

L'économie des réseaux

(cours au CNAM)

Michel Volle

4 mars 2014

Table des matières

1	Le réseau général	3
1.1	Architecture du réseau	3
1.2	Les composants techniques	7
1.3	Informations relatives au trafic	8
1.4	Évolution des technologies	9
1.5	Le coût du réseau	10
1.6	Résultats	12
2	L'Internet	13
2.1	Principes du modèle	16
2.2	Structure géographique	18
2.3	Le trafic	20
2.4	Conventions du modèle	22
2.5	Unités d'œuvre	25
2.6	Dimensionnement	27
2.7	Les résultats	30

3	Économie du dimensionnement	31
3.1	Approche probabiliste	34
3.2	Effet de l'incertitude sur le coût	35
3.3	Volume du trafic servi	36
3.4	Profit et surplus social	39
4	Le cyberspace	41
4.1	Mobilité ou ubiquité ?	41
4.2	Architecture du cyberspace	43
4.3	Changements par rapport à aujourd'hui	44
4.4	Outils du cyberspace	46
4.5	Cyberspace et géographie	48
4.6	Représentation mathématique	50
5	Questions stratégiques	55
5.1	Réseau, monopole et concurrence	55
5.1.1	Qu'est-ce qu'un monopole naturel ?	55
5.1.2	Le réseau est-il un monopole naturel ?	56
5.2	Stratégie extensive ou intensive	58

Les évaluations qui figurent dans ce document résultent des travaux réalisés en 1995-1996 avec Christophe Talière sur la fonction de coût du réseau général et de l'Internet. Il se peut que certains de ces résultats soient désuets en raison des innovations survenues depuis. La démarche suivie pour modéliser la construction du réseau reste cependant convenable.

1 Le réseau général

On appelle « réseau général » l'équipement qui permet l'exploitation du service téléphonique et l'acheminement de communications d'autres types (liaisons louées, transmission de données) dont le débit est transporté sur les mêmes faisceaux de transit que le trafic téléphonique.

La topologie du réseau définit les relations entre nœuds du réseau. Elle reflète à la fois sa géographie (localisation des nœuds et arcs du réseau) et les règles de routage et d'acheminement des communications.

Le coût du réseau dépend de deux paramètres :

- la taille du réseau (nombre des raccordements d'abonnés, intensité du trafic),
- les coûts unitaires des équipements et des charges induites (maintenance, exploitation, etc.). Ces coûts unitaires évoluent dans le temps.

1.1 Architecture du réseau

L'architecture du réseau général est schématisée en trois niveaux fonctionnellement indépendants : transport interurbain, réseau intra-ZAA (ressources partagées) et réseau de distribution (ressources dédiées à l'abonné). Le territoire est donc découpé en zones géographiques de façon hiérarchique.

Réseau de distribution (ou « boucle locale »)

Ce niveau sert à raccorder chaque abonné au réseau général. Il correspond aux terminaisons du réseau et comprend des équipements individualisés par abonné :

- unités de raccordements des centres de rattachement des usagers,
- lignes de rattachement individuelles,
- équipements terminaux.

Seuls les terminaux faisant partie intégrante de l'infrastructure (et non d'une offre commerciale de service) sont à considérer. Les terminaisons numériques de réseau sont donc à prendre en compte, mais non les terminaux téléphoniques (mobiles ou non).

La distribution utilise des paires torsadées, mais aussi la fibre optique, la boucle radio et les techniques ADSL/HDSL qui permettent de transmettre des hauts débits sur de « vieilles » paires torsadées.

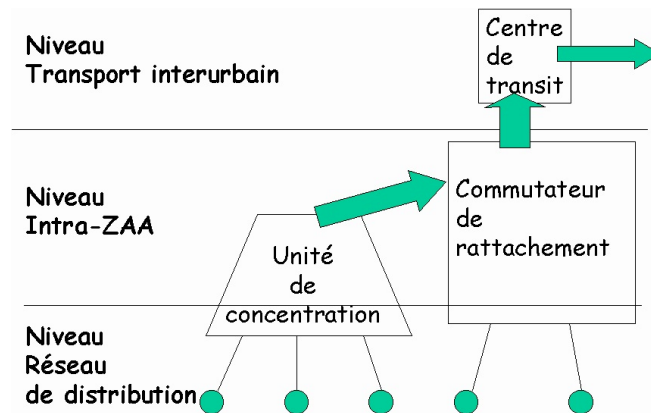


FIGURE 1 – Structure du réseau
Réseau intra-ZAA

Le niveau intra-ZAA (zone à autonomie d'acheminement) comprend l'ensemble des ressources communes desservant les « premiers points de rattachement » (ou points de concentration) des usagers. Il se compose des nœuds de concentration et de raccordement, des nœuds de commutation de circuits et des systèmes de transmission reliant ces deux types d'équipements.

On nomme ZL (zone locale) le périmètre où les abonnés sont raccordés à un même autocommutateur. Ces autocommutateurs sont des CL (Commutateurs Locaux) qui établissent les connexions entre les lignes d'abonnés et leur CAA (Commutateurs à Autonomie d'Acheminement) de rattachement (chaque CL dépend d'un CAA).

La ZAA (Zone à Autonomie d'Acheminement) est un ensemble de ZL appartenant à une même zone de taxation et de numérotage. Elle est équipée de CAA, nœuds d'interconnexion entre des CL, des abonnés directement raccordés au CAA, et des faisceaux de transport vers d'autres ZAA (ou d'autres CAA d'une même ZAA). Ces commutateurs analysent la numérotation reçue, choisissent entre plusieurs faisceaux de circuits selon les règles d'acheminement et commutent les circuits.

Le réseau de transport interurbain

La ZTS (Zone de Transit Secondaire) est équipée de CTS (Commutateurs de Transit Secondaire), qui interconnectent les CAA. Ils assurent le transit à l'intérieur d'une même ZTS et les relations avec d'autres ZTS.

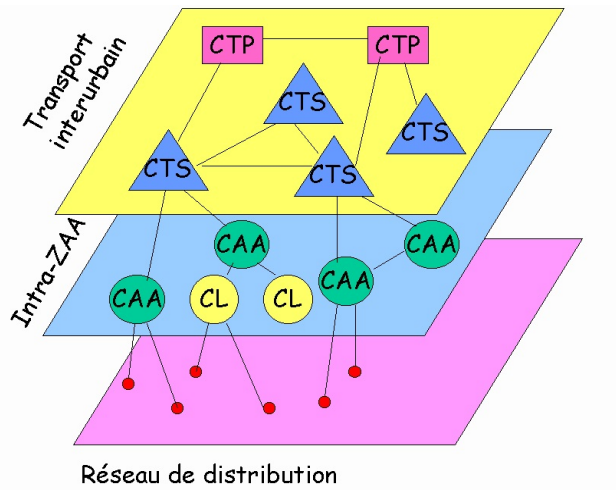


FIGURE 2 – Structure du réseau (suite)

La ZTP (Zone de Transit Principale) : c'est un ensemble de ZTS qui correspond à une zone d'entraide et de concentration. Elle est équipée de CTP (Commutateurs de Transit Principaux). Un CTP est relié à tous les autres CTP et aux CTS de sa ZTP, assurant ainsi la connexité du réseau (existence d'un chemin entre deux CAA quelconque).

Aucun usager n'est raccordé directement à ce niveau, dont les pôles sont destinés uniquement au transit. Le transport est réalisé sur des arcs de grande capacité et il est transparent à la nature des flux transportés (voix, donnée, image, liaisons louées) : les flux issus des CAA sont rassemblés sur des artères de transmission pour être transportés entre les centres de transit.

Pour le trafic international, le territoire est découpé en ZTI (Zones de Transit International). Chaque ZTI est équipée d'un CIA (Commutateur International Automatique) qui assure le transit du trafic international entrant et sortant de la ZTI. Un niveau hiérarchiquement supérieur concentre et ventile le

trafic des CIA ; c'est le CITP (Centre International de Transit Principal).

Contrairement aux réseaux de transport qui sont maillés, et en raison de leur fonction, les réseaux de distribution téléphonique sont arborescents (étoilés dans le cas le plus simple). L'optimisation d'un réseau d'accès suppose un compromis entre :

- la taille des points de concentration : l'économie d'échelle porte sur la mise en commun de ressources par le plus grand nombre possible d'utilisateurs (il faut donc réduire le nombre de ressources communes et avoir des ressources de grande taille) ;
- le nombre de points de concentration pour une population donnée, donc l'étendue de la zone desservie par chaque point de concentration : cette zone est en effet composée de ressources dédiées à chaque usage.

1.2 Les composants techniques

Le réseau de transport interurbain est fortement maillé. Sa structure est définie par des règles fondées sur la recherche d'une économie d'échelle. Par exemple, le coût d'un système de transmission (une artère inter-CTS) ne varie pas linéairement en fonction du trafic écoulé car il croît par paliers avec la modularité des équipements. La capacité installée est donc généralement différente de la capacité utilisée en moyenne.

Ceci s'explique par le caractère aléatoire des communications : pour un taux de blocage donné, le taux de remplissage moyen d'une grosse artère est plus élevé que celui d'une petite artère. Cependant cette économie d'échelle est contrecarrée par un autre phénomène : plus le trafic est regroupé sur des artères de taille importante, plus le chemin moyen à parcourir est long (ce phénomène s'accroît s'il y a débordement).

La structure topographique du réseau de transport est

donc un compromis, pour une qualité de service donnée, entre taille et densité des équipements (artères de transmission, centres de transit). La longueur moyenne des arcs, et donc la structure géographique du réseau, résulte de cette optimisation. Une représentation spatiale du réseau est nécessaire pour aboutir au compromis qui déterminera sa structure optimale.

Toute variation importante de la demande est susceptible d'impliquer une redéfinition de la structure du réseau ; de même, toute modification du coût d'un des composants du réseau implique un compromis différent ; et toute modification partielle (suppression ou ajout d'un élément), de la structure du réseau met en cause l'ensemble de la structure.

Les caractéristiques géographiques du réseau de transport ont une forte influence sur le coût global du réseau. La représentation spatiale du réseau permet d'étudier la variation des coûts selon la répartition géographique des communications : on peut ainsi évaluer l'écart entre le coût d'une communication entre deux zones à forte densité d'usagers et le coût d'une communication entre deux zones faiblement peuplées.

1.3 Informations relatives au trafic

Les informations relatives au trafic sont fournies par des matrices de trafic carrées et non symétriques représentant les flux entre ZAA selon la zone origine et la zone destination. Si l'on considère la matrice de trafic d'une zone géographique continentale, par exemple au niveau européen, on constate qu'elle est bloc-diagonale car le trafic international est moins intense que le trafic national ; c'est ce qui explique que les opérateurs télécoms historiques se soient bâtis chacun sur un territoire national.

