EP2IC3IV - ECUE Introduction au web

Introduction aux Réseaux

Nous avons précédemment travaillé sur la structuration de pages Web à l'aide des feuilles de styles CSS. Cependant, pour que ces pages Web soient accessibles par le public, elles doivent être hébergées dans un serveur. Ainsi, par le biais du protocole HTTP et d'une infrastructure réseau, le serveur enverra nos pages Web aux navigateurs souhaitant y accéder.

Nous allons donc maintenant nous intéresser brièvement à certains concepts réseau, et au protocole HTTP plus tard, ce qui va être utile pour permettre de comprendre finalement comment la "magie" des services Web opère.

Adresses IP et ports

Lorsqu'on souhaite se connecter à une machine distante (pour télécharger ou envoyer un fichier, parcourir une page web, lire ses emails, etc.), il faut connaître au minimum, les 3 informations suivantes :

L'identité de la machine distante (« le serveur ») qui fournit le service recherché;
Le port d'écoute du serveur;
Le protocole de transport.

1.1. Identité d'une machine

Fréquemment, nous connaissons l'identité de la machine distante par son nom canonique (« humain »). Par exemple, www.google.fr (ou google.fr uniquement), smtp.unice.fr, etc. Cependant, pour se connecter à une machine distante, ce nom de machine doit toujours être traduit en son ID, qui est une adresse IP, version 4 (adresse IPv4) ou version 6 (IPv6).

Actuellement, en France, la plupart des serveurs utilisent une adresse IPv4 (selon le **rapport de l'Arcep d'octobre 2018**, seulement 26% des sites les plus populaires en France utilisent le protocole IPv6). Une adresse IPv4 est composée de 4 octets. Une adresse IPv4 est représentée par la valeur numérique de chaque octet en forme décimale, avec un point de séparation entre chaque octet : « w.x.y.z ».

Par conséquence, lorsqu'on visite un site web avec un navigateur, le nom du site entré sur la barre de navigation est toujours traduit en adresse IP par le navigateur web. De ce fait, on pourrait ne pas connaître le nom canonique de la machine distante ou tout simplement vouloir travailler avec une machine distante sans nom canonique, et donc la désigner uniquement avec son adresse IP et pouvoir néanmoins travailler dessus.

La traduction du nom canonique en adresse IP et vice-versa se fait grâce à l'utilisation d'un protocole appelé DNS (Domain Name System - Système de Noms de Domaines).

Important : Un client (votre ordinateur par exemple) doit également avoir une adresse IP pour pouvoir envoyer et recevoir des données par le réseau.

1.2. Relation numéro de port <-> service

Un serveur est à l'attente des demandes, appelées « requêtes », des clients à travers un numéro de port d'écoute. Un seul serveur peut donc offrir plusieurs services (e.g. serveur de pages web, serveur de messagerie, etc) en ouvrant plusieurs ports.

Un numéro de port est un entier non-signé (i.e. numéros positifs uniquement) écrit sur 2 octets. Il y a donc au total 65536 ports, divisés de la manière suivante :

les numéros de port inférieurs ou égaux à 1023 sont des ports dits « bien connus ». Ils sont utilisés par de services standardisés (HTTP, SMTP, DNS, ...).
Dans un système d'exploitation, l'utilisation de ports bien connus requiert des droits d'administration (droits de root dans les systèmes basés sur UNIX, ou Administrateur dans les systèmes Windows).
Le numéro de port donne donc un premier indice sur le service offert. Par exemple, le port 80 est utilisé par le protocole HTTP (le service Web), le port 443 est utilisé par le protocole HTTPS (service Web sécurisé), le port 25 par le protocole SMTP (le service de transmission d'e-mails), etc.
les ports compris entre 1024 et 49151 sont des ports « enregistrés ». Le succès de l'Internet a favorisé le déploiement d'une multitude de services. Certains de ces services sont tellement utilisés que les organismes de standardisation ont décidé de leur assigner un port spécifique.
Il n'y a pas besoin d'être root ou administrateur de la machine pour écrire un programme ouvrant un port enregistré.
Entre 49152 et 65535 nous avons les ports « dynamiques » ou éphémères
tout utilisateur peut les ouvrir avec ou sans droits d'administration ;
généralement, un client (e.g. Firefox) ouvre l'un de ces ports (avec une politique de *round robin* sur Linux) pour entamer un dialogue (envoyer des requêtes) avec un serveur.

