



QoS et Multimédia

SIR / RTS

Gestion de la congestion

Isabelle Guérin Lassous

Isabelle.Guerin-Lassous@ens-lyon.fr

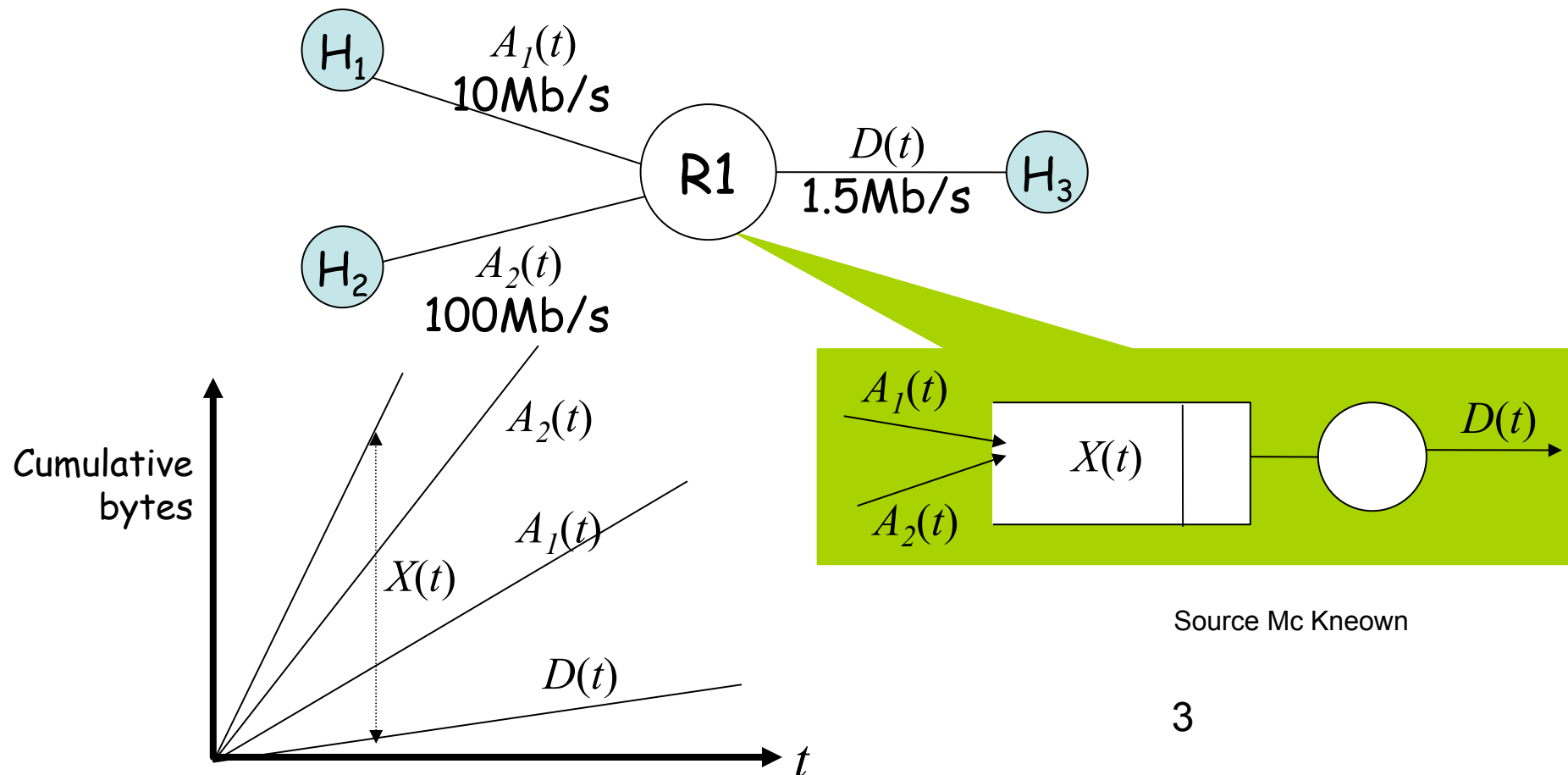
<http://perso.ens-lyon.fr/isabelle.guerin-lassous>

Introduction

- Cours précédent
 - Mécanismes effectifs si pas de congestion
- Actions préventives pour éviter la congestion
- Mais difficile de l'éviter complètement
 - Il faut alors être en mesure de retourner à un état normal dans le réseau
- Anticiper et contrôler

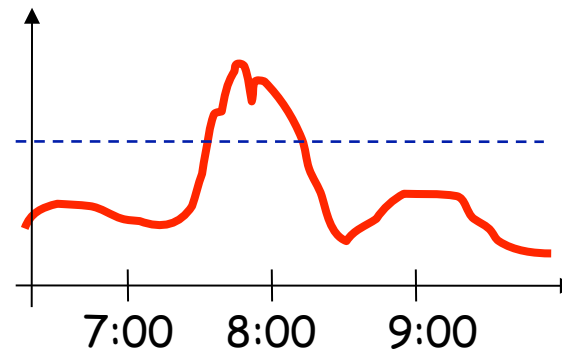
Congestion

- Quand les demandes d'utilisation des ressources sont plus grandes que les capacités de ces ressources
 - Perte de paquets

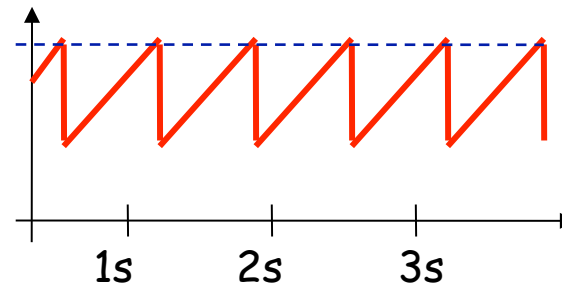


Causes de congestion

Trop d'utilisateurs (qui génèrent trop de trafic)



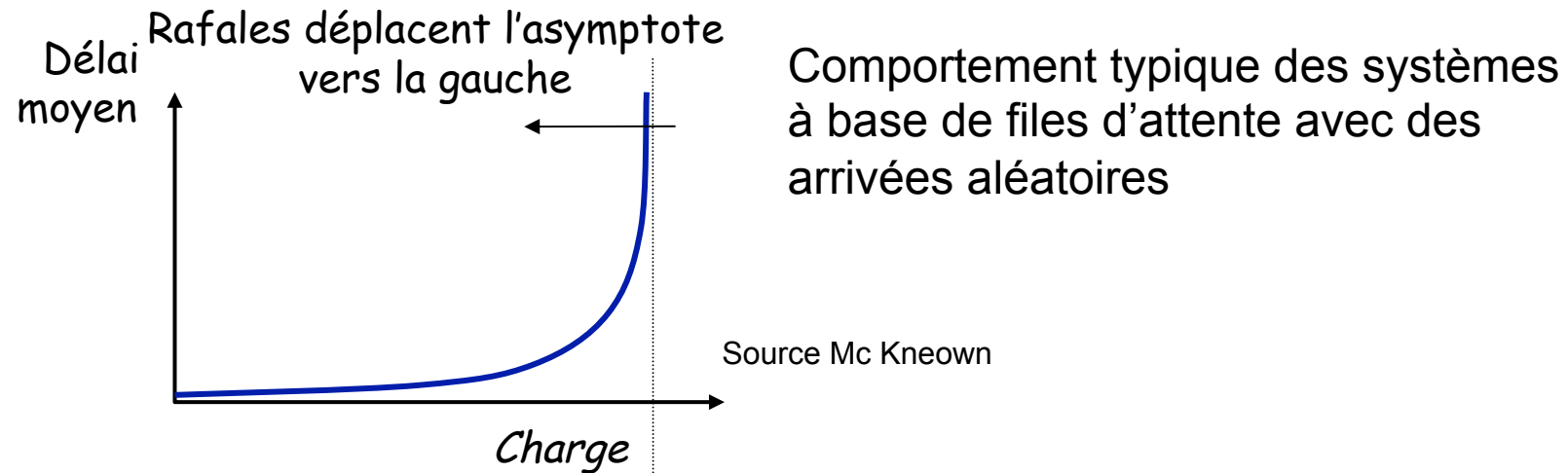
Flux TCP utilisent toute la bande passante



Source Mc Kneown

Panne sur un lien

Délai et taux de pertes



Impact de la taille du buffer sur le délai et le taux de pertes

Règle historique au dimensionnement des buffers

- Papier de 94
 - Mesures expérimentales d'au plus 8 flux TCP sur un lien de $C = 40 \text{ Mb/s}$
- $B = \text{RTT} \times C$
 - $\text{RTT} = 250\text{ms}$
- Recommandation dans les routeurs d'accès et de cœur
 - Si $C = 10$ ou 40 Gb/s , buffer de 2,5 à 10 Gb
 - Pas si simple

Dimensionnement des buffers

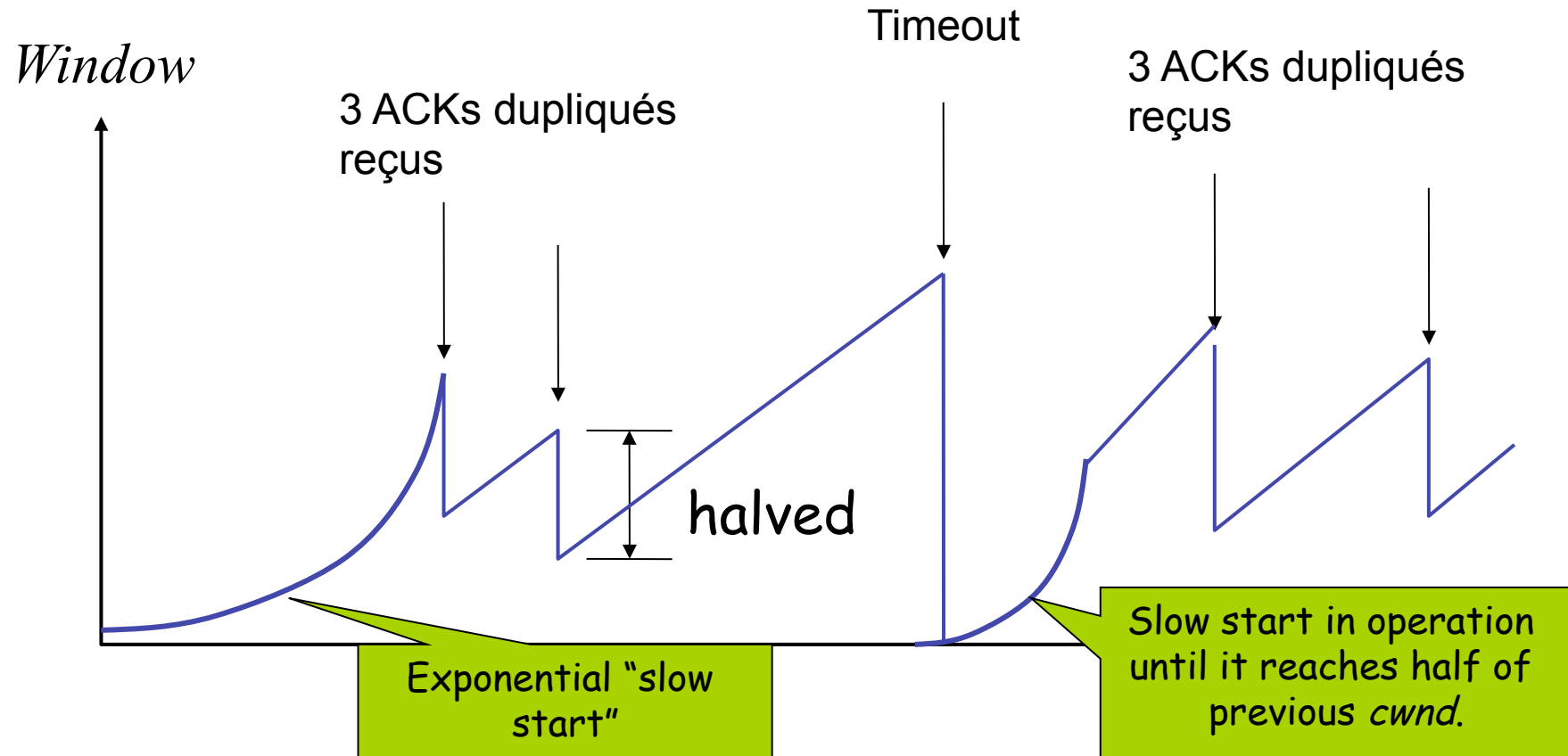
- Question d'actualité
- Partisans des grands buffers
 - Trafic dynamique qu'il faut absorber
 - Changement de route, panne, variation du trafic
 - RTT difficile à estimer
- Partisans des petits buffers
 - Beaucoup de flux non synchronisés
 - Il faut des délais de bout-en-bout raisonnable
 - Quitte à contrôler le trafic en entrée

TCP

Contrôle de congestion le plus populaire
dans l'Internet

TCP

Hôte, boucle fermée
Fenêtre de congestion



TCP

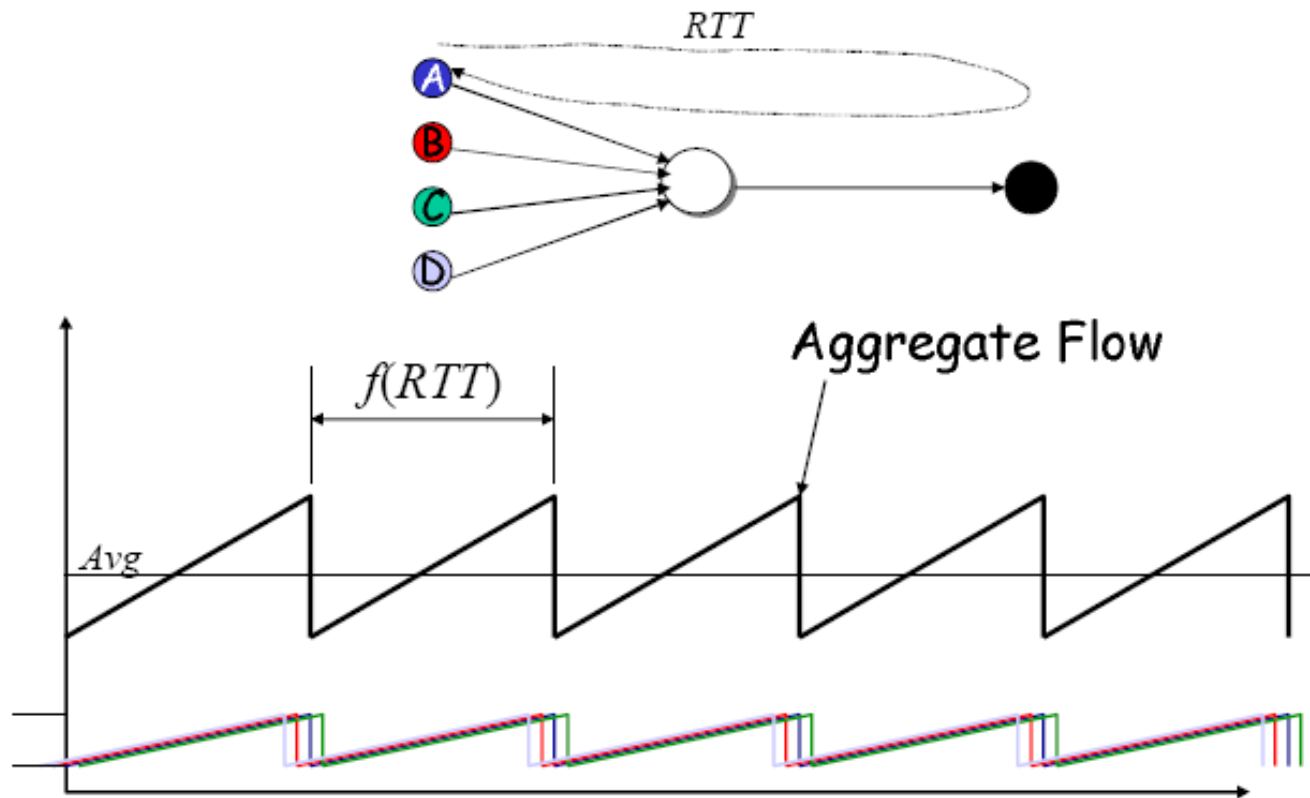
- Plein de versions de TCP
 - Souvent adaptées à chaque type de réseau
 - Couche transport implantée dans les systèmes d'exploitation a peu évolué
- A mi-chemin entre TCP et UDP
 - Contrôle de congestion au-dessus d'UDP au niveau applicatif
 - e.g. RTMFP d'Adobe

Active Queue Management

Gestion passive de files d'attente

- On attend que la file se remplisse complètement puis on se pose la question de quoi faire
- Drop tail ou tail-drop
- Drop front on full
- Rejeter n'importe quel paquet dans la file aléatoirement

Synchronisation des sources TCP



Source Mc Kneown

AQM

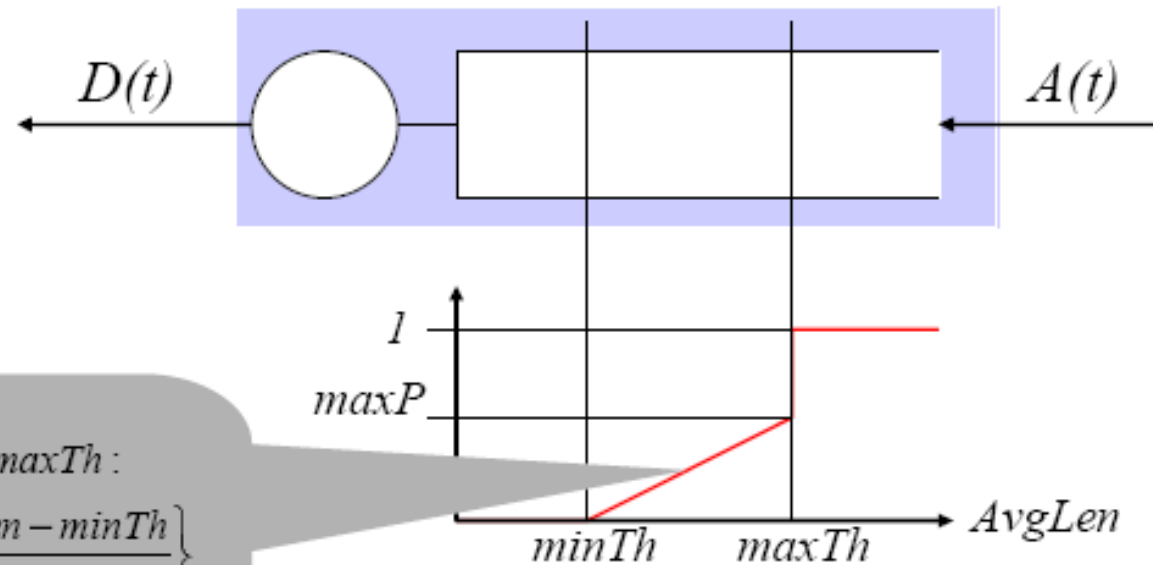
- On tente de traiter proactivement le problème
- Techniques basées sur
 - Des informations implicites
 - Destruction de certains paquets
 - Des informations explicites
 - Dans les paquets de données ou via des paquets dédiés

RED

- Random Early Detection/Discard/Drop
 - IETF – RFC 2309
- Rejet de paquet avant que la file soit pleine
 - Information implicite
 - Aléatoire
- Probabilité de rejet fonction de l'évolution de l'occupation moyenne de la file
 - $AvgLen_i = u.AvgLen_{i-1} + (1-u).l$
 - l occupation instantanée de la file

RED

Principes



If $minTh < AvgLen < maxTh$:

$$\hat{P}_{AvgLen} = maxP \left\{ \frac{AvgLen - minTh}{maxTh - minTh} \right\}$$

$$Pr(\text{Drop Packet}) = \frac{\hat{P}_{AvgLen}}{1 - count \times \hat{P}_{AvgLen}}$$

count counts how long we've been in $minTh < AvgLen < maxTh$ since we last dropped a packet. i.e. drops are spaced out in time, reducing likelihood of re-entering slow-start.

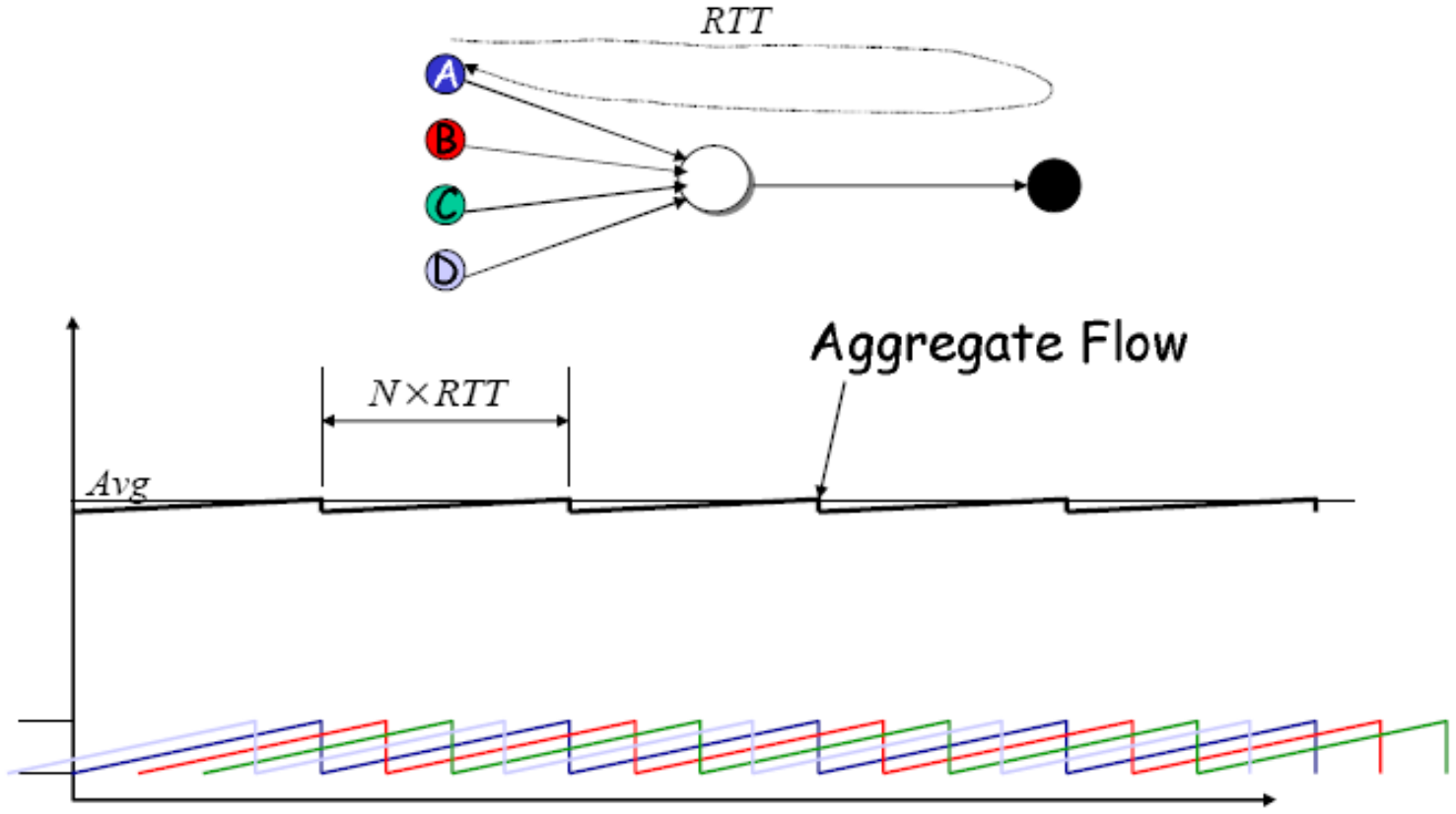
Source Mc Kneown

RED

Propriétés

- Rejet des paquets avant que la file ne soit pleine
 - On espère ainsi diminuer le débit de certains flux
 - Et anticiper la congestion
- Basé sur une évaluation moyenne de l'occupation de la file
 - On ne réagit pas trop vite aux rafales
- Rejet aléatoire
 - On espère toucher les flux consommateurs en bande passante
- Rejets sont espacés dans le temps
 - Évite de repartir à zéro pour certains flux
 - Désynchronisation des sources TCP

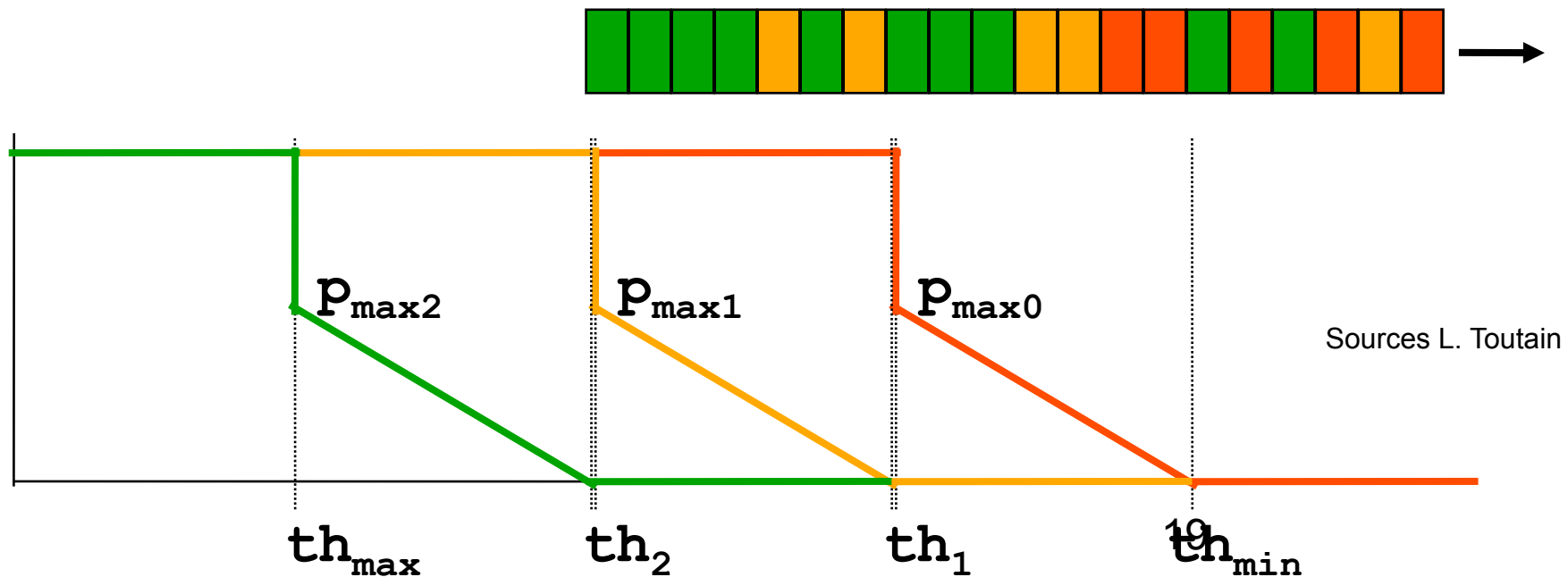
Désynchronisation des sources TCP



Source Mc Kneown

WRED

- RED with In and Out
 - Exécution de plusieurs RED en parallèle ayant des paramètres différents
- Permet des politiques plus complexes



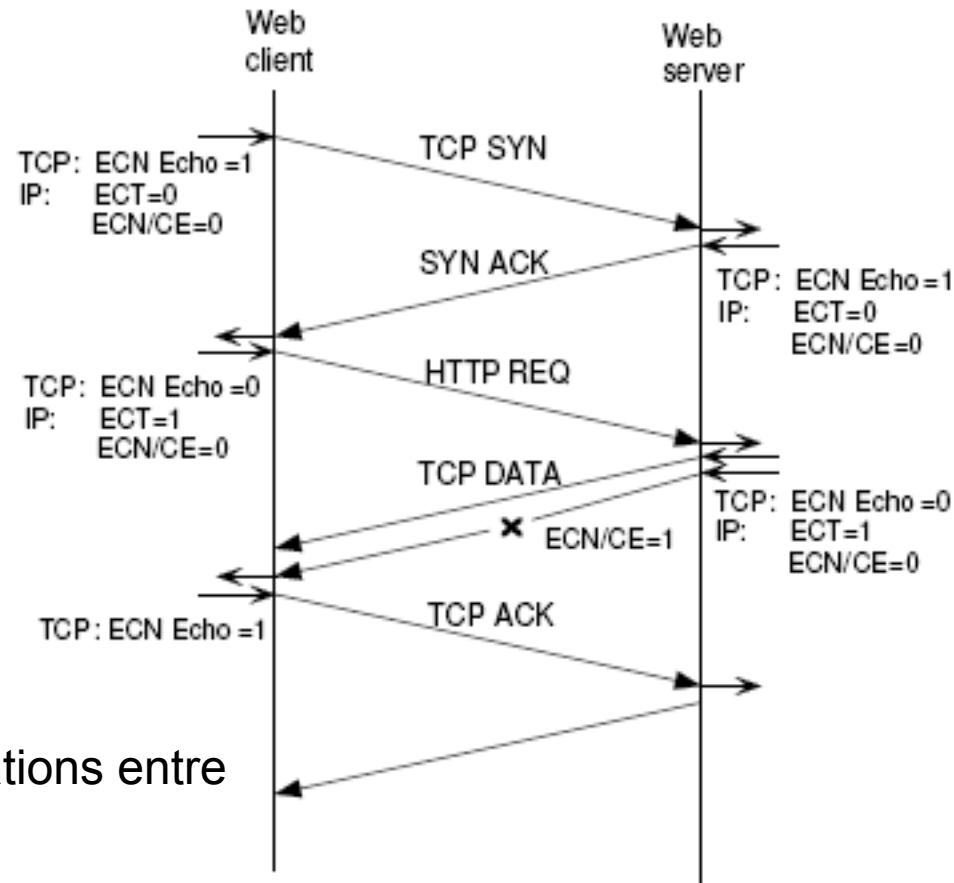
Problèmes de RED

- Paramétrage compliqué
- Stabilité
 - Surtout si le paramétrage n'est pas bien fait
- Au final on rejette des paquets
- Souvent implanté dans les routeurs
 - Mais option désactivée

Notification explicite de congestion

- ECN - RFC 3168
 - Plutôt que de rejeter des paquets, on notifie clairement les hôtes qu'il y a un problème de congestion
 - Information explicite
- Champs de congestion
 - Contenu dans le champ ToS de l'en-tête IP
 - Deux derniers bits
 - ECT (ECN capability) et CE (Congestion Experienced)
- Signification
 - ECT mis à 1
 - La source peut réagir aux notifications de congestion
 - CE mis à 1
 - par un équipement qui détecte une congestion
- Indication de congestion
 - au niveau IP
- Traitement de l'indication
 - par les couches supérieures

ECN et TCP



Passage d'informations entre
IP et TCP

Figure 1: Negotiating ECN capabilities

ECN et TCP

- Différentes recommandations possibles pour l'utilisation de ECN avec TCP
- ECN ne doit pas modifier le comportement de TCP si un timeout est déclenché
 - Même réaction que pour des paquets perdus
- Si ECN indique une congestion
 - La réaction de TCP peut se faire plus légère qu'avec une perte de paquet
 - Mais elle peut se faire dès qu'un paquet est marqué avec ECN à 1
 - Pas obligé de réagir aux autres paquets marqués avec ECN à 1 si la source a déjà réagi à un premier tel paquet
- Quand mettre le bit CE à 1?
 - Encore une histoire de paramétrage...

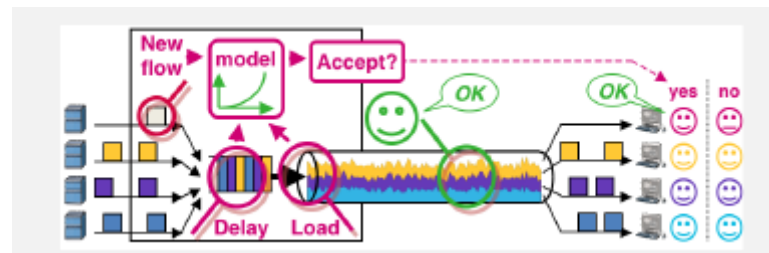
Utilisation de ECN

- Très faible
- En 2000
 - 1,1% des serveurs Web pouvait réagir aux ECN (bit ECT) (84000 serveurs étudiés)
- En 2004
 - 2,1%
 - Dans certaines expériences, aucun paquet marqué par les routeurs dans ce champ
 - Probabilité de routeurs utilisant ECN encore plus faible ?
 - Et pourtant ECN implantée dans les OS (GNU/Linux, MacOSX, Windows) et dans des routeurs (e.g. CISCO)
 - Mais option désactivée

Contrôle d'admission

Objectifs

- Objectif principal
 - Empêcher la congestion dans le réseau
- Objectif dérivé
 - Offrir de bonnes performances aux flux acceptés



Thèse D. Ammar

Différents CA

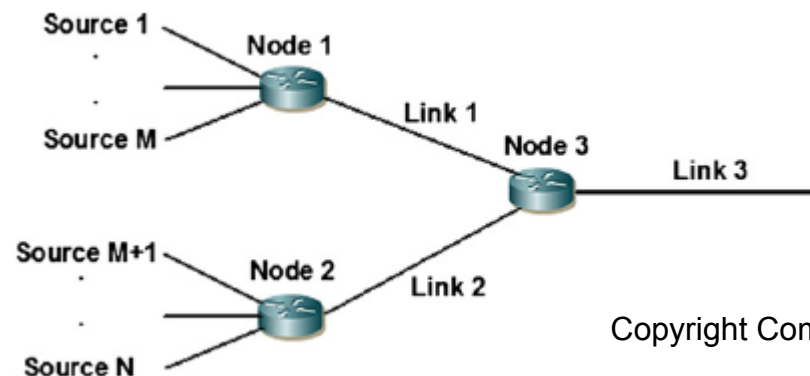
- Par lien
 - Basé sur la description du trafic
 - Basé sur des mesures passives
- De bout-en-bout
 - Basé sur des mesures actives

