

## L'inextricable simplicité de l'internet

Le *web*, le courrier électronique, le commerce en ligne, le pair-à-pair, les discussions en direct et beaucoup d'autres applications qui sont en train de révolutionner notre société reposent sur une infrastructure souvent méconnue : l'internet. Enchevêtrement inextricable de millions d'ordinateurs et de câbles de toutes sortes, l'internet, le réseau des réseaux, est une des constructions les plus complexes ayant émergé de l'activité humaine. Et pourtant, il repose sur quelques concepts extrêmement simples. Nous allons voir comment, en combinant ces principes élémentaires mais subtils, on obtient finalement un réseau d'une complexité telle que ses concepteurs eux-mêmes ne sont plus en mesure d'en comprendre toutes les facettes.

### Les ordinateurs se parlent

Si vous voulez organiser une discussion de façon à éviter la cohue, il faut en premier lieu faire en sorte que plusieurs personnes ne parlent pas en même temps.

Une solution, souvent utilisée dans les émissions télévisées, est de confier l'attribution de la parole à un *animateur*. Il faut toutefois s'assurer que tout le monde respecte l'autorité de cette personne et que celle-ci est en mesure de gérer correctement le droit de parole.

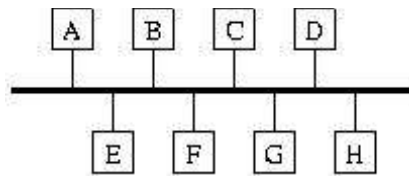
Une autre méthode, qui évite ces difficultés, consiste à voir la parole comme un objet que les personnes se passent : la personne ayant la parole s'exprime (sans toutefois dépasser un temps de parole prédéfini), puis passe la parole à son voisin. Si celui-ci a quelque chose à dire, il s'exprime à son tour (avec la même contrainte de temps de parole limitée), sinon il passe à son tour la parole à son voisin. Et ainsi de suite. Toute personne souhaitant s'exprimer est ainsi certaine de pouvoir le faire dans un laps de temps raisonnablement court, qui dépend du nombre de personnes participant à la réunion et du temps de parole maximal fixé.

Il existe encore une autre méthode, généralement utilisée dans le contexte de réunions d'amis. Lorsqu'on veut parler, on écoute afin de savoir si quelqu'un parle déjà. Si c'est le cas, on attend ; sinon, on s'exprime. Toutefois, si quelqu'un a fait le même raisonnement, il peut commencer à parler en même temps. C'est ce qu'on appelle une *collision*. C'est pourquoi il est important d'écouter en même temps qu'on parle, ce qui permet de se rendre compte qu'il y a collision, et de s'arrêter de parler dans ce cas. Il faut ensuite décider qui reprend la parole (éventuellement une troisième personne), ce qui peut être fait suivant des critères plus ou moins complexes ou arbitraires.

Tout ça paraît bien simple et naturel. Nous en voyons des exemples tous les jours lors de réunions de travail, lorsque nous regardons la télévision, ou lorsque nous retrouvons des amis. Et tout ça semble bien loin des obscures problématiques de l'informatique. Pourtant, l'essentiel du fonctionnement des réseaux locaux est contenu dans ces remarques.

Un *réseau local* est typiquement constitué des ordinateurs d'un laboratoire, d'une petite entreprise, d'une agence bancaire, etc. Les ordinateurs se trouvent dans quelques salles, éventuellement quelques bâtiments, et sont tous connectés à un même câble (Figure 1). Lorsqu'un ordinateur veut émettre un message, il l'envoie, accompagné du nom du destinataire, sur le câble (sous forme d'impulsions électriques). Bien sûr, si un autre ordinateur fait la même chose au même moment, les deux messages émis sur le câble entrent en collision, et ce qui passe sur le câble ne signifie plus rien. Il se passe exactement la même chose que lorsque deux personnes parlent en même temps : on ne comprend plus ce qui est dit. Pour éviter cette situation dans un réseau local, on applique une des méthodes décrites ci-dessus. Typiquement, lorsqu'un ordinateur envoie un message, il surveille ce qui passe sur le câble et s'arrête d'émettre s'il détecte une collision. Il attend alors un moment (dont la durée est tirée au hasard) puis il recommence au début : écoute, attente éventuelle, puis tentative d'envoi. Pour la réception du message, c'est encore plus simple : tous les ordinateurs surveillent en permanence ce qui passe sur le câble et, dès qu'ils voient passer un message qui leur est destiné (il est accompagné de leur nom), ils le réceptionnent.

