UNIVERSITÉ TOULOUSE III PAUL SABATIER

2020/2021

Travaux Pratiques : TCP IP

Dans la suite du document, vous devrez rédiger un rapport répondant à chacune des questions posées à envoyer par mél à
Patrice.Torguet@irit.fr ou **Philippe.Truillet@irit.fr**
avant le **13 juin 2021 23h55 UTC**

1. Rappel : adressage dans Internet

1.1 Identification d'une machine

Une machine (appelée aussi hôte ou host) est identifiée dans l'Internet par son adresse. L'adresse Internet d'une machine correspond à un numéro qui est unique dans le monde.
Pour des raisons mnémoniques, il est possible de donner un nom à une machine (ex : `Toto`, `Garonne`, `Mimosa`...). Attention, ce nom n'est pas compris par le réseau pour le routage. Certains hôtes ont plusieurs noms.
Les applications (`ftp`, `http`, `ssh`, ...) acceptent aussi bien les adresses que les noms.

1.2 Structure de l'adresse IPv4

L'adresse utilisée par le protocole IP (adresse IP), comporte deux champs : le champ adresse réseau (Network) dans Internet et le champ adresse hôte (Host) dans le réseau. Sa taille est de quatre octets. Elle est souvent donnée en notation décimale pointée (ex : 127.95.35.54).
Comme l'adresse IP contient l'adresse réseau, une station changeant de réseau change d'adresse. D'autre part, une station multi-domiciliée (qui dispose de plusieurs interfaces réseau) ou un routeur ont plusieurs adresses.
L'adresse réseau est donnée par un organisme officiel (le **NIC** : **N**etwork **I**nformation **C**enter) garantissant l'unicité de cette adresse. En France il faut contacter l'INRIA (`fr-NIC@inria.fr`) ou le GIP Rénater (`rensvp@renater.fr`). Le champ adresse hôte est donné par l'administrateur réseau.

1.3 Les classes d'adresses IPv4

Il existe à l'origine trois classes d'adresses IP qui permettent de gérer des réseaux de tailles diverses (RFC 790).

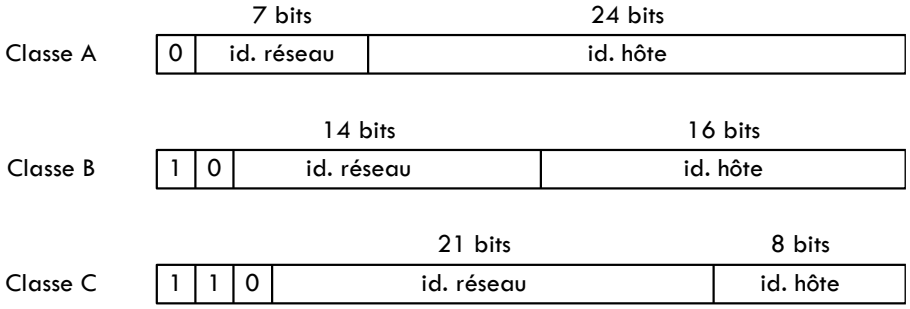


Figure 1 : Les trois classes d'adresses ipv4 (RFC 790)

Ainsi, les adresses de classe **A** ont leur premier octet compris entre **1** et **127**, les adresses de classe **B** entre **128** et **191** et les adresses de classe **C** entre **192** et **223**.
Les adresses qui ont leur octet supérieur à **223** sont des adresses spéciales qui ne désignent pas une machine.

Remarques :
Les adresses IP commençant par 127 sont réservées pour des tests en local. Par exemple 127.0.0.1 correspond à la machine locale (localhost).

Une nouvelle version de IP (**IPv6** pour IP version 6) défini les adresses sur 128 bits au lieu de 32 de façon à pouvoir gérer un plus grand nombre d'adresses et de niveaux hiérarchiques d'adresses. Cela dit les adresses IP sur 128 bits ne sont pas encore tout à fait généralisées.

