Examen Pair-à-pair

23 mars 2017

Remarques

- L'examen comprend trois parties indépendantes (DHTs, bande passante, streaming).
- Il n'y a pas de barême officiel, mais la troisième partie comptera vraisemblablement pour plus de points que les deux autres.
- La somme des points obtenus va donner une somme sur vingt, et il y a plus de vingt points en jeu : il n'est pas nécessaire de répondre à toutes les questions pour avoir une très bonne note → si vous séchez complètement sur une question, essayez d'avancer !
- La difficulté d'une question n'est pas forcément proportionnelle à sa longueur.

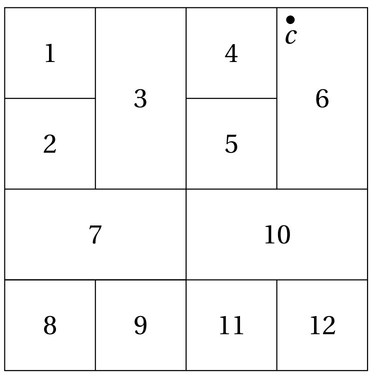
Anti-sèches Attention, toutes les formules ci-dessous ne sont pas utiles pour l'examen.

- $\ln(n) \leq \sum_{k=1}^{n} \frac{1}{k} \leq \ln(n) + 1$ (ln : logarithme naturel)
- $a^b = e^{b \ln(a)}$
- $\log_k(x) = \frac{\ln(x)}{\ln(k)}$ ($\log_k$: logarithme en base k)
- $(1 + \frac{x}{n})^n \to e^x$ quand $n \to \infty$
- Loi de Little : si des entités arrivent avec intensité λ dans un système et restent un temps T, il y a λT entités dans le système en moyenne.

Tables de hachage

On considère une table de hachage de type CAN à 12 pairs, dont les zones sont indiquées sur la figure ci-contre. On identifiera chaque pair à la zone dont il est responsable. Les règles sont les suivantes :

- chaque zone connaît les zones adjacentes à la sienne (un côté commun, se toucher par les angles ne compte pas), qui forment sa table de routage ;
- le routage se fait en choisissant dans la table la zone la plus proche de la destination. La distance entre un point x et une zone Z se définit comme la plus petite distance euclidienne entre x et un point de Z.



Question 1

La zone 8 recherche un contenu c dont les coordonnées de hachage sont situées en haut à gauche de la zone 6 (voir figure). Quelle est la route suivie pour que 8 contacte la zone en charge de c ?

Question 2

Le pair responsable de la zone 2 veut quitter le système. Quelles sont les opérations à réaliser pour que le départ soit fait correctement ?

Bande passante

Soit un système pair-à-pair idéal qui vérifie les hypothèses suivantes :

— Les messages de contrôles et d'acknowledgement ont un coût négligeable par rapport au transfert des données. Il ne sont pas pris en compte dans le calcul de bande passante.

— Le *download* est supposé infini et l'allocation est optimale. En d'autres termes, tout l'upload disponible est utilisé pour transférer des données.

— On suppose qu'il n'y a pas de serveurs, ou du moins que la bande passante qu'ils fournissent est négligeable devant celle des pairs.

— Le système est ouvert : les nouveaux pairs arrivent selon un processus d'intensité λ (exprimé en nombre d'arrivées par seconde).

— Le régime considéré est le régime fluide stationnaire : on suppose que le nombre de leechers et de seeders reste constant dans le temps.

— L'objectif est de télécharger un fichier de taille F.

— Il n'y a pas de seeder : dès qu'un pair a fini de récupérer le fichier, il se déconnecte du système.

Question 3

Nous allons considérer que parmi les pairs qui rentrent dans le système, une proportion p_f sont des *free-riders*, qui offrent une bande passante montante (upload) $u_f = 0$. Les autres pairs sont des *leechers* classiques. Ils sont en proportion $p_l = 1 - p_f$ parmi les nouveaux arrivants et offrent tous le même upload $u_l > 0$.

La bande passante est répartie suivant une technique de *tit-for-tat* à c connexions (1 connexion généreuse pour $c-1$ connexions réciproques). Un free-rider ne peut recevoir de contenu que via les connexions généreuses des free-riders, c'est-à-dire qu'ils ne peuvent accéder qu'à une fraction $\frac{1}{c}$ de la bande passante disponible. On admettra que la vitesse de téléchargement d'un pair d'upload u est donné par

$$d(u) = \frac{1}{c}\bar{u} + \frac{c-1}{c}u, \text{ où } \bar{u} \text{ est l'upload moyen des pairs présents dans le système.}$$

Le but est de montrer dans ce cas particulier un résultat admis en cours sur la proportion tolérable de free-riders.

(a) L'analyse du système serait très simplifiée si l'on connaissait $\bar{u}$. Peut-on écrire $\bar{u} = p_l u_l$?