Les paramètres à renseigner sont :

- le nombre des usagers (c'est-à-dire de raccordements),
- l'intensité du trafic. Pour mesurer cette intensité, on considère le trafic à l'heure de pointe, entouré d'une plage d'incertitude.
- la répartition géographique du trafic, l'unité géographique considérée étant la ZAA. Il faudra notamment définir les axes principaux de trafic (axes chargés).

Il existe en fait deux heures de pointe, l'une pour le trafic d'affaires l'après-midi, l'autre pour le trafic résidentiel en fin de journée. Chacune de ces heures de pointe correspond à une matrice de trafic différente, et à des exigences également différentes en matière de dimensionnement.

1.4 Évolution des technologies

On distingue dans un réseau les équipements terminaux ou de raccordement, les serveurs (y compris les équipements de commutation) et enfin les systèmes de transmission constitués par l'ensemble des équipements et supports de routage des communications qui gèrent le chemin physique emprunté entre commutateurs ou avec les équipements terminaux.

On identifie deux types d'évolution :

a) évolution de la technologie et des équipements : évolution des fonctionnalités, des performances, des services etc. pour les équipements de commutation et les serveurs ainsi que pour la performance et la maintenabilité des systèmes de transmission.

b) évolution de l'architecture du réseau : topologie, règles de routage, d'acheminement, de sécurisation, etc.; introduction de techniques ayant des répercussions sur la structure du réseau : réseau intelligent, HDSL/ADSL, distribution par

radio, ATM etc.

Les évolutions techniques ont des répercussions tant sur les dépenses d'investissement que sur les charges de fonctionnement. Elles influenceront la sécurité du réseau, sa flexibilité, sa disponibilité, d'où une répercussion difficilement quantifiable sur la qualité du service.

Les évolutions qui ont été des facteurs de rupture potentiels sont :

- l'extension de la téléphonie cellulaire (notamment vers la boucle radio) pour la distribution des bas débits ;
- l'introduction de la technique SDH pour la transmission des hauts débits, l'utilisation des techniques HDSL et ADSL ainsi que de la fibre optique pour leur distribution ;
- la poursuite de l'évolution des techniques de compression.

1.5 Le coût du réseau

Le coût d'un réseau de télécommunication comprend :

- un coût de fonctionnement annuel (par fonctionnement, on entend maintenance, exploitation technique, dans une certaine mesure exploitation commerciale - coût de facturation par exemple - et diverses autres dépenses),
- un coût d'investissement dont le cycle dépend de la durée de vie des équipements considérés.

Il faut donc calculer un coût annuel représentant la somme du coût de fonctionnement et d'un équivalent annuel du coût d'investissement. Cette conception du coût, qui correspond à celle d'un flux de dépense, est bien adaptée au raisonnement tarifaire.

Les paramètres du calcul de coût sont :

- les prix unitaires des divers composants du réseau,

- les durées de vie,
- les coûts annuels de fonctionnement, éventuellement augmentés d'une quote part de charges indirectes.

Notons τ le taux d'actualisation (nous le supposons égal à 9 %).

L'annuité équivalente à un investissement de durée de vie d et de montant i est a tel que :

$$a = i(1 - \frac{1}{1 + \tau}) / (1 - \frac{1}{(1 + \tau)^d})$$

Cette formule doit être enrichie pour tenir compte de l'évolution du coût des unités d'œuvre, car elle correspond à un équipement dont le coût de renouvellement serait égal au coût initial.

On peut décomposer chaque unité d'œuvre en une partie dont le prix reste constant, et une partie dont le prix baisse exponentiellement. La loi d'évolution du prix est donc de la forme :

$$P = P_0[(\alpha + (1 - \alpha)e^{-\beta(t-t_0)}]$$

où α et β dépendent du type d'équipement choisi.

L'obsolescence, qui résulte en général de la mise en œuvre d'une nouvelle génération technique, a pour conséquence une durée de vie inférieure à la durée de vie physique. Selon qu'un opérateur prend en compte ou non l'effet de l'obsolescence il pratiquera des politiques d'amortissement différentes et l'évaluation de son résultat annuel en dépend. S'il néglige d'anticiper l'obsolescence de ses équipements et retient de longues durées de vie, il sous-estimera les amortissements nécessaires au maintien du réseau à l'état de l'art, surestimera son résultat annuel, et plus dure sera la chute.

1.6 Résultats

Le modèle de calcul de coût a fourni trois types de résultats :

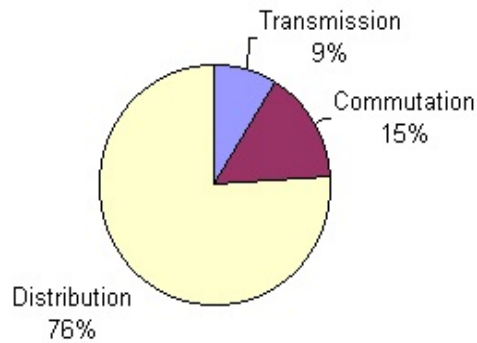
- le coût total du réseau général,
- le coût moyen par abonné,
- le coût moyen de production d'une communication.

Le coût de construction à neuf du réseau général s'obtient en simulant la construction d'un réseau de topologie et de taille équivalentes à celles du réseau réel à partir d'équipements conformes à l'état de l'art.

Ce coût diffère de la valeur du réseau réel, car les équipements en service ne sont pas neufs. Si l'âge moyen des équipements est égal à la moitié de leur durée de vie, la valeur du réseau est égale à la moitié du coût de construction à neuf du réseau.

Le coût du réseau neuf pour un pays de la taille de la France a été estimé à 60 milliards d'euros. Il en résulte une « valeur du réseau » de l'ordre de 30 milliards d'euros.

Les résultats ont été établis en distinguant les niveaux fonctionnels du réseau : réseau de distribution, transmission et commutation. On a ainsi obtenu un résultat qui peut surprendre ceux qui connaissent mal les réseaux de télécommunications : ramené au coût annuel (donc en faisant la part des durées de vie), le réseau de distribution représente 76 % du coût total du réseau ; la part de la commutation est de 15 %, celle de la transmission 9 %.



Or le coût du réseau de distribution est un coût fixe par rapport au trafic (il ne dépend que de la densité des abonnés). Il est utile de garder en mémoire le graphique ci-dessus : si la part la plus « noble », la plus « intelligente » d'un réseau est constituée par la commutation et la transmission, où se concentrent les technologies les plus récentes, la grande masse du capital réside dans le raccordement des abonnés où les travaux font un large appel au génie civil et donc à des techniques anciennes dont le coût n'est pas décroissant.

Par ailleurs toute innovation concernant la « boucle locale » peut diminuer par obsolescence la durée de vie des équipements du réseau de distribution, donc dévaloriser brutalement le capital d'un opérateur de télécommunications.

2 L'Internet

Nota Bene : cette section résume la présentation publiée à la page

<http://www.volle.com/travaux/coutinternet.htm>.

La pérennité de l'Internet était en 1995 problématique. L'administration américaine venait d'interrompre le financement du backbone de l'Internet, devenu dès lors une affaire purement commerciale. L'Internet ne risquait-il pas de s'effondrer, comme une salle de bal au plancher fragile, en précipitant dans le vide les affaires qui avaient été construites en s'appuyant sur lui ?

Les recettes reçues par l'Internet proviennent de la vente d'accès aux utilisateurs par les « Internet Access Providers », qui paient eux-mêmes à un opérateur un lien vers l'Internet, et cet opérateur rémunère aussi les ressources qu'il va utiliser. Ainsi les divers acteurs qui contribuent à l'Internet sont rémunérés en cascade.

Pour savoir si l'Internet était réaliste sur le plan économique, il fallait donc modéliser sa fonction de coût. Il serait ensuite possible, en divisant le coût total par le nombre des utilisateurs, de voir si ceux-ci pourraient payer le réseau ou si, au contraire, l'ordre de grandeur était tel qu'un équilibre commercial serait hors de portée.

Notre intuition penchait vers la seconde hypothèse et nous étions sur la piste d'un « scoop » : quel vacarme aurait suscité une annonce de l'effondrement prochain de l'Internet !

Nous avons construit le modèle laborieusement, en schématisant l'architecture et son dimensionnement (lui-même fonction de la demande exprimée par le nombre des utilisateurs et le débit nécessaire), puis en déterminant selon les bonnes règles d'ingénierie la nature et la taille des équipements. Après quoi il suffisait, pour évaluer le coût total, de multiplier le nombre des unités d'œuvre par leur coût unitaire et d'additionner le tout. Il fallait aussi, bien sûr, prendre en compte la baisse tendancielle du coût des unités d'œuvre qui caractérise

les équipements informatisés et poser quelques hypothèses sur l'évolution de la demande.

Le résultat se condenserait en une courbe montrant l'évolution du coût de l'Internet rapporté à chaque utilisateur. Notre intuition était que cette courbe ferait apparaître une forte hausse : alors l'Internet serait condamné en l'absence d'une subvention publique, parce que ses utilisateurs ne pourraient pas équilibrer son coût.

Après les calculs, et à notre grande surprise, c'est l'inverse qui apparut : la courbe de coût moyen était décroissante. L'explication du phénomène résidait bien sûr dans le modèle, et nous avons dû vérifier le réalisme des hypothèses qui le fondaient.

A l'issue de cette démarche, nous sommes parvenus à une conclusion contraire à notre intuition initiale : oui, l'Internet est économiquement robuste. Lorsqu'une conclusion dûment vérifiée et contrôlée s'impose ainsi, l'intuition s'enrichit d'une certitude nouvelle.

Cette conclusion n'a pas été du goût de ceux qui croyaient, comme nous l'avions fait d'abord, que l'Internet n'avait pas d'avenir.

* *

Le modèle s'appuie sur une représentation simplifiée de la topologie du réseau. Pour estimer le coût d'un réseau pris dans son ensemble, il faut en effet s'écarter des méthodes comptables habituellement utilisées (et qui reposent sur le constat des coûts que celui-ci supporte actuellement), et produire une simulation.

Le dimensionnement des liaisons et des équipements est déterminé à partir d'informations relatives au trafic que devra supporter le réseau et en utilisant des règles d'ingénierie

simples, ce qui permet de dénombrer les unités d'œuvre d'investissement et d'exploitation dont le coût unitaire sera enfin introduit.

Des variantes ont permis d'observer, au delà du résultat central, la sensibilité de la fonction de coût à ses paramètres (dont les principaux sont le nombre des utilisateurs de l'Internet et le volume du trafic total) et d'apprécier l'effet des incertitudes qui entourent les hypothèses.

Nous avons retenu la définition de l'Internet par l'ISOC (Internet Society) : l'ensemble des réseaux de réseaux interopérant via l'Internet Protocol (IP).

Nous considérons l'architecture qui permet la connexion d'utilisateurs par un maillage de couverture mondiale. Le niveau applicatif, ou ce que l'on appelle la toile (« the Web »), en est donc exclu.