Classe	Bits de départ	Début	Fin	Notation CIDR	Masque de sous-réseau par défaut
Classe A	**0**	1.0.0.0	127.255.255.255[2]	/8	255.0.0.0
Classe B	**10**	128.0.0.0	191.255.255.255	/16	255.255.0.0
Classe C	**110**	192.0.0.0	223.255.255.255	/24	255.255.255.0
Classe D (multicast)	1110	224.0.0.0	239.255.255.255	/8	non défini
Classe E (réservée)	1111	240.0.0.0	255.255.255.255		non défini

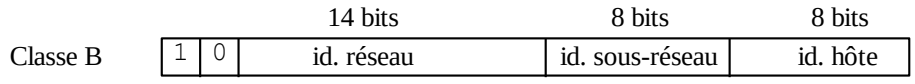
Figure 2 : résumé des plages d'adresses IPv4

Notion de sous-réseaux

Comme on vient de le voir il y a un très grand nombre possible de machines par réseau de classe A et B. Même un réseau de classe C, avec 254 machines possibles, est souvent trop important pour être contenu sur un même réseau physique (ex. un réseau Ethernet). Il est donc nécessaire de fractionner un réseau en groupes plus petits.
La RFC 917 créé le concept de sous-réseaux qui introduit une hiérarchie complémentaire.
Un réseau partitionné en sous-réseau est vu de l'extérieur du réseau (par exemple de l'Internet) comme un seul réseau.
Pour créer les sous-réseaux on va utiliser une partie de l'identificateur d'hôte. Ainsi on aura 3 niveaux hiérarchiques d'identificateurs :

- l'identificateur de réseau dans Internet
- l'identificateur de sous-réseau dans le réseau
- l'identificateur d'hôte

Exemple de partitionnement d'un réseau de classe B en 254 sous-réseaux (les valeurs avec que des 1 ou que des 0 sont interdites) :



On peut choisir de réserver un nombre quelconque de bits pour l'identificateur de sous-réseau en fonction du nombre de sous-réseau et du nombre maximum de machines par sous-réseau que l'on désire. La seule restriction est que chaque identificateur ne doit pas être composé que de 1 ou que de 0 et donc l'identificateur de sous-réseau (par convention uniquement et pour respecter d'anciennes utilisations de ces valeurs) tout comme l'identificateur d'hôte final ne peuvent pas être codés sur 1 bit.

Masque de sous-réseau

Pour pouvoir communiquer correctement avec une machine il faut savoir si elle est connectée au même réseau (et donc si on peut lui "parler" directement) ou si il faudra s'adresser à un équipement réseau (un routeur) qui pourra la contacter.

2 machines sont connectées sur le même réseau physique si leurs identificateurs de réseau et leurs identificateurs de sous-réseau sont identiques. C'est à dire si les seuls bits qui différent entre les adresses sont contenus dans la partie identificateur d'hôte.
Compte tenu de la présence de sous-réseaux, on ne peut donc pas déterminer directement à partir d'une adresse IP et de sa classe quels sont les bits qui sont réservés aux identificateurs de machine.
Il faut donc une autre information : le nombre de bits utilisés pour l'identificateur de sous-réseau.
Pour indiquer cette information on utilise un masque binaire qui permet de déterminer la partie identificateur de réseau + identificateur de sous-réseau d'une adresse IP avec un simple et binaire. Ce masque binaire est appelé : masque de sous-réseau.

La RFC 1338 propose d'abolir la notion de classe et le CIDR (**C**lassless **I**nter-**D**omain **R**outing) est mis au point en 1993. L'utilisation de masques de longueur variable découpe l'espace d'adressage en blocs (/x) de taille variable permettant une utilisation plus efficace de l'espace d'adressage. Néanmoins, le **25 novembre 2019 à15h35**, le RIPE NCC (Registre Régional d'attribution des adresses IP Europe/Asie) a attribué les derniers blocs /22 (les derniers blocs /8 ont été attribués en 2012).
Il reste encore quelques adresses IPv4 par bloc /24 en récupérant des adresses inutilisées ou récupérées.

La transition vers IPv6 ne devrait que s'accélérer (taux d'utilisation de 36% en France fin 2019 et 25% au niveau mondial).

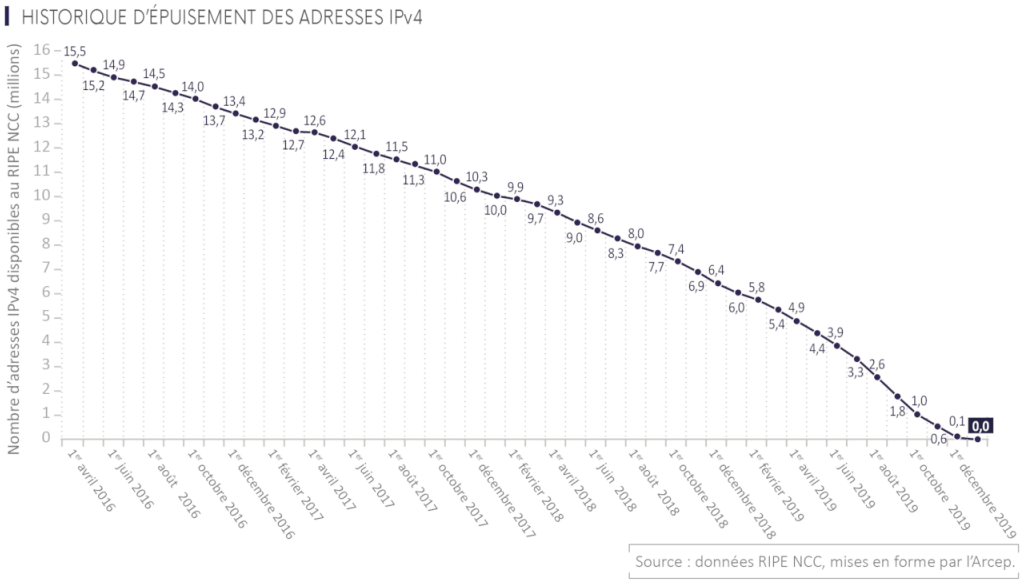


Figure 3 : Epuisement des adresses IPv4 en Europe

2. Wireshark

2.1 Introduction

Tout au long du reste du TP, nous allons utiliser Wireshark (`https://www.wireshark.org`), un outil dit de « *capture réseau* » afin d'observer des données qui transitent sur le réseau.