(b) Sachant que le régime est stationnaire, quelle est la bande passante totale consommée par les free-riders ?

(c) Combien de temps au maximum est-ce qu'un leecher classique reste dans le système ? En déduire une valeur que ne peut pas dépasser la bande passante totale disponible pour les free-riders.

(d) En déduire une condition nécessaire pour que le régime stationnaire puisse exister. Sinon, que se passe-t-il ?

Diffusion Live

On s'intéresse au délai de diffusion d'un chunk dans un système pair-à-pair contenant N = 1023 pairs. On négligera l'overhead, la latence réseau et les délais de synchronisation (le délai de transmission d'un chunk ne dépend donc que de la taille du chunk et de la bande passante). Un pair ne peut transmettre un chunk que s'il l'a complètement reçu. Les pairs savent se synchroniser de sorte qu'un pair ne reçoit jamais 2 copies du même chunk. À l'instant $t = 0$, un seul pair possède le chunk. Les autres paramètres vont varier suivant les questions. La question 7 peut être traitée sans avoir fait les questions 5 et 6.

Question 4

On reprend un scénario vu en cours. Les règles de diffusion sont les suivantes :

— Chaque pair a une capacité 1, i.e. il peut envoyer un chunk par unité de temps : si à l'instant t, p_1 commence à distribuer le chunk à p_2, alors p_2 possède le chunk à l'instant $t+1$.

— Il n'y a pas de limite au nombre de copies qu'un pair envoie au total sur toute la durée de l'expérience.

— Un pair ne peut recevoir le chunk que d'un seul autre pair à la fois.

On note $X(t)$ le nombre de pairs pouvant posséder le chunk à l'instant t dans le meilleur des cas.

(a) Donner $X(t)$ aux instants $t = 0, t = 1, \ldots, t = 5$.

(b) Au bout de combien de temps au minimum (en fonction de N et en valeur numérique) tous les pairs possèdent le chunk ?

Question 5

On reprend le scénario de la question 4, mais avec deux classes de pairs : 511 pairs sont des free-riders qui n'envoient jamais rien. Les 512 pairs restant peuvent distribuer 2 chunks par unité de temps (si à l'instant t, p_1 commence à distribuer le chunk à p_2 et p_3, alors p_2 et p_3 possèdent le chunk à l'instant $t+1$). À l'instant $t = 0$, un seul pair de capacité 2 possède le chunk. $X(t)$ désigne toujours le nombre de pairs pouvant posséder le chunk à l'instant t dans le meilleur des cas.

(a) Donner $X(t)$ aux instants $t = 0, t = 1, \ldots, t = 5$.

(b) Au bout de combien de temps au minimum (en fonction de N et en valeur numérique) tous les pairs possèdent le chunk ?

(c) On suppose maintenant que quand un pair de capacité 2 possède un chunk, il choisit pour envoyer, parmi les pairs qui n'ont pas encore le chunk, un pair de capacité 2 et un freerider [1]. Au bout de combien de temps (en fonction de N et en valeur numérique) tous les pairs possèdent le chunk ?

(d) Comparer avec les réponses à la question 4 et commenter.

Question 6

En pratique, ce n'est pas un seul chunk, mais un flot de chunks qui doit être diffusé. On reprend donc le scénario de la question 5, mais cette fois la source injecte un nouveau chunk (et pas une copie des précédents) à chaque unité de temps. On supposera que c'est toujours le même pair, de capacité 2, qui reçoit les nouveaux chunks.

(a) Dire comment tous les pairs peuvent recevoir en délai fini tous les chunks émis par la source, et donner ce délai (en fonction de N et en valeur numérique).

(b) Comparer avec les réponses aux la question 4 et 5 et commenter.

Question 7

On reprend le scénario de la question 4 (tous les pairs ont capacité 1), mais on autorise maintenant la collaboration, i.e. plusieurs pairs peuvent collaborer pour l'envoi d'un chunk à un pair en mutualisant leur bande passante. Par exemple, si p_1, p_2 et p_3 possèdent le chunk, mais pas p_4, p_1, p_2 et p_3 peuvent envoyer le chunk à p_4 en un temps $\frac{1}{3}$. On remarque au passage que le temps t n'est plus forcément entier.

(a) Si X pairs possèdent le chunk à un instant donné, les autres ne possédant rien du tout, combien de temps faut-il au minimum pour qu'un nouveau pair possède le chunk ?

(b) En déduire le temps minimum nécessaire pour que tous les pairs possèdent le chunk (en fonction de N et en valeur numérique, à 1 près).

(c) En comparant avec les réponses à la question 4, discuter de l'intérêt de la collaboration pour l'envoi d'un chunk donné à un pair donné. D'après vous, cette solution n'a-t-elle que des avantages ?

1. Sauf bien sûr s'il ne reste à fournir que des free-riders ou que des pairs de capacité 2, auquel cas il choisit parmi ce qui reste.