Nous ne prenons donc en compte pour la détermination du coût de l'Internet que les ressources utiles à la fonction d'interconnexion de réseaux (et donc accessibles à tout moment). Le coût du PC d'un utilisateur connectable en mode dial-up n'est pas pris en compte (il n'est pas accessible à tout moment). De même, les serveurs relèvent du niveau applicatif (plus précisément, du niveau des applications ou services fournis sur le réseau d'interconnexion) et leur coût ne sera donc pas pris en compte. Nous ne considérons donc pas l'économie des services, ni l'économie des réseaux d'accès (ou « boucles locales »).

2.1 Principes du modèle

Le déploiement de l'Internet a abouti à un tissu complexe maillant le monde entier. Ses constituants sont des réseaux de taille diverse : réseaux locaux, réseaux régionaux, réseaux

nationaux, réseaux internationaux, interconnectés sans hiérarchie.

Néanmoins, on identifie plusieurs couches qui se composent d'unités d'oeuvre différentes :

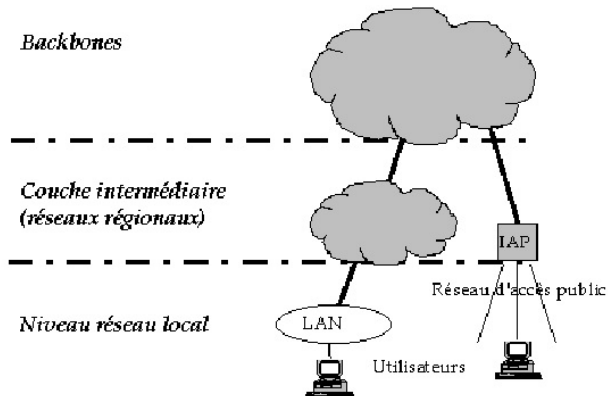
- les réseaux locaux : LAN d'un organisme (université, entreprise, etc), boucle locale du réseau téléphonique public utilisée pour le raccordement des PC des utilisateurs individuels, LL (liaisons louées) rattachant certains utilisateurs. Le coût de ce niveau, étant à la charge de l'utilisateur, est hors du périmètre du modèle.

- les couches intermédiaires : ce sont l'équivalent des réseaux régionaux qui étaient initialement gérés par des organismes d'Etat ou par des consortiums universitaires et qui sont maintenant pour la plupart exploités par des sociétés commerciales.

Les « Internet Access Provider » (IAP) jouent un rôle d'intermédiaire entre les utilisateurs individuels (individus ou entreprises utilisatrices) et les réseaux régionaux et/ou directement un backbone.

Ces couches intermédiaires ont pour fonction de rattacher les utilisateurs (niveau réseau local), mais aussi pour certain d'eux d'interconnecter des réseaux régionaux (fonction de transit).

- les backbones. Leur prototype a été le backbone américain NSFNET. Il a été démantelé en avril 1995, la fonction de réseau fédérateur étant alors confiée à divers opérateurs privés.



Cette modélisation n'impose pas une hiérarchie car des réseaux régionaux peuvent être directement interconnectés sans passer par un backbone.

2.2 Structure géographique

Le modèle découpe le territoire en mailles en fonction de la densité des utilisateurs.

Chaque maille abritera un nœud d'interconnexion (soit le routeur d'un backbone, soit le routeur d'un réseau régional). La surface des mailles est inversement proportionnelle à la densité géographique des utilisateurs (i. e. le nombre d'utilisateurs par nœud sera constant); on peut en effet approcher une topologie optimum en considérant qu'il existe une densité optimale par maille, ou zone de couverture d'un routeur.

Chaque maille est schématisée géométriquement par un triangle équilatéral avec un nœud (i. e. un routeur) en son centre. Les nœuds de deux zones voisines sont reliés directement entre eux.

Le schéma ainsi obtenu reflète les caractéristiques principales du « réseau de réseaux » tout en simplifiant sa topologie :

- les routeurs d'un backbone, ou les réseaux régionaux, sont interconnectés de proche en proche,
- le maillage minimal (contrainte de sécurité) est assuré : pour chaque flux entrant dans un routeur, il existe au moins deux sorties différentes,
- le nombre de mailles croît avec le nombre des utilisateurs : plus la densité des utilisateurs augmente, plus les routeurs recherchent la proximité des utilisateurs.

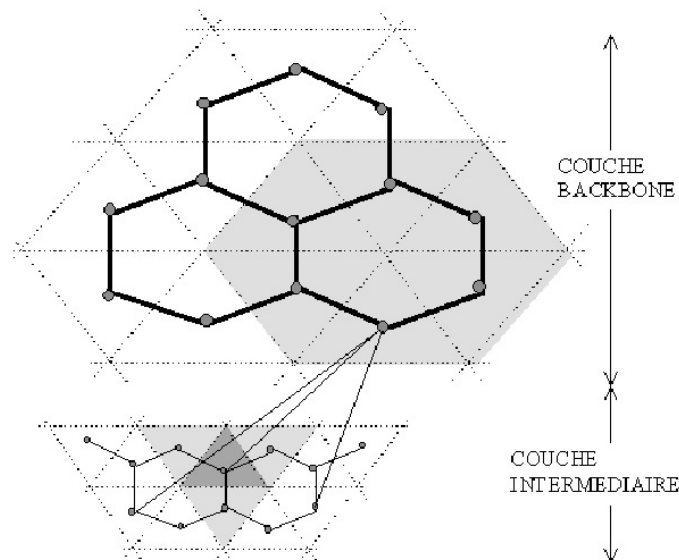
Remarque 1 : on suppose dans cette première approche que chaque maille élémentaire constitue une zone de monopole local de fait alors qu'une aire géographique peut être partagée entre plusieurs acteurs (« exploitants de routeurs »).

Remarque 2 : le maillage total des superficies considérées n'est pas l'exact reflet de la réalité.

Remarque 3 : ce principe de modélisation suppose une équirépartition des utilisateurs sur les superficies considérées, or en fait la densité des utilisateurs n'est pas homogène. On retiendra néanmoins l'hypothèse d'une densité homogène sur chaque région étudiée car cela n'altère pas sensiblement le calcul du coût moyen par utilisateur.

Deux niveaux de découpage sont définis :

- un découpage au niveau des couches intermédiaires : mailles représentant la couverture des différents réseaux régionaux (interconnectés entre eux et connectés au backbone),
- un découpage au niveau backbone : mailles correspondant aux zones de desserte de chaque routeur d'un (ou de plusieurs) backbones.



La densité des utilisateurs varie selon la région du globe. On a découpé le monde en grandes régions géographiques : USA, Canada, Europe de l'ouest, Japon, Océanie (dont notamment l'Australie), Amérique (hors USA et Canada), Europe de l'est, Asie (hors Japon), Afrique. À l'époque où ce modèle a été confectionné il n'était pas nécessaire de distinguer la Chine.

Chaque région comporte une couche « backbone(s) » et plusieurs couches « réseaux régionaux ».

2.3 Le trafic

Le trafic émis par un routeur peut être soit un trafic de proximité (trafic entre deux mailles adjacentes) et dans ce cas il ne transite pas par des nœuds intermédiaires, soit un trafic lointain (à destination d'une maille non adjacente) : le

trafic devra alors transiter par un ou plusieurs nœuds intermédiaires appartenant soit à une (des) couche(s) intermédiaire(s) de rang élevé, soit à la couche backbone.

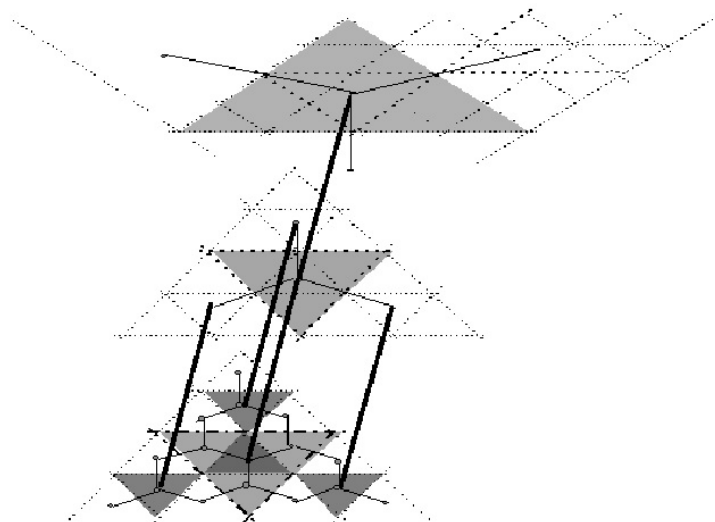
On définit ainsi plusieurs niveaux de couche intermédiaire :

- le niveau le plus bas : mailles élémentaires du modèle, elles contiennent les utilisateurs de l'Internet et leur nœud de rattachement ;

- les niveaux plus élevés : mailles superposées au niveau inférieur et qui contiennent des nœuds de transit.

Il s'agit de niveaux logiquement distincts, mais dont les mailles sont géographiquement superposées et dont les nœuds de transit sont physiquement identiques aux nœuds des niveaux inférieurs.

Les mailles élémentaires (mailles de la couche la plus basse) ne possèdent pas toutes un nœud de transit. Par contre, le trafic se répandant de proche en proche, chaque maille élémentaire est voisine d'une maille possédant un nœud de transit. Il faut pour cela qu'une maille élémentaire sur quatre possède un nœud de transit.



On modélise ainsi les routages pouvant exister entre réseaux régionaux. Pour un niveau de transit donné, la règle est que chaque triangle est relié directement et uniquement avec ses triangles adjacents ; cette règle s'applique à chaque niveau de transit. Un triangle sur quatre sera donc nœud de transit de premier niveau ; un triangle sur seize sera nœud de transit de second niveau, etc.

2.4 Conventions du modèle

Mailler totalement la couche la plus basse ne serait pas le reflet exact de la réalité. Nous supposons donc dans le calcul des coûts qu'un nœud de la couche la plus basse n'est relié qu'à deux nœuds adjacents (au lieu des trois nœuds adjacents figurés sur le schéma de principe ci-dessus).

Le territoire de chaque grande région est découpé en triangles équilatéraux représentant chacun la zone de couverture

d'un nœud d'accès à l'Internet.

La surface des mailles est inversement proportionnelle à la densité géographique des utilisateurs (c'est-à-dire que le nombre d'utilisateurs par nœud est constant). La règle de proportionnalité (nombre d'utilisateurs par nœud) est calculée chaque année en divisant le nombre total d'utilisateurs par le nombre des réseaux accédant à l'Internet.

Ce ratio n'est pas stable car il évolue d'une année sur l'autre, mais cette évolution reste faible sur la chronique considérée. Le paramètre retenu est la valeur moyenne sur l'ensemble de la chronique considérée.

La surface d'une maille est bornée inférieurement par la surface d'une « zone tarifaire », qui est un paramètre du modèle. Par « zone tarifaire », nous entendons toute aire géographique où les utilisateurs peuvent accéder à un nœud de rattachement à un tarif identique quelle que soit leur position dans cette aire. Il peut s'agir par exemple d'une ZLE, d'une zone locale américaine dans laquelle le téléphone est gratuit, ou de la zone de couverture d'un CD de réseau câblé. .