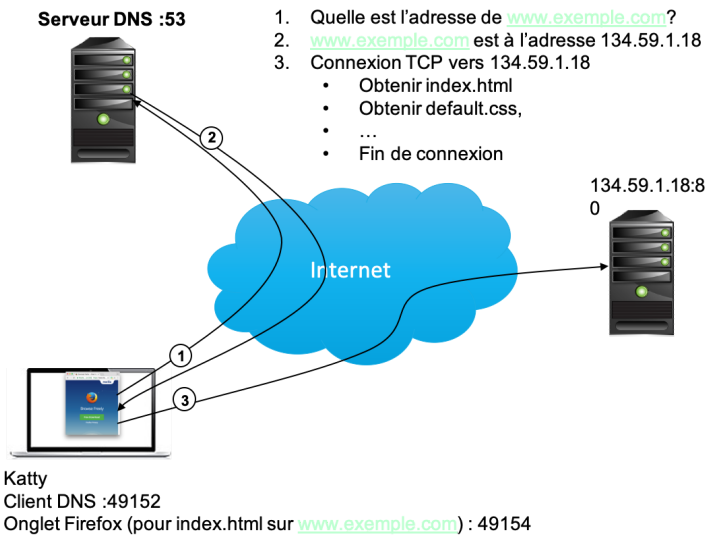
1.3. Le protocole de Transport

L'objectif d'un protocole de transport est de fournir un chemin entre deux applications distantes. Il existe 2 principaux protocoles de transport : TCP (*Transport Control Protocol*) et UDP (*User Datagram Protocol*).

Le protocole de transport UDP est un protocole non-fiable. Un client ou serveur utilisant le protocole UDP envoie un message à la machine distante, mais sans aucune aide de l'application distante, ce premier n'a aucun moyen de savoir si son message a été bien reçu ou non. Le protocole de transport TCP est un protocole fiable et très complexe. Une machine envoyant un message à une autre par le protocole TCP peut savoir si son message a été bien reçu ou non, grâce à un système d'acquittements mis en place par TCP.

Étant donné que le service Web doit être un service fiable, le service Web utilise par conséquence le protocole de transport TCP.

La figure ci-dessous illustre graphiquement les grandes étapes à accomplir pour télécharger une page web à partir d'un nom canonique de serveur.



Quelques outils réseau

La commande « `ip` »

Il arrive parfois que notre navigateur n'arrive pas à télécharger une page et nous conseille de vérifier notre connectivité réseau. Notre premier réflexe doit être donc de vérifier que notre machine (cliente ici) posséde bien une adresse IP. Pour cela, nous pouvons utiliser la commande « `ip` » avec ces paramètres « `a s` » (comme « address » et « show »).

Voici un exemple de sortie de la commande :

```
$ ip a s
2: enp0s3:<BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP group
default qlen 1000
    link/ether 08:00:27:8d:ac:03 brd ff:ff:ff:ff:ff:ff
```

```
    inet 10.0.2.15/24 brd 10.0.2.255 scope global dynamic noprefixroute enp0s3
       valid_lft 86392sec preferred_lft 86392sec
    inet6 fe80::7606:5464:4e87:101a/64 scope link noprefixroute
       valid_lft forever preferred_lft forever
```

Dans l'exemple, nous pouvons voir que l'adresse IPv4 est la 10.0.2.15. Notre interface réseau (avec ID « 2 ») possède le nom « enp0s3 ».

