

Les requêtes HTTP et HTTPS

Texte de Patrick Sentinel – 29A

Votre navigateur web effectue des milliers de requêtes lorsque vous l'utilisez. Évidemment, tout se passe en dessous du capot, sans que vous en ayez connaissance. Mais lorsque vous tapez `www.amazon.ca`, il y a plusieurs actions qui s'exécutent. Afin de bien comprendre la sécurité des requêtes, il faut saisir le fonctionnement de celles-ci. Voici donc en détail comment fonctionnent les requêtes HTTP.

Dans votre navigateur, lorsque vous effectuez une requête, vous tapez une URL dans la barre d'adresse (URL signifie *Universal Resource Locator*). Cette URL permet d'identifier de façon unique un accès à une ressource. On utilise souvent l'expression « adresse web » plutôt que URL. Une URL est formée de trois composantes :

1. Le protocole (le type de connexion) :
 - a. HTTP : HyperText Transfert Protocol;
 - b. HTTPS : HyperText Transfer Protocol Secure.
2. Le domaine :
 - a. C'est l'entité par rapport à laquelle vous désirez effectuer la requête. Par exemple, `google.com`, `29a.ca` et `amazon.ca` sont toutes des entités;
 - b. Tous les domaines appartiennent à un TLD (*Top Level Domain*) comme `.ca`, `.com`, `.net` et autres;
 - c. L'adresse du IP (*Internet Protocol*) du domaine est fournie par un service DNS (les DNS sont un mécanisme qui sert à identifier une adresse IP par un nom plutôt que par une série de chiffres et de point telle que `xxx.xxx.xxx.xxx`).
3. Le document :
 - a. L'URL peut se référer à un document spécifique tel qu'un PDF.

Lorsqu'on tape l'URL dans la barre d'adresse du navigateur et qu'on appuie sur Enter, on effectue une requête auprès d'un serveur. Cette requête prend la forme suivante : *Request/Response*. Expliquons en détail le **fonctionnement** de chacun d'eux :

1. *Request* :
 - a. Le navigateur obtient l'adresse IP du domaine (une série de chiffres et de points comme `xxx.xxx.xxx.xxx`);
 - b. Le navigateur prépare une requête;
 - c. Le navigateur ajoute des *headers* dans la requête. Les *headers* (en-têtes) contiennent l'information sur :
 - i. Le protocole;
 - ii. La communication;
 - iii. Le contenu de la requête.

Response :

- d. Le serveur formule une réponse en retournant un document;
- e. La réponse contient également des *headers*;
- f. Le navigateur détermine comment présenter le ou les documents;
- g. Le serveur retourne un code de réponse.

Constitution des *Request/Response*

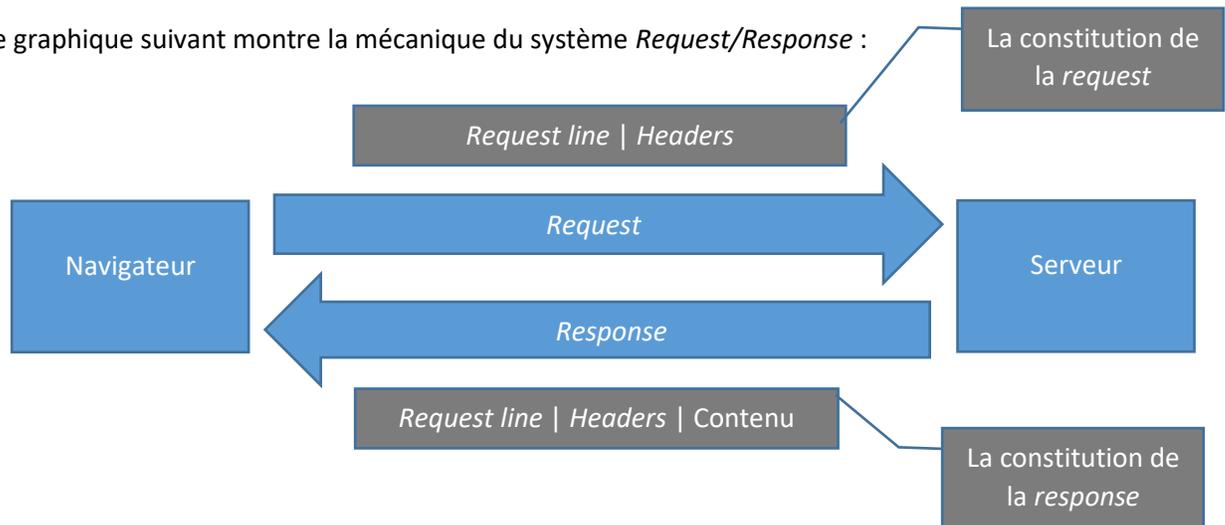
Les *requests* sont constituées des deux éléments suivants :