Lorsque la surface des mailles est bornée inférieurement, le nombre d'utilisateurs par maille augmente, ce qui entraîne une augmentation du nombre de routeurs dans la maille.

Il existe donc deux paramètres pour décrire les mailles de la couche basse :

- le nombre d'utilisateurs par nœud, qui détermine la surface des mailles lorsque le nombre d'utilisateurs est faible, donc lorsque la surface des mailles est grande.

- si la maille se réduit à une « zone de tarification », la surface des mailles reste fixe et égale à cette valeur qui est le deuxième paramètre de dimensionnement des mailles.

Il existe un nœud de transit de niveau N pour quatre mailles de niveaux $N - 1$. La surface des mailles de niveau N

est donc quatre fois supérieure à celles des mailles de niveau $N-1$. Ceci reste vrai quel que soit le paramétrage de la couche la plus basse.

Le nombre des niveaux de transit est tel que le plus haut niveau de transit soit composé de triangles égaux à (ou du même ordre de grandeur que) la zone de couverture d'un nœud du niveau « backbone » (voir ci-dessous la modélisation de ce niveau). Ainsi, le dernier niveau de transit est pris en compte dans la modélisation du « niveau backbone ».

Le nombre des niveaux de transit est déterminé dynamiquement par le modèle.

Chaque maille représente un « réseau » qui se compose de :

- un routeur (situé au centre du triangle) ;
- une batterie de modems ;
- des LL permettent de raccorder le routeur aux routeurs des mailles adjacentes d'un même niveau.

On retient deux LL par nœud de la couche la plus basse. On retient pour les autres couches intermédiaires (les différents niveaux de transit possible) trois LL par nœud pour le raccordement aux nœuds adjacents d'une même couche : un triangle sur quatre possède un nœud de transit de premier niveau ; un triangle sur seize possède un nœud de transit de second niveau etc. Les LL de raccordement à un niveau de transit ne concerneront donc qu'un réseau sur quatre pour le premier niveau de transit etc.

On observe le volume de trafic concentré au niveau de chaque niveau de transit pour déterminer la capacité nécessaire des LL. Le débit des LL est calculé selon les règles de dimensionnement ci-dessous.

Remarque : nous considérons que les liaisons de raccordement des LAN (ou directement des stations utilisateurs) au

routeur d'un nœud d'une maille de niveau le plus bas sont du niveau « réseau local » et donc à la charge de l'utilisateur.

2.5 Unités d'œuvre

Les frais de personnel concernent la maintenance et l'exploitation technique des équipements. S'agissant de calculer le coût de l'architecture, l'exploitation commerciale n'est pas prise en compte à ce niveau.

En octobre 1994, 17 000 organismes étaient recensés. En moyenne un quart de temps d'un technicien par organisme est employé à l'exploitation technique spécifique à l'Internet.

Le nombre de routeurs des backbone a augmenté avec la croissance du nombre des réseaux locaux et régionaux en se dispersant sur le territoire. Ainsi, les points d'accès se sont rapprochés des réseaux régionaux au fur et à mesure que la densité de ceux-ci s'est accrue. Cette croissance peut être considérée comme valide jusqu'à ce que les mailles des backbones atteignent une surface minimum en deçà de laquelle la notion de backbone perd de son sens (la maillage du backbone devient trop fin, les mailles trop nombreuses, et alors un nouveau backbone doit se superposer pour fédérer l'ancien backbone).

Nous considérons dans notre modèle que la couche backbone est composée de mailles qui ont une surface fixe dans le temps.

La taille de ces mailles (ou le nombre de nœud de la couche backbone) est un paramètre du modèle qui prend des valeurs différentes selon les grandes régions géographiques.

Unités d'oeuvre - nœuds de rattachement et de routage : calculateurs modifiés pour les besoins de l'Internet commandant des interfaces avec des LL hauts débits et des interfaces

FDDI ou Ethernet pour la connexion de serveurs locaux.

- LL assurant le maillage des nœuds : ce sont des LL à haut débit louées aux opérateurs télécom. Leur prix unitaire sera donc le tarif moyen sur chaque région de location de ces LL (il existe de fortes différences entre les tarifs des diverses régions).

Remarque : on prend en compte les distances réelles (et non les distances à vol d'oiseau).

Nous considérerons deux catégories de personnel :

- le personnel d'exploitation des backbones : son coût annuel équivaut typiquement à 10 % du coût d'investissement des équipements (les LL étant louées, leurs frais de fonctionnement sont inclus dans le prix de location).

- les organismes de standardisation (ISOC, IETF, etc) : on évalue à 1000 personnes l'effectif total de ces organismes, réparti dans les diverses régions au prorata du nombre d'utilisateurs par région.

Chaque unité d'oeuvre est décomposée en deux parties dont les caractéristiques d'évolution de prix sont différentes : une partie dont le prix reste constant (typiquement, les composants mécaniques ou d'alimentation d'un équipement) et une partie dont le prix évolue exponentiellement à la baisse (typiquement les composants électroniques ou logiciels).

La loi d'évolution des coûts unitaires est donc de la forme :

$$P = P_0[\alpha + (1 - \alpha)e^{-\beta(t-t_0)}]$$

Les coefficients α et β diffèrent selon le type d'équipement. Nous supposons qu'ils sont les suivants :

Pour les routeurs, $\alpha = 25 \%$ et $\beta = 30 \%$

Pour les LL aux USA, $\alpha = 60 \%$ et $\beta = 20 \%$

Pour les LL en Europe, $\alpha = 50 \%$ et $\beta = 50 \%$

2.6 Dimensionnement

Le dimensionnement du réseau concerne le nombre des unités d'œuvre (routeurs et liaisons spécialisées) et leur taille.

Le nombre des unités d'œuvre ne dépend que du dimensionnement de la couche la plus basse et de celui de la couche backbone. Leur nombre dans les couches intermédiaires est calculé automatiquement selon le modèle triangulaire. Le dimensionnement dépend essentiellement des paramètres suivants :

- nombre d'utilisateurs par nœuds de rattachement (tant que la surface des mailles élémentaires est supérieure à celle des « zones tarifaires ») ;
- surface d'une « zone tarifaire » ;
- nombre de nœuds des backbones.

La taille des LL est déterminée par les règles de dimensionnement définies ci-dessous.

L'Internet Protocol découpe les flux de données en paquets de taille variable. Un en-tête est ensuite ajouté par TCP-IP, ce qui induit un overhead moyen de 17%.

Le rythme d'émission d'un utilisateur est sporadique : de brefs pics de trafic (lors du téléchargement d'une page HTML ou d'un fichier) entrecoupés de silences. La hauteur des pics est limitée par la capacité en débit de l'équipement de l'utilisateur.

Nous dimensionnons l'infrastructure sur le trafic moyen à l'heure chargée, la définition de l'heure chargée pouvant varier suivant que l'on considère un réseau régional ou un backbone international (ou intercontinental).

L'heure chargée est difficile à définir à un niveau global car :

- le réseau étant mondial, les plages horaires chargées dif-

fèrent d'une région à une autre (en temps universel) ;

- on n'a pas identifié, comme pour le téléphone, des habitudes sociales concentrant les communications sur certaines heures de la journée (les applications de l'Internet n'étant généralement pas des applications interactives de personne à personne, il est indifférent de le consulter à n'importe quelle heure du jour ou de la nuit) ;

- pour certaines applications (téléchargement, transfert ...), les utilisateurs peuvent programmer les communications dans des plages horaires localement non chargées (de façon à bénéficier notamment des tarifs réduits de la boucle locale) ;

- les flux de trafic de l'Internet sont chaotiques.

On suppose que l'heure chargée est une plage horaire de 8 heures (en temps universel) et que 50% du trafic total est écoulé dans cette plage horaire

Nous supposons que le trafic sortant de chaque routeur connecté est isotrope tant en volume qu'en destination :

- chaque routeur génère le même flux. On raisonnera donc sur un trafic moyen par maille : si N est le nombre total de mailles de niveau 1, chaque maille émet vers l'extérieur un trafic égal à $1/N$ fois le trafic total.

- chaque maille de niveau 1 émet la même quantité de trafic à destination de chacune des autres mailles de même niveau. Chaque maille reçoit donc un trafic égal à $1/N$ fois le trafic total. Chaque maille disposant de deux liaisons avec les mailles voisines, chaque liaison écoulera la moitié de son trafic sortant, soit $1/2N$ du trafic total.

Une maille sur quatre remplit une fonction de transit de premier niveau. Le trafic de transit obéit aux hypothèses ci-dessus, en considérant que chaque nœud de niveau dispose de trois liaisons avec les nœuds voisins de même niveau. Le trafic de transit de premier niveau est égal à $4 * (N - 3)/N$ fois le

trafic total. Celui de deuxième niveau est égal à $4 * ((N/4) - 3)/(N/4)$ fois le trafic de transit de premier niveau, etc. Par récurrence, on voit que le trafic de transit de niveau i est égal à $4 * ((N/4i - 1) - 3)/(N/4i - 1)$ fois le trafic de transit de niveau $i - 1$.

Les flux de trafic sont (ou deviendront) majoritairement régionaux ou nationaux. Ceci répond à un usage souvent régional des services proposés et à des orientations économiques (mise en place des serveurs miroirs pour économiser les liaisons à longue distance lorsque les services réclament de forts débits). Pour considérer ce phénomène, nous introduirons un paramètre limitant la part de trafic remontant à la couche backbone par rapport au trafic qui aurait transité par cette couche si la répartition du trafic avait été uniforme.

Remarque : ce paramètre devrait en fait dépendre du service considéré. Pour le téléphone, il est proche de 100%. Pour des services large bande, il est sans doute nettement inférieur.

Le rendement des LL (rapport entre débit utile et capacité des LL) est dans l'hypothèse de base égal à 30%.

Le dimensionnement d'un réseau de paquets diffère de celui d'un réseau à commutation de circuit ; il repose sur la théorie des files d'attente. De ce fait, ce n'est pas uniquement la capacité (en bit/s) qui détermine la dimension du réseau mais également des paramètres de qualité dont le temps de traversée du réseau, fonction de la capacité des équipements et du chemin parcouru (nombre d'équipements traversés).

Il n'existe pas d'assurance de qualité sur l'Internet. L'Internet garantit la mise à disposition de ressources pour permettre une connexion entre deux points quelconques, mais ne garantit pas le résultat : il n'est pas en mesure de garantir qu'un paquet émis arrivera à destination. Ceci résulte notamment du fait qu'un paquet est susceptible de traverser des

réseaux exploités par des acteurs différents, et qu'un opérateur n'est pas en mesure de garantir la qualité des réseaux auxquels il est interconnecté.

Nous dimensionnons le modèle en nous assurant uniquement qu'il n'y a pas de blocage, c'est à dire que le trafic moyen, équi-réparti géographiquement, peut s'écouler à l'heure de pointe sans blocage, étant données les hypothèses faites sur l'heure de pointe et sur la répartition du trafic. Le nombre de LL des backbones est notamment calculé de cette façon.

