

Réseaux Haut-débit et Qualité de Service, Université de Pau

Cours de C. Pham

EXAMEM du mardi 8 janvier 2008

Durée 2h - Aucun document autorisé, calculatrice autorisée.

Répondre de manière concise aux questions. Les durées sont indicatives. **Vous pouvez répondre à certaines questions sur le sujet, n'oubliez alors pas d'indiquer votre numéro d'anonymat.**

Exercice 1 (15pts)

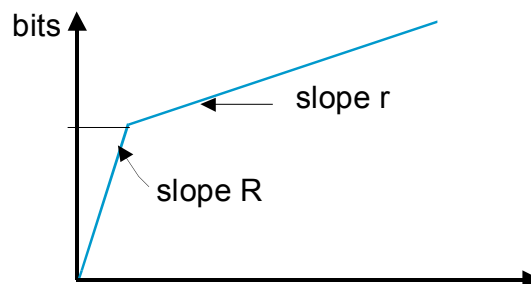
Une entreprise paoise travaillant dans le domaine du génie pétrolier doit effectuer pour ces clients des simulations numériques très gourmandes en puissance de traitement. Elle utilise pour cela une grille de calcul constituée de 3 centres nationaux : Paris, Lyon et Marseille. Dans les années 90, cette entreprise avait été la première à utiliser l'infrastructure ATM naissante pour distribuer géographiquement ses calculs. À l'époque, l'utilisation répartie des ressources de calcul se faisait essentiellement en parallélisant les exécutions des simulations avec des paramètres différents afin d'obtenir plusieurs points de simulation. Le lien que possédait l'entreprise était un lien à 155Mbits/s entre Pau et les autres sites.

(0.25) Pour quels types d'utilisation ATM a-t-il été conçu au départ ? **(0.5)** Expliquer les avantages d'avoir choisi une taille de cellules de 53 octets, ce qui est petit par rapport à la taille des paquets IP. **(0.25)** Pourquoi une taille de cellule plus petite aurait été pénalisante pour des applications informatiques ?

ATM réutilise la notion de circuit-virtuel déjà utilisé par les réseaux X.25. **(0.5)** Pourquoi l'acheminement par voie logique (commutation d'étiquette) présente des temps de traitement plus faibles que le mode datagramme IP ?

En 2007, cette entreprise est actuellement connectée à l'Internet par le biais du réseau fibre de l'agglomération de Pau (PBC) à 100Mbits/s d'une part, et d'autre part par une liaison Equant IP VPN 300 Mbits/s basée sur le réseau IP/MPLS d'Orange Business Service¹. Le lien sur le PBC sert de lien de secours.

Les 300 Mbits/s sur le lien Equant IP VPN sont fournis avec un *Token Bucket* pour réguler le trafic. L'entreprise a négocié un *token bucket* décrit par les paramètres (b, r, R) où b est la capacité en jetons, r le taux de génération des jetons et R le taux d'émission maximum (e.g., $R =$ capacité du lien). On considère que 1 bit envoyé consomme un jeton.



¹ L'offre Equant IP VPN est une offre d'Orange Business Service. Il est réellement déployé sur le réseau IP/MPLS de France Télécom (voir http://www.orange-business.com/fr/lna/reseaux/manages/equant_ipvpn/). Le reste de l'exercice est fictif.

(1) Quel est l'avantage pour l'utilisateur d'avoir un accès régulé par un token bucket sur un lien IP par rapport à une ligne spécialisée de type multiplexage temporel à 300Mbits/s (avantage autre que financier bien sûr) ?

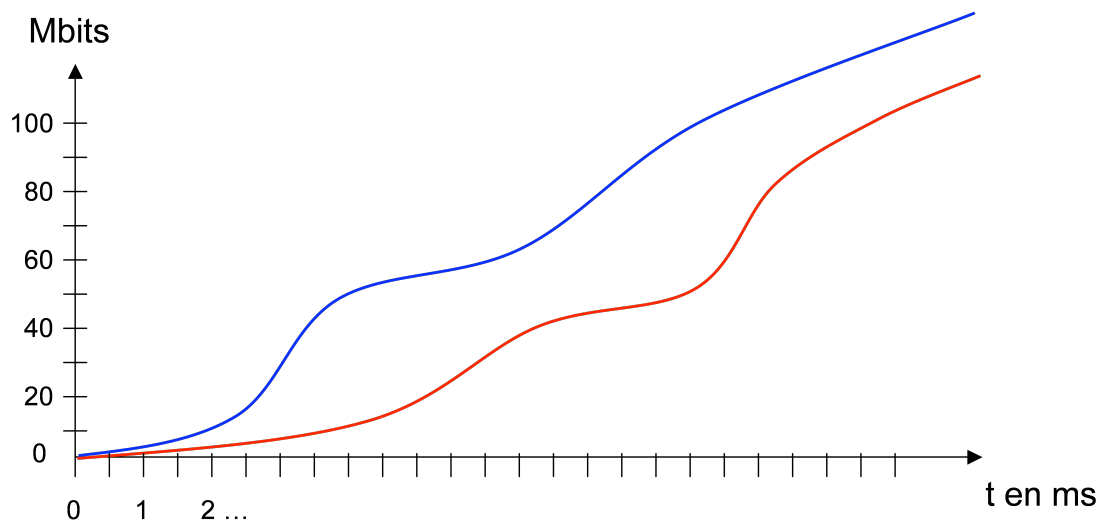
L'avantage pour l'opérateur est de pouvoir borner la quantité d'information envoyée par un client et ainsi pouvoir prédire les ressources qui seront consommées. Le trafic utilisateur est contenu dans ce que l'on appelle une enveloppe (figure ci-dessus).

