

Réseau - Protocole STP/RSTP



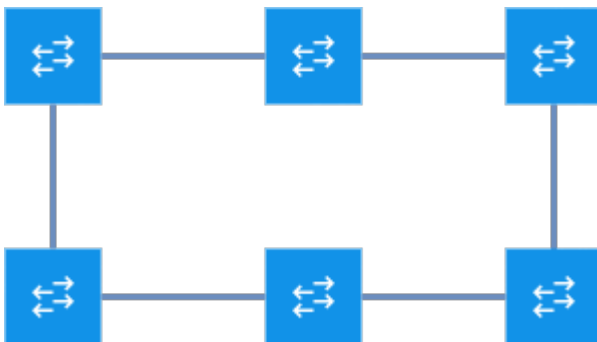
Difficulté : Confirmé

Notions : Réseau, protocoles, résilience réseau.

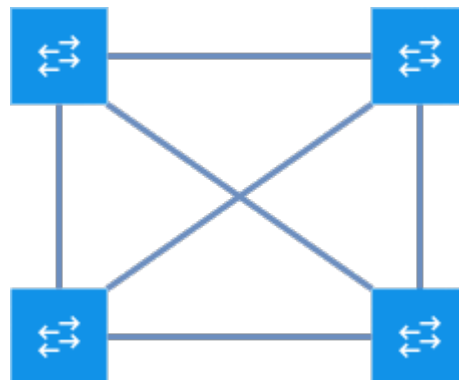
I. Introduction

Dans le cadre de la résilience d'une topologie réseau, il est essentiel de redonder des liens afin de palier à une rupture éventuelle.

Plusieurs topologies réseaux spécifiques existent. Dont les deux suivantes :



Topologie Boucle (ring)



Topologie meshée

Cependant, ces boucles créent également deux problèmes majeurs :

- les tempêtes de broadcasts, dues aux boucles réseaux présentes dans la topologie.
- Des instabilités dans les tables MAC.
- Des trames en doublons.

Ce qui peut mener à des dysfonctionnements applicatifs (rupture de session), des pertes de performances (collision de paquets, réadressage MAC) voire à un effondrement du réseau.

Pour éviter ces problèmes et rendre ce type de topologies possibles, le protocole **STP** (**S**panning **T**ree **P**rotocol) a été créé.

Cependant, dû à ses problèmes de lenteur de cicatrisation réseau (30-50 secondes de délai de convergence), une version plus moderne en a été créée : le protocole **RSTP** (**R**apid **S**panning **T**ree **P**rotocol).

Celui-ci réduit les temps de cicatrisation entre 1 et 3 secondes et apporte quelques améliorations :

- Simplification des rôles des ports
- Meilleure stabilité
- rétrocompatible avec STP

II. Fonctionnement STP/RSTP

2.1 Principe général

Le fonctionnement de base du STP / RSTP est le suivant :

2.1.1 Election de la passerelle racine (root bridge)

Chaque switch envoie une trame d'un type spécifique appelée **BPDU** (**B**ridge **P**rotocol **D**ata **U**nit).

Cette trame contient les informations de priorité, d'adresse MAC et de coût du chemin racine.

Une fois ces trames reçues, le switch ayant la **priorité** avec la valeur la plus basse devient le switch racine.

Info : En l'absence de priorité définies sur les switch, c'est la valeur des adresses MAC qui est prise en compte et fait office de valeur de priorité.

Une fois ce processus d'élection achevé, les trames continuent à être envoyées de manière continue. Si un switch apparaît avec une priorité plus élevée, ce rôle lui sera transféré.

Conseil : C'est ce qui rend la résilience possible et l'architecture scalable. Si le root bridge tombe, celui avec la meilleure priorité disponible est automatiquement élu. Si un switch avec une meilleure priorité est ajouté à la topologie, il est également automatiquement élu et récupère le rôle de root bridge.