1. Une *request line*;
2. Des *headers* (en-têtes).

Les *responses* sont constituées des trois éléments suivants :

1. Un *status line*;
2. Des *headers*;
3. Du contenu (le contenu peut être un document, plusieurs documents, une page web, etc.).

Le graphique suivant montre la mécanique du système *Request/Response* :

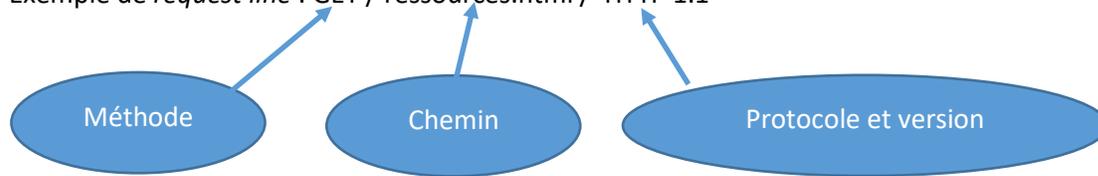


Composition de la *request*

Voyons en détail la constitution des deux éléments de la *request*.

1. ***Request line*** : Elle possède les trois éléments suivants :
 - Méthode : représente le type de la requête (*get, post, head*);
 - Chemin : contient une partie de l'URL à laquelle on veut accéder. C'est ce qui vient après le domaine;
 - Protocole : le protocole utilisé ainsi que la version.

Exemple de *request line* : GET / ressources.html / HTTP 1.1



2. **En-têtes** : Ils sont constitués de paires « Nom et Valeur », comme le montre le résultat suivant :

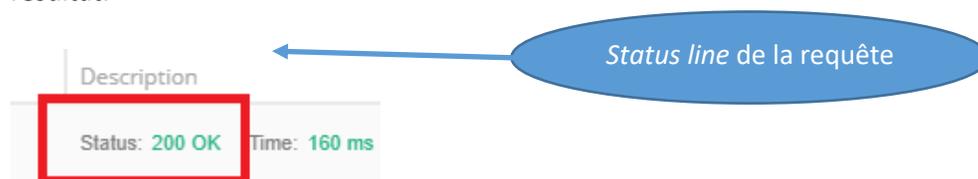


Il existe plusieurs autres types de couples « Nom et Valeur » non présentés dans l'image précédente. L'objectif est de présenter sa structure et non toutes les valeurs possibles. Faites quelques essais avec un logiciel effectuant des requêtes IP. Vous constaterez les résultats ☺

Constitution de la *response*

Voyons en détail les trois composantes de la *response*.

1. **Status line** : Elle exprime le résultat de la requête en donnant de l'information sur le résultat.



La *status line* est composée des trois éléments suivants :

1. Protocole;
2. *Status code*;
3. *Reason line*.

Vous pouvez consulter la liste de tous les codes à l'adresse suivante : https://fr.wikipedia.org/wiki/Liste_des_codes_HTTP

2. **En-têtes** : Les en-têtes de la réponse sont basés sur la même structure que les en-têtes des *requests*.
3. **Contenu** : Le contenu peut être un statut, un document, plusieurs documents, etc.

En résumé, sans les *Request/Response*, il est impossible pour un navigateur d'obtenir de l'information contenue dans un serveur. Elles représentent la base à partir de laquelle les informations circulent dans la Toile!

Tous droits réservés – www.29A.ca – Patrick Sentinel @ 2019