Remarque : cette règle n'implique aucune contrainte quant au principal critère de qualité du réseau qui est le temps moyen de traversée du réseau.

Lorsque le débit moyen à l'heure de pointe ne nécessite que des LL moyen débit (cas des couches intermédiaires pour lesquelles le débit moyen à l'heure de pointe par réseau est de l'ordre de quelques dizaines de kbit/s), la taille des liaisons sera dimensionnée de façon suffisante pour laisser passer le trafic à l'heure chargée en prenant en compte les divers paliers de la hiérarchie de débit de LL.

2.7 Les résultats

Le résultat principal du modèle est l'évolution du coût moyen par utilisateur et par an. Il était estimé à 450 F en 1995, et selon le modèle le coût prévu pour 2005 était de 200 F, l'évolution entre ces deux valeurs suivant une trajectoire exponentielle décroissante.

Ce résultat découlait des hypothèses, notamment celle concernant la baisse tendancielle du coût des unités d'oeuvre. Nous avons réalisé plusieurs variantes en modifiant les hypothèses tout en restant dans l'intervalle des valeurs plausibles : évolution du trafic, définition de l'heure chargée, densité des utili-

sateurs etc. Ces calculs n'ont pas altéré de façon significative le résultat du modèle.

L'un des résultats les plus intéressants concerne la structure de coût de l'Internet : la structure de l'Internet est simple et on peut classer ses composants en trois types : équipements des nœuds de routage, liaisons de transmission, personnel.

Le poste « liaisons de transmission » est le plus important. Leur coût est le tarif de location des LL par les opérateurs télécoms : c'est un prix de détail souvent éloigné du coût de revient de l'opérateur télécom.

Répartition du coût total



Les variantes ont fait apparaître une grande sensibilité du coût de l'Internet aux charges de personnel : une faible variation de la charge de personnel nécessaire pour un réseau régional aurait un effet significatif sur le coût total de l'Internet.

3 Économie du dimensionnement

Les réseaux sont soit des infrastructures dont le coût d'exploitation est quasi nul (routes), soit des automates dont le

coût d'exploitation est faible (télécommunications), soit des plates-formes de services dont le coût d'exploitation est élevé mais dépend peu du volume du service produit (transport aérien), soit une juxtaposition de moyens de production au coût de plus en plus élevé mis en œuvre progressivement lorsque la charge du réseau augmente (énergie).

La fonction de production d'un réseau est « à coût fixe », une fois celui-ci construit, car il peut fournir des services dont le coût est nul ou négligeable (communication téléphonique en dehors de l'heure de pointe, sièges libres d'un avion etc.). Le coût de sa construction dépend de la dimension que lui donnent ses promoteurs et sera pratiquement une fonction linéaire de cette dimension : nombre de lignes principales du réseau téléphonique, nombre et taille des avions d'un transporteur aérien, largeur des routes etc.

Le coût du dimensionnement pour une règle d'ingénierie r donnée est fonction du trafic anticipé en période de pointe t^a :

$$c = f(r, t^a)$$

Le trafic en dehors de la période de pointe a un coût de moyen terme nul (puisqu'il ne conduit pas à réviser l'anticipation t^a) et un coût de court terme soit nul (trafic télécoms en dehors de l'heure de pointe), soit très faible (passager qui occupe un siège d'avion qui autrement serait resté libre).

Le trafic qui excède la capacité en heure de pointe est refoulé : il a donc un coût de court terme nul mais son coût de moyen terme ne l'est pas car sa prise en compte induit une révision de l'anticipation t^a .

La définition de l'heure de pointe, ainsi que du taux de blocage du trafic jugé admissible à l'heure de pointe, est un enjeu important. Les opérateurs télécoms définissent deux périodes

de pointe pour la téléphonie : l'une s'étale sur les heures de bureau du matin et de l'après-midi et correspond à un taux de blocage faible ; l'autre concerne la pointe du soir, provoquée par le trafic résidentiel, et correspond à un taux de blocage plus élevé. Le trafic d'affaires et le trafic résidentiel n'ayant pas la même répartition géographique, leurs matrices de trafic sont différentes.

Le dimensionnement du réseau se fait donc en deux fois, d'abord de façon à satisfaire la demande d'affaires, puis en tenant compte du trafic résidentiel. Les taux de blocages sont établis de façon à minimiser le coût tout en fournissant une qualité de service socialement admissible. L'arbitrage entre ces deux objectifs est fait de façon empirique.

La fonction de coût apparaît ainsi sous deux angles différents, selon la chronologie considérée :

Le coût de dimensionnement est fonction linéaire de la dimension D du réseau, mesurée par le débit offert. Il comprend le salaire des personnels nécessaires pour faire fonctionner les installations, et ne se réduit donc pas au coût d'investissement en capital fixe.

Le coût d'exploitation est un coût fixe, indépendant du trafic transporté jusqu'au seuil D , et qui devient infini si le trafic demandé dépasse D .

Une fois le réseau dimensionné, son coût a été payé et il sera le même quel que soit le nombre de communications pourvu que celui-ci soit inférieur ou égal à D . Au delà de D le réseau est incapable d'acheminer davantage de communications (les appels supplémentaires sont refoulés, les trames supplémentaires sont retardées ou perdues), et donc le coût d'exploitation de court terme est infini.

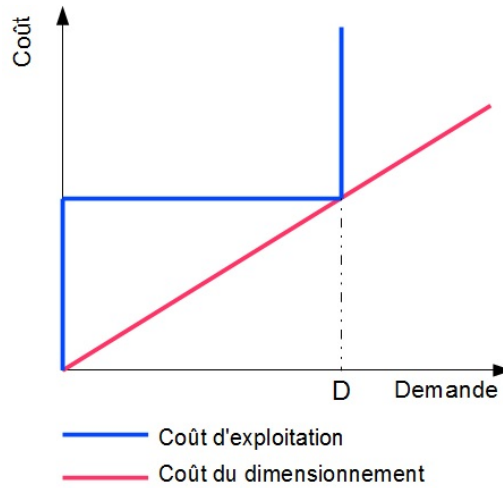


FIGURE 3 – Fonction de coût d'un réseau

On peut alors représenter la fonction de coût en associant deux fonctions :

- coût de dimensionnement, mesuré par a , équivalent quotidien (en valeur actualisée) du coût des investissements et des frais de personnel nécessaires au dimensionnement, rapporté à une unité du dimensionnement : $C = aD$
- coût d'exploitation (quotidien) : c est égal à C si la demande est inférieure à D , c est infini si la demande est supérieure à D .

3.1 Approche probabiliste

Des modèles prévoient la demande adressée à un réseau à un instant donné en fonction des facteurs qui l'expliquent. Ces modèles sont probabilistes : même si l'incertitude est limitée par la prise en compte de tous les facteurs explicatifs, la demande est aléatoire par nature.

Le modèle fournit ainsi non la prévision de la demande, mais celle des paramètres d'une loi statistique à laquelle la demande se conforme à chaque instant (il s'agit le plus souvent d'une loi de Poisson, qui peut être approchée par une loi de Laplace-Gauss).

Supposons que la demande x obéit à chaque instant à une distribution statistique représentée par la loi de Laplace-Gauss :

$$x = \mathcal{N}(m, \sigma), \text{ soit :}$$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right]$$

Nota Bene : $f(x)$ est ici l'approximation d'une loi Log-normale, car x ne peut pas prendre de valeurs négatives. On peut dire, de façon équivalente, que m est assez grand par rapport à σ pour que la probabilité des valeurs négatives de x puisse être considérée comme nulle.

Supposons aussi que réseau doive être dimensionné de sorte que le taux de blocage moyen soit inférieur à θ . Son dimensionnement est D tel que $D = m + k\sigma$, où k doit être tel que, en notant $\Phi(k)$ la probabilité d'une valeur inférieure à k pour une variable aléatoire qui suit la loi $\mathcal{N}(0, 1)$, $\Phi(k) = 1 - \theta$, soit :

$$k = \Phi^{-1}(1 - \theta)$$

k est fonction décroissante de θ . Si $\theta < 50\%$, $k > 0$.

3.2 Effet de l'incertitude sur le coût

Supposons que le coût du réseau puisse être approché au premier ordre par une fonction affine de son dimensionnement, soit :

$$C = \alpha + \beta(m + k\sigma), \text{ avec } \beta > 0.$$

Si $\theta < 50\%$ (c'est le cas le plus fréquent),

$$\partial C / \partial \sigma = \beta k > 0.$$

Si l'incertitude augmente, le réseau doit donc pour assurer une qualité de service égale être dimensionné plus largement.

De nombreux réseaux (télécommunications, transport aérien, chemin de fer etc.) sont passés dans les dernières décennies d'un régime de monopole (de droit ou de fait) à un régime de concurrence.

Ce passage accroît toutes choses égales d'ailleurs le coût du réseau car il dégrade l'information disponible sur le marché (l'entreprise en situation de monopole connaît toute la demande, l'entreprise en concurrence ne connaît que celle qui lui est adressée et sa part de marché connaît des fluctuations en partie aléatoires), donc la qualité de l'estimation des coefficients et l'incertitude σ qui en résulte.

3.3 Volume du trafic servi

Notons x_t la demande à l'instant t , et D le dimensionnement du réseau.

Si $x_t < D$, tous les clients peuvent être servis.

Si $x_t > D$, il faut renoncer à servir une partie des clients.

Le nombre moyen M des clients servis à chaque instant est non pas m , puisque cette moyenne recouvre aussi les instants où $x_t > D$, mais, en notant $F(x)$ la fonction cumulative de $f(x)$:

$$M = \int_{-\infty}^D x f(x) dx + D[1 - F(D)],$$

qui est inférieur à m .

Nota Bene 1 : l'expression de M additionne la moyenne du nombre des clients lorsque $x_t < D$ au nombre des passagers lorsque $x_t \geq D$ (l'avion étant alors à plein dimensionnement, le nombre de passagers est égal à D).

Nota Bene 2 : l'intégrale est prise depuis $-\infty$, mais par hypothèse m est assez grand par rapport à σ pour que l'on puisse considérer que les valeurs de $f(x)$ sont nulles pour $x < 0$.

Posons

$$G = \int_{-\infty}^D x f(x) dx \text{ et } x = m + \sigma U, \text{ soit :}$$

$$G = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{D-m}{\sigma}} (m + \sigma U) \exp\left(-\frac{U^2}{2}\right) dU$$

$$G = mF(D) + H, \text{ avec :}$$

$$H = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\frac{D-m}{\sigma}} U \exp\left(-\frac{U^2}{2}\right) dU. \text{ On obtient :}$$

$$H = -\frac{\sigma}{\sqrt{2\pi}} \exp\left(-\frac{(D-m)^2}{2\sigma^2}\right), \text{ soit :}$$

$$H = -\sigma^2 f(D). \text{ Donc :}$$

$$G = mF(D) - \sigma^2 f(D) \text{ et finalement :}$$

$$M = D[1 - F(D)] + mF(D) - \sigma^2 f(D)$$

Pour obtenir les représentations graphiques ci-dessous, nous posons $m = 10$ et $\sigma = 2$.

