Cours 46 : Quality of Service (Partie 1)

Dans ce cours nous verrons ce qu'est la qualité de service ou en Anglais QoS (Quality of Service). La qualité de service est utilisé pour prioriser certain type de trafique réseau pour minimiser des choses comme les délais et la perte de paquets.
QoS est souvent utilisé pour prioriser la Voix à travers le trafique de l'IP depuis un téléphone IP pour s'assurer que la qualité de l'audio est correct.
Nous verrons donc tout d'abord une introduction sur les téléphones IP et sur les la voix à travers les VLANs, puis nous verrons ce qu'est le Power over Ethernet (PoE) et nous donnerons une introduction à la qualité de service (QoS) pour mieux comprendre comment cela fonctionne.

Lest téléphones IP sont des téléphones standards qui fonctionnent à travers le réseau des téléphones publique ou en anglais : public switched telephone network (PSTN).
Certaines fois cela est appelé POTS (Plain Old Telephone Service)
Les téléphones IP utilisent VoIP (Voice over IP) qui est une technologie qui permet au téléphones d'appeler à travers un réseau IP comme Internet.
Les téléphones IP sont connectés à un Switch comme n'importe quelle hôte. Il existe cependant une autre option pour connecter ces appareils à un Switch.
Un téléphone IP Cisco ressemble à cela :

Les téléphones IP on 3 port Switch Internes :
- 1 port est le uplink qui connecte au switch externe.
- 1 port est le downlink vers le PC
- 1 port est connecté en interne au téléphone lui même

Il y a donc à l'intérieur du téléphone IP une sorte de petit Switch interne avec 3 ports.
L'un est utilisé pour se connecter le Switch en Ethernet, un autre pour se connecter au PC avec un câble Ethernet. Et le dernier port sert à se connecter en interne au téléphone lui même.
Cela permet au PC et au téléphone IP de partager un seul port Switch. Le trafique depuis le PC passe par l'IP du téléphone vers le Switch.
Il est recommandé de séparer le trafique de la « voix » (depuis le téléphone IP) et le trafique de donnée (depuis le PC) en les plaçant dans des Vlans séparés. Cela peut être fais en utilisant un voice VLAN.
Le trafique depuis le PC sera non balisé, mais le trafique depuis le téléphone sera balisé avec un ID de VLAN.
Nous utiliserons ce réseau pour configurer le Switch :

Voici les commandes à exécuter pour configurer le Switch :

```
SW1(config)#interface gigabitethernet0/0
SW1(config-if)#switchport mode access
SW1(config-if)#switchport access vlan 10
SW1(config-if)#switchport voice vlan 11

SW1#show interfaces g0/0 switchport
Name: Gi0/0
Switchport: Enabled
Administrative Mode: static access
Operational Mode: static access
Administrative Trunking Encapsulation: negotiate
Operational Trunking Encapsulation: native
Negotiation of Trunking: Off
Access Mode VLAN: 10 (VLAN0010)
Trunking Native Mode VLAN: 1 (default)
Administrative Native VLAN tagging: enabled
Voice VLAN: 11 (VLAN0011)
![output omitted]
```

```
SW1(config)#interface gigabitethernet0/0
SW1(config-if)#switchport mode access
SW1(config-if)#switchport access vlan 10
SW1(config-if)#switchport voice vlan 11
```

Le PC1 enverra ici le trafique sans balise. Le SW1 utilisera CDP pour avertir PH1 de baliser le trafique PH1 dans le Vlan 11