Tout cela est donc très simple : les ordinateurs communiquent comme des personnes qui discutent autour d'une table.



**Figure 1. Un réseau local.** Les ordinateurs (représentés par des carrés contenant des lettres) sont tous connectés à un même câble (le trait épais).

### Mettre les réseaux en réseau

Mais que se passe-t-il si plusieurs milliers de personnes veulent discuter les unes avec les autres, échanger de l'information ? Il n'est plus question de s'asseoir autour d'une table et d'attendre que personne ne parle pour prendre la parole. En effet, s'il y a beaucoup de personnes, la probabilité pour que deux d'entre elles veuillent parler en même temps devient très grande. A ce moment là, plus personne ne parle : si une personne veut parler, elle écoute, et il y a beaucoup de chances que quelqu'un parle déjà. Si jamais ce n'est pas le cas, elle commence à parler, mais il y a alors de fortes chances qu'une des nombreuses personnes qui attendaient de prendre la parole en fasse autant. Et donc il y a collision, les deux personnes s'arrêtent de parler, et tout est à recommencer. Non seulement l'information ne passe plus aussi bien qu'avant mais, en fait, toutes les personnes s'empêchent mutuellement de parler et presque plus aucune information ne passe. De plus, si beaucoup de personnes s'assoient autour d'une table, il faut une très grande table. Et les personnes qui sont à des extrémités opposées ne s'entendent pas. On peut leur demander de parler plus fort, ou même leur donner des porte-voix, mais ces approches ont des limites...

Là aussi, il se passe la même chose dans un réseau d'ordinateurs. Si on connecte trop d'ordinateurs comme décrit ci-dessus, le réseau est saturé par les collisions et les ordinateurs passent leur temps à attendre sans jamais arriver à communiquer. De plus, eux aussi rencontrent le problème de la distance : les impulsions électriques par lesquelles ils codent leurs messages sur le câble commun se dissipent et finissent par être indéchiffrables à longue distance. Un réseau local typique ne comportera donc que quelques dizaines ou quelques centaines d'ordinateurs, répartis sur une distance de quelques centaines de mètres. C'est pourquoi la structure de réseau local est typique des ordinateurs localisés dans un même bâtiment, correspondant en général à l'infrastructure informatique d'une petite entreprise.

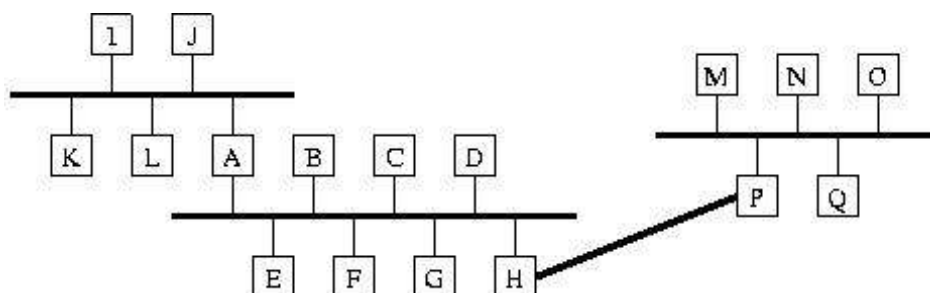
Mais alors, comment connecter des ordinateurs en plus grand nombre, et surtout plus distants ? La plupart des grandes entreprises possèdent des dizaines de milliers d'ordinateurs, répartis sur un territoire qui s'étend sur plusieurs pays, et même plusieurs continents. Sont-elles condamnées à voir leurs ressources informatiques cloisonnées en réseaux locaux étrangers les uns aux autres ?

