Les réseaux ont la fibre de l'information

Laurent Viennot

▶ **To cite this version:**

Laurent Viennot. Les réseaux ont la fibre de l'information. DocSciences, 2008, 5, pp.34-39. inria-00471725

HAL Id: inria-00471725
https://inria.hal.science/inria-00471725

Submitted on 9 Apr 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Les réseaux ont la fibre de l'information

Laurent Viennot
Chercheur en algorithmique des réseaux à l'INRIA
Centre de Paris - Rocquencourt

Passer un coup de fil, surfer sur Internet, rien de plus banal aujourd'hui. Mais que sait-on de la machinerie qui opère en coulisse. Que se passe-t-il à l'autre bout du fil ?

Si un pan de science s'ouvre devant la question : « Que peut-on calculer ? », un autre tente de répondre à : « Que peut-on télécommuniquer ? » Concrètement, ce sont les *réseaux* qui acheminent l'information. Deux modèles ont aujourd'hui atteint une échelle planétaire : le réseau téléphonique qui transporte nos voix et Internet qui transporte nos données numériques.

Le réseau de voix

Avec le téléphone, nous sommes devenus connectés, branchés, câblés... En effet, pour discuter avec son *alter ego*, il faut se connecter à lui par un circuit électrique, c'est la phase de connexion si poétique à l'époque où l'on demandait « Montmartre 22 12 » ou « Elysée 15 15 ». Encore de nos jours, il vous faut savoir où se trouve un correspondant dans le réseau pour le joindre. Par exemple, au numéro 46062212, s'il se trouve au bout du fil 2212 du central téléphonique de Montmartre qui a pour numéro (de central) le 4606. Vous pouvez alors activer votre téléphone en le décrochant ou en tournant la manivelle, selon l'époque. Dans votre central téléphonique, à l'autre bout du fil, une opératrice ou un commutateur automatique constate que votre ligne s'active et attend que vous énonciez l'adresse de votre correspondant. Vous demandez donc « Montmartre 22 12 » ou « tic, tic-tic-tic-tic-tic-tic... » ou « do si do si mi mi ré mi » ou le 46062212 qui s'affiche sur l'écran du combiné et circule en douce derrière le souffle de l'appareil. L'autre bout de votre fil est alors connecté à l'un des fils qui relient votre central à celui de Montmartre. De même, une opératrice (ou un commutateur automatique) connecte, à Montmartre, l'autre bout du fil au poste 2212 qui se met à sonner. Quand votre correspondant décroche, vous êtes connectés l'un à l'autre par un circuit transportant les signaux électriques qui codent votre voix.

Bien sûr, au fur et à mesure que le réseau grandit, on peut rajouter des super-centraux de numéros 01, 02, 03, 04, 05, par exemple, si l'on découpe le pays en cinq régions. Ce numéro se rajoute alors au début du numéro de votre correspondant en préfixe. Notre numéro à Montmartre devient ainsi 0146062212. On relie alors chaque central téléphonique au super-central de sa région par de nombreux fils et on relie entre eux les supers-centraux par des quantités encore plus importantes de fils. Le numéro de téléphone est ainsi une sorte d'adresse construite selon un découpage hiérarchique de la géographie du monde similaire à celui de la poste. Il s'agit donc d'une *adresse électronique*.

Dans les premiers réseaux téléphoniques, le circuit établi est physique : une suite de fils sont connectés les uns aux autres. Dans les réseaux plus évolués, le circuit est logique, et l'on parle de *connexion*. Avec l'augmentation du trafic, les fils entre centraux deviennent si demandés qu'il faut inventer un moyen de faire transiter plusieurs connexions sur le même fil.