CA basé sur une description de trafic

- Description de trafic
 - Déterministe
 - Probabiliste
 - Donnée par l'application
- Connaissance sur les trafics acceptés et entrant
 - Donne une connaissance sur l'utilisation actuelle du réseau

CA basé sur une description de trafic

- **Avantage**
 - Bonne utilisation du réseau si les descriptions sont correctes
- **Inconvénients**
 - Difficile de bien caractériser précisément le trafic
 - Encore plus difficile de caractériser l'agrégation du trafic



CA basé sur des mesures

- Caractérisation du trafic par le réseau
 - Via des mesures temps réel
 - e.g. taux d'utilisation, taux d'occupation, délai, taux de pertes
- Avantages
 - Ne nécessite pas une caractérisation a priori
 - Proche de l'utilisation réelle
 - Plus facile de prendre en compte l'agrégation de trafic
- Inconvénients
 - Déterminer la bonne échelle de temps pour les mesures
 - Fonction de la variabilité du trafic

CA par lien

- CA qui doit être répété sur chaque lien traversé
 - Un refus peut se faire tardivement
- Application de ce CA seulement sur les liens critiques

CA basé sur des mesures de bout-en-bout

- CA à l'entrée (aux entrées) du réseau
 - Envoi de paquets de sonde pour caractériser la charge sur un chemin
 - Mesures faites aux destinations
- Avantages
 - CA à l'entrée et non par lien
 - Refus ou acceptation dès le début
 - Idée de la charge sur un chemin
 - Pas besoin de garder des informations sur le réseau
- Inconvénients
 - Ajout de paquets supplémentaires
 - Idée de la charge à un instant donné
 - Ne mesure pas l'état du réseau sur le long terme

Conclusions sur le CA

- Beaucoup étudié dans les années 90 et 2000
 - Beaucoup de solutions proposées
- Revient sur le devant de la scène
 - Avec l'augmentation du trafic
 - e.g. RFC PCN (Pre-Congestion Notification – IETF – RFC 5559)
 - Combinaison de CA avec des rejets de flux en cours

Routage avec QoS

Principe

- Jusqu'ici, on travaille avec les routes qu'on a
 - Très souvent, routes basées sur le plus court chemin
- Chercher des routes qui offrent certaines garanties
- Routes répondant à une ou plusieurs contraintes
 - Objectif initial
 - Répondre aux performances attendues
 - Objectifs secondaires
 - Prévention de la congestion
 - Répartition de la charge dans le réseau

Métriques

- Métriques additives
 - Poids d'un chemin = somme des poids des liens de ce chemin
- Métriques multiplicatives
 - Poids d'un chemin = multiplication des poids des liens sur ce chemin
 - Peut être transformée en une métrique additive
- Métriques concaves
 - Poids d'un chemin = minimum (ou max) des poids des liens sur ce chemin

Une seule métrique

- Topologie du réseau connue avec les poids associés
- On enlève les liens qui ne répondent pas à la contrainte
- Dijkstra suivant le type de métrique
 - e.g. IGRP, EIGRP

Routage multicontraint

- Satisfaire plusieurs contraintes à la fois
- De tels chemins peuvent ne pas exister
- NP-complet si N métriques additives, M métriques multiplicatives et $N+M \geq 2$
- Heuristiques
 - Pas simples

Routage multicontraint

- Nécessite souvent
 - Une connaissance (au moins partielle) de la topologie
 - Route souvent calculée à la source
 - Source routing
 - Une évaluation des ressources
 - Bande passante libre
 - Délai rencontré sur un lien
 - Taux d'utilisation
 - Taux de perte
- Problème de l'évaluation des ressources encore difficiles
 - Problème de fraîcheur d'information

Conclusions sur le routage QoS

- Probablement peu déployé
 - Solutions souvent compliquées
 - Mais on ne connaît pas bien les métriques utilisées dans les routages actuels
- Exemple
 - QOSPF
 - Quality of Service Extension to OSPF
 - Prise en compte de la bande passante et du délai

Conclusion sur la gestion de la congestion

- Qui fait du contrôle ?
 - Hôtes en bout de chaîne
 - Réseau
 - Les deux
- Quand le contrôle est appliqué ?
 - Avant la congestion
 - Une fois la congestion apparue
- Contrôle
 - En boucle ouverte
 - En boucle fermée
- **LE PROBLEME DUR**
 - Le dimensionnement