

Compte rendu – Ordonnancement (Cours 7)

1 Introduction

L'ordonnancement (scheduling, queueing) désigne l'ensemble des mécanismes permettant de choisir le prochain paquet à transmettre. Il repose sur l'utilisation :

- d'une ou plusieurs files d'attente (buffers),
- éventuellement de temporisateurs.

L'ordonnancement est généralement implanté au niveau réseau, juste avant la couche liaison, à l'endroit où les paquets sont mis en attente.

2 Types d'ordonnanceurs

On distingue deux grandes catégories :

- **Ordonnanceurs à conservation de travail** : transmission obligatoire dès qu'un paquet est présent dans le buffer.
- **Ordonnanceurs sans conservation de travail** : la transmission peut s'interrompre même si des paquets sont en attente.

3 FIFO – First In First Out

- Le paquet arrivé en premier est transmis en premier.
- Un seul buffer est utilisé.
- L'ordre des paquets d'une même connexion est respecté.
- Le délai d'attente d'un paquet correspond au temps de transmission des paquets situés devant lui.
- Ordonnanceur à conservation de travail.

4 LIFO – Last In First Out

- Le dernier paquet arrivé est transmis en premier.
- Les paquets présents dans le buffer ne sont transmis que s'il n'y a plus d'arrivées.
- Usage pratique limité.

5 Priority Queueing (PQ)

- Plusieurs files d'attente, chacune avec une priorité.
- La file de plus haute priorité est toujours servie en premier.
- Une file n'est servie que si toutes les files de priorité supérieure sont vides.
- Risque de **starvation** pour les flux de faible priorité.
- Le délai maximal d'un paquet prioritaire en tête de file est :

$$\frac{L_{\max}}{C_l}$$

où L_{\max} est la taille maximale d'un paquet non prioritaire et C_l la capacité du lien.

6 Ordonnanceur idéal : GPS

Le **Generalized Processor Sharing (GPS)** est :

- un modèle fluide idéal,
- à conservation de travail,
- assurant un partage maîtrisé de la bande passante,
- garantissant la séparation des flux et l'absence d'effet négatif des flots agressifs.

Chaque connexion i reçoit un poids f_i , lui garantissant un débit :

$$g_i = \frac{f_i}{\sum_j f_j} C_l.$$

7 WFQ – Weighted Fair Queueing

WFQ est une approximation paquet-par-paquet du GPS (PGPS). Caractéristiques :

- chaque file d'attente possède un poids,
- un temps de fin de service $F(i, k)$ est attribué à chaque paquet,
- le paquet avec le plus petit F est transmis en premier,
- assure la séparation des flux,
- l'écart entre GPS et WFQ est :

$$F_{WFQ} - F_{GPS} \leq \frac{L_{\max}}{C_l}.$$

Les formules fondamentales :

$$S(i, k) = \max\{F(i, k - 1), V(a(i, k))\},$$

$$F(i, k) = S(i, k) + \frac{L(i, k)}{f_i}.$$

$V(t)$ représente le nombre de tours de service dans le modèle GPS.

8 Weighted Round Robin (WRR)

- Plusieurs files, chacune avec un poids.
- Le nombre de paquets servis à chaque tour dépend du poids.
- Approximation grossière de GPS.
- Inéquitable si les tailles des paquets varient.

9 Deficit Round Robin (DRR)

- Chaque file reçoit un quantum Q_i (en bits).
- À chaque tour, une file peut envoyer autant de paquets que la somme de son quantum et du déficit accumulé.
- Si un paquet ne peut être transmis, son dépassement est reporté au tour suivant.
- Approche plus équitable que WRR lorsque les tailles de paquets varient.

10 Courbes de service

Une courbe de service décrit la quantité minimale de trafic qui peut être transmise par un nœud en fonction du temps. Caractéristiques :

- dépend de l'ordonnancement,
- liée aux périodes d'activité (busy periods),
- combine les effets liés à la transmission, au buffer, etc.,
- utilisée avec les courbes d'arrivée pour fournir des garanties.

11 Modèle IETF (RFC 2212)

Le modèle ajoute deux termes d'erreur au modèle fluide :

- C/R : délai lié aux limitations de transmission (ex. paquetisation),
- D : délai indépendant du débit (temps de transit, non-préemption, etc.).

$$\text{Service minimal garanti} = R(t - (C/R + D)).$$