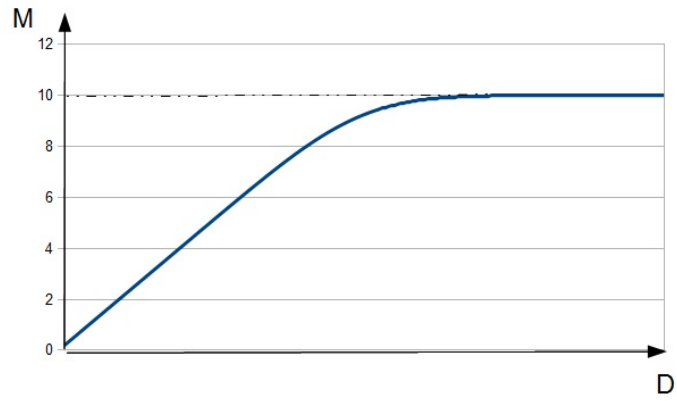


FIGURE 4 – Trafic moyen en fonction du dimensionnement

Commentaire du graphique 2 : lorsque le dimensionnement est faible, le réseau est saturé : le trafic est égal au dimensionnement. Lorsque le dimensionnement est très élevé, le trafic moyen est égal à m .

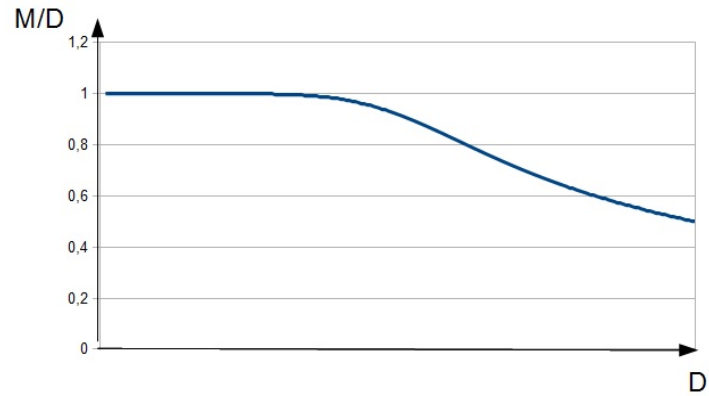


FIGURE 5 – Taux de remplissage moyen en fonction du dimensionnement

Commentaire du graphique 3 : lorsque le dimensionnement est faible, le remplissage est égal à 1. Le taux de remplissage décroît lorsque le dimensionnement augmente.

3.4 Profit et surplus social

Notons p le prix de vente d'une unité de trafic, supposé donné, et a le coût actualisé d'une unité de dimensionnement rapporté à une période de trafic. Le profit lors de cette période est $\Pi = pM - aD$ et sa dérivée par rapport à D est :

$$\frac{\partial \Pi}{\partial D} = p \frac{\partial M}{\partial D} - a$$

$$\text{Or } \frac{\partial M}{\partial D} = 1 - F(D) - Df(D) + mf(D) - \sigma^2 f'(D)$$

Il est facile de vérifier que pour toute valeur de x $\sigma^2 f'(x) = mf(x) - xf(x)$. Il en résulte que :

$$\frac{\partial M}{\partial D} = 1 - F(D)$$

Le dimensionnement D^* qui maximise le profit est donc tel que le taux de blocage soit :

$$\theta^* = 1 - F(D^*) = a/p$$

Le taux de blocage θ^* qui maximise le profit est donc égal au quotient du coût unitaire du dimensionnement par le prix unitaire. On trouve par exemple que $D^* = m$ si $a/p = 50\%$ (dans ce cas, $\theta^* = 50\%$: la moitié du trafic est refoulée).

Supposons que $p = 10$ et $a = 5$; le profit est alors maximal pour $D^* = 10$.

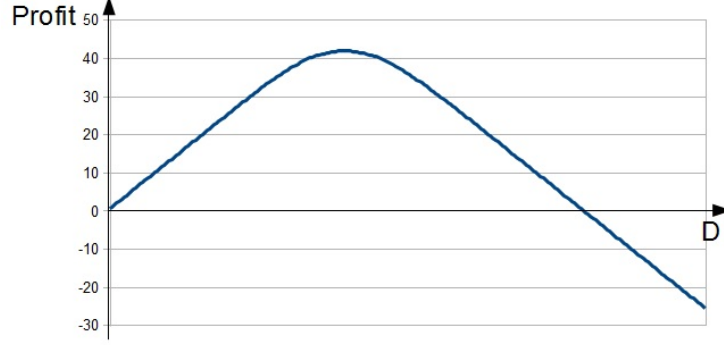


FIGURE 6 – Profit selon le dimensionnement

Certains clients sont rationnés. Supposons leur perte d'utilité égale à k fois le prix du service avec évidemment $k > 0$. La moyenne du trafic qui est bloqué est $B = m - M$.

Le surplus est $\Sigma = pM - aD - pkB = p(1+k)M - aD - km$. Il est maximal si le dimensionnement est D^{**} tel que :

$$\theta^{**} = 1 - F(D^{**}) = \frac{a}{p(1+k)}$$

Dans le cas considéré, $\theta^{**} = 25\%$: le taux de blocage est deux fois moins élevé que celui qui permet de maximiser le profit. $D^{**} > D^*$ (le calcul numérique donne en l'occurrence $D^{**} = 11,3$). On retrouve ici un résultat classique : le dimensionnement qui maximise le surplus social est donc supérieur à celui qui maximise le profit du réseau.

Le régulateur, porteur de l'intérêt général, doit donc intervenir pour que le réseau soit dimensionné plus largement que ne l'implique la maximisation du profit.

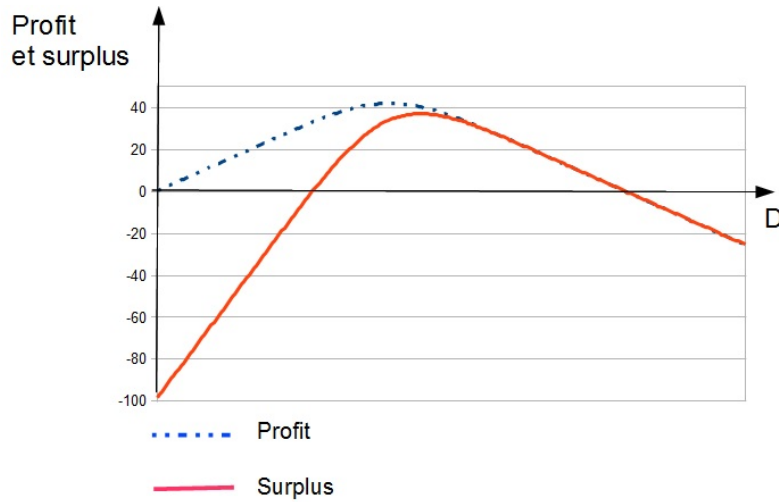


FIGURE 7 – Surplus selon le dimensionnement

4 Le cyberspace

4.1 Mobilité ou ubiquité ?

Depuis 1876 il n'est plus indispensable, pour converser avec une personne, de se trouver à portée de voix : le téléphone a conféré l'ubiquité au signal vocal. Lorsque seuls des téléphones fixes étaient disponibles, il fallait cependant pour pouvoir téléphoner se trouver près d'un poste téléphonique (dans l'appartement, au bureau, dans une cabine). A l'ubiquité du signal, le téléphone mobile a ajouté l'ubiquité de l'accès : il équipe le corps de l'utilisateur et non plus un local. L'ubiquité de la téléphonie est ainsi devenue totale, aux zones blanches du réseau près.

Depuis 1991 l'utilisateur du Web accède à des serveurs

dont la localisation lui importe peu : que le serveur soit proche ou à l'autre bout du monde, le délai d'affichage est analogue. Le Web confère ainsi l'ubiquité aux ressources informatiques¹, celles-ci étant également accessibles à partir de tout ordinateur connecté au réseau. Depuis que l'ordinateur est devenu lui aussi mobile, qu'il équipe le corps de l'utilisateur, l'informatique bénéficie elle aussi de l'ubiquité totale.

Pour décrire cette évolution on utilise par analogie avec la téléphonie mobile le terme de mobilité. Mieux vaut dire *ubiquité* : l'utilisateur, qu'il soit mobile ou non, se trouve en effet plongé dans un espace où la distance géographique n'existe pas.

Les routes, les rues des villes, les portes des immeubles et des appartements délimitent, dans l'espace géographique, des voies d'accès licites. Il en est de même dans le cyberspace où les habilitations délimitent les ressources informatiques auxquelles un utilisateur particulier a accès et jouent donc le même rôle que les murs, portes et serrures de l'espace géographique. Entre un utilisateur et une ressource on peut en outre définir une distance : celle de l'intelligibilité de cette ressource pour cet utilisateur, ou encore (et cela revient au même, car nous ne comprenons que ce qui nous intéresse), celle de l'intérêt qu'elle présente pour lui. On retrouve ainsi dans le cyberspace, sous des formes certes différentes, les limites et la distance qui sont familières dans l'espace géographique.

On peut aussi y définir d'autres distances : une distance entre utilisateurs, d'autant plus grande que leurs centres d'intérêt sont plus éloignés ; une distance entre les ressources, d'autant plus grande que leur contenu est plus différent. C'est

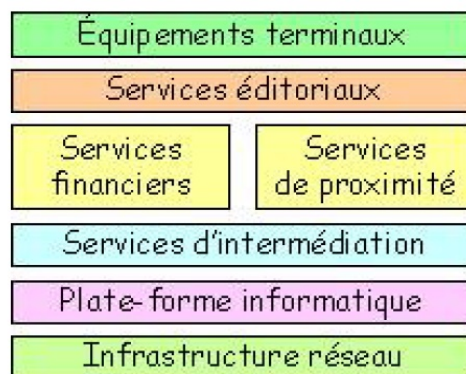
1. Les « ressources informatiques », que nous appellerons « ressources » tout court, sont composées (1) de textes, données, sons et images ; (2) des logiciels utilisés pour les classer, trouver, traiter ; (3) de la mémoire et de la puissance des serveurs.

après avoir défini ces deux distances que l'on peut évaluer, en s'appuyant sur l'observation statistique, la distance entre un utilisateur et une ressource.

L'ubiquité logique que procure l'Internet est absolue (sous réserve des droits d'accès). L'ubiquité physique qu'il procure est par contre, comme en téléphonie, limitée aux signaux : le réseau peut transmettre le signal vocal, ou l'image d'une personne, ou encore le codage de la forme géométrique d'un objet physique que le récepteur pourra reproduire (impression 3D), mais il ne peut pas transporter les objets physiques eux-mêmes.