2.2.2 Détermination des ports racine (root ports)

Chaque **switch non-racine** détermine ensuite son **root port**. Il s'agit du port qui a le chemin le plus court pour accéder à la racine.

Ce chemin est calculé en faisant la somme des **coûts** pour chaque chemin.

Le coût est calculé en fonction de la rapidité des liens. Voici un tableau avec les valeurs de références.

Débit du lien	Coût STP (802.1t)
10 Mb/s	2 000 000
100 Mb/s	200 000
1 Gb/s	20 000
10 Gb/s	2 000
100 Gb/s	200
1 Tb/s	20

Info : En cas d'égalité entre deux chemins, c'est le voisin avec le bridge ID le plus faible qui l'emporte. Si l'égalité persiste, c'est le Port ID le plus faible qui est choisi.

Ce port sera donc tagué comme **actif**.

Exemple : Switch A est le Root Bridge. Switch B peut atteindre A via deux chemins :

- **Chemin 1 : B → C → A**

- B-C : 1 Gb/s → coût 20 000
- C-A : 1 Gb/s → coût 20 000
- **Total = 40 000**

- **Chemin 2 : B → D → A**

- B-D : 10 Gb/s → coût 2 000
- D-A : 1 Gb/s → coût 20 000
- **Total = 22 000**

Le Root Port de B sera le port vers D, car 22 000 < 40 000.

2.2.3 Désignation des switches et ports désignés (*designated ports*)

Sur chaque segment de la topologie, un switch devient **Designated Bridge** et ces ports non root participants à la topologie deviennent des **designated ports**. Ces switches et ces ports serviront de **forwarders** aux switches suivants sur la topologie, ceux qui sont plus loin que la racine.

2.2.4 Blocage des ports redondants (*Redundant ports*)

Les ports participant à la topologie mais n'étant désignés ni comme **root**, ni comme **designated** seront alors bloqués et deviendront des **blocked ports**. Ces ports pourront être activés en cas de dysfonctionnement de l'un des ports actif.

En attendant, ceux-ci :

- ne transfèrent plus le trafic,
- n'apprennent plus d'adresses MAC,
- n'écoutent plus de trafic, hormis les trames BPDU.

C'est comme cela que les switches de la topologie se prémunissent des boucles réseaux et c'est ce fonctionnement qui explique également les latences de cicatrisation.

On a donc 5 états possibles pour un port :

1. **Disabled**
2. **Blocking** (20 seconds)
3. **Listening** (15 seconds)
4. **Learning** (15 seconds)
5. **Forwarding**

2.2 Les améliorations de RSTP

Pour simplifier le fonctionnement du protocole et accélérer les temps de cicatrisation réseau, le RSTP va apporter quelques changements à ce fonctionnement.

- L'ajout d'un nouveau type de port dit **edge port** (PortFast), permettant de relier des équipements actifs ou des terminaux sans que cela ne déclenche un recalcul de topologie.
- Des négociations plus rapides lors des élections de switch et de détermination des status des ports basés sur un mécanisme de **Proposal/Agreement** :
 - Un switch propose un port comme Designated port.
 - Les voisins directs approuvent.
 - Les deux ports passent automatiquement en mode forward.

Pour vulgariser, lorsqu'un switch reçoit un BPDU de son voisin, il peut décider :

- "Ce voisin a un meilleur chemin vers le root que moi" --> Je deviens Root port
- "J'ai un meilleur chemin que lui" --> Je deviens Designated Port

Détail de la négociation

1. **Un switch reçoit une BPDU "meilleure"** : Il comprend que son port doit devenir 'root port'.
2. **Il bloque temporairement tous ses 'designated ports'** : Ce qui évite les boucles.
3. **Il envoie un 'agreement' à son voisin** : Si il peut garantir que cela ne crée pas de boucle.

4. **Le voisin passe son port en forwarding** : Pas de délai d'attente comme en stp classique.
5. **Le switch débloque tous ses 'designated ports'** : La topologie converge ainsi en quelques millisecondes.