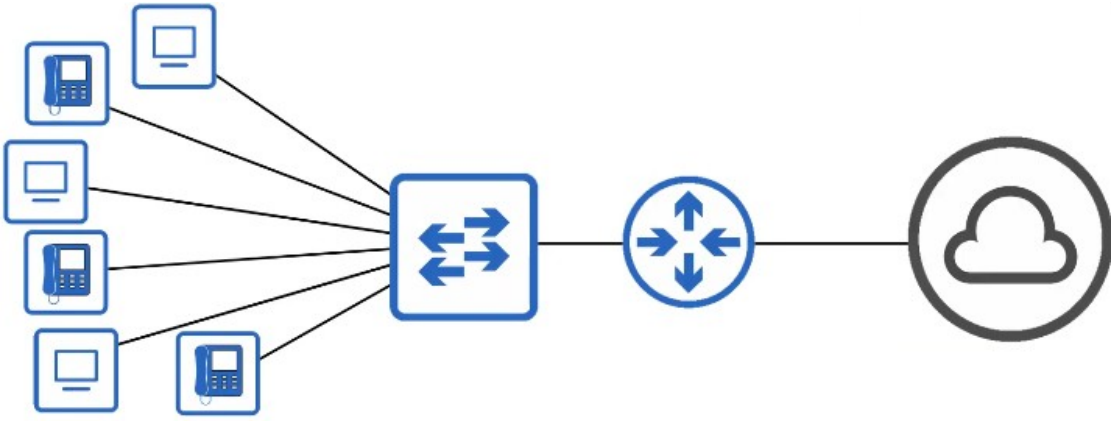
On affiche ensuite la configuration avec la commande :
```
SW1#show interfaces G0/0 switchport
```

On peut voir afficher la vlan utilisé pour le port 0/0 est la Vlan 10
la Vlan utilisé pour la voix est la Vlan 11
Les mode administratif et opérationnel sont en statique.

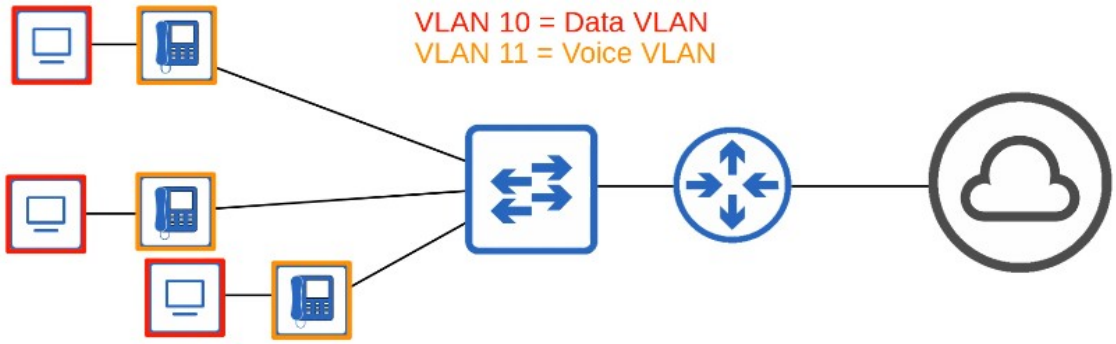
Il est aussi possible d'afficher la configuration trunk avec la commande :
```
SW1#show interfaces trunk
SW1#show interfaces g0/0 trunk
```

```
SW1#show interfaces trunk
SW1#
SW1#show interfaces g0/0 trunk

Port        Mode             Encapsulation  Status        Native vlan
Gi0/0       off              negotiate      not-trunking  1

Port        Vlans allowed on trunk
Gi0/0       10-11

Port        Vlans allowed and active in management domain
Gi0/0       10-11

Port        Vlans in spanning tree forwarding state and not pruned
Gi0/0       10-11
```

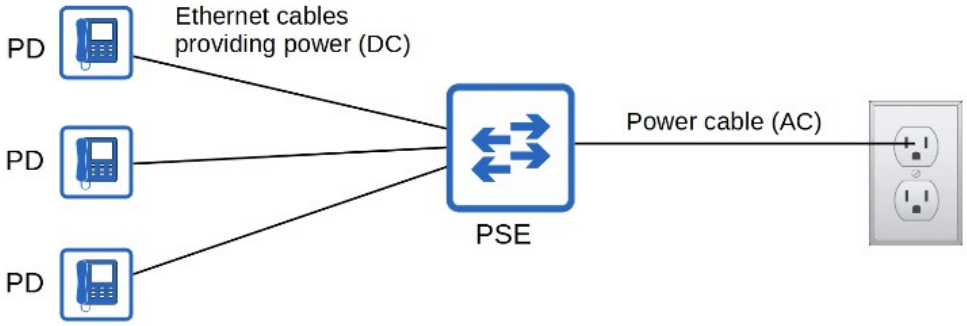
Donc au lieu d'utiliser un câble pour chacun des PC et téléphones IP comme ceci :



Il est possible d'utiliser moins de ports Ethernet sur le Switch en connectant les téléphones IP au PC et d'utiliser une Vlan pour séparer les trafique par exemple VLAN 10 pour les données et le VLAN 11 pour la VLAN de la voix comme sur le réseau suivant :



PoE permet à des équipements avec sources d'énergie, en Anglais (Power Sourcing Equipment (PSE) de fournir de l'énergie aux appareils à alimenter (Powered Devices ou PD en Anglais) à travers un câble Ethernet.
Le PSE est un Switch et les appareils à alimenter sont des téléphones IP, des caméras IP, des points d'accès, etc...
Le PSE reçoit l'énergie depuis une sortie AC, il le convertit en énergie DC et le fournit cette énergie DC vers les appareils à alimenter.



Il faut tout de même faire attention lorsque l'on utilise cette solution, parfois trop d'électricité peut endommager l'appareil électrique.
PoE possède donc une procédure déterminer si l'appareil connecté à besoin d'énergie et combien d'énergie il en aurait besoin.

Lorsqu'un appareil est connecté à un port PoE, le PSE (Switch) envoie de très faible signaux d'énergie, et gère la réponse afin de déterminer combien de puissance l'appareil aurait besoin.
Si l'appareil à besoin d'énergie, le PSE envoie l'énergie et permet à l'appareil de démarrer.
Le PSE continue de gérer les appareils et le montant requis de puissance énergique.
Il y a une fonction appelé Power policing qui peut être configuré pour empêcher un appareil de prendre trop d'énergie.

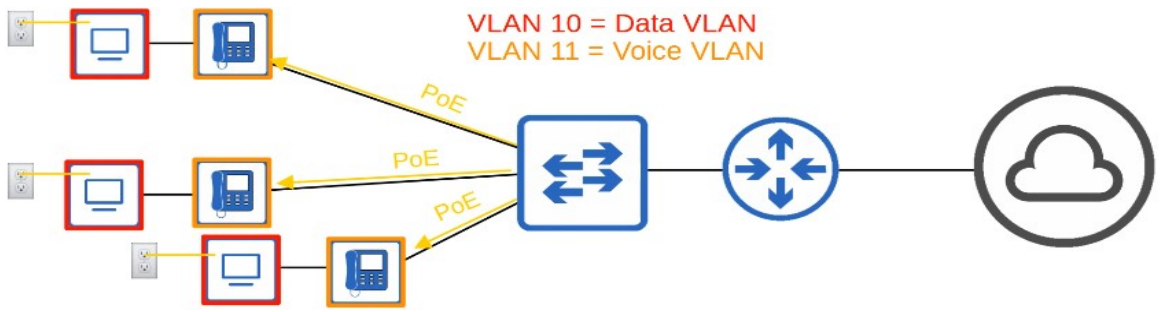
```
SW1# conf t
Enter configuration commands, one per line. End with CNTL/Z.
SW1(config)# int g0/0
SW1(config-if)# power inline police
SW1(config-if)# end
SW1# show power inline police g0/0
Available:800(w) Used:32(w) Remaining:768(w)
Interface Admin  Oper       Admin      Oper       Cutoff Oper
          State  State      Police     Police     Power  Power
--------- ------ ---------- ---------- ---------- ------ -----
Gi2/1     auto   on         errdisable ok          17.2  16.7
```

Voici quelques commandes qui permettent la gestion de l'énergie avec la politique d'énergie :
`power inline police` permet de configurer l'énergie avec des paramètres par défaut : en désactivant le port et en envoyant un message Syslog si l'appareil reçoit trop d'énergie.
Ceci est équivalent à la commande `power inline police action err-disable`
L'interface sera mise en état « error-disabled » et pourra être réactivé avec un `shutdown` suivi d'un `no shutdown`.
`power inline police action log` n'éteint pas l'interface si l'appareil reçoit trop d'énergie. Cela redémarre l'interface et envoie un message Syslog.

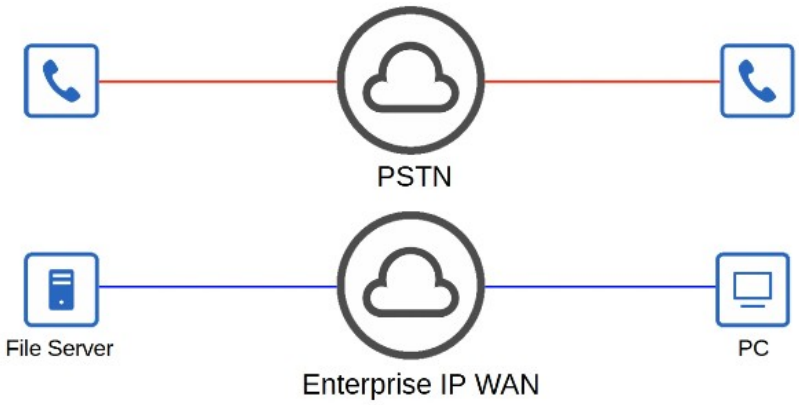
Voici quelques standard de PoE :

Nom	Standard	Watts	Puissance par pair
Cisco Inline Power (ILP)	Crée par Cisco, n'est pas un Standard	7	2
PoE (Type 1)	802.3af	15	2
PoE+(Type 2)	802.3at	30	2
UpoE (Type 3)	802.3bt	60	4
UpoE+ (Type 4)	802.3bt	100	4

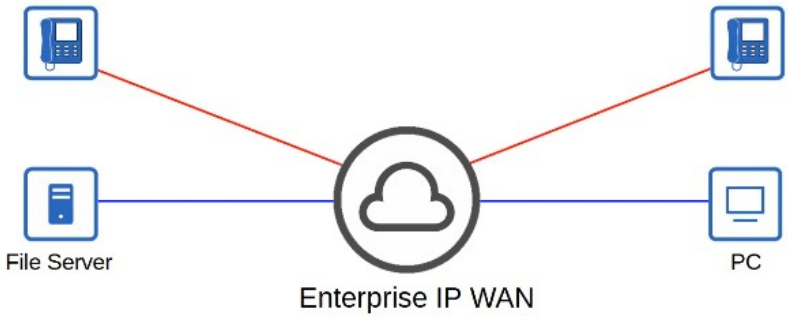
Sur le réseau précédent on peut utilise le PoE pour les téléphones IP et une prise mural pour les PC.



Voyons à présent ce qu'est QoS. Auparavant :
Le trafique de Voix et le trafique des données étaient utilisés par des réseaux totalement séparés.
Le trafique réseau utilisait le PSTN (Public Switch Telephone Network)
Le trafique des données utilisait le réseau IP (WAN entreprise, Internet, etc.)
QoS n'était pas nécessaire car les différents types de trafiques ne se partageaient pas la ban,de passante. Comme



Dans les réseaux modernes les réseaux partagent le même réseau dans lequel passe le téléphone IP, le trafique vidéo, le trafique de données, etc...



La qualité de service permet donc de gagner de l'argent pour des fonctionnalités avancés pour la voix et le trafique vidéo, par exemple l'intégration avec la collaboration de logiciels (Cisco WebEx, Microsoft Teams, etc...)
Les différents types de trafique doivent maintenant se partager la bande passante.
QoS est un ensemble d'outils utilisés par les appareils réseau pour appliquer différents traitements à différents paquets. Par exemple ajouter une priorité à certain type de trafique et réduire la priorité sur un autre.
QoS est utilisé pour gérer les caractéristiques suivantes du trafique réseau :
1. La bande passante : la capacité de la ligne est mesuré en bits par secondes (Kbps, Mbps, Gbps, etc.) Les outils QoS permettent de réserver un certain montant de bande passante pour un certain type de trafique. Par exemple 20% de trafique de voix, 30% de trafique pour un certain type de données, et laisser 50% pour tout les autres types de trafique.

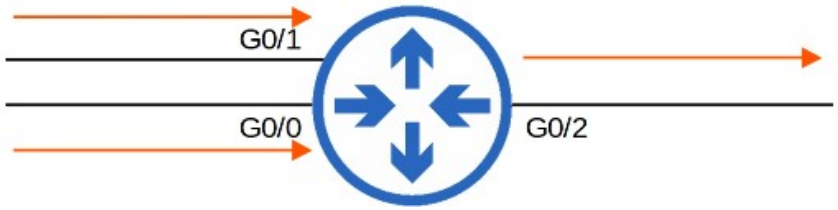
2. Le délai : Il y a plusieurs façons de mesurer le délai, avec le temps que le trafique a pris pour aller depuis la source vers la destination est appelé le « one-way delay »
Le temps que le trafique prend pour aller depuis la source vers la destination et retourner est appelé le « two-way delay »

3. Jitter : est la variation du one-way delay entre le paquet envoyé par la même application.
Les téléphones IP ont un « jitter buffer » pour fournir un délai fixe aux paquets audio.

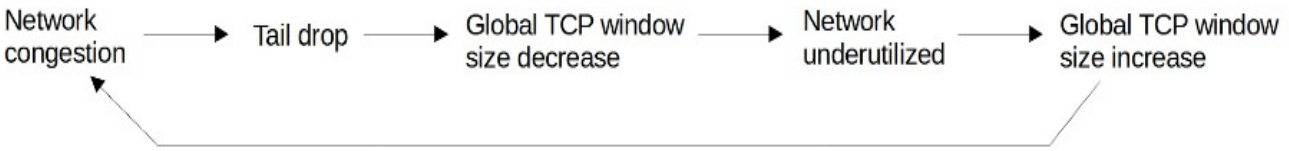
4. Perte : est le pourcentage de paquets envoyés qui n'atteignent pas leurs destinations, cela peut être dû à des câbles endommagés.
Peut aussi être causé lorsque la queue de paquets d'un appareil est pleine et que l'appareil commence à bloquer la réception des paquets.

Il y a plusieurs standard recommandés pour une qualité audio correcte (par exemple les appels audio) :
Le One-way delay : doit être de 150ms ou moins
Le Jitter : de 30ms ou moins
La perte : de 1% ou moins
Si ces standard ne sont pas remplis, il peut y avoir une réduction de la qualité de l'appel audio.

Si le réseau reçoit des messages plus rapidement qu'il ne les repartage à l'interface approprié, les messages sont alors placés dans une queue.
Par exemple sur ce routeur, le trafique est réceptionné plus rapidement depuis les interface G0/1 et G0/0 pour être redirigé vers l'interface G0/2
Par défaut, un message en queue est redirigé dans un First In First Out (FIFO)
Les messages sont envoyés dans l'ordre duquel ils sont reçus. Il n'y a pas de traitement spécial, pour le type de trafique, la queue peut alors devenir pleine et les nouveaux paquets peuvent être perdus, ceci est appelé le tail drop.



Tail drop peut être problématique car cela peut mener à TCP global synchronisation.
Les hôtes qui utilisent le « sliding window » augmentent ou diminuent le taux duquel ils envoient le trafique en fonction du besoin.
Lorsqu'un paquet est perdu il sera alors retransmit. Lorsqu'une perte se passe, l'envoyeur réduira le taux de bande passante à l'interface auquel il envoie le trafique, il augmentera ensuite petit à petit le taux de bande passante à nouveau.
Lorsque la queue est pleine et les tails drop se passent, tous les hôtes TCP qui envoient un trafique vont ralentir le taux de bande passante auquel ils envoient le trafique.
Ils vont ensuite augmenter le taux duquel ils envoient le trafique qui conduira rapidement à plus de congestion, plus de perte de paquets, et ce même processus recommencera à nouveau.
Voici un schéma qui résume le processus :



Une solution qui permet d’empêcher le tail drop et le TCP global synchronisation est le Random Early Detection (RED), lorsque le montant de trafique dans la queue atteint un certain seuil, l’appareil commence à envoyer aléatoirement des paquets perdus depuis le flux TCP.
Ce flux TCP qui perdent des paquets vont réduire le taux auquel le trafique est envoyé, mais cela provoquera la TCP synchronisation, dans lesquelles tous les flux TCP se réduisent et puis augmentent le taux de transmissions au même moment dans des ondes.
Dans le standard RED, tous les types de trafiques sont traités de la même manière.
Dans une version amélioré appelé le Weighted Random Early Detection (WRED) permet de contrôler quelles paquets sont perdus en dépendant de la classe de trafique.