

Réseau I : introduction

Juliusz Chroboczek

7 septembre 2020

1 Commutation

Un *réseau informatique* est un ensemble de machines qui peuvent communiquer entre elles. Il n'est pas possible de construire un réseau en connectant directement toutes les paires de hôtes — il faudrait un nombre quadratique de connexions. Il est donc nécessaire d'utiliser une technique qui permet de communiquer là où il n'y a pas de connexion directe — c'est la *commutation*.

1.1 Commutation de circuits

Avant l'Internet, il existait déjà des réseaux à l'échelle planétaire. Le plus visible était le *réseau téléphonique*, basé sur la *commutation de circuits*, une technique qui consiste à connecter bout-à-bout des « circuits » permanents pour obtenir un circuit connectant temporairement les deux participants de la communication.

Lorsqu'il désire communiquer, l'utilisateur contacte le *central téléphonique*, et demande poliment à parler au 22 à Asnières. La « demoiselle des téléphones » insère alors une fiche dans le connecteur correspondant, et la communication s'établit.

On peut aussi effectuer plusieurs étapes de commutations de circuit. Si l'utilisateur désire contacter un numéro à New York, il contacte le central de Paris, qui établit un circuit vers le central à New York, qui le prolonge à son tour jusqu'au destinataire.

La commutation de circuits requiert l'établissement d'un circuit électrique de bout-en-bout, ce qui demande que le matériel soit fiable et soigneusement calibré. De plus, elle ne permet pas le *multiplexage* d'un lien — si le 22 à Asnières participe déjà à un circuit, l'utilisateur est branché sur la tonalité occupée. Enfin, elle n'a pas de bonnes propriétés de résilience aux pannes — si un central téléphonique disparaît, la communication est interrompue.

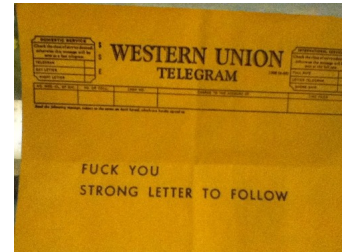


1.2 Commutation de messages

Un autre réseau planétaire était le réseau télégraphique, qui utilisait la *commutation de messages*. Un utilisateur du télégraphe se rend au plus proche bureau des PTT, et remplit un formulaire avec

son message. Le message est alors transmis par le télégraphiste à un bureau des PTT plus proche du destinataire, où il est recopié puis communiqué à un autre télégraphiste. Celui-ci, à son tour, émet le message à un autre bureau de poste. De proche en proche, le message finit par arriver à destination où il est recopié sur papier une dernière fois et remis au destinataire par le facteur.

Le message transite donc à travers plusieurs liens qui ne sont jamais interconnectés physiquement; c'est cette technique qui s'appelle la *commutation de messages*. La commutation de messages est relativement fiable et résiliente aux pannes. Sa principale limitation est qu'elle impose une limite sur la taille maximale du message : il est impossible d'envoyer le texte complet de *Guerre et Paix* par le télégraphe sans le découper en morceaux.



1.3 Commutation de paquets

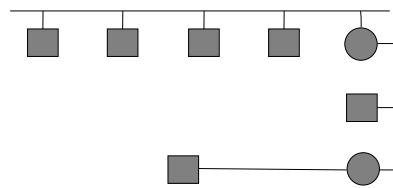
La façon naturelle d'envoyer des messages volumineux est de *ségmenter* le message en morceaux de taille bornée, appelés *paquets*, les numéroter, et les envoyer indépendamment. Du côté du récepteur, les paquets sont reordonnés, les paquets perdus sont réclamés au récepteur (qui les réemet), puis le message est *reassemblé*. Ce type de communication s'appelle la *commutation de paquets*, et tous les réseaux modernes sont basés sur cette technique.

La commutation de paquets s'est avérée avoir de nombreux avantages. Comme chaque paquet est envoyé indépendamment, la commutation de paquets est naturellement résiliente aux pannes : si un câble se rompt, ou si un télégraphiste meurt au travail, il suffit de choisir un autre chemin (« rerouter ») et réémettre le paquet qu'il était en train de transférer. Mais surtout, elle permet des modes de communication variés, notamment la communication simultanée dans les deux sens (pensez à la voix sur IP) et la communication à plusieurs (pensez BitTorrent ou IRC).

Par contre, la commutation de paquets induit en pratique des délais imprévisibles — il est difficile de donner des garanties de temps réel en commutation de paquets (les *gamers* comprendront).

2 Internet

Un *internet* est un réseau constitué de plusieurs *liens* interconnectés à travers des *routeurs* (*router*¹ en anglais). On appelle *hôte* (*host*) une machine qui n'est pas un routeur (et qui sert donc à exécuter des applications); un *nœud* (*node*) est soit un hôte soit un routeur.



L'Internet Global La dynamique naturelle des internets est de fusionner : s'il existe deux gros internets, ils finiront probablement par être interconnectés et ne constituer plus qu'un réseau. De ce fait, il existe un Internet qui est beaucoup plus grand que tous les autres : c'est l'*Internet Global*, ou simplement l'*Internet* (avec une majuscule).

1. Les prononciations [ju:təɪ] et [jæʊtəɪ] sont toutes les deux correctes.

3 Protocoles

Un *protocole* spécifie les règles de la communication entre deux (ou plus) pairs. La définition d'un protocole spécifie le format des messages ainsi que les règles de l'interaction (quels sont les messages autorisés dans un état donné).

Il semblerait évident que les protocoles informatiques doivent être spécifiés formellement. Si c'est en grande mesure le cas du format des messages (qui est généralement spécifié à l'aide de *diagrammes de paquets* ou par une grammaire sans contexte), ce n'est pas le cas des règles de communication, qui sont définies par de la prose plus ou moins précise (parfois cependant accompagnée d'un automate définissant une partie du protocole), et rarement vérifiés automatiquement. En pratique, il est rare que deux implémentations indépendantes du même protocole puissent communiquer (« interopérer ») avant d'avoir été testées.

Suite de protocoles Les règles de la communication sur un réseau ne sont normalement pas définies par un protocole monolithique, mais par une *suite de protocoles*. Une suite de protocoles consiste d'une famille de protocoles organisés en *couches*, et un protocole de la couche n ne peut se servir que des services fournis par la couche $n - 1$. Lorsqu'un protocole n'obéit pas à cette discipline, on parle de *violation* de la structure.

Une couche n est donc définie par :

- le service qu'elle utilise (fourni par la couche $n - 1$);
- le service qu'elle fournit (à la couche $n + 1$).

Un ensemble de couches s'appelle un *modèle*; l'ensemble des protocoles qui implémentent un modèle s'appelle une *suite* de protocoles (*protocol stack*).

Le fait d'obéir à une discipline aussi stricte permet en principe de remplacer un protocole indépendamment des autres, et même d'utiliser plusieurs protocoles simultanément : il est habituel d'utiliser plusieurs protocoles de couche haute sur le même réseau (vous écoutez de la musique transmise à travers le réseau tout en lisant votre mail) au dessus de plusieurs protocoles de couche basse (le document que vous lisez en ce moment a probablement transité par le réseau filaire de votre fournisseur puis par votre lien *wifi*). Cependant, les problèmes de la transition à IPv6 (voir ci-dessous) semblent indiquer que notre technologie de modularité n'est pas encore complètement au point.

Le modèle OSI simplifié Les protocoles qui nous concernent — ceux de la suite TCP/IP, utilisée sur l'Internet — sont structurés selon le modèle *OSI simplifié*, qui consiste de 5 couches numérotées de 1 à 7 :

Application	(7)
Transport	(4)
Internet ou Réseau	(3)
Lien	(2)
Physique	(1)

Les numéros 5 et 6 correspondent aux couches *Session* et *Présentation* du modèle OSI d'origine, et ne sont plus utilisées aujourd'hui.

4 La suite de protocoles TCP/IP

La suite TCP/IP est la suite de protocoles utilisée sur l'Internet Global. Elle est structurée selon le modèle OSI simplifié.

NTP, DNS, FTP, SMTP, HTTP, ed2k, Bittorrent etc.	(7)
UDP, TCP	(4)
IP	(3)
SLIP, PPP, Ethernet, 802.11 etc.	(2)
RS-232, 10Base2, 100BaseTX, radio 2,4 GHz, etc.	(1)

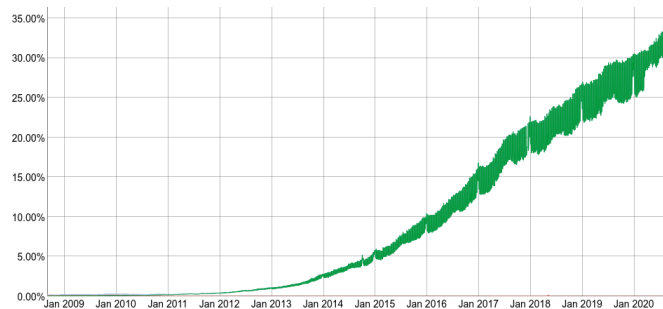
La suite TCP/IP a une structure en sablier. Aux bords, il y a de nombreux protocoles de couche physique et de nombreux protocoles de couche application. Au milieu, à la couche réseau, il n'y a (en principe) qu'un seul protocole, IP : c'est la *couche de convergence*. Les couches intermédiaires, lien et transport, servent à factoriser des fonctionnalités communes à plusieurs protocoles de couche physique et application respectivement.

Les couches basses (physique et lien) sont des couches de proche en proche (*hop-to-hop*) : elles s'occupent de la communication locale au lien, et ne savent rien de l'Internet. Les couches hautes (transport et application) sont de bout en bout (*end-to-end*) : elles ne savent rien des liens individuels. La couche de convergence (réseau) recolle les morceaux, et est donc la seule à avoir une connaissance aussi bien des liens individuels que de l'Internet.

Transition à IPv6 Le protocole IP ordinaire, dit *IPv4*, limite le nombre d'adresses possibles à moins de 2^{32} . Ces adresses sont maintenant épuisées : l'IANA, organisme gérant les adresses IP, a distribué ses derniers blocs d'adresses en février 2011².

Dans les années 1990, on a développé un successeur à IPv4 nommé *IPv6*. IPv6 utilise des adresses de 128 bits, ce qui fournit un espace d'adressage illimité en pratique. Le déploiement d'IPv6 a été très lent jusqu'au début des années 2010, quand il a commencé à décoller. On estime qu'environ un tiers des utilisateurs de Google utilisent IPv6³.

Il n'y a pas de traduction automatique entre IPv4 et IPv6 : pour participer aux deux protocoles, un hôte doit implémenter une double pile. Malheureusement, cette double implémentation est en grande partie à charge de l'application.



2. Et ils en ont profité pour faire une fête.

3. <https://www.google.com/intl/en/ipv6/statistics.html>