De l'information circule de fait sur tout signal physique qui se propage dans l'espace : un courant dans un fil électrique ou une onde lumineuse dans une fibre optique ou même une onde radio dans l'espace ambiant. La théorie de l'information naît de cette question cruciale : « Combien de connexions vont pouvoir transiter sur mon signal ? » En 1948, Claude Shannon* met en équations les liens entre les caractéristiques physiques d'un signal et la quantité d'information que celui-ci peut transmettre. Il promeut l'utilisation du mot *bit* pour désigner l'unité atomique d'information : 1 ou 0, vrai ou faux, oui ou non... On peut ainsi quantifier l'information transmise en bits par seconde, voire en kilo-, méga-, giga- ou térabits par seconde, ce qui correspond à 1, 1000, 1000 000, … et ainsi de suite, chaque terme indiquant une quantité mille fois plus grande de bits par seconde que le précédent.

Mais revenons à l'invention de machines capables de traiter l'information : comment peuvent-elles communiquer ?

Le réseau de réseaux

Dès les années 1950, on sait échanger des données entre deux ordinateurs, mais comment relier un ensemble d'ordinateurs ? Le concept de réseau en tant que nœuds reliés par des liens de communication* apparaît à cette époque avec la naissance de la théorie probabiliste des files d'attente. Leonard Kleinrock* identifie en 1961 le point-clé pour pouvoir appliquer ces théories : le concept de *routeur* soit un « nœud capable de stocker un message en attendant que le lien sur lequel il doit être retransmis se libère ». Cette idée va à l'inverse des réseaux téléphoniques reposant sur l'établissement de circuits où l'information circule sans jamais être stockée. On trouve, en filigrane, derrière le concept de message, celui de *datagramme*, c'est-à-dire un paquet élémentaire d'information qui circule de manière autonome dans le réseau. Le premier réseau utilisant des routeurs est l'Arpanet, Kleinrock en expérimente le premier routeur.

Quand les routeurs ne sont pas programmés par les mêmes personnes, il faut s'entendre sur la manière de les faire communiquer. Ainsi naît le concept de *protocole* qui spécifie comment sont codés les messages que s'échangent les routeurs et quels sont les algorithmes à appliquer lors de l'envoi et de la réception de tel ou tel message. Un protocole est une sorte d'algorithme à l'échelle du réseau. Par exemple, l'établissement d'une connexion dans un réseau téléphonique suit un certain protocole. Celui-ci résout d'ailleurs le problème critique du *routage* : par quelle suite de liens faire cheminer les données, autrement dit, quelle *route* doivent-elles suivre. Dans un réseau téléphonique, le réseau est pré-configuré pour répondre à cette question. Dans un réseau de routeurs, le problème se pose pour chaque datagramme. Concrètement, quand un routeur reçoit un datagramme, un calcul doit lui permettre de décider sur quel lien le retransmettre. Pour cela, un protocole de routage permet aux routeurs, via l'échange de messages spécifiques, de détecter comment ils sont connectés les uns aux autres et d'acquérir suffisamment d'information pour pouvoir mener à bien un tel calcul.

Dans les années 1970, on cherche à inter-connecter les réseaux avec la conception des protocoles IP et TCP (« Internet Protocol » et « Transfer Control Protocol »). Le protocole IP définit les adresses des machines et le format général des datagrammes. L'adresse IP reprend le principe de préfixe des numéros de téléphone sous forme d'une suite de 32 bits. La hiérarchie s'organise en réseaux identifiés par des préfixes d'adresses. L'organisation n'a donc plus de lien avec la géographie : on entre dans le *cyberespace*. Le protocole TCP définit comment établir des connexions logiques au moyen d'échanges de datagrammes. Un flux de données est envoyé par une suite de datagrammes indépendants. Une partie critique de ce

protocole réduit le débit de datagrammes en cas de congestion dans un des réseaux traversés. Ce concept d' « inter-netting » en anglais donnera son nom à Internet qui naît au début des années 1980 quand Arpanet est relié au réseau académique CSNET. Très vite, d'autres réseaux européens puis japonais viennent s'inter-connecter à ce réseau originel.

Vingt-cinq ans plus tard, Internet est devenu aussi imposant que le réseau téléphonique qui s'est lui-même considérablement développé. Les deux géants sont maintenant inter-connectés mais restent distincts. L'un n'a toujours pas supplanté l'autre. Les progrès les plus récents dans les deux cas concernent l'utilisation généralisée de liens radio qui permettent de communiquer en se déplaçant. Les routeurs sont entrés jusque dans nos maisons sous la forme de boîtiers Wi-Fi (pour « Wireless », sans fil) qui connectent par radio tous les ordinateurs de la famille à Internet via une liaison ADSL sur le fil du téléphone. Les premiers réseaux radio remontent pourtant au début des années 1970 avec le réseau ALOHA dans les îles hawaïennes. Mais il n'était pas facile à l'époque de se promener avec un ordinateur et un émetteur-récepteur radio dans la poche !

Le nuage

Très vite, on se rend compte que les réseaux vont aider à calculer. Dans les années 1970, ils permettent d'accéder aux super-calculateurs sans avoir besoin de se déplacer. Dans les années 1980, les super-calculateurs eux-mêmes deviennent des machines parallèles, c'est-à-dire composées de plusieurs processeurs en parallèle reliés par un réseau très rapide. De nouveaux paradigmes de langages de programmation sont inventés pour pouvoir lancer des ordres en parallèle. À la fin des années 1990, on obtient un super-calculateur en connectant des milliers, voire des dizaines de milliers d'ordinateurs personnels par un réseau haut débit. On parle alors d'une *grille* de PCs. Plus que la puissance de calcul, c'est le volume de données que peut stocker et manipuler une telle machine qui devient considérable. Une compagnie devenue célèbre depuis, construit fin 1998 un moteur de recherche sur ce principe de machine. Cela lui permet ainsi de stocker en mémoire vive un index d'une grande partie de toutes les pages web accessibles par Internet avec un temps d'accès record. Après avoir équipé les quatre coins d'Internet de telles grilles, ses ingénieurs administrent une machine surpuissante de calcul et de stockage qui se trouve diffuse dans le réseau. Fascinés par leur invention, ils l'appellent « le nuage ».

Zoom

Des rivières et des lacs

On peut voir les réseaux comme un tissu de rivières où s'écoule l'information, et les grilles comme des lacs stockant de grandes quantités d'information. Quelle est la capacité de ces rivières et de ces lacs aujourd'hui ? Combien de temps une rivière met-elle à remplir un lac ?

La rivière qui arrive aujourd'hui chez nous s'appelle ADSL et délivre typiquement 20 mégabits par seconde. Pour donner un ordre de grandeur, un livre de 200 pages contient environ un million de caractères, soit 1 mégaoctet ou encore 8 mégabits. Le débit de votre liaison ADSL est donc légèrement supérieur à deux livres par seconde. En comparaison, les liaisons qui transportent l'information au cœur d'Internet sont de véritables fleuves constitués de câbles de quelques fibres optiques, voire plusieurs dizaines. Chaque fibre y transporte, par une centaine de signaux lumineux, quelques térabits par seconde. Une seule fibre peut acheminer tout le trafic téléphonique de la France, même si tout le monde téléphone en même temps.

Les plus grands lacs d'information sont des grilles de plusieurs dizaines de milliers d'ordinateurs bon marché. Un ordinateur pouvant stocker 1 téraoctet sur ses disques durs, un tel lac peut contenir environ 10 000 téraoctets. En comparaison, appelons BNF la taille du texte intégral des 13 millions de livres de la Bibliothèque Nationale de France, soit environ 10 téraoctets. Une grille peut donc stocker quelques milliers de BNF. Cependant, une seule fibre optique peut transporter toute cette information en 10 heures environ. Un câble d'une centaine de fibres, en 5 minutes environ. Si impressionnants soient-ils, les lacs d'Internet paraissent petits devant des rivières capables de les remplir si vite.