Voici un comparatif des deux protocoles permettant de se faire un ordre d'idée

Fonctionnalité	STP (802.1D)	RSTP (802.1w)
Convergence	30-50s	1-3s
Etat de ports	5	3
Rôle de ports	3	4
BPDU	seul le root envoie	tous les switchs envoient
Détection de dysfonctionnement	Lent	Fast Handshake
Edge ports	Optionnel	Concept Natif

III. Caractéristiques des ports

3.1 Caractéristiques générales

Les ports peuvent être identifiés sur les switches de la façon suivante :

Identification sur le switch	Correspondance	Caractéristiques
RP	Root Port	• Un seul par switch. • Non applicable au root bridge. • Représente le chemin le plus court vers le root bridge.

DP	Designated Port	• Un par segment. • Transmettent le trafic vers l'aval depuis le root bridge.
AP	Alternate Port	• port de backup pour le root port. • Si le RP est coupé, il prends le relai.
BP	Backup Port	• port de backup pour un designated port (sur le même switch). • Rare dans les topologies réseaux modernes.
EP	Edge Port (PortFast)	• Ne participe pas à la topologie. • Relie un équipement actif.

En RSTP, un port peut avoir trois états fonctionnels (Au lieu de 5 en STP) :

- **Discarding** : Ne forward pas le trafic, n'apprends pas les MAC.
- **Learning** : Apprends les MAC address.
- **Forwarding** : Forward le trafic, port actif.

Les ports ne participant pas à la topologie où les équipements actifs sont connectés sont les **Edge Ports** (PortFast).

Pas de recalcul ou de changement d'état sur ces ports car ils ne participent pas à la topologie. Ils passent directement en mode **forward**. Cela permet d'annuler les délais pour les équipements actifs branchés dessus.

2.2 BPDU Guard & Loop Guard

En plus de ce qui a été vu ci-dessus, il existe deux fonctionnalités supplémentaires applicables aux ports en STP/RSTP.

La première est le **BPDU Guard** : C'est une fonctionnalité qui sert à protéger les **Edge Ports** des trames BPDU non désirées.

Ces ports ne participant pas à la topologie, ils ne doivent pas traiter de trames BPDU.

Si toutefois, suite à une erreur de manipulation (mauvais branchement par exemple), ces ports reçoivent des trames BPDU, cela veut dire qu'un port RSTP a été branché sur un port ne participant pas à la topologie et donc naturellement non protégé contre les boucles réseaux.

Dans ce cas, le BPDU Guard va automatiquement couper le port pour prévenir la boucle. Le port passe alors en état **Err-disable**.

La seconde est **Loop Guard** : C'est une fonctionnalité qui sert à protéger les liens **Trunk** et les **uplinks** de boucles unidirectionnelles.

2.2 Rapid-PVST

Le rapid-PVST ou (**P**er **V**lan **S**panning **T**ree) est un RSTP appliqué par VLAN.

Ainsi si plusieurs VLAN co-existent au sein d'une même topologie physique, il est possible de définir une instance de RSTP par VLAN.

Chaque VLAN dispose ainsi de :

- Son Root Bridge,
- Ses chemins favoris,
- Ses propres root ports,
- Ses propres designated ports.

Conseil : Cela peut se révéler particulièrement intéressant dans de vastes structures dans lesquelles il est important voire crucial d'optimiser ses chemins réseaux et d'équilibrer la charges sur les différents liens. Ou dans le cas d'usages intensifs de bande passante sur des vlan concurrentiels (voip / Vidéo)...

Cependant pour des réseaux à très grande échelle, il sera préférable d'utiliser le protocole **MSTP**.

Il sera également bon de prendre en compte que vu que le Rapid-PVST est instancié par VLAN, il consomme également plus de CPU.

IV. Les trames BPDU

Comme vu jusqu'ici, le RSTP repose sur l'utilisation de trames spéciales appelées BPDU.