Attention : toute machine possède par défaut une adresse IPv4 débutant par 127, très souvent la 127.0.0.1. Cet adresse, dite de « boucle locale », permet à 2 processus dans une même machine de communiquer via le protocole IPv4, mais elle ne permet pas la communication entre 2 processus distants.

La commande « `host` »
Si nous possédons bien une adresse IP, l'erreur pourrait venir d'un problème du nom de domaine (nom mal écrit) ou même, d'un serveur DNS défaillant. Dans ces 2 cas précédents, il serait impossible d'obtenir une adresse IP de serveur auquel se connecter pour télécharger des pages d'un site web.

La traduction d'un nom canonique en adresse IP (IPv4 ou IPv6) se fait automatiquement par la plupart des programmes réseau. Mais si nous avons besoin de valider l'existence d'un nom de serveur ou du serveur DNS, nous pouvons exécuter manuellement un client DNS avec la commande `nslookup` ou `host`.

Pour vérifier qu'une adresse IP existe pour une URL donnée, nous utiliserons la commande host de la manière suivante. Dans une URL, le nom du serveur se trouve entre le nom du protocole suivi de « :// » et le caractère « : » ou le prochain « / ». Par exemple

pour l'URL `https://lms.univ-cotedazur.fr/pluginfile.php/151410/mod_resource/content/2/08-09%20Redirections.pdf` le nom du serveur sera *lms.univ-cotedazur.fr*
pour l'URL `https://myserver.domaine.fr:8080/index.php` le nom du serveur sera *myserver.domaine.com*.
Voici l'exécution de la commande host pour déterminer l'adresse IP des 2 serveurs précédents :

```
$ host www.univ-cotedazur.fr
www.univ-cotedazur.fr is an alias for univ-cotedazur.fr.
univ-cotedazur.fr has address 149.202.195.14
univ-cotedazur.fr mail is handled by 10 ip-nice06.unice.fr.
univ-cotedazur.fr mail is handled by 20 ip-sophia06.unice.fr.
```

```
$ host myserver.domaine.fr
Host myserver.domaine.fr not found: 3(NXDOMAIN)
```

Dans le premier cas, il y a bien une adresse IP pour le serveur, mais pas pour le 2ème.

La commande « `ping` »
Si le serveur DNS est opérationnel et le nom du serveur bien écrit, il nous reste à vérifier qu'il existe un chemin (une route) entre notre ordinateur et le serveur Web. Pour cela, nous pouvons utiliser la commande « ping », qui est un outil permettant d'envoyer un type de requêtes spécifique à une machine distante et attendre la réponse pour vous dire si la machine est joignable ou pas.

Étant donné que les réponses aux requêtes ping peuvent être utilisées par un hacker pour avoir un aperçu de l'architecture réseau (attaque appelée *netwok scanning*), certains administrateurs décident de donner une configuration spéciale aux serveurs afin qu'ils ne répondent pas aux requêtes ping, ou bien, de configurer l'infrastructure réseau pour bloquer ce type de requêtes. Actuellement, par défaut Windows bloque par son firewall toute requête ping. Une réponse négative de la commande ping ne signifie donc pas forcément que la machine distante n'existe pas ou que le chemin vers le serveur est cassé.

En revanche, un ping positif est un signal clair que la machine distante existe et qu'il existe un chemin entre votre machine et la cible. Par conséquence, un ping positif pour le serveur Web indique tout simplement que l'erreur de navigation se trouve dans le serveur distant, due très probablement à une application défaillante.

Pour exécuter la commande ping, vous devez lui donner en argument l'adresse IPv4 ou le nom canonique de la machine. Par défaut, ping envoie un nombre infini de requêtes. Pour terminer son exécution, vous pouvez utiliser Ctrl+C. Vous pouvez aussi indiquer un nombre fixe de requêtes à envoyer, avec le paramètre « -c ». Voici un exemple de la commande ping pour tester la présence de la machine 10.0.2.4 en lui envoyant 3 requêtes uniquement, et les réponses positives correspondantes :

```
$ ping -c3 10.0.2.4
PING 10.0.2.4 (10.0.2.4) 56(84) bytes of data.
64 bytes from 10.0.2.4: icmp_seq=1 ttl=64 time=0.284 ms
64 bytes from 10.0.2.4: icmp_seq=2 ttl=64 time=0.299 ms
64 bytes from 10.0.2.4: icmp_seq=3 ttl=64 time=0.230 ms
```

Chaque ligne commençant par "64 bytes..." est un rapport du résultat de la requête ping. Sans aller dans le détail, sachez que la colonne "time=" indique le temps écoulé entre l'émission de la requête ping et la réception de sa réponse. La commande ping peut à ce titre être utilisée pour savoir si la machine distante est très éloignée ou pas de notre machine (e.g. si nous sommes en France, un ping vers une machine aux USA donnera un temps d'aller-retour plus long que si nous ciblons une machine en France). Et voici maintenant un exemple avec des réponses négatives du ping

```
$ ping -c3 10.0.2.5
PING 10.0.2.5 (10.0.2.5) 56(84) bytes of data.
From 10.0.2.15 icmp_seq=1 Destination Host Unreachable
From 10.0.2.15 icmp_seq=2 Destination Host Unreachable
From 10.0.2.15 icmp_seq=3 Destination Host Unreachable
```

La commande « nmap »
Si la commande ping envoie des requêtes généralement filtrées et souvent bloquées, il existe un autre outil que vous pouvez installer et utiliser si vous êtes administrateur de l'ordinateur. Il s'agit de la commande nmap. Contrairement à la commande ping, nmap n'est pas disponible par défaut dans les systèmes Linux. Mais vous pouvez l'installer avec la commande "sudo apt install nmap".

nmap est un outil d'audit de sécurité pour les réseaux. Nous pouvons utiliser nmap pour exécuter un ping TCP. Sachant que le protocole TCP n'est pas bloqué et que, un serveur Web doit en principe écouter sur le port 80 (HTTP) ou 443 (HTTPS), voici comment utiliser nmap pour savoir si la machine distante est joignable et disponible sur le port 80 :

```
$ nmap -p 80 www.univ-cotedazur.fr

Starting Nmap 7.60 ( https://nmap.org ) at 2019-11-09 21:36 CET
Nmap scan report for www.univ-cotedazur.fr (149.202.195.14)
Host is up (0.086s latency).
rDNS record for 149.202.195.14: ns3018279.ip-149-202-195.eu

PORT STATE SERVICE
80/tcp open http

Nmap done: 1 IP address (1 host up) scanned in 0.89 seconds
```

Voici un « ping TCP » avec nmap pour une machine existante, dont le serveur Web n'est pas opérationnel

```
$ nmap -p 80 10.0.2.2

Starting Nmap 7.60 ( https://nmap.org ) at 2019-11-09 21:55 CET
Nmap scan report for _gateway (10.0.2.2)
Host is up (0.00083s latency).

PORT STATE SERVICE
80/tcp closed http

Nmap done: 1 IP address (1 host up) scanned in 0.08 seconds
```

Les Uniform Resource Locators (URLs)

Une URL est un type spécifique d'URI (Uniform Resource Identifier), qui est utilisé pour désigner une ressource Web. Elle composée de plusieurs éléments, qui indiquent le type de protocole à utiliser pour accéder à une telle resource, l'identité de la machine et emplacement à l'intérieur de cette machine. Pour connaître chaque élément d'une URL, vous devez lire cette page **https://fr.wikipedia.org/wiki/Uniform_Resource_Locator**.

Exercices

Pour ce TP, vous devez ajouter dans votre rapport les captures d'écran montrant l'exécution des commandes et les résultats affichés, en plus des explications demandées dans les questions.

1. Toute machine connectée au réseau a une adresse IP associée. Si vous êtes capable de naviguer sur Internet, c'est que votre machine est connectée au réseau et qu'elle possède une adresse IP. Avec la commande `ip`, listez les interfaces réseaux de votre machine et indiquez quelle adresse IPv4 vous permet de naviguer sur Internet.
2. Utilisez la commande `host` pour trouver l'adresse IP de la machine www.i3s.unice.fr. Ensuite, ouvrez votre navigateur web et écrivez dans sa barre d'adresses l'adresse IP trouvée. Vérifiez que vous téléchargez en effet la page web du site. Essayez de faire de même avec d'autres sites web.
3. Ecrire une adresse IP dans la barre d'adresse et réussir à télécharger une page web n'est plus une pratique courante. En effet, les serveurs web tournent sur des machines très puissantes et « servir » les pages pour un seul domaine gaspillerait la puissance d'une telle machine. Si on utilise un navigateur Internet, il faut écrire dans la barre d'adresses le nom de la machine. De cette manière, un seul serveur (et donc une seule adresse IP) peut être utilisé pour « servir » les pages de plusieurs domaines. Explorez avec www.unice.fr et www.mit.edu les 2 réponses distinctes que les serveurs actuels donnent lorsque l'on les contacte par leur adresse IP plutôt que par un nom de domaine.
4. Expliquez chaque composante de l'URL de cette page Moodle. Attention : certains navigateurs « cachent » l'emplacement de la ressource et ne montrent que le nom de la machine (+ le nom du domaine). Pour être sûr que vous obtenez l'URL dans sa totalité, clickez sur la barre d'adresse et explorez la totalité de la chaîne de caractères.
5. Faites un ping TCP avec la commande `nmap`, à installer si nécessaire, vers les serveurs suivants **par leur adresse IPv4** : www.unice.fr, www.mit.edu (serveur web du MIT aux USA), www.njcu.edu (serveur Web du New Jersey City University aux USA), ufrj.br (serveur Web de l'université de Rio de Janeiro, au Brésil).
   - Que constatez vous par rapport au temps d'aller-retour des réponses ? Expliquez vos observations
   - Copiez/collez dans votre rapport les sorties des commandes ping et nmap.
   - Expliquez comment avec un ping (TCP ou normal), il est possible d'avoir une idée de la proximité géographique d'un serveur.
6. **(Optionnel)** Utilisez la commande `ping` afin de vérifier la disponibilité de certaines machines dans le réseau. Si vous travaillez sur une VM, faites plusieurs pings pour les adresses IP comprises entre 10.0.2.1 et 10.0.2.14 et dites quel temps d'aller-retour est rapporté et quelles machines sont disponibles. Si vous êtes sur le poste physique, vous pouvez activer le mode « Point d'Accès » de votre smartphone et connecter votre machine à votre smartphone. Ensuite, pingez l'adresse IP qui se trouve juste après le mot « via », après avoir exécuté la commande « ip r s | grep "via" ». Copiez/collez le résultat du ping dans votre rapport de TP.

Votre progression

Compte Rendu ☐

Corrections ☐

◄Responsive Web Design

Le Protocole HTTP/1.1►

Tableau de bord
Mes cours 2021-2022
Mes cours 2020-2021
Mes cours 2019-2020
Recherche de cours
Espaces enseignant
Documentation Moodle
Espace de test Moodle
Pédagothèque
Syllabus mode d'emploi
Construire et enrichir son cours
Adopter les compétences
Formations
Boite à suggestions
Espaces étudiant
Kit de (sur)vie étudiant
Hub pour rebondir
Utiliser son portfolio
BU - Metoda
Boite à suggestions
Assistance

Résumé de conservation de données