Photo 1

Le cœur d'Internet est constitué de câbles contenant des fibres optiques capables de transporter des quantités impressionnantes d'information par-delà les océans. Pour pouvoir irriguer la planète entière, ces fibres sont reliées en réseaux, sortes de filets très lâches posés sur le globe. De fait, Internet est constitué d'une multitude de réseaux connectés les uns aux autres. Chaque réseau est administré par une entreprise ou une organisation indépendante des autres. La collaboration de ces réseaux repose sur une collection de règles techniques publiques (les protocoles) et d'accords commerciaux bilatéraux. Les réseaux les plus importants s'étendent sur la planète entière et irriguent des réseaux de moindre importance comme ceux des fournisseurs d'accès à Internet, qui eux-mêmes se ramifient dans la multitude de micro-réseaux qui se trouvent dans nos maisons. Voici donc la nature physique du cyberespace dans lequel nous évoluons de plus en plus, un espace non plus géographique mais logique, formé de connexions. L'important n'est plus où vous êtes, mais à qui vous êtes connecté.

Photo 2

La puissance réside dans le nombre. Les plus grands centres de calcul sont constitués d'armoires regroupant des milliers d'ordinateurs qui sont reliés par un réseau rapide. On appelle *grille* la machine virtuelle ainsi obtenue par union des ordinateurs d'un ou plusieurs centres. Plus que la maîtrise du calcul, c'est la maîtrise des données et de l'information qui devient un enjeu. La grille est devenue un outil industriel pour manier et servir des masses de données énormes. Les plus grosses entreprises actrices d'Internet en installent près de barrages hydroélectriques tant le coût de la consommation électrique l'emporte sur celui d'une connexion Internet très haut débit. Alternativement, des communautés d'internautes obtiennent des grilles en se connectant, via Internet, les uns aux autres, entre pairs, (en « peer-to-peer »). Si de tels systèmes sociaux n'ont pas la force de calcul et le temps de réponse de centres dédiés, ils peuvent effectuer certains calculs de grande ampleur, ouéchanger de larges masses de données, ébranlant ainsi certains intérêts commerciaux et présageant de nouveaux horizons démocratiques.

Lexique

Protocole : Une spécification de comment doivent fonctionner et communiquer les nœuds d'un réseau. Un protocole spécifie en particulier les formats des messages échangés (au bit près) et comment ces messages doivent être traités. Les protocoles d'Internet sont labellisés par l'IETF.

Routeur : Un ordinateur dédié à la retransmission d'information. Généralement, un routeur est relié à plusieurs autres routeurs et permet d'acheminer les données des uns vers les autres.

Lien de communication : Une laison numérique entre deux nœuds du réseau. Par exemple, une fibre optique, une liaison radio, une connexion ADSL...

Routage : Le processus qui, dans un réseau, permet d'acheminer de proche en proche un paquet de données jusqu'à une destination.

Réseau : Un réseau est constitué de nœuds reliés par des liens, ce qui se modélise souvent par un graphe. Par exemple, des routeurs peuvent être reliés par des fibres optiques.

Adresse IP : Sur Internet, chaque ordinateur connecté est identifié par une adresse qui lui est propre : l'adresse IP (Internet Protocol). Elle se compose de 32 bits, soit encore 4 octets qui peuvent se lire comme des nombres entre 0 et 255, par exemple « 125.32.217.4 ».

Datagramme : Un paquet de données qui transite de manière indépendante dans le réseau.

Pair à pair (ou « peer-to-peer » en anglais) : le paradigme qui consiste à se connecter directement de client à client, entre pairs, plutôt que de passer par un serveur.

Claude Elwood Shannon (1916-2001) est un ingénieur électricien et mathématicien américain. Il est l'un des pères, si ce n'est le père fondateur, de la théorie de l'information.

Leonard Kleinrock (1934-…) est un informaticien américain. Il est à la fois l'inventeur du concept de routeur et l'un des premiers à expérimenter une telle machine.

Savoir plus

Réseaux, de Andrew Tanenbaum.

http://www.ietf.org/ (Les textes des protocoles d'Internet.)