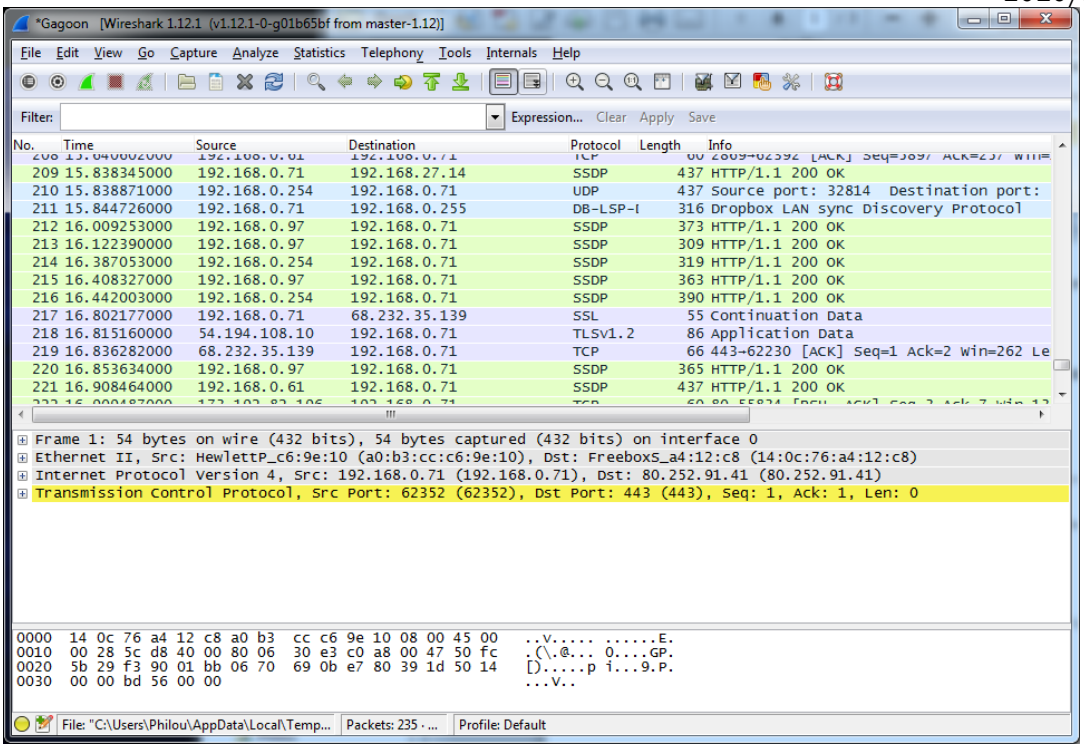


Figure 4 : interface de Wireshark

L'interface utilisateur se compose de plusieurs parties :

- **Barre de menus/icônes**
- **Barre de filtrage** : cette zone permet de filtrer a posteriori une capture pour isoler une partie du trafic réseau.
- **Fenêtre de la liste des messages capturés.** On retrouve sur chaque ligne le numéro du message, le moment de sa capture (en secondes depuis le début), son adresse source et destination, le protocole de plus haut niveau décodé et le résumé des champs caractéristiques de ce protocole.
- **Fenêtre de la pile de protocoles décodés pour le message sélectionné.** On trouve ici la pile des protocoles décodés allant du niveau liaison de donnée jusqu'au niveau le plus haut reconnu. On a aussi accès aux différents champs de chaque protocole.
- **Fenêtre d'affichage brut du message sélectionné.** Tous les octets du message sont affichés en hexadécimal et sous forme textuelle (ce qui n'est intéressant que s'il s'agît de messages textuels).

Démarrez le logiciel Wireshark. **Capturez du trafic réseau** (**seulement avec des droits administrateur**) et familiarisez-vous avec l'interface utilisateur en identifiant les différentes parties exposées précédemment.

2.2 Concept de protocole

- Rappelez ce qu'est un « **protocole** »
- Nous allons observer le trafic HTTP, protocole applicatif utilisé dans Internet entre un client (un navigateur par exemple) et un serveur web hébergeant le site. Enregistrez le trafic réseau et naviguez sur le site **http://www.redirection-web.net**. Pour n'afficher que le trafic *http*, utilisez la zone **Filter** en saisissant le mot-clé **http**.
  Pensez à consulter la 2ème partie de la fenêtre pour avoir accès aux différents champs des unités de données des protocoles (PDU).
- Quelle est la commande HTTP utilisée par le client pour demander au serveur une ressource (fichier ou page web). Quels sont les types de réponse qui sont mis en évidence ? Donnez leur code de retour.

- Observez le contenu de la réponse du serveur. Quelle est la taille (en octets de la réponse HTTP (message HTTP incluant les données) ? Que représente le texte décrit en langage HTML ?

Nous allons vérifier la validité du fichier HTML. Pour cela, sélectionnez le champ **Line-based text data** et exporter la sélection dans un fichier index.html (menu | fichier « **Export Selected Packet Bytes** ») Ouvrez ce fichier dans un navigateur afin de confirmer votre hypothèse de la question précédente.

2.3 Concept de couche et d'encapsulation

Nous venons d'illustrer le concept de protocole à travers le protocole HTTP permettant à un navigateur d'obtenir une page web. Cette communication n'est en réalité pas directe entre les deux processus applicatifs (client et serveur) mais s'appuie pour fonctionner sur un certain nombre de services sous-jacents. Les systèmes de communication sont construits traditionnellement selon un modèle en couches.

- rappelez la définition du concept de « ***couche*** »
- sur la capture effectuée précédemment, observez les différents niveaux d'encapsulation d'un message.

