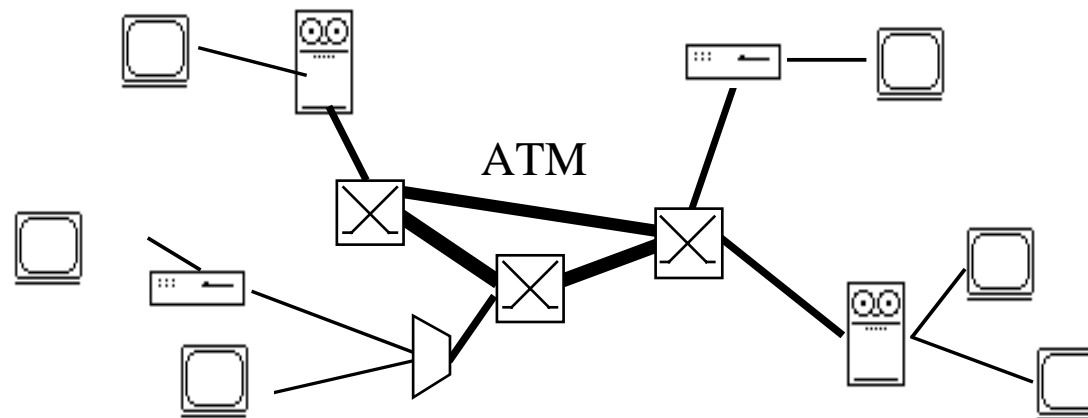


Les mécanismes de contrôle de congestion dans ATM

©b

par Bernard Cousin

(/home/kouna/d01/adp/bcousin/Cours/ATM-congest.fm- 11 Mars 1999 14:25)



Plan

- Introduction : la congestion
- Les techniques de contrôle de congestion
- Le contrôle d'établissement des connexions
- Le contrôle de conformité du trafic
- La notification de congestion
- Le rejet sélectif de cellules
- Conclusion

1. Introduction

- . Les ressources du réseau sont limitées :
débit des liens, capacité de stockage des commutateurs, etc.
- . Les applications soumettent des trafics variés :
temporellement et quantitativement.
- . L'utilisation des ressources doit être optimisée :

⇒ **le multiplexage statistique** :

on alloue à chaque connexion un débit inférieur à son débit crête en supposant que la probabilité que toutes les sources transmettent en même temps soit faible (plus le nombre de connexions multiplexées est grand plus cela est statistiquement probable).



Congestion

La congestion des liens : Impossible !



contrôle d'accès réglé par la politique d'ordonnancement (scheduling) des cellules par les commutateurs.

La congestion des commutateurs : de leurs espaces de stockage (buffer).



encombrement



retard des cellules



débordement



pertes des cellules

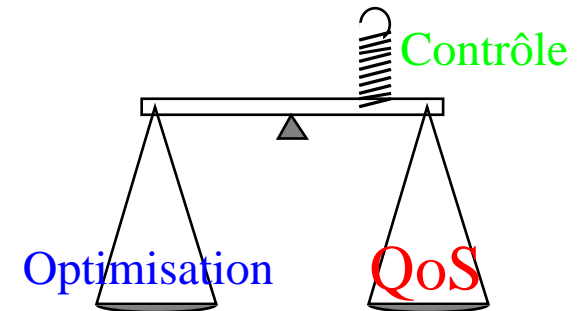
1.1. Le contrôle de congestion

Besoins contradictoires :

- . pour les usagers et leurs applications :
 - ⇒ garantir la qualité du transfert de leurs données (QoS : taux de perte, délai, débit, etc).
- . pour les opérateurs :
 - ⇒ optimiser l'utilisation des ressources.



Le contrôle de congestion gère ce compromis.



Propriétés des mécanismes de contrôle :

- flexibilité (s'adaptent à tous les types de trafics)
- efficacité (faible complexité, peu de ressources)
- robustesse (permanence du service en toutes circonstances)

1.2. Difficultés du contrôle de congestion

Haut débit :

. les contrôle réactifs sont peu efficaces :

⇒ pendant le délai d'aller et retour une quantité gigantesque de données a le temps d'arriver (de submerger le réseau).