Supposons que vous vouliez réunir deux cents personnes qui veulent pouvoir échanger des messages, mais que vous ne disposiez que de salles pouvant accueillir cent personnes à la fois. Chaque salle s'organise comme avant, mais la communication entre des personnes de salles différentes pose problème. Comment pouvez-vous vous organiser ? Il y a bien sûr de nombreuses solutions possibles. Par exemple, chaque personne peut se lever pour apporter son message au destinataire, éventuellement dans l'autre pièce, ou encore systématiquement émettre son message dans les deux pièces. Cependant, ces méthodes multiplient les déplacements et les risques de collisions, et donc elles ne peuvent pas être appliquées en général. Il faut réfléchir à des alternatives plus efficaces.

Il y a principalement deux façons de faire. La première, c'est de mettre une personne entre les deux salles, plus précisément une personne qui peut se déplacer de salle en salle. Cette personne fait le lien entre les deux salles ; nous l'appellerons *passerelle*. De cette façon, lorsque quelqu'un a quelque chose à dire, il agit comme avant sans se soucier de savoir si le destinataire du message est dans la salle ou non. La personne-passerelle écoute tous les messages et intervient quand le message qu'elle entend est destiné à une personne de l'autre salle : elle le transmet alors à l'autre salle. L'autre méthode consiste à installer un moyen de communication entre deux personnes n'étant pas dans la même salle, qui se font passer les messages qui doivent changer de salle. Ces deux méthodes ont de gros avantages. Tout d'abord, elles évitent qu'il y ait un trop grand nombre de collisions : un message émis dans une pièce n'est transmis à l'autre pièce que si elle est destinée à quelqu'un s'y trouvant. De plus, ces méthodes peuvent être utilisées à très grande échelle, pour relier des milliers de personnes dans des centaines de salles.

Encore une fois, les réseaux d'ordinateurs fonctionnent exactement de la même façon. Si on veut permettre à un grand nombre d'ordinateurs de communiquer ensemble, on regroupe géographiquement quelques dizaines ou centaines d'ordinateurs reliés au sein de réseaux locaux que l'on interconnecte. Pour ce faire, on peut

mettre un câble entre deux ordinateurs appartenant à des réseaux distincts, ou alors relier un même ordinateur aux deux réseaux (Figure 2). Les ordinateurs permettant le passage d'un réseau à l'autre s'appellent des *passerelles*.



**Figure 2. Interconnexion de réseaux locaux.** Les deux réseaux de gauche sont connectés par l'intermédiaire de la passerelle A, tandis que ceux de droite sont interconnectés par un câble entre les ordinateurs H et P. De cette façon, les ordinateurs des trois réseaux peuvent échanger des messages. Par exemple, si L veut envoyer un message à N, celui-ci transitera d'abord par A puis par H et P.

### L'internet : le réseau des réseaux

Historiquement, les entreprises et les diverses institutions se sont dotées d'ordinateurs, de plus en plus nombreux, qu'elles ont ensuite connectés par des réseaux locaux. Ceci s'est fait indépendamment dans des milliers de contextes différents, dans de nombreux pays. Certaines fois, la priorité était donnée à la rapidité, d'autres fois à la sécurité. D'autres fois encore, le confort d'utilisation primait. Et, à chaque fois, des contraintes d'optimisation des coûts, qui varient en fonction des objectifs et des technologies disponibles, ont conduit les responsables à effectuer des choix technologiques divers. Lorsque le besoin d'interconnecter les ordinateurs à grande échelle s'est fait sentir, il existait donc une très grande diversité de réseaux locaux indépendants et incompatibles.