4.2 Architecture du cyberspace

L'architecture du cyberspace peut se représenter selon un modèle en couches :



- *Infrastructure du réseau* : deux réseaux d'accès (mobile à haut débit, fixe à très haut débit) et un réseau de transport, l'Internet, avec les outils informatiques de gestion et supervision (adressage, réplication, dimensionnement, etc.) ;

- *Services de plate-forme informatique* : serveurs (processeurs, systèmes d'exploitation, outils de supervision), mémoires, sécurité (chiffrement, back up, protection contre les intrusions, antivirus etc.). Ces services traitent la physique informatique et assurent la fiabilité et la disponibilité du service ainsi que la protection des ressources contre les indiscretions et les attaques ;

- *Services d'intermédiation* : aide à la recherche des ressources qui correspondent aux besoins de l'utilisateur (moteurs de recherche, classification documentaire, dissémination sélective, publish and subscribe etc.). L'intermédiation est associée, pour les ressources payantes, à des services financiers ; pour la livraison et la maintenance des produits matériels (livres, équipements etc.), à des services de proximité ;

- *Services financiers* : traitement automatique des effets de commerce (dettes, créances, ordres de virement, répartition entre ayants droit) échangés par les utilisateurs et les fournisseurs ;

- *Services de proximité* : livraison, installation, formation des utilisateurs, entretien, dépannage etc.

- *Services éditoriaux* : ressources gratuites ou payantes mises à disposition sur le réseau (logiciels, textes, musique, audiovisuel, données etc.) ;

- *Gamme des terminaux*, divisée en deux familles (fixes et mobiles).

4.3 Changements par rapport à aujourd'hui

On retrouve dans ce monde certains services que nous connaissons déjà : le commerce électronique est en place avec les outils de paiement informatisés et il s'articule avec des services de livraison pour les produits pondéreux. Le futur

apportera donc non pas de l'émergence de services entièrement nouveaux, mais l'exploitation systématique, plus puissante aussi parce que disposant de moyens nouveaux, des services qui existent déjà aujourd'hui.

Cette intensification « change la vie » tout comme le téléphone mobile, qui n'avait disait-on pourtant « rien de nouveau », l'a changée lorsqu'à partir de 1995 son prix a baissé, que sa couverture géographique s'est densifiée et qu'il est devenu un produit de masse.

Des services dont le « business plan » aurait été auparavant peu crédible seront rentables. Des équipements qui n'avaient jamais été mis en réseau (systèmes d'arrosage, équipements ménagers, chauffage etc.) lui seront raccordés, ce qui facilitera leur télécommande, leur télémaintenance et le téléchargement de nouvelles versions de leur logiciel. L'informatisation des transactions facilitera l'accès des PME à des services de comptabilité, de gestion d'agenda etc., ainsi que le partage des centres d'appel.

La ressource informatique d'une personne contiendra son dossier médical, l'historique de ses comptes bancaires, de ses revenus et de ses dépenses, celui de son cursus scolaire et des formations qu'il a suivies etc. L'accès à ces informations doit être sécurisé² mais l'utilisateur pourra s'il le souhaite l'ouvrir de façon permanente ou temporaire à des personnes qu'il autorise (médecins, administration fiscale etc.)³.

2. Le chiffrement est un enjeu crucial : pour que les entreprises, les personnes, puissent déposer leurs ressources sur le réseau, il faut que la confidentialité soit inviolable fût-ce même pour un organisme qui, comme la NSA, dispose des outils les plus puissants.

3. Un rapide calcul montre que l'ensemble des conversations (horodatées et localisées) auxquelles un être humain participe pendant sa vie tient, après traitement par un logiciel de reconnaissance vocale et codage en mode caractère, sur un CD-Rom. Que cette possibilité soit utilisée ou

Cependant tout ne sera pas automatisé : parallèlement à la croissance de la part prise par l'automate dans les services, croîtra pour l'utilisateur le besoin d'une assistance et d'un dialogue avec un être humain, que ce soit par réseau (centre d'appel, visiophonie) ou en face à face, afin de traiter les problèmes qui ne sont pas du ressort de l'automate (expliquer, comprendre). Il en résultera un changement de la structure de l'emploi.

L'ubiquité du cyberspace permettra des usages aujourd'hui inédits. Il se peut par exemple qu'une personne qui habilite l'administration fiscale à consulter automatiquement l'historique de ses revenus bénéficie d'une réduction de l'impôt sur le revenu en contrepartie de la fiabilité de l'information et de la commodité d'accès. Une personne qui consulte un médecin pourra, pendant la durée de la consultation, lui donner accès à son dossier médical, à l'historique des prescriptions et traitements etc.

4.4 Outils du cyberspace

Les outils nécessaires à la conquête de l'ubiquité sont disponibles : réseaux à haut débit, outils de classement et de recherche, de chiffrement et de gestion des habilitations etc.

A chaque utilisateur on peut associer dans le cyberspace un domaine comportant l'ensemble des ressources susceptibles de l'intéresser. Ce domaine sera défini (1) par segmentation, le domaine étant commun à tous les utilisateurs appartenant à un même segment ; (2) par personnalisation, l'observation des requêtes et du comportement d'un utilisateur permettant de préciser l'information sur ses besoins. On pourra lui fournir

non, elle illustre ce qu'une personne pourra indexer et stocker dans une mémoire de taille relativement modeste.

des outils qui :

- facilitent la recherche des ressources intéressantes (moteurs de recherche, mais aussi synthèses, critiques, paratexte etc.) ;
- indiquent, de façon proactive, les changements survenus dans le contenu de son domaine (nouvelles ressources etc.) ;
- permettent de prendre connaissance du contenu de son domaine, de ses frontières, des relations qu'il entretient avec les domaines connexes, des extensions possibles de son champ d'intérêt.

Des communautés d'utilisateurs se forment à l'exemple de ce qui se passe sur Flickr où l'on stocke, documente et partage des photographies, l'index documentaire amorçant le fonctionnement communautaire en facilitant la recherche.

* *

Les possibilités ne sont pas immédiatement utilisées : si la ressource informatique personnelle semble par exemple le lieu naturel où l'on peut stocker un dossier médical, il faut du temps pour que la corporation médicale, le milieu hospitalier, acceptent de s'y adapter. De façon générale les institutions, confrontées à une transformation profonde de leurs procédures et organisations, sont lentes à en tirer parti.

Certaines personnes résistent, car l'informatisation de la vie personnelle n'est pas bien vécue par tout le monde ; elle pose d'ailleurs, tout comme l'a fait le téléphone mobile des problèmes de savoir-faire et de savoir-vivre nouveaux.

Les prédateurs cherchent les failles des systèmes de chiffrement et de sécurité : le cyberspace devra donc, pour être habitable, être équipé de dispositions juridiques, surveillé par une gendarmerie vigilante, sanctionné par un appareil judiciaire qualifié.

Les entreprises, enfin, rencontreront dans le cyberspace de nouvelles formes de commerce et d'organisation de la force de travail auxquelles elles devront elles aussi s'adapter.

Les efforts d'organisation personnelle et collective, la transformation des institutions, prennent plus de temps que l'innovation technique.

4.5 Cyberspace et géographie

Le cyberspace est la négation de l'espace géographique. Déjà, dans les entreprises, l'Intranet a permis de combler (voire de retourner) l'écart de compétence professionnelle qui existait entre la direction générale et les directions régionales⁴. Une part des travaux de l'entreprise s'effectue dans le cyberspace : interprétation des textes, classement comptable, programmation informatique etc. Le télétravail sera donc mis en œuvre pour les activités pour lesquelles il est économiquement efficace (c'est-à-dire : non pas à 100 % ni pour toutes les activités, mais de façon significative tout de même).

Le cyberspace est aussi en relation avec l'espace géographique : le commerce des produits physiques sur le Web se boucle par une livraison ; la localisation d'appel est à la base de plusieurs des services offerts sur les réseaux ; les automobilistes utilisent des systèmes de navigation qui s'appuient sur le GPS. Voici une exploration (non exhaustive) des services qui concrétisent l'articulation entre cyberspace et espace géographique :

La **localisation**, fournie de façon précise par le GPS incorporé dans le terminal mobile, est un des attributs de l'uti-

4. Les « gens des DR » sont des lecteurs plus attentifs de la documentation technique, désormais accessible et à jour sur l'Intranet, que ne le sont les « gens de la DG » qui comptent pour s'informer sur des conversations entre experts.

lisateur et contribue à la définition des ressources qui l'intéressent : il s'intéresse potentiellement à sa région, sa commune, aux manifestations culturelles locales, aux commerçants du coin, aux services de proximité qui peuvent lui être offerts. Il peut être utile aussi de localiser des personnes dépendantes (enfants, personnes âgées).

Le GPS permet à l'utilisateur de savoir où il se trouve pendant ses déplacements. Les systèmes de **navigation** équipent le piéton aussi bien que l'automobile.

Le **calcul d'itinéraire** peut se faire en temps réel, c'est-à-dire tenir compte de la situation présente du trafic (chantiers, embouteillages, manifestations etc.). Le piéton peut localiser les taxis et transports en commun disponible : emplacement des stations et arrêts, durée d'attente, durée prévisible du voyage.

Le terminal mobile peut :

- informer l'utilisateur sur les **ressources physiques** proches et ouvertes (pharmacies, hôpitaux, distributeurs automatiques de billets, agences bancaires, bureaux de poste, restaurants, cinémas etc.) ;
- afficher des **vitrines virtuelles** (quand on passe devant un immeuble, indique des magasins dont certains ne se voient pas depuis la rue) ;
- donner une **information météorologique** localisée ;
- signaler au touriste des monuments intéressants, des expositions etc., et lui donner des explications historiques sur les lieux qu'il parcourt ;
- indiquer les noms des habitants d'un immeuble, etc.

L'utilisation locale d'un **système d'information géographique (SIG)** fournit des informations diverses et à des échelles diverses ; elle permet de localiser les réseaux qui se trouvent en sous-sol, de visualiser le cadastre etc.

4.6 Représentation mathématique

Les utilisateurs du Web disposent actuellement de moteurs de recherche dont Google est le plus fameux ; d'encyclopédies en ligne dont Wikipedia est la plus connue. Les liens hypertexte leur permettent de « surfer » pour trouver des ressources inconnues et, d'aventure, intéressantes.

Les outils disponibles, aussi puissants et utiles qu'ils soient, ont des limites. Google indexe des chaînes de caractère et les résultats qu'il fournit sont altérés par des homonymies. Le surf est aléatoire. Les outils de dissémination sélective, censés apporter automatiquement à l'utilisateur des informations qui l'intéressent, sont dans les limbes.

Cependant des outils sémantiques visent à limiter les homonymies dans les moteurs de recherche ; le contenu du Web est soumis à des outils de classification automatique délimitant des clusters de textes ; des outils de datamining examinent les co-occurrences lexicales et détectent les tendances de l'évolution des contenus.