⇒ **capacité du réseau !** :

- LFN (“Long fat network”),
- débit x délai.

	débit (Mbit/s)	longueur (km)	capacité (Mbit)
Ethernet	10	2	0,0004
ATM	155	10000	16
X25	0,048	1000	0,0005

Services multiples :

. les applications ont des besoins très variés :

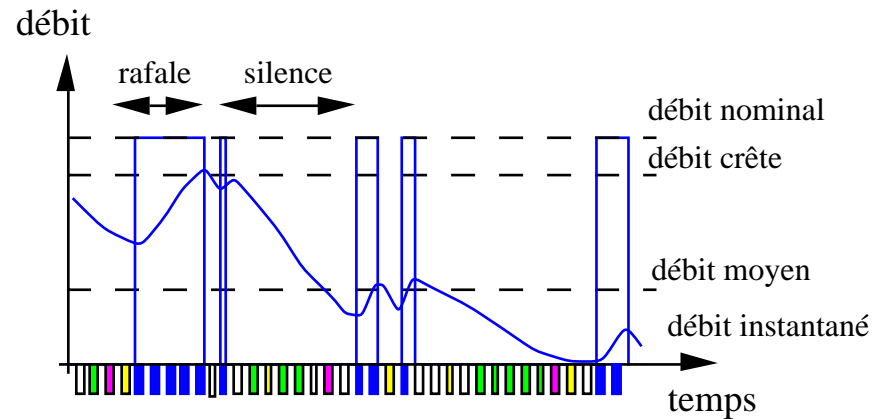
- taux de perte nul, faible, quelconque, etc.
- délai de transmission constant, variable, infini, etc.

Types de trafic multiples :

. contant, périodique, sporadique, continûment variable, variable par palier, quelconque, etc.

1.3. Les débits

- . Débit crête (PCR : peak cell rate) :
 - débit maximum atteint,
- . Le débit nominal de la liaison.
- . Débit moyen (MCR : mean cell rate) :
 - débit moyen sur un intervalle de temps
- . Le débit instantané.
 - ⇒ Débit cellulaire.
- . On distingue:
 - des périodes d'activité (burst)
 - des périodes de silence.



1.4. Gigue

Gigue : variation du délai de transmission,

- . **gigue d'insertion** :
 - l'instant d'arrivée des données de celui de la cellule,
 - retard d'une durée cellulaire.
- . **gigue de multiplexage** :
 - les cellules de différentes connexions sont multiplexées sur la même liaison,
 - à tout moment une seule cellule est émise sur la liaison, les autres attendent !
 - politique d'ordonnancement des cellules des différentes connexions.
- . **gigue de charge** :
 - les délais introduits par le réseau dépendent de sa charge (longueur des files d'attente, durée des traitements, etc).
- . **gigue de routage** (actuellement pas de re-routage dans ATM) :
 - si la route empruntée par les cellules est modifiée, le délai de transmission est modifié.

⇒ la gigue influe sur la forme du trafic :

le trafic reçu n'est plus le trafic émis !



conformité du trafic (mesure du débit)

1.5. Niveaux d'analyse du trafic

⇒ Connexion :

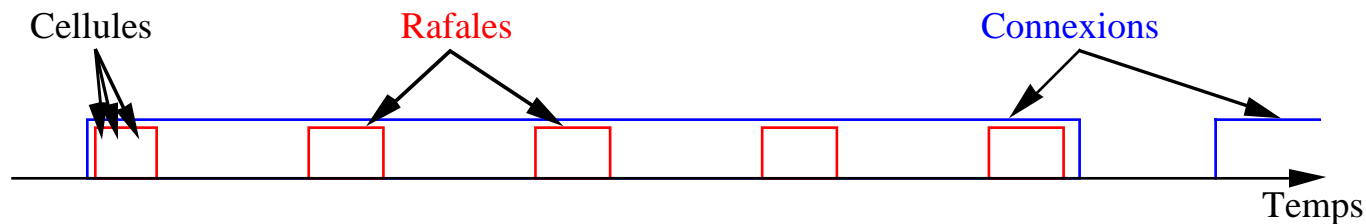
- sa nature (variable, en rafale, constante, etc),
- la bande passante requise,
- la qualité de service (QoS : Quality of Service):
taux d'erreur admissible, délai maximum, variation du délai (gigue), etc.
- échelle de temps : quelques secondes à plusieurs jours.

⇒ Rafale (burst) :

- fréquence, longueur, intensité (sporadicité : burstiness),
- un message ⇒ des cellules !
- échelle de temps : la milliseconde.

⇒ Cellule :

- échelle de temps : la microseconde.



2. Les techniques de contrôle de congestion

2.1. Contrat de trafic

Les applications et le réseau **négoient un contrat** :

- le réseau ATM utilise le mode connecté,
- lors de la phase d'établissement de la connexion (ou pendant le transfert des données),

⇒ **type de trafic** :

. CBR : constant bit rate, VBR-rt or -nrt : variable bit rate (real time), ABR : available bit rate, UBR : unspecified bit rate.

⇒ **descripteur de trafic** (débits et QoS) :

- . PCR, SCR, MCR : peak, sustainable and minimum cell rates.
- . CLR : cell loss ratio.
- . CTD : cell transfer delay.
- . BT : burst tolerance, CDV : cell delay variation.

⇒ pour les 2 niveaux de trafic : CLP=0 et CLP=0+1.

⇒ pour les 2 sens : aller et retour.

2.2. Contrôle de flux

Gestion de la disponibilité du récepteur :

- . occupation des tampons de stockage.
- . capacité de traitement des données.

Protocole Xon/Xoff

- . peu précis ou trop contraint !

Sliding window (fenêtre coulissante)

- . utilisé par de nombreux protocoles :
 - HDLC, X25.3, TP, TCP, TPX, SSCOP, etc.
- . numérotation des données (modulo la capacité maximale du champ)
- . acquittement,
- . largeur de la fenêtre : nombre de données pouvant être émises par anticipation
 - implicitement : fenêtre de largeur fixe),
 - explicitement : crédit.
- . couplé au contrôle d'erreur.

2.3. Sliding Window

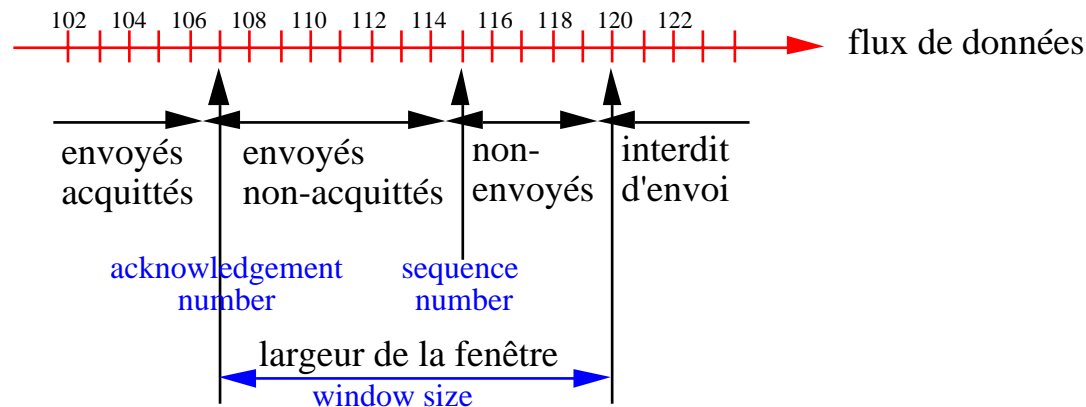
Mécanisme permettant à la fois :

- Le contrôle de flux et de congestion.
- Le contrôle des pertes, duplication, déséquentialité.
- La récupération des erreurs par retransmission.
- L'optimisation de l'utilisation de la connexion par l'envoi anticipé de paquets

(avant que les octets des paquets précédents soient acquittés).

Basé sur l'identification des données (octets ou des paquets) :

⇒ leur numérotation (modulo) .



2.4. Méthodes de contrôle de congestion

Défini par la recommandation I.371 de l'ITU-T.

Principales méthodes :

. **Préventives** :

- le contrôle d'admission (d'établissement des connexions),
- le contrôle de trafic.

. **Réactives** :

- la notification de congestion.
- le rejet sélectif de cellules.

temps de réaction,
techniques d'évitement et correctives.

3. Le contrôle d'établissement des connexions

3.1. Le CAC

CAC (Connexion admission control) :

☞ Contrôle les établissements de nouvelles connexions

. Analyse de la demande (descripteur de trafic)

⇒ Evaluation de la bande passante équivalente

. Recherche de chemin optimal,

. **Réservation des ressources.**

☞ Optimisation :

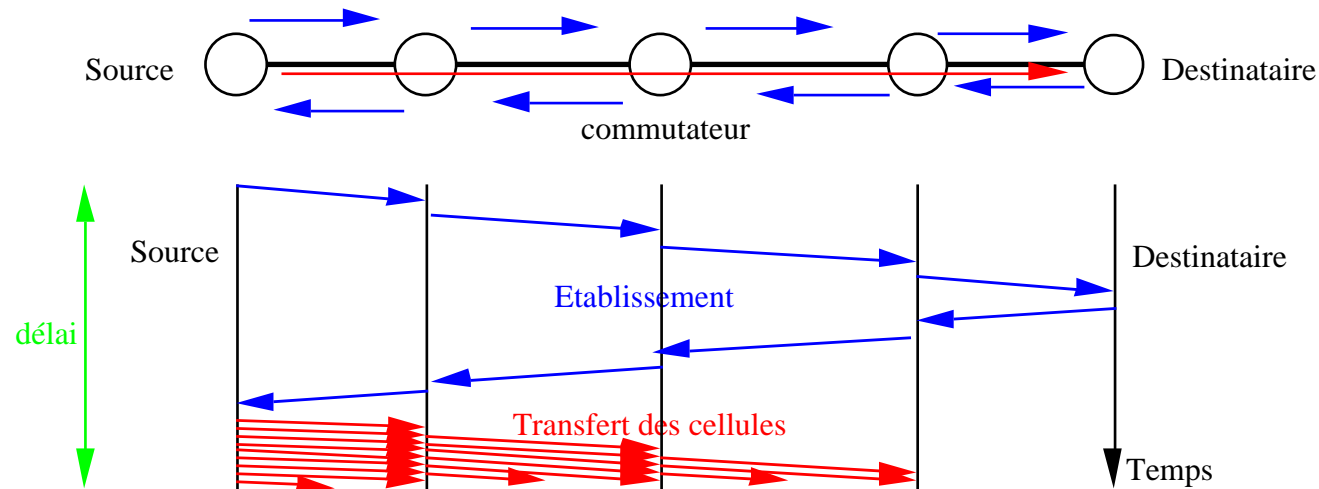
. multiplexage statistique = surallocation

⇒ faible probabilité d'un grand nombre de rafales simultanées

⇒ mais pas nulle ⇒ **perte de cellules !**

3.2. L'établissement

- La durée d'établissement (de négociation) de la connexion est longue :
- . délai de propagation de la demande de la source au destinataire aller et retour.
 - . réservation des ressources au sein de chaque commutateur traversé.



3.3. Le type de connexion

Connexions permanentes ou semi-permanentes :

- . besoins de communication entre partenaires en nombre réduit et stable.

Utilisation de VPC:

- . conduits virtuels préétablis,
- . entre deux points de trafic importants,
- . domaine de la gestion du réseau (configuration générale).

FRP : Fast reservation protocol

- . Les données accompagnent la demande d'établissement de la connexion.
- . Si la connexion est refusée, les données sont détruites !

4. Le contrôle de conformité du trafic

4.1. Contrôle de trafic

Usage parameter control (UPC) (source policing , bandwidth enforcement):

- **surveillance** des paramètres du contrat,
- durant la phase de transfert des données.
- **protection** des ressources du réseau contre une inadéquation entre les paramètres du contrat et le comportement réel du trafic :

- ☞ utilisation malveillante,
- ☞ erreurs involontaires.

Idéalement :

- capable de détecter toute situation illégale,
- réaction rapide,
- transparent au trafic conforme,
- simple et efficace.

Existe aussi entre deux réseaux d'opérateurs différents :

- . NPC (network parameter control)

4.2. Techniques

Contrôle de conformité :

- à l'accès du réseau public (UPC),
- entre les réseaux (NPC).

Actions sur les cellules non-conformes :

. **destruction**

- les cellules non-conformes sont détruites,
- c'est trop tôt !

. **marquage** (tag)

- les cellules non-conformes sont marquées :

☞ CLP bit de la cellule

- en cas de congestion les cellules marquées sont prioritairement détruites.

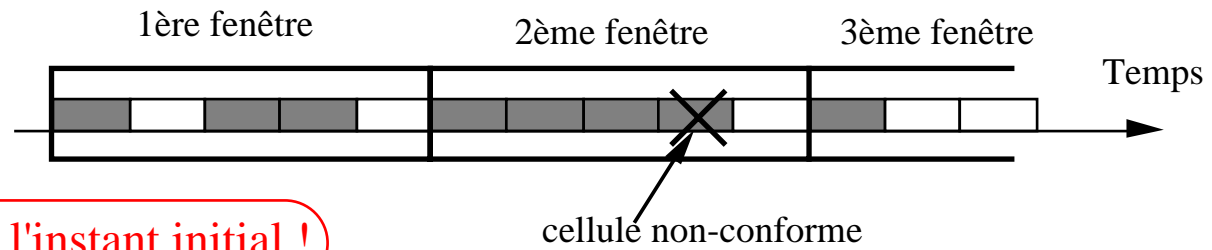
. **réordonnancement temporel** (lissage : shaping)

- les cellules non-conformes sont retardées,
- accumulation dans un tampon du contrôleur.

4.3. Contrôle par fenêtre

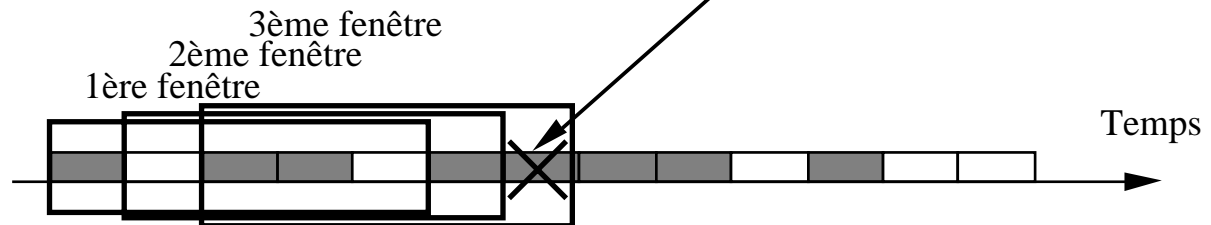
- . La fenêtre est un intervalle de durée fixe :
 - W : la largeur de la fenêtre en durée cellulaire.
- . Le nombre de cellules autorisées par intervalle de temps :
 - N , $N < W$: taille de la rafale.
- . débit cellulaire conforme :
 - $N/W * D$ (avec D : débit nominal)

Jumping window :
 $W=5, N=3$.



Dépend de l'instant initial !

Moving Window :



Trop de calcul !

[EWMA] : Exponentially weighed moving window, [TRJ] : Triggered jumping window

4.4. Leaky Bucket

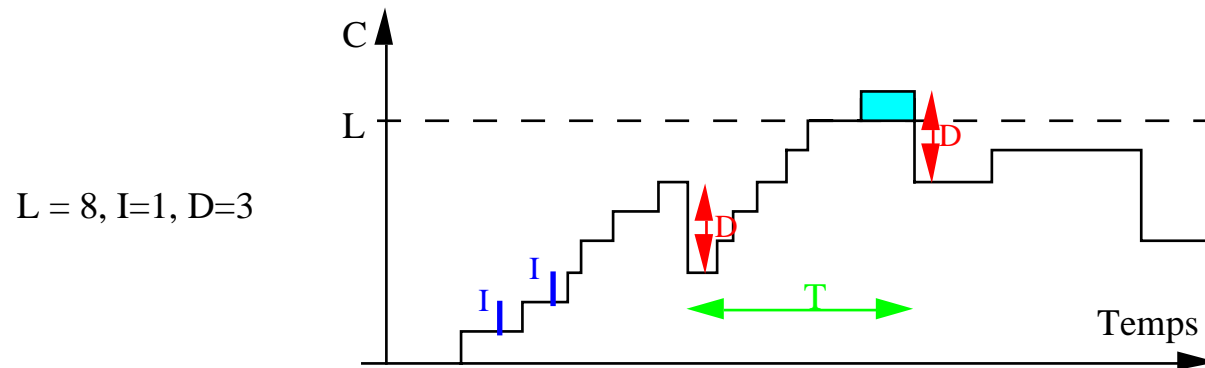
Seau percé [J.Turner 88]:

- variable C : contenance instantané du seau
- constante L : capacité maximum du seau
- constante D : taille de la fuite
- constante I : taille du verre

Débit cible : D/I

Fonctionnement :

- initialement : $C=0$,
- arrivée d'une cellule : si $C < L+I$ alors $C = C+I$
sinon débordement,
- périodiquement (T) : $C = C - D$ (si $C > 0$).



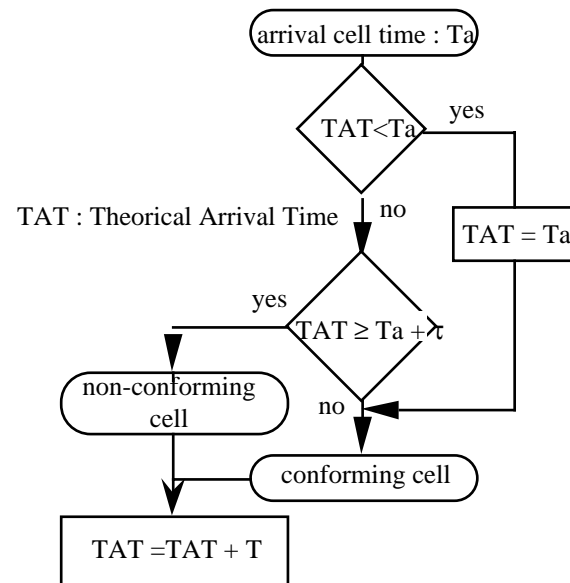
4.5. CGRA

Generic cell rate algorithm : GCRA(T, τ)

- . normalisé l'ATM Forum et l'ITU_T [I.371]
- . identique à "**Continuous state leaky bucket**" (à valeurs réelles)

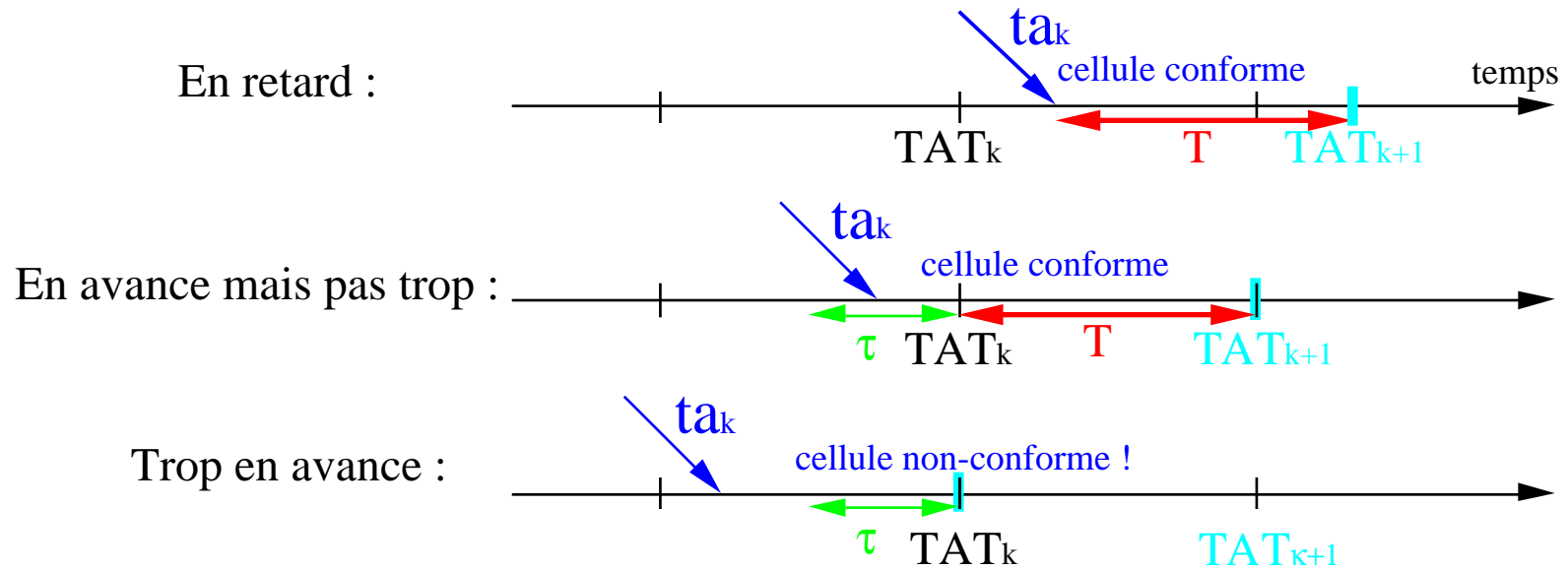
$1/T$: débit cellulaire contrôlé,

τ : tolérance sur le temps de propagation des cellules.



