

Miage 2 Miage

Projet Applications réseau M1 Miage Apprentissage Paris-Dauphine 2012/2013

F. Sikora (D'après un sujet d'E. Duris)

1 Dates importantes

- Rendu de la RFC : 28 avril 2013, 23h59.
- Rendu du projet : ~~26 mai 2013~~ ~~9 juin 2013~~ 16 juin 2013, 23h59.
- Soutenances : ~~27 mai 2013, à partir de 13h45.~~ 17 juin 2013, à partir de 16h.

2 Versions

Ce sujet est susceptible d'être modifié.

- 3 juin 2013 : Changement de date de rendu.
- 23 mai 2013 : Changement de dates de rendu suite à un lobby puissant.
- 25 avril 2013 : Un pair peut mourir violemment.
- 24 avril 2013 : Ajout d'autres questions/réponses.
- 20 avril 2013 : Ajout de questions réponses.
- 17 avril 2013 : Mise en ligne du sujet.

3 En quelques mots

Le but du projet est la conception d'un protocole d'échange de fichiers partiellement décentralisé entre pairs (pair à pair ou peer to peer). Un *pair* souhaitant télécharger un certain fichier demandera à un *annuaire* les informations sur ce fichier. Celui-ci lui proposera une liste de pairs possédant (au moins une partie de) ce fichier afin qu'il puisse obtenir le fichier désiré. Une fois le (une partie du) fichier téléchargé, le pair devient un potentiel distributeur de ce fichier à d'autres pairs. Il est demandé d'écrire une RFC décrivant votre protocole, puis dans un second temps, une implémentation de ce protocole.

4 Description

Chaque fichier devra être découpé en *morceaux*. Ce protocole met en jeu deux entités. Des *annuaires* et des *pairs*, ces derniers sont à la fois des clients et des serveurs pour les

morceaux qu'ils possèdent.

Un annuaire possède l'information sur les pairs possédants certains morceaux de certains fichiers. Il doit également savoir comment joindre les autres annuaires. Un fichier ne doit être répertorié que par un seul annuaire.

Un pair doit être lancé avec une liste initiale d'adresses d'annuaires afin de se connecter au réseau. Chaque morceau de fichier qu'il partage est indexé par un seul annuaire.

Un pair doit pouvoir partager un fichier avec le reste du réseau et devenir serveur pour ces morceaux. Pour cela, il doit demander si le fichier (référéncé selon son nom, ou son contenu¹, ou une combinaison des deux) existe déjà sur le réseau, *i.e.* s'il est déjà indexé par un annuaire. Si oui, alors il doit informer l'annuaire gérant le fichier qu'il en possède des morceaux. Sinon, un annuaire doit indexer ce fichier.

Un pair doit pouvoir télécharger un fichier (il joue le rôle de client). Pour cela, il demande à un annuaire l'adresse de l'annuaire gérant le fichier désiré. Ensuite, il demande à cet annuaire l'adresse des pairs possédant les morceaux du fichier désiré, puis contacte les pairs afin de télécharger les morceaux désirés. Après le téléchargement d'un morceau, il devra prévenir l'annuaire qu'il est en sa possession.

Un annuaire peut être ajouté au réseau. Les autres annuaires doivent en être informés.

Un annuaire peut être déconnecté (par le FBI, par votre mère passant l'aspirateur...). Or, un fichier n'est indexé que par un unique annuaire. Les pairs possédant des morceaux gérés par un annuaire mort doivent alors choisir un autre annuaire (mais l'unicité d'un annuaire par fichier reste de mise).

Lorsqu'un pair meurt, les annuaires indexant les morceaux que le pair possédait doivent gérer cette disparition.

5 Conditions de rendu

Le projet est à effectuer en binôme, *i.e.* par 2 personnes. En cas de nombre impair d'élèves, un seul groupe sera autorisé à effectuer le projet en trinôme, en me prévenant par email florian.sikora@dauphine.fr.

Première étape : description du protocole. Une première RFC de votre protocole est à rendre avant le 28 avril 2013, 23h59 sur l'espace MyCourse dédié. Les noms des deux personnes du binôme devront apparaître dans la RFC, dans le nom du fichier ainsi que dans les commentaires de l'espace MyCourse. Attention, l'espace est fermé automatiquement après l'horaire précisé. La RFC doit être dans le même format que les autres RFC, c'est à dire un document texte ASCII rédigé en anglais (voir la RFC 2223 pour le format d'une RFC). Elle doit seulement décrire les méthodes de communications (format des messages, protocoles utilisés...) entre les différentes entités intervenant dans votre protocole. Aucun détail d'implémentation ne doit être donné. La RFC 1350 (TFTP) peut-être prise en exemple.