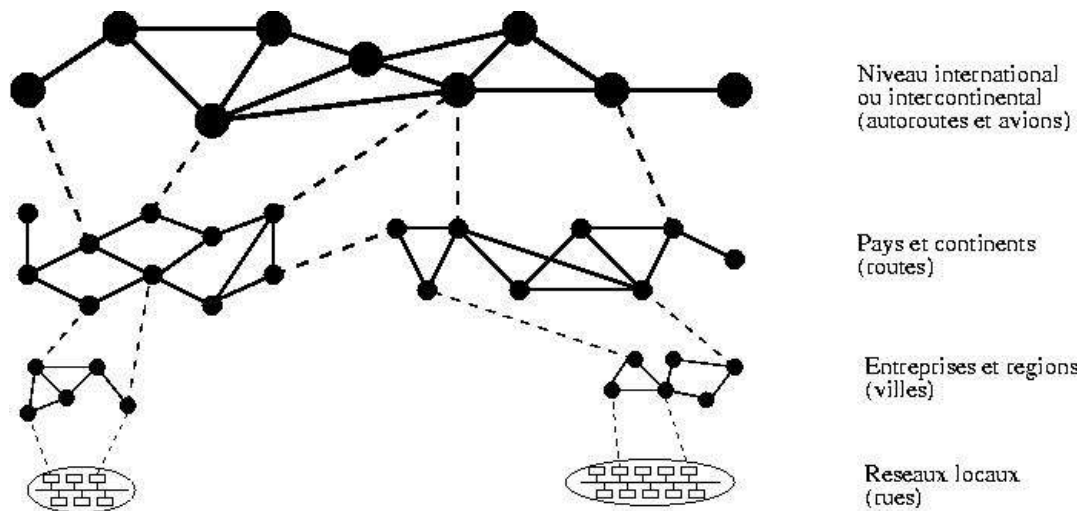
De la même façon, lorsque les chemins de fer se sont développés, chaque pays a mis en place son propre réseau de voies, en faisant ses propres choix, parfois arbitraires. Les personnes voyageant de pays en pays devaient changer de train aux frontières afin de passer d'un réseau de chemin de fer à l'autre. Afin d'éviter ces changements fastidieux (en particulier pour le transport de marchandises), il a semblé naturel de fusionner les réseaux de chemins de fer, simplement en mettant les voies bout à bout aux frontières. Mais que s'est-il passé alors ? Il s'est avéré que les systèmes étaient incompatibles. Par exemple, l'écartement entre les voies variait d'un pays à l'autre. La solution est passée par l'adoption de normes, issues de longues et fastidieuses concertations, qu'il a ensuite fallu appliquer. De nombreux pays ont été obligés de reconstruire entièrement leur réseau de chemin de fer, ainsi que leurs locomotives et leurs wagons.

En informatique, et dans le domaine des réseaux en particulier, la notion de norme, ou de standard, est aussi extrêmement importante. Quand on veut faire communiquer ensemble des matériels divers, construits à des époques différentes dans des pays différents et par des industriels concurrents, il s'agit de mettre au point des normes de communication entre les différents appareils. Il faut trouver un langage commun, que tout le monde comprenne, indépendamment de son langage d'origine. Dans le contexte des réseaux, c'est ce qu'on appelle un *protocole*. Quel que soit le fonctionnement interne d'un réseau local, pour pouvoir s'insérer dans un plus grand ensemble contenant plusieurs réseaux locaux interconnectés, il doit être capable de traduire l'information dans une langue que comprennent tous les réseaux.

Heureusement, un ordinateur est une machine beaucoup plus souple qu'une locomotive, et les protocoles permettant l'interconnexion de réseaux peuvent être mis en oeuvre sans qu'il faille reconstruire quoi que ce soit. En effet, les ordinateurs qui font la jonction entre deux réseaux, les passerelles, sont capables de traduire l'information venue d'un réseau dans une forme compréhensible par l'autre réseau, avant de l'envoyer sur celui-ci. Là encore l'idée est simple : les réseaux locaux sont limités en taille et parlent des langages différents ? Qu'à cela ne tienne, on met en place des passerelles qui jouent en plus le rôle d'interprètes.

C'est ainsi que l'internet s'est construit. Des milliers de réseaux locaux ont d'abord été mis en place. Puis diverses institutions ont souhaité interconnecter leurs réseaux locaux pour créer des réseaux à l'échelle de pays. Finalement, ces grands réseaux régionaux, nationaux ou continentaux ont été connectés ensemble, toujours par le moyen de passerelles, dans un immense ensemble qui constitue l'internet d'aujourd'hui.