Un double CGRA par connexion

4.6. Exemple

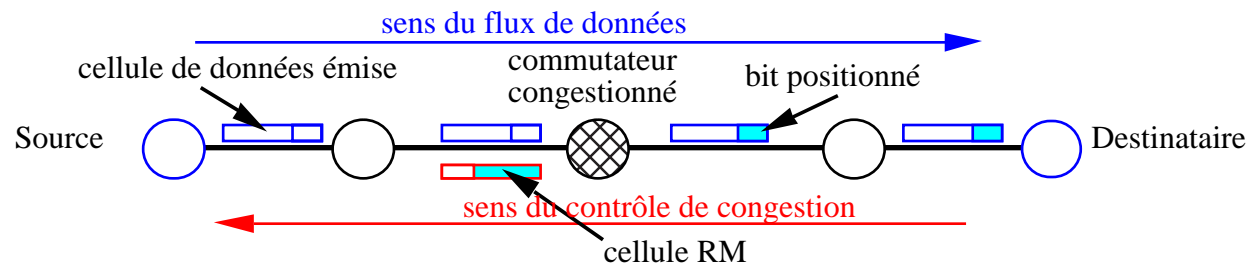


Ta : date d'arrivée de la cellule
 TAT : date théorique de la cellule
 $1/T$: débit de référence
 τ : tolérance

5. La notification de congestion

5.1. Présentation

- . Envoi d'une indication explicite de congestion :
- . Par les commutateurs :
 - ☞ lors d'une perte de cellules
 - c'est trop tard !
 - ☞ dépassement de seuils :
 - taux de perte cellulaire,
 - taux d'occupation (des tampons),
 - débit, etc.
- . Demande de diminution du débit de la source
 - ☞ prise en compte optionnelle).



5.2. Impémentation

☞ EFCI : explicit **forward** congestion indication

- dans les cellules de données :

[bit 3 du 4ème octet (bit EFCI du champ PTI) pour les cellules de données (dont le bit 4 est à 0)].

- utilisation possible des mécanismes de contrôle de congestion des couches supérieures :

. messages spécifiques (destinataire -> source)

☞ **Backward** !

- réduction du temps de réaction.

- traitements complexes au sein de chaque commutateur.

- utilisation de cellules spécifiques (RM: resource management cell)

[code 110 du champ PTI dans l'entête de cellule].

- plus précise : les cellules contiennent plus d'informations (débit explicite, actuel, minimum, longueur des files d'attente, numérotation des cellules RM, etc).

Robustesse

- . La perte de cellules :
 - ☞ de données : perte de précision
 - ☞ RM : perte de détection
- . Emission périodique de cellules de contrôle de congestion
(PRCA : Proportional rate control algorithm)

6. Le rejet sélectif de cellules

6.1. Introduction

Lors d'une congestion effective, il faut choisir les cellules à détruire :

- . les cellules les moins prioritaires (CLP=1),
- . les cellules marquées par l'UPC ou le NPC ne respectant pas le trafic (CLP=1).

Les cellules des connexions les moins importantes :

- . connexion de trafic UBR,
- . connexions ayant le CLR le plus élevé,
 - ☞ RUV : relative usage value.

EPD : early packet discard

- . généralement les cellules successives d'un même connexion sont sémantiquement liées.
- . perte d'une cellule --> perte de la totalité du message.
- . destruction de toutes les cellules du message à partir de la congestion.
- . marque de fin de message

[bit 2 du 4ème octet (bit ATM_user-to-user.indication du champ PTI) pour les cellules de données (dont le bit 4 est à 0)]

7. Conclusion

Tous les mécanismes décrits précédemment ne sont pas utiles pour tous les types de trafic :

☞ Ils sont utilisés par ABR.

Autres mécanismes :

☞ **Adaptation** des paramètres du trafic (Bandwidth renegotiation)

- ABT : ATM block transfer,
- des demandes de modification des paramètres précèdent les blocs de

données.

☞ **Re-routage**

- contournement des pannes et des congestion,
- gestion du réseau.

☞ **Credit-based Flow control**

- fenêtre coulissante entre commutateurs adjacents,
- temps de réaction plus court,
- complexité des commutateurs.