2.4 Concept de point d'accès au service

Cette mise en œuvre d'un protocole pose plusieurs questions : en particulier de savoir comment le réseau sait à quelle machine est destiné un message et comment chaque couche sait à quelle entité de la couche supérieure remettre le contenu d'une unité de protocole. En ce qui concerne la communication entre deux couches adjacentes, celle-ci se fait grâce à un **point d'accès au service** (SAP)

1. L'identifiant du point d'accès au service de la couche transport (T-SAP) est un numéro de port. Quel est le numéro de port pour l'entité HTTP de la machine **www.redirection-web.net** ? Et de la machine communiquant avec cette entité ?
2. A partir d'un des messages HTTP de la capture, trouvez les autres identifiants de point d'accès aux services des couches réseaux et liaison (acheminement des unités de données entre les couches Transport-Réseau et Réseaux-Liaison)

2.5 Concept d'adresse

Chaque machine est identifiée par une adresse qui permet de la distinguer des autres. Une adresse assure donc l'identification des correspondants d'une communication. Il existe plusieurs types d'adresses en fonction du protocole utilisé.

Une adresse liaison de type Ethernet suit un format standardisé par l'IEEE. En particulier, les **trois premiers octets** identifient de manière unique le constructeur/vendeur à qui est attribuée l'adresse. Cet identifiant est appelé OUI (*Organizationally Unique Identifier*)
Vous avez par exemple accès à la liste ici : **https://macvendors.com**

1. Quelle est l'adresse Ethernet (MAC) de votre machine ? Utilisez la commande **`/sbin/ifconfig`** pour en savoir plus sous Unix/MacOS et **`ipconfig /all`** sous windows.
2. Déterminez à quelle organisation appartient cette adresse.

Une adresse IP comporte **deux parties** (voir partie **2.**) Le nombre de bits dédiés à la partie réseau peut être connu implicitement à partir des classes d'adresses ou explicitement par un masque réseau.

1. Quelle est l'adresse IP de votre machine ?
2. A partir du masque de réseau, déterminez la partie réseau et la partie équipement de votre adresse IP

Votre adresse est-elle une adresse privée ? (consultez **https://www.arin.net/knowledge/address_filters.html** pour en savoir plus)

L'IANA (*http://www.iana.org*) est l'organisme responsable de la coordination du système d'adressage d'internet. Il alloue des adresses IP à des registres régionaux (*RIR – Regional Internet Registry*)

1. Donnez 5 registres régionaux d'adresses IP ainsi que la zone géographique qu'ils couvrent

2. Consultez le registre d'allocation des plages d'adresses IPv4 sur le site de l'IANA (**http://www.iana.org/numbers**)
    - A quel RIR est alloué le préfixe **193** ?
    - Utilisez la commande **`whois`** ou le site **http://ripe.net/whois** pour savoir à qui appartient l'adresse **193.48.251.195**
    - Qui gère les adresses ayant le préfixe **130** ?
    - Interrogez le serveur **whois** pour avoir des informations sur l'adresse **130.120.84.5**
    - Que pouvez-vous en déduire sur les plages d'adresses **130.120.0.0 – 130.120.255.255**
    - A qui appartient l'adresse **130.120.84.5** ?
    - Utiliser l'outil `nslookup` pour connaitre le ou les noms relatifs à l'adresse **130.120.84.5**

3. commandes et outils réseau

Nous avons utilisé les commandes **`ifconfig`**, **`whois`** et **`nslookup`**. Rappelez brièvement la fonction de ces trois commandes.
La commande **`ping`** permet de tester la connectivité au niveau IP. Testez cette commande entre votre machine et celle de votre voisin. Testez la connectivité avec l'adresse **www.univ-tlse3.fr** et une machine extérieure au réseau de l'université (exemple : **www.google.com**) Quelle différence pouvez-vous observer entre les différents *ping* ?

La commande **`traceroute`** permet de tracer l'itinéraire qu'emprunte un paquet IP entre un hôte source et destination.

1. Testez et observez les résultats de la commande. Faites un **`traceroute`** vers **www.univ-tlse3.fr**. Représentez sur un schéma l'interconnexion (au niveau IP) entre votre PC et le serveur web de l'UPS.