Contrairement au STP classique où seul le root bridge envoie des trames, en RSTP, chaque switch envoie ces trames en plus de forwarder celles du root bridge. C'est cela qui permet des délais de cicatrisation si courts.

Ces trames se déclinent en deux types.

4.1 Les Configuration BPDU

Il s'agit des paquets BPDU normaux, envoyés toutes les 2 secondes par le Root Bridge.

Celles-ci contiennent :

- l'ID du root Bridge
- Le coût du chemin vers le root
- l'ID du bridge émetteur
- l'ID du port émetteur
- les différents timers (Hello, Max Age, Forward Delay) si ceux-ci sont utilisés

Leur rôle est de maintenir la topologie. Elles servent également à élire ou ré-élire un root bridge et déterminer les root ports et designated ports.

4.2 Les TCN BPDU (Topology Change Notification) en STP

En STP classique, ce sont les trames spéciales envoyées par un switch qui détecte un changement dans la topologie.

(par exemple : Un port qui passe de forwarding à down ou inversement, un lien qui tombe, un switch ajouté ou retiré de la topologie, etc...)

Cela fonctionne de la façon suivante :

1. Un switch non-root détecte un changement de topologie.
2. Il envoie une TCN BPDU vers son root port.
3. Chaque switch intermédiaire accuse réception et relaie vers le root bridge.

Le root bridge active alors le **Topology Change Flag** dans ses configuration BPDU et diffuse cette information sur l'ensemble du réseau.

Chaque switch va alors réduire le temps de mise en cache des adresse MAC (aging time) afin de les purger au plus vite.

Info : En RSTP, ces trames sont supprimées et leur mécaniques sont intégrées dans les BPDU normaux, ce qui est beaucoup plus rapide.

4.3 Détection des changements en RSTP

En RSTP, plutôt que d'utiliser des trames dédiées sur un changement de topologie, le Change Flag est intégré dans les trames classiques. Et comme chaque switch émet ses propres trames en permanence en plus de forward celles du root bridge, l'information se diffuse donc bien plus rapidement.

Ainsi, lorsqu'un port change d'état (Up / Down / Forwarding), le RSTP :

- Met immédiatement à jour son rôle (RP, DP, AP, ...)
- Envoie une BPDU avec le bit **Topology Change** activé
- Informe directement ses voisins, et ce
 - sans attendre le Max Age

- sans envoyer de TCN BPDU à destination du root bridge
- Ses voisins mettent immédiatement à jour leurs tables MAC
 - Rendant ainsi la propagation de l'information immédiate et bidirectionnelle.

La propagation devient ainsi locale plutôt que globale et les délais de cicatrisation sont ainsi réduits à leur minimum.

Un changement de topologie met désormais entre 1 et 3 secondes à être totalement opérationnel.

V. Bonnes pratiques

Il est important, compte tenu des mécanismes ci-dessus, de prendre en considération l'ensemble des fonctionnalités dans la construction d'une topologie à la fois résiliente et optimisée.

- Toujours définir un Root Bridge soi-même.
- Définir les priorités des switches manuellement pour définir des routes "préférées".
- Bien déclarer les ports ne participant pas à la topologie en Edge Ports.
- Sur ces ports, activer BPDU Guard pour éviter les erreurs de manipulation.
- Utiliser préférentiellement un mode 'Point-to-point' sur les trunks.
- Activer Loop Guard sur les trunks
- Éviter de mixer STP et RSTP.

VI. Conclusion

Le protocole RSTP est primordial dans les architectures réseaux modernes critiques où le réseau doit à tout prix être stable et fonctionnel. Rendant ainsi centraux les aspects de redondance et de

convergence rapide.

Mais il fait entrer beaucoup de notions différentes complexes qu'il est essentiel de bien comprendre et maîtriser afin de rendre son infrastructure la plus optimisée possible.

Revision #2

Created 2026-07-17 20:31:40 UTC by Olivier

Updated 2026-07-17 20:39:52 UTC by Olivier