
- Interconnection compliquée
 - Chemins parallèles pour la fiabilité
 - STP et RSTP offrent l'interconnection sans boucle
 - VLANs aident à séparer les domaines de collision (broadcast domain)
 - Extensibilité limitée
- Produits
 - Switch/Router – portes intégrées : Couche 2 et Couche

3

QUESTION