Seconde étape : implémentation d'un pair et d'un annuaire pour le protocole en langage Java. Théoriquement, deux implémentations de différents binômes pourraient commu-

1. voir la classe MessageDigest et le hachage par md5 par exemple

niquer entre elles si elles suivent la même RFC (théoriquement toujours, une implémentation en langage C++ lancée sur une machine personnelle pourrait communiquer avec une implémentation Java sur un téléphone android). Votre projet est à rendre avant le 26 mai 2013, 23h59 sur l'espace MyCourse dédié, également fermé automatiquement. Une fois votre fichier envoyé, vous devez avoir un écran de confirmation avec votre fichier et la date de l'envoi. **Il y aura un point en moins par heure de retard, une heure entamée est due.** Le format de rendu est une archive au format ZIP contenant :

- Un répertoire *src* avec les sources de votre implémentation java.
- Un répertoire *classes* vide dans l'archive, devant contenir les classes du projet une fois compilé.
- Un répertoire *docs* contenant :
 - La première version de la RFC, au format précisé plus haut.
 - Une version éventuellement mise à jour de la RFC, au même format que la première.
 - Une documentation pour l'utilisateur *user.pdf* décrivant à un utilisateur comment se servir de votre projet.
 - Une documentation pour le développeur *dev.pdf*, devant justifier les choix effectués, indiquer les différences entre les 2 RFC, les avantages et inconvénients du protocole, présenter l'architecture choisie, indiquer quelles ont été les difficultés rencontrées au cours du projet ainsi que la répartition du travail entre les membres du binôme.
 - Un répertoire vide *doc* pour la javadoc.
- Deux jars exécutables, *pair.jar* et *annuaire.jar*, pouvant être lancés au moyen de la commande `java -jar`.
- Un fichier ant *build.xml* permettant de :
 - Compiler les sources (target compile).
 - Créer les jars exécutables.
 - Générer la javadoc dans docs/doc (target javadoc).
 - Nettoyer le projet pour ne garder que les éléments demandés (target clean).

L'archive aura pour nom Nom1Nom2.zip, où Nom1 et Nom2 sont les noms des membres du binôme par ordre alphabétique. L'extraction de l'archive devra créer un dossier Nom1Nom2 contenant les éléments précisés ci-dessus.

Aucune interface graphique n'est demandée.

Il va sans dire que les différents points suivants doivent être pris en compte :

- Uniformité de la langue utilisée dans le code (anglais conseillé).
- Uniformité des conventions de nommage et de code (convention Java obligatoire).
- Projet compilant sans erreur et fonctionnant sur les machines de l'université.
- La documentation ne doit pas être un copié-collé du code source du projet.
- Les sources doivent être commentées, dans une unique langue, de manière pertinente (pas de commentaire "fait un test" avant un if.).
- Le nom des variables, classes et méthodes doivent être choisis judicieusement.
- Le code devra être propre, les exceptions correctement gérées, les classes correctement organisées en packages. La visibilité des méthodes et champs doit être perti-

nente (privée ou non...).

- La RFC et le projet doivent évidemment être personnels à chaque binôme. Un détecteur automatique de plagiat sera utilisé.

L'utilisation d'un gestionnaire de version type CVS ou SVN est fortement conseillée.

5.1 Soutenance

Une soutenance d'un quart d'heure aura lieu binôme par binôme, à partir de 13h45 le lundi 27 mai 2013 selon un planning communiqué ultérieurement. Elle doit être préparée et menée par le binôme.

6 Questions / Réponses

- **Est-ce qu'un pair possède des fichiers lorsqu'il entre sur le réseau ?**
- Éventuellement. Ceci est de l'ordre de l'implémentation et ne doit pas figurer dans la RFC.
- **Est qu'un pair peut ajouter des fichiers sur le réseau au cours d'une connexion ? Autrement dit, est-ce qu'un pair au bout de 10 minutes de connexion au réseau peut ajouter un fichier à partager avec les autres pairs ?**
- Oui, cf. le sujet "Un pair doit pouvoir partager un fichier avec le reste du réseau ". Dans la RFC, il faut spécifier le protocole à suivre pour ce cas.
- **Doit-on gérer plusieurs téléchargement en même temps ?**
- C'est mieux si oui - il faut toutefois faire attention aux ressources. Ceci est de l'ordre de l'implémentation et ne doit pas figurer dans la RFC (on veut juste savoir comment 2 pairs peuvent s'échanger un fichier, mais pas le nombre de threads créés par un pair).
- **Est-ce que différents morceaux d'un même fichier peuvent être indexés par différents annuaires ?**
- Non, tout un fichier est géré par un seul annuaire.
- **Doit-on réfléchir à une approche similaire à celle du P2P traditionnelle (obligation pour un nouveau pair de passer par un tracker qui posséderait un annuaire "d'initialisation" avec l'association pair/fichier) ou peut-on réfléchir à un système où chaque pair serait serveur pour ses fichiers ET pour un annuaire (réduit) de l'ensemble des autres pairs ? En d'autres termes, doit-on obligatoirement intégrer un serveur (dont la fonction se réduirait à être serveur-annuaire) dans notre protocole ?**
- Première proposition. Annuaires et pairs sont bien deux entités distinctes. C'est pour cela qu'il est demandé 2 jans différents dans le rendu. En d'autres termes, un pair n'indexe aucun fichier.
- **Il est dit "Un annuaire peut être ajouté au réseau. Les autres annuaires doivent en être informés." Tel que nous l'avons compris, un annuaire va**