(0.75) A quel instant t va t-on passer du débit R à r ? Donner d'abord l'expression littérale, puis passer à l'application numérique. A.N. donner t pour $b=100\text{Mbit}$, $r=150.10^6$ jetons/s et $R=300\text{Mbit/s}$.

(0.75) Si on considère que l'on commence à t_0 à envoyer les paquets, à quel instant t_1 aura t-on envoyé $n=6000$ paquets de taille $S=4096$ octets? Même données numérique que pour la question précédente.

L'opérateur du réseau Equant connaît le modèle de service de son réseau entre Pau et Paris.

(0.5) Expliquez ce qu'est un modèle de service en quelques lignes. (0.5) Donner un moyen « simple » pour l'opérateur de déterminer un tel modèle de service.



Dans le schéma ci-dessus, il est représenté la courbe du trafic en entrée du lien Pau-Paris et la courbe du modèle de service de l'opérateur Equant sur ce lien. **(0.25) Indiquez quelle courbe est logiquement celle du modèle de service de l'opérateur Equant. (0.75) Dans ce schéma pouvez-vous rapidement dire si le délai maximum de 8ms pour une application de calcul distribué pourra être respecté ? Faites un dessin sur le schéma montrant la manière dont vous avez procédé.**

Le trafic sur la grille de calcul peut être décomposé en 3 catégories : une à très forte contrainte temporelle (messages de synchronisation et de contrôle des simulations numériques, résultats temps-réels pour la visualisation) notée type I, une avec les mêmes contraintes qu'un FTP interactif, c'est-à-dire que le plus vite possible est le mieux pour transférer les données, notée type II, et enfin une autre similaire à un FTP programmé pour transférer la nuit les résultats dans des bases de données, notée type III.

L'opérateur Equant gère la qualité de service avec un mécanisme DiffServ. L'opérateur installe donc sur chaque site des routeurs qui vont marquer les paquets provenant des sites. Assez naturellement, le trafic de type I est associé à un PHB de type EF (Expedited Forwarding, premium) avec un *token bucket* de paramètre (100Mbits, 150.10^6 , 300Mbits/s), et les trafics de type II et III sont associés à un PHB de type AF (Assured Forwarding) avec respectivement un *token bucket* de

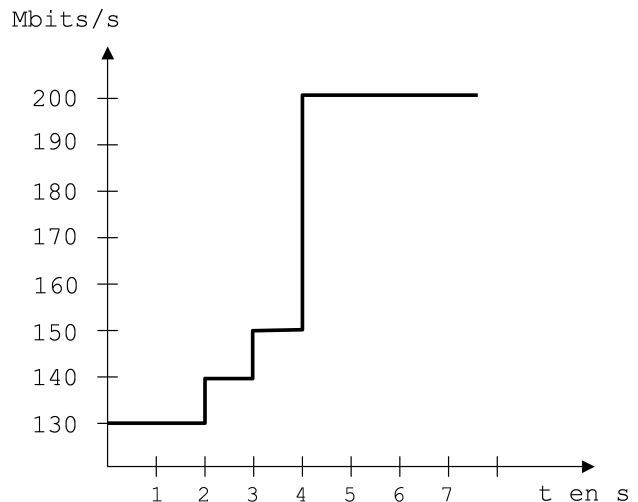
paramètres (50Mbits, $200 \cdot 10^6$, 300Mbits/s) et (30Mbits, $250 \cdot 10^6$, 300Mbits/s). Dans ce dernier cas, le trafic de type II possède un taux de pertes moins grand que celui de type III.

(0.5) Pourquoi effectue-t-on le marquage et la régulation DiffServ dans les routeurs de périphérie uniquement, et pas dans les routeurs du cœur de réseau?

(0.75) Quel est l'avantage d'avoir un paramètre b plus grand pour le trafic de type I ?