L'internet peut être vu comme un réseau de transports entre des habitations. Il y a des rues bordées d'habitations, qui sont similaires aux ordinateurs d'un réseau local, tous connectés à un même câble. Ces rues sont reliées ensemble par des carrefours, et forment des villes qu'on peut voir comme l'interconnexion des différents réseaux locaux d'une même entreprise. Ensuite, les villes sont reliées entre elles par des routes plus importantes, puis le tout est relié par des autoroutes, qui sont des voies rapides de dimensions nationales. Enfin, on peut se déplacer d'une grande ville à une autre par avion, et même passer d'un réseau routier à un autre (par exemple en changeant de continent). Les voies aériennes constituent des voies encore plus rapides, qui peuvent être vues comme les grandes artères des réseaux de transports. De la même façon, l'internet est constitué de petits ensembles indépendants, les réseaux locaux, qui sont regroupés en paquets (à l'échelle d'une grande entreprise ou d'un pays par exemple) par des connexions plus rapides, ces ensembles plus importants étant eux-mêmes reliés par des connexions encore plus rapides, finalement complétées par des connexions très rapides qui relient les continents entre eux (ce sont généralement des gros câbles déroulés sur le fond des océans). Cette structure hiérarchique est illustrée Figure 3.



**Figure 3. La structure hiérarchique l'internet.** Sur le dessin, les différents niveaux de la hiérarchie sont représentés (les connexions entre réseaux indépendants sont en pointillés). L'analogie avec le réseau mondial des transports est indiquée entre parenthèses. Soulignons que la plupart des réseaux ne sont connectés qu'à quelques réseaux du niveau immédiatement inférieur et à quelques réseaux du niveau immédiatement supérieur, mais qu'il peut également en être autrement. C'est par exemple le cas au niveau "Pays et continents" dans le dessin, où on peut voir une liaison (pointillée) entre deux réseaux d'un même niveau.

### Retrouver son chemin.

Lorsque vous voulez vous rendre chez un ami à partir de chez vous, vous empruntez un certain chemin dans le réseau mondial des transports. Si votre ami habite dans la même ville que vous, *a priori* vous allez seulement emprunter les routes de la ville (s'il est dans la même rue, vous irez même à pieds). S'il habite loin de chez vous, vous emprunterez certainement les routes de votre ville pour rejoindre une route départementale, qui vous amènera à une autoroute (ou à un aéroport). Vous pourrez ainsi vous rapprocher de votre destination à grande vitesse. Ensuite, vous emprunterez les routes régionales et enfin les routes de la ville de votre ami, jusque chez lui. C'est une première façon de faire. Il y a une autre possibilité, que vous emploierez si les autoroutes sont saturées, ou si elles coûtent trop cher pour votre budget, ou encore tout simplement si c'est plus pratique pour rejoindre votre destination. Cette deuxième possibilité consiste à n'emprunter que des (relativement) petites routes, départementales et régionales, et à traverser ainsi un grand nombre de villes, jusqu'à arriver de proche en proche à destination.

Lorsqu'un ordinateur envoie un message à un autre ordinateur sur l'internet, le message doit trouver un chemin qui le mènera à destination exactement comme vous lors de votre visite à votre ami. Il va passer d'ordinateur en ordinateur, de réseau en réseau, jusqu'à l'ordinateur destination. En général, il va monter dans la structure hiérarchique l'internet, que nous venons de décrire, afin d'emprunter les liaisons les plus rapides. Par exemple, un message partant d'un ordinateur personnel en France vers un ordinateur personnel aux États-Unis sera émis sur le réseau local d'une entreprise, dont la passerelle le transmettra à un ordinateur d'un

fournisseur d'accès à l'internet auprès duquel l'entreprise a souscrit un abonnement, puis il sera envoyé sur le réseau universitaire français, pour ensuite employer une connexion rapide transatlantique et enfin refaire le même type de chemin jusqu'à destination. C'est une des possibilités, mais il en existe beaucoup d'autres, qu'elles emploient la structure hiérarchique ou qu'elles n'utilisent que les réseaux locaux et les passerelles entre eux.

De même que vous avez plusieurs chemins possibles pour aller chez un ami en partant de chez vous, un message sur l'internet a donc généralement un grand nombre de choix possibles pour atteindre sa destination. En fonction des accidents de la route, des grèves, des embouteillages ou d'autres facteurs changeants, la meilleure route pour vous rendre à destination varie au cours du temps. De même, le meilleur chemin à suivre pour un message sur l'internet n'est pas fixe : il y a des ordinateurs qui tombent en panne, des câbles qui sont rompus accidentellement, de nouvelles installations, des ordinateurs saturés de messages, etc. Toute la difficulté consiste à trouver un chemin en essayant d'arriver à destination le plus rapidement possible. C'est ce qu'on appelle le *routage* car il s'agit de trouver une route de l'expéditeur au destinataire. Le routage est une des principales problématiques des réseaux d'ordinateurs.

La similitude avec ce que vous faites quand vous vous rendez chez un ami s'arrête là. En effet, lors de votre déplacement, vous disposez d'un outil extraordinairement puissant : vous avez une *carte routière*. C'est sur cette carte que vous allez vous baser pour choisir votre itinéraire, en vous aidant éventuellement d'informations sur le trafic, les grèves, etc. Vous disposez d'un arsenal qui vous semble naturel car il est disponible, mais qui n'existe absolument pas sur les réseaux d'ordinateurs : ils changent vite, les embouteillages s'y déplacent encore plus vite, ils sont immenses et ils sont constitués d'une multitude de parties indépendantes. Du fait de tous ces facteurs, et d'autres encore, un message allant d'ordinateur en ordinateur vers sa destination n'a absolument pas de carte du réseau, ni d'information sur les ordinateurs qui sont saturés, en panne, etc. En fait, tout ce qu'il sait c'est l'adresse de l'ordinateur destination. La route à suivre lui est indiquée par les ordinateurs qu'il traverse pour arriver à destination : l'ordinateur sur lequel le message se trouve à moment donné lui indique le prochain ordinateur sur sa route. C'est comme si vous alliez chez votre ami en allant de ville en ville et que dans chacune d'elles il y ait une personne vous disant quelle est la ville suivante que vous devez traverser. C'est même pire que ça : à chaque carrefour où vous arrivez, quelqu'un vous indique quelle route vous devez ensuite emprunter. Et vous arrivez ainsi à destination sans savoir du tout à quoi ressemble la réseau routier dans son ensemble, ni même comment la route que vous avez suivie a été choisie.

### **Cartographier l'internet ?**

Nous avons vu l'internet est constitué d'un immense ensemble de réseaux interconnectés, tous ces réseaux étant gérés localement par des entreprises ou des administrations, indépendamment les uns des autres. Les messages s'y déplacent d'ordinateur en ordinateur jusqu'à leur destination, sans avoir de vue globale du réseau. *En effet, il n'existe aucun organisme central ayant connaissance de l'ensemble des ordinateurs connectés à l'internet, ou de la structure de leurs interconnexions.* Et pourtant, ces informations jouent un rôle crucial dans un grand nombre de contextes. Tout d'abord, il est important d'avoir de l'information sur la structure de l'internet pour en assurer le bon fonctionnement, ainsi que pour mettre au point des technologies nécessaires à son évolution. De plus, la structure de l'internet joue un rôle central dans sa stabilité et sa fiabilité : est-ce que des pannes peuvent déconnecter des parties entières de l'internet ? Est-ce que des attaques malveillantes, ciblées sur certains ordinateurs, peuvent altérer son fonctionnement ? Si c'est le cas, comment y remédier ? Chercher des réponses à ces questions passe par l'étude de la structure de l'internet, c'est-à-dire de la façon dont les ordinateurs sont connectés les uns aux autres, au-delà de la description globale que nous avons donnée dans cet article. Ceci passe par la construction d'une carte de l'internet, mais cette construction n'est pas du tout aisée.

Supposez que vous vouliez dresser une carte du réseau de chemins de fer du monde entier. Comment procéderiez-vous ? Vous iriez certainement à la gare la plus proche et lui demanderiez à un guichet la carte du réseau de chemin de fer de votre pays. Puis, vous feriez de même dans tous les pays et vous fusionneriez toutes les cartes. C'est fastidieux, mais vous y arriveriez.

Maintenant, supposez qu'au guichet on refuse de vous donner la carte, mais qu'on vous réponde : « par contre, si vous voulez savoir par quelles gares vous passerez pour aller d'ici à une ville donnée, nous pouvons vous répondre. » Vous pouvez donc avoir pas mal d'information : il suffit de prendre le nom de toutes les villes du monde, et de demander par quelles gares on passe pour s'y rendre. Toutefois, cette méthode a deux défauts :

■ C'est affreusement fastidieux car il y a énormément de villes de par le monde. De plus, vous ne

connaissez pas le nom de toutes les villes, et il n'y a pas vraiment de moyen de tous les obtenir. Ou plutôt si, il y a un moyen : il faut essayer toutes les suites de lettres, ce qui fait un nombre immense de possibilités.

- ▣ Il y a des liaisons de gares que vous ne verrez jamais, car aucun train partant de l'endroit où vous êtes ne les utilise pour aller à un autre endroit. La seule solution, c'est de changer d'endroit, et de recommencer l'opération en espérant que cette fois les chemins partant de cette gare emploient les voies que vous n'aviez pas vues à partir de la première gare.

Si vous vous rendez dans toutes les gares du monde et si vous demandez pour chacune les chemins employés pour aller dans toutes les autres gares, vous aurez une carte complète. Bien sûr, auparavant, vous devez dresser la liste de toutes les villes du monde en essayant toutes les suites de lettres.

Il est toutefois impossible d'appliquer cette méthode : elle prendrait bien trop de temps. Le temps que vous ayez fini ce travail, vous pouvez être sûr qu'un grand nombre de gares auront fermé, que d'autres auront été créées, et de même des voies auront été ajoutées et supprimées. Et de toutes façons, cette méthode est affreusement redondante : il y a beaucoup d'informations que vous obtiendrez, inutilement, de nombreuses fois. La question à se poser est alors la suivante : à combien d'endroits devez vous aller, et à chaque endroit combien de destinations devez vous choisir, pour obtenir une carte fidèle ? Cette question est difficile et pour l'instant personne ne sait y répondre.

Comme vous vous en doutez peut-être, cette façon bizarre de construire une carte du réseau de chemins de fer correspond en fait à ce qu'il se passe pour l'internet. Comme nous l'avons vu, l'internet est composé de millions de réseaux interconnectés, et les messages se déplacent d'un émetteur à un destinataire en passant d'un ordinateur à l'autre. Les administrateurs des différents réseaux, pour des raisons stratégiques, ne livrent pas de carte de leurs réseaux, mais il est possible de connaître la suite d'ordinateurs par lesquels passe un message. A partir de votre ordinateur, vous pouvez donc connaître les routes suivies par les messages que vous envoyez aux autres ordinateurs, et ainsi dresser une carte *partielle* de l'internet. A partir d'un autre ordinateur, vous pouvez améliorer votre carte en effectuant la même opération et en fusionnant les informations obtenues. Sur combien d'ordinateurs devez-vous répéter l'opération, et combien de destinataires devez-vous choisir à chaque fois, afin de construire une carte relativement précise de l'internet ?

Des cartes partielles de l'internet obtenues de cette façon sont actuellement disponibles et sont étudiées par les chercheurs. Elles contiennent des centaines de milliers d'ordinateurs et plusieurs millions de liens, et des explorations plus vastes sont en cours. Toutefois, personne n'est en mesure de dire aujourd'hui à quel point ces cartes sont complètes, et comment notre façon d'explorer le réseau influence la vision qu'on en obtient.

Matthieu Latapy  
<http://liafa.jussieu.fr/~latapy>