- gérer le référencement de pairs possédant les différents morceaux d'un et un seul fichier (exemple : le film "La guerre des étoiles épisode 1").
- Oui.
 - **Etant donné la citation ci-dessus, nous avons du mal à saisir l'intérêt d'une visibilité d'un annuaire sur les autres annuaires du réseau. (En suivant l'exemple, quel serait l'intérêt que l'annuaire gérant les morceaux du fichier "Star Wars" connaisse l'adresse de l'annuaire gérant les morceaux du fichier "Indiana Jones et la dernière croisade").**
 - Pair 1 ne connaît que l'adresse de l'annuaire A gérant Star Wars (seul cette adresse a été donnée à son démarrage). Il veut télécharger Indiana Jones. Il indique à l'annuaire A qu'il souhaite télécharger Indiana Jones. Annuaire A ne connaît pas les pairs possédant Indiana Jones. Il demande alors aux autres serveurs "qui gère Indiana Jones?", disons Annuaire B. Il renvoie alors au Pair 1 l'adresse d'annuaire B. Pair 1 peut alors interroger Annuaire B et obtenir la liste des pairs ayant des morceaux d'Indiana Jones...
 - **En effet, nous pensons qu'un seul annuaire ("annuaire-maitre" en quelque sorte) serait chargé de référencer les annuaires existants en réalisant une association "Adresse de l'annuaire - Id du fichier géré par l'annuaire.**
 - Non, les annuaires sont tous "au même niveau".
 - **il est bien précisé qu'un fichier ne doit être indexé que par un et un seul annuaire. Cela signifie t'il qu'il ne doit exister physiquement sur le réseaux qu'un seul annuaire référençant un fichier, ou alors que plusieurs annuaires référençant le même fichier peuvent exister sur le réseau, mais ne peuvent pas être actifs au même moment ?**
 - Je ne pense pas qu'il soit nécessaire que vous perdiez du temps avec cet histoire d'actif / non actif. Dites vous simplement que si le processus / le programme "tourne", alors l'annuaire est actif. Si le programme est stoppé, l'annuaire est "mort".
 - **Pour illustrer, si le FBI décidait de déconnecter l'annuaire gérant le fichier "Star Wars" du réseau, un autre annuaire gérant le fichier "Star Wars" pourrait-il être créée ?**
 - Comme dit dans le sujet, les pairs possédant des morceaux de Star Wars voyant que l'annuaire gérant ce fichier est "mort" doivent trouver un autre annuaire pour gérer ces morceaux. Donc, ils vont choisir parmi leur liste d'annuaires un annuaire pour indexer ce fichier (en faisant attention qu'un seul annuaire indexe ce fichier sur tout le réseau). Donc, en effet, sur le réseau, il y aura un autre annuaire gérant le fichier star wars si des pairs en possèdent des morceaux.
 - **Lorsqu'un pair souhaite connaître l'adresse IP de l'annuaire référençant le fichier qui l'intéresse, est-ce : - l'annuaire à qui il le demande qui doit lui retourner l'adresse - ou l'annuaire qui référence le fichier qui le contacte ?**
 - Il est écrit dans le sujet : " [Si un pair veut télécharger un fichier], il demande à un annuaire l'adresse de l'annuaire gérant le fichier désiré." Autrement dit, il contacte un annuaire qu'il connaît, et cet annuaire lui renvoie l'adresse de l'annuaire qui gère le fichier demandé.

- **Est-ce qu'un morceau de fichier peut se trouver chez plusieurs pairs ?**
- Oui.
- **Quelle est la taille d'une Socketaddress ? en bits ?**
- A vous de voir, en sachant qu'une adresse IP est une suite de 4 nombres entre 0 et 255 ($= 2^8$), et qu'un port est un nombre entre 1 et 65 536 ($= 2^{16}$).