(0.5) Que se passe-t-il pour un paquet du trafic de type II au niveau du routeur de bordure s'il n'y a pas suffisamment de jetons ?

Dans la figure ci-contre, on a représenté le débit d'un flux appartenant au type I. Par exemple, entre 1s et 2s, le débit demandé a été 130Mbits/s (soit 130Mbits d'information envoyés). **(1,5) Dites si ce scénario est possible jusqu'à l'instant $t=5$ avec les paramètres de *token bucket* et DiffServ associés au flux de type I. Expliquez pourquoi en indiquant à peu près le nombre de jetons excédentaires ou manquants selon votre réponse et à quel moment.** On considère que le token bucket est plein à l'instant $t=0$.



Chaque PHB utilise l'ordonnancement pour garantir dans chaque routeur les contraintes de chaque type de flux.

(0.5) L'ordonnancement Weighted Round Robin permet d'éviter qu'un flux trop gourmand n'affecte les autres flux. Expliquez comment cela est réalisé.

(0.25) Pourquoi l'ordonnancement GPS (Generalized Processor Sharing) qui est a priori optimal est-il difficilement implémentable? (0.25) Quelle approche permet de réaliser approximativement la même chose ?

(0.75) Comment l'opérateur peut-il faire pour que le taux de perte du trafic de type II soit plus petit que le trafic de type III alors qu'ils sont de la même catégorie ?

L'offre de connexion Equant IP/VPN utilise un réseau IP/MPLS. En fonction du modèle de service de chacun de ses liens, l'opérateur est capable de définir quel « chemin » dans le réseau permet de garantir tel ou tel contrainte de qualité de service.

(1) Expliquer les principes de bases de MPLS, comment les étiquettes peuvent-elles être distribuées et surtout l'intérêt de cette approche (environ 10 lignes). (0.25) Expliquer aussi pourquoi dit-t-on que MPLS est « multi-protocole » et dites où se situe la « couche » MPLS dans l'architecture OSI?

(0.5) Dans ce scénario d'utilisation d'un réseau IP/MPLS, combien de « labels » (étiquettes) MPLS l'entreprise pourra se voir assigner si on suppose que les trafics de type I, II et III sont très différents les uns des autres.

Coût des sites : Soit L_j le coût du site j . C'est le coût d'achat ou de location des locaux dans lesquels se trouveront les équipements réseaux permettant de concentrer le trafic venant des nœuds clients.

On définit la variable de décision u_j , $j = 1, \dots, K$ de la façon suivante :

$$u_j = \begin{cases} 1 & \sum_{i=1}^N x_{ij} \geq 1 \\ 0 & \text{sinon} \end{cases}$$

La variable u_j indique si le site j est utilisé ou non dans la solution x .

Le coût total des sites associé à la solution x s'écrit alors,
$$\Gamma_S(x) = \sum_{j=1}^K L_j \cdot u_j$$

Coût des liens : Pour la solution x , si $x_{ij} = 1$, cela signifie qu'il faut raccorder le nœud client i au site j par un lien de communication. Le coût de ce lien sera d'autant plus fort que la distance (euclidienne) entre i et j sera grande et que la demande d_i du client i sera grande. Notons l_{ij} la distance à vol d'oiseau en km entre le nœud client i et le site j . Typiquement, une offre opérateur va définir des seuils de distance $l_1 < l_2 < \dots < l_M$ tels que le coût du lien (i, j) sera donné par,

$$c_{ij} = \alpha_k(d_i) + \beta_k(d_i) \cdot l_{ij} \quad \text{si } l_{k-1} < l_{ij} < l_k$$

$\alpha_k(d_i)$: coût d'ouverture du service,

$\beta_k(d_i)$: coût kilométrique.

Avec par convention $l_0 = 0$ et $l_{M+1} = \infty$.

Les distances entre les nœuds clients et les sites potentiels étant connues à l'avance, les coûts c_{ij} , pour $i = 1, \dots, N$ et $j = 1, \dots, K$ sont des constantes du problème.

Le coût global des liens dans la solution x s'écrit,
$$\Gamma_L(x) = \sum_{i=1}^N \sum_{j=1}^K c_{ij} \cdot x_{ij}$$

Coût des équipements : Si le site j est sélectionné dans la solution x , il faut y installer des équipements permettant de supporter le trafic en émission et réception des clients raccordés à ce site. Le coût des équipements installés au site j dépend du nombre de clients raccordés à ce site et des demandes de ces clients.

Notons $E_j(x)$ le coût des équipements à installer sur le site $j = 1, \dots, K$ pour l'affectation x .

Le coût global des équipements s'écrit,

$$\Gamma_E(x) = \sum_{j=1}^K E_j(x)$$

Question : (2,25) En déduire le coût global d'une solution. Que remarquez vous ?

Exercice 2 (3pts): La qualité de service en pratique

Un de vos amis vous propose d'améliorer les performances de votre connexion ADSL au moyen de quelques scripts de son crû. Il omet seulement de documenter sa solution miracle. Afin de briller en société vous décidez à votre tour de distribuer cette solution en y ajoutant les informations nécessaires à la compréhension de ces scripts.

1. Ecrivez cette documentation : il ne s'agit pas de faire un commentaire ligne par ligne des commandes, mais de décrire la fonction des outils utilisés et le fonctionnement général de la solution proposée, le tout illustré de quelques exemples pertinents.
2. Afin que vos amis soient convaincus de votre solution, proposez une méthode pour illustrer l'efficacité de vos scripts.

Les commandes :

```
CEIL=240
tc qdisc add dev eth0 root handle 1: htb default 15
tc class add dev eth0 parent 1: classid 1:1 htb rate ${CEIL}kbit ceil ${CEIL}kbit
tc class add dev eth0 parent 1:1 classid 1:10 htb rate 80kbit ceil 80kbit prio 0
tc class add dev eth0 parent 1:1 classid 1:11 htb rate 80kbit ceil ${CEIL}kbit prio 1
tc class add dev eth0 parent 1:1 classid 1:12 htb rate 20kbit ceil ${CEIL}kbit prio 2
tc class add dev eth0 parent 1:1 classid 1:13 htb rate 20kbit ceil ${CEIL}kbit prio 2
tc class add dev eth0 parent 1:1 classid 1:14 htb rate 10kbit ceil ${CEIL}kbit prio 3
tc class add dev eth0 parent 1:1 classid 1:15 htb rate 30kbit ceil ${CEIL}kbit prio 3
tc qdisc add dev eth0 parent 1:12 handle 120: sfq perturb 10
tc qdisc add dev eth0 parent 1:13 handle 130: sfq perturb 10
tc qdisc add dev eth0 parent 1:14 handle 140: sfq perturb 10
tc qdisc add dev eth0 parent 1:15 handle 150: sfq perturb 10

tc filter add dev eth0 parent 1:0 protocol ip prio 1 handle 1 fw classid 1:10
tc filter add dev eth0 parent 1:0 protocol ip prio 2 handle 2 fw classid 1:11
tc filter add dev eth0 parent 1:0 protocol ip prio 3 handle 3 fw classid 1:12
tc filter add dev eth0 parent 1:0 protocol ip prio 4 handle 4 fw classid 1:13
tc filter add dev eth0 parent 1:0 protocol ip prio 5 handle 5 fw classid 1:14
tc filter add dev eth0 parent 1:0 protocol ip prio 6 handle 6 fw classid 1:15

iptables -t mangle -A PREROUTING -p icmp -j MARK --set-mark 0x1
iptables -t mangle -A PREROUTING -p icmp -j RETURN

iptables -t mangle -A PREROUTING -m tos --tos Minimize-Delay -j MARK --set-mark 0x1
iptables -t mangle -A PREROUTING -m tos --tos Minimize-Delay -j RETURN
iptables -t mangle -A PREROUTING -m tos --tos Minimize-Cost -j MARK --set-mark 0x3
iptables -t mangle -A PREROUTING -m tos --tos Minimize-Cost -j RETURN
iptables -t mangle -A PREROUTING -m tos --tos Maximize-Throughput -j MARK --set-mark 0x6
iptables -t mangle -A PREROUTING -m tos --tos Maximize-Throughput -j RETURN

iptables -t mangle -A PREROUTING -p tcp -m tcp --sport 22 -j MARK --set-mark 0x1
iptables -t mangle -A PREROUTING -p tcp -m tcp --sport 22 -j RETURN

iptables -t mangle -A PREROUTING -j MARK --set-mark 0x6
```

Exercice 3 (2pts): Transport de données sur réseaux haut-débit

Une fois tous ses liens configurés, l'entreprise paloise doit transférer une base de données de taille $T=10\text{Go}$ (trafic de type II) entre Pau et Paris sur le lien Equant au débit $D=300\text{Mbits/s}$. Le protocole applicatif utilisé est FTP pour le transfert. Théoriquement, le temps de transfert est d'un peu moins de 5mins. Pratiquement, l'entreprise constate des durées de transfert de plus 40 minutes alors que le réseau n'est pas chargé.

(0.5) Quel(s) protocole(s) de la pile de communication est(sont) en cause ?

(0.75) Expliquez les causes possibles de ces différences en argumentant.

(0.75) Pour cette entreprise, quelle(s) solution(s) d'optimisation au niveau des protocoles vous permettrai(en)t d'améliorer la situation ?