* *

Entre un utilisateur et une ressource, on peut définir une distance qui est celle de l'intelligibilité ou, ce qui revient au fond au même, de l'information que ce texte peut apporter au lecteur en prenant « information » non au sens qu'a ce mot dans la théorie de Shannon mais au sens que lui donne Simon-don : une information, c'est quelque chose qui vous in-forme, qui modifie ou complète la forme intérieure de votre représentation du monde, qui vous forme vous-même. L'information ainsi conçue a une signification : elle suscite une action de la part de celui qui la reçoit ou du moins elle modifie (transforme) les conditions de son action future. Plus un texte peut

apporter d'information à un lecteur, plus il présente d'intérêt pour lui : intelligible, intéressant, informatif sont des synonymes.

Pour pouvoir recevoir une information, il faut avoir été formé, et c'est en recevant de l'information que l'on se forme. L'amorce de ce cycle est enfouie dans les origines de la personne tout comme l'amorce du cycle de la poule et de l'œuf est enfouie dans les origines de la vie.

* *

Notons U l'ensemble des utilisateurs, R l'ensemble des ressources. Supposons définie une distance $d(u, r)$ entre le lecteur u et le texte r , d'autant plus petite que ce texte est plus intelligible pour ce lecteur. Nous appellerons « domaine du lecteur u » l'ensemble $D(u)$ des ressources dont la distance au lecteur est inférieure à un seuil conventionnel s :

$$D(u) = \{r | d(u, r) < s\}.$$

Nous appellerons « lectorat de la ressource r » l'ensemble $L(r)$ des utilisateurs dont la distance à la ressource est inférieure à s :

$$L(r) = \{u | d(u, r) < s\}.$$

Considérons un sous-ensemble D de R . Nous noterons $L(D)$ l'ensemble des lecteurs intéressés par toutes les ressources que contient D :

$$L(D) = \{u | D \subset D(u)\}.$$

Considérons un sous-ensemble L de U . Nous noterons $D(L)$ l'ensemble des ressources qui intéressent tous les membres de L :

$$D(L) = \{\cap D(u) | u \in L\}.$$

Nous dirons que L et D sont en correspondance si :

$$L(D(L)) = L \text{ et } D(L(D)) = D.$$

Nous allons esquisser la démarche qui permet de construire deux segmentations (l'une des utilisateurs, l'autre des ressources) dont les segments mutuels sont statistiquement en correspondance.

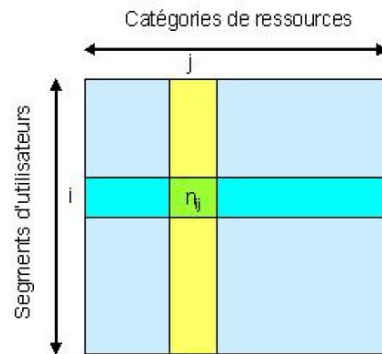
Les outils utilisent des techniques statistiques (notamment les Support Vector Machines) mais à une échelle et sur des volumes qui exigent des algorithmes puissants. Pour classer des textes, ils considèrent la taille, la structure formelle, les métadonnées et les co-occurrences lexicographiques. Pour classer des images et les enregistrements audio ou vidéo, ils considèrent les métadonnées (index etc.) et, plus difficilement, les formes représentées ainsi que les paramètres techniques de l'image ou du son.

Dans tous les cas, il faut définir une distance entre les ressources, puis une distance entre agrégats de ressources ; on peut alors, par classification ascendante hiérarchique, construire une ultramétrie sur l'ensemble des ressources puis obtenir une segmentation en « coupant les branches les plus longues » de l'arbre qui représente l'ultramétrie.

Si l'on veut que le Web puisse signaler à l'utilisateur de façon proactive les ressources susceptibles de l'intéresser, il faut segmenter la population des utilisateurs en utilisant d'abord des données « intrinsèques » (c'est-à-dire indépendantes de l'usage qu'ils font de la ressource) observées par enquête auprès d'un échantillon représentatif : âge et sexe, catégorie socioprofessionnelle, lieu de résidence, lieu de travail, métier, fonction dans l'entreprise, statut matrimonial, nombre et âge des enfants, hobbies etc.

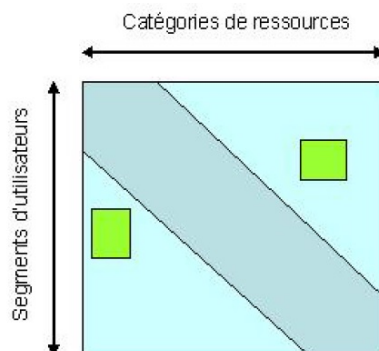
Puis on observe, toujours sur le même échantillon, les pratiques dans l'utilisation du Web et on demande éventuelle-

ment aux utilisateurs de noter l'intérêt des ressources qu'ils consultent. On établit un tableau des fréquentations en croisant la segmentation des utilisateurs et celle des ressources (figure 2).



Une analyse factorielle des correspondances, réalisée sur ce tableau, met en évidence la corrélation entre les deux segmentations. L'ordre des classes sur le premier axe de l'analyse fournit un tableau aussi proche que possible de la forme diagonale.

Il est alors possible d'associer statistiquement à chaque segment des utilisateurs une catégorie de ressources et inversement. Cela fournit une segmentation globale, définie à la fois selon les paramètres propres aux utilisateurs et selon ceux des ressources : les ensembles L et D sont ainsi mis en correspondance. On peut de la sorte, ayant identifié les ressources qui intéressent le plus un segment d'utilisateurs, amorcer une politique de diffusion proactive.



On trouvera en dehors de la diagonale quelques cases bien remplies faisant exception à la logique d'ensemble. C'est un des résultats les plus précieux de l'analyse : ces cases contiennent des utilisateurs dont le comportement s'écarte du comportement majoritaire, qu'ils soient « en avance » ou « en retard » dans la consultation des ressources ou pour tout autre raison. Les examiner donne un aperçu sur la dynamique de la demande.

Lorsque l'outillage est ainsi amorcé, on peut mettre les segmentations à jour en recherchant, parmi les combinaisons linéaires des données intrinsèques et des ressources consultées par un utilisateur, celles qui sont les mieux corrélées : il faudra ici recourir aux techniques de l'analyse canonique.

Cette analyse permet de :

- placer un nouvel utilisateur dans l'espace documentaire, une fois connu le segment auquel il appartient (par la suite, la personnalisation permet de faire évoluer ce qu'on lui propose) ;
- signaler aux utilisateurs d'un segment les nouveaux documents qui font partie du segment documentaire qui a priori les intéresse ;
- fournir à chacun une vue du cyberspace, de la façon dont il est structuré et de l'endroit où on le situe lui-même (il

peut ainsi organiser des voyages dans cet espace) ;

- identifier les segments « voisins », potentiellement intéressants ceux des utilisateurs d'un segment donné qui souhaitent élargir leur horizon.

5 Questions stratégiques

5.1 Réseau, monopole et concurrence

5.1.1 Qu'est-ce qu'un monopole naturel ?

Un secteur de l'économie obéit au régime du *monopole naturel* lorsque la fonction de coût d'une entreprise typique de ce secteur est à rendement d'échelle croissant.

Le rendement d'échelle croissant peut se manifester de plusieurs façons :

- *économie d'échelle* : le coût moyen de production décroît lorsque le volume produit augmente ;
- *économie d'envergure* : le coût de la production de divers produits par une même entreprise est inférieur à la somme des coûts lorsque la production est réalisée par plusieurs entreprises.
- *économie d'innovation* : les résultats d'une R&D restant internes à l'entreprise, l'innovation est d'autant plus intense (en principe) que l'entreprise est plus importante.

Dans un secteur où le rendement d'échelle est croissant, la plus grosse entreprise bénéficie si elle est bien gérée (cette condition est importante !) de coûts de production plus bas que ceux des autres entreprises. Elle est en mesure de les évincer du marché en pratiquant des prix inférieurs aux leurs : alors le monopole s'instaure *naturellement*.

La fonction de coût résulte des conditions techniques, phy-

siques, de la production ; elle se trouve donc en amont de la décision économique ou politique, qu'elle *conditionne*. Instaurer par une décision politique la concurrence dans un secteur qui relève *physiquement* du monopole naturel suscite une inefficacité, un gaspillage des ressources utilisées pour produire.

Toutefois un monopole est toujours tenté d'abuser de sa force envers les consommateurs en pratiquant un prix excessivement élevé ; l'excès de profit, de sécurité, risque par ailleurs de l'endormir et d'inhiber l'innovation. L'État, garant de l'intérêt collectif, doit donc créer une institution, le *régulateur*, qui impose au monopole des contraintes de prix et d'efficacité analogues à celles que susciterait la concurrence.

5.1.2 Le réseau est-il un monopole naturel ?

Le réseau télécoms est un automate qui assure la communication à la demande entre les équipements terminaux - téléphone, ordinateur etc. - des utilisateurs, ménages ou entreprises.

La demande qui lui est adressée est une *matrice de trafic* probabiliste. Pour des raisons historiques elle a une forme approximativement bloc-diagonale, chaque bloc correspondant au territoire d'une nation : le trafic international est en effet moins intense que le trafic interne à un pays.

La fonction de coût du réseau recouvre :

- le coût de la *boucle locale* qui raccorde les utilisateurs à un commutateur local et représente les trois quarts du coût du réseau ;
- le coût du *coeur du réseau*, qui assure la commutation et le transport.

La qualité du service s'exprime en termes de blocage (pour le téléphone), de débit (pour les ordinateurs), de priorité et de sécurité (pour le transport et le stockage des données). Le

coeur du réseau est dimensionné (débit des liaisons, puissance des commutateurs) en tenant compte du caractère probabiliste du trafic et selon un arbitrage entre coût et qualité (cf. ci-dessus « économie du dimensionnement »).

On peut considérer les faits suivants :

- le coeur du réseau, intégralement informatisé, est depuis les années 70 l'objet d'un flux d'innovations qui implique la mise à jour continue de ses protocoles, logiciels et équipements ;
- la boucle locale, qui exige un lourd investissement en génie civil, doit elle aussi être périodiquement mise à niveau (ADSL sur la paire de cuivre, puis fibre optique ou voie hertzienne) ;
- la diversification des services (téléphonie fixe, puis mobile ; télécopie ; commutateurs privés (PABX) et réseaux des entreprises ; Web, transfert de fichiers et autres services sur l'Internet etc.) complique la vie de l'utilisateur : ingénierie et maintenance de son installation, choix dans une grille tarifaire compliquée ;
- comme tout automate, le réseau est sujet à des pannes qu'il faut traiter : supervision, reprise automatique en régime dégradé etc.

En regard de la matrice de trafic et des possibilités techniques, les équipements et techniques (filaire, hertzienne, satellitaire etc.) que le réseau utilise sont soumis à une contrainte de *cohérence* portant sur les protocoles, le dimensionnement, le cadencement des investissements et – *last but not least* – la relation avec les utilisateurs.

Par ailleurs l'économie d'innovation est particulièrement forte dans les télécoms, tant pour les techniques du réseau que pour les services qui l'utilisent.

L'exigence de cohérence et l'économie d'innovