



Haute école d'ingénierie et d'architecture Fribourg  
Hochschule für Technik und Architektur Freiburg

---

## Réseaux IP

### 446. Qualité de Service (QoS)

# Réseaux IP

## 446. Qualité de Service (QoS)

Qualité de service (QoS), Congestion, Files d'attente, Priorités

### Références:

- Les Réseaux (Edition 2005, Pujolle)
- Computer Networks (4th Edition, Peterson, Davie)
- Cisco CCNA course, Connecting Network Module

# Qualité de service

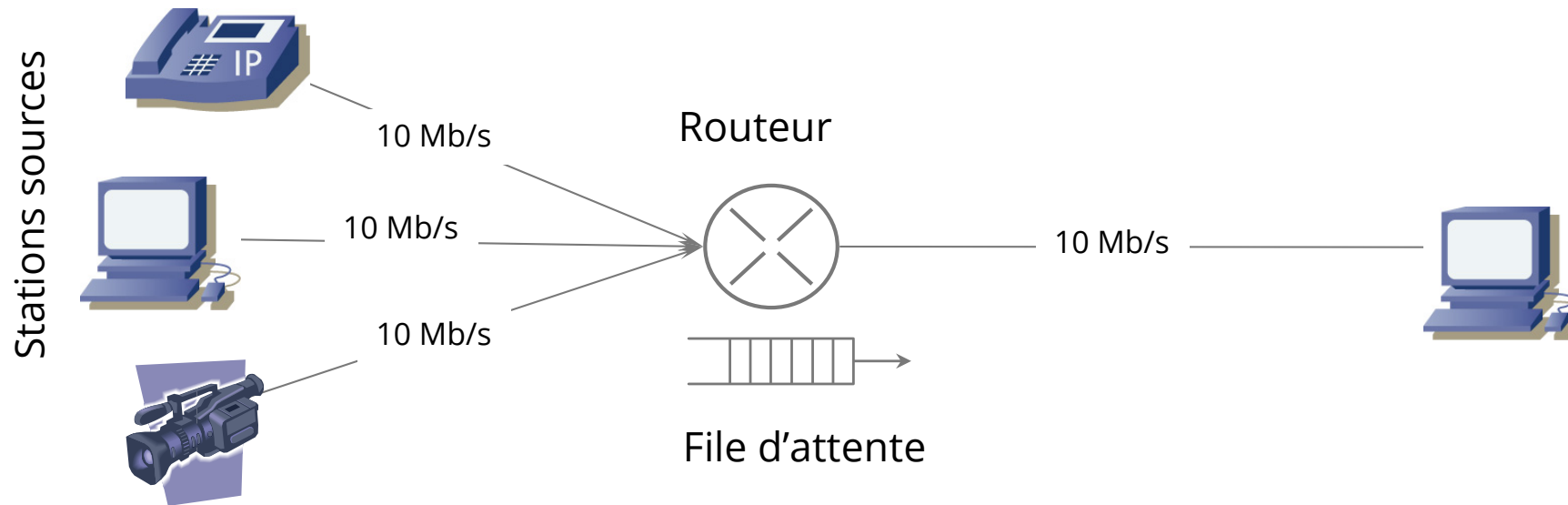
Pendant très longtemps, les réseaux informatiques et Internet ont eu pour unique but de transporter de façon fiable des **données** avec le moins d'erreurs possible. Les données sont le plus souvent transportés avec des protocoles de type TCP garantissant une séquence correcte et une transparence sémantique. La transparence temporelle et le temps réel ne faisait pas partie des attentes sur ce type de réseaux.

La situation actuelle est différente, car ces réseaux transportent souvent simultanément plusieurs types de services aux caractéristiques et exigences très différentes

Services	Caractéristiques et exigences
Téléphonie	débits faibles, contraintes temporelles élevées
Commandes et contrôle	débit faible, contraintes temporelles élevées
Vidéo et TV	très hauts débits, faibles taux d'erreurs, délai de transmission approximativement constant
Sessions interactives (vidéoconférence, Telnet, ...)	débits variables, délais courts et si possible constants
Transfert de données	Débits variables, tolère généralement un délai variable, requiert un très faible taux d'erreurs

La **qualité de service** d'un réseau de transports de paquets est généralement définie en fonction du **taux d'erreurs**, du **délai de bout en bout** et des **variations de ce délai**.

# Congestions



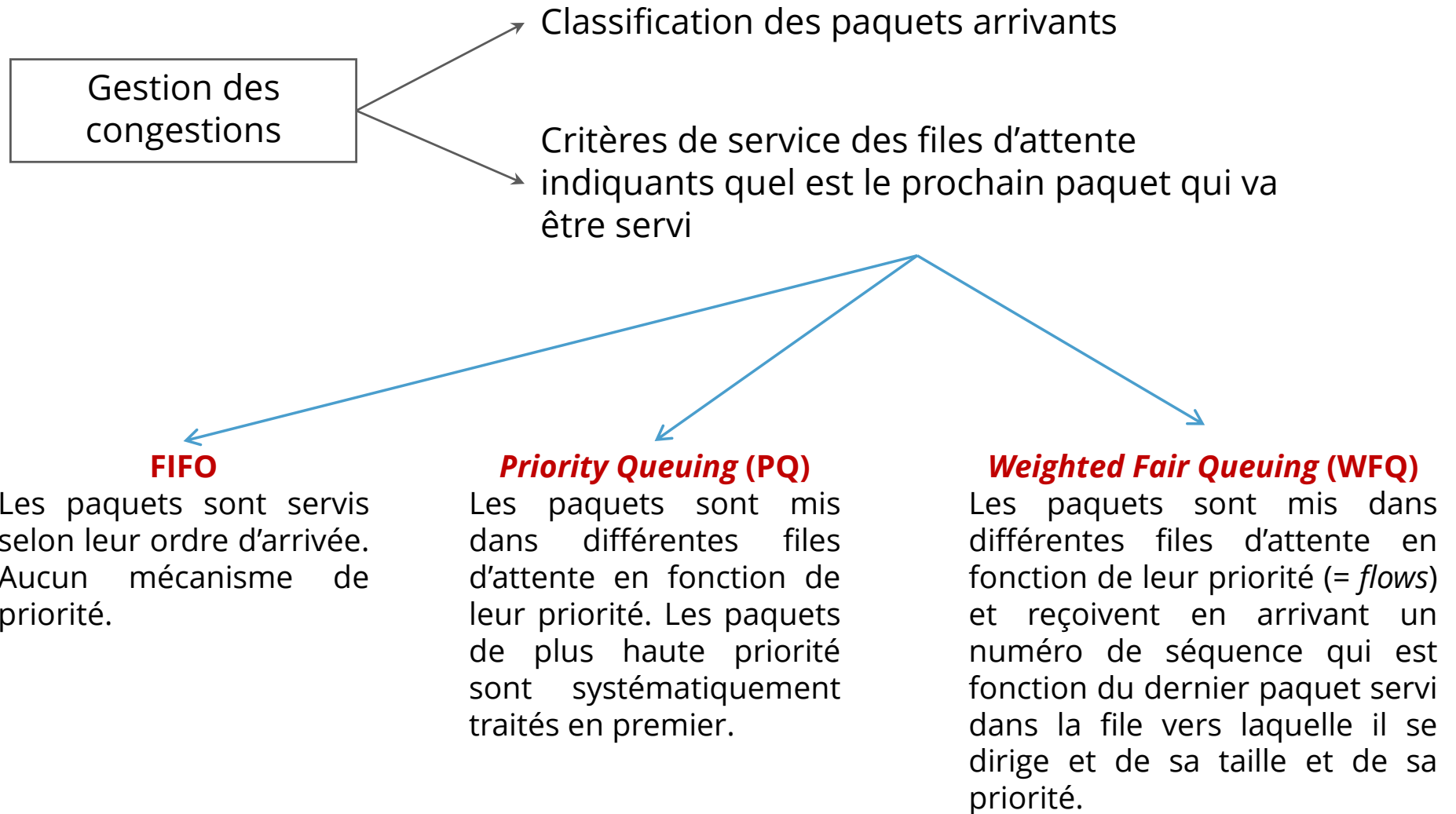
Si toutes les stations sources génèrent 10 Mbit/s simultanément pendant un certain temps, le routeur (ou le switch), même si il est capable de traiter tout ce trafic, va devoir mettre en attente les paquets.

Il est donc probable que les applications en temps réel vont être affectées. Le contrôle des congestions devient alors essentiel pour maintenir la qualité de service. Deux types de techniques sont utilisées en général en parallèle.

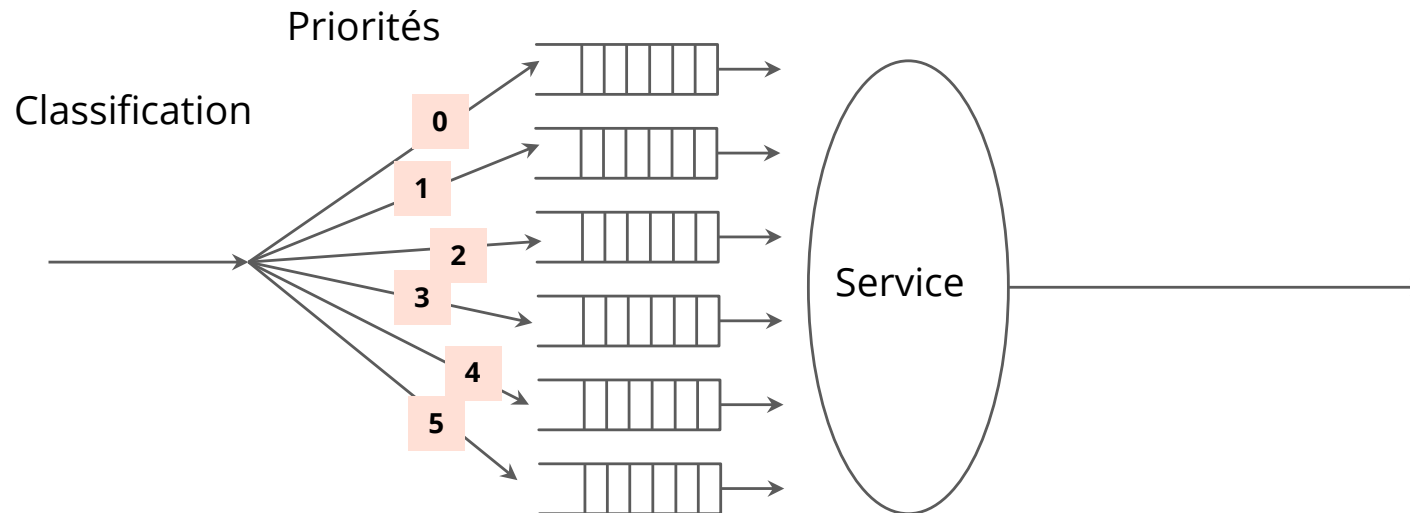
La **gestion des congestions** essaye de **servir les files d'attente** selon des critères de priorités de façon à maintenir la qualité de service.

L'**évitement des congestions** essaye de diminuer le nombre de paquets en attente en demandant aux sources de réduire leur trafic (p.ex. en utilisant les mécanismes de TCP)

# Gestion des congestions



# Priority queuing



## **Priority Queuing (PQ)**

Les paquets sont mis dans différentes files d'attente en fonction de leur priorité. Les paquets de plus haute priorité sont systématiquement traités en premier.

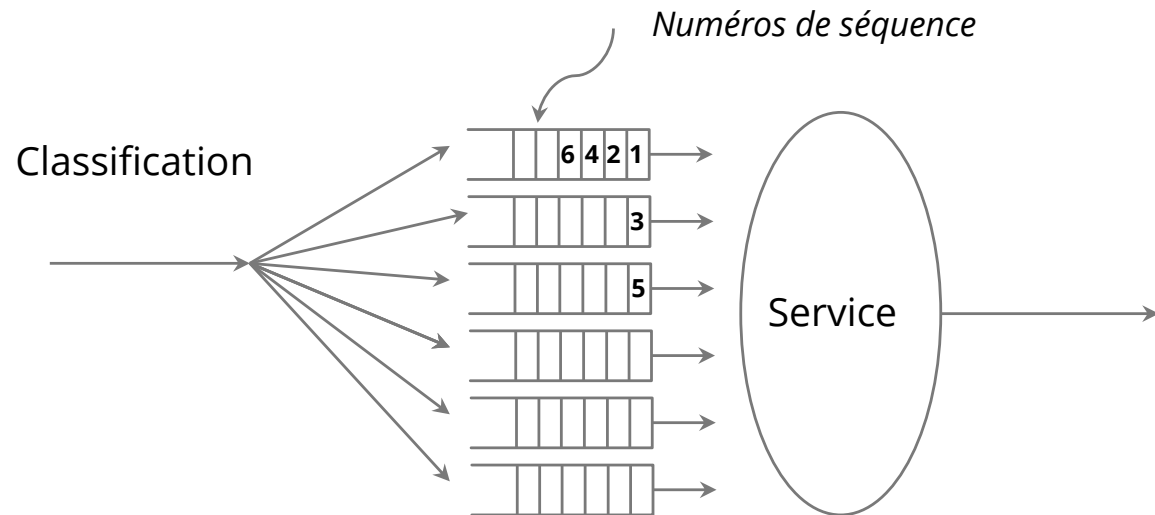
Un paquet ne pourra donc jamais être servi si il y a un autre paquet de plus haute priorité qui est en attente.

Ce type de service basé sur la priorité peut conduire à complètement « affamer » (*starve*) une file d'attente.

# Weighted Fair Queuing

## Critères de classification (IP)

- Bits de type de service dans l'entête IP (*ToS field*)
- Type de protocole
- Adresse IP de source ou destination
- Ports TCP ou UDP de source ou destination



## Weighted Fair Queuing (WFQ)

Les paquets sont mis dans différentes files d'attente en fonction de leur priorité (= *flows*) et reçoivent en arrivant un numéro de séquence qui est fonction du numéro de séquence du dernier paquet servi dans la file vers laquelle il est dirigé et de sa taille et de sa priorité.

Ce mécanisme, plus complexe à mettre en oeuvre que le priority queuing, permet d'éviter qu'un paquet de basse priorité ne passe trop de temps en attente, même si il y a du trafic de plus haute priorité.

Ainsi, un paquet de haute priorité recevra un numéro de séquence avantageux qui ne le fera cependant pas forcément passer devant un paquet de plus basse priorité arrivé avant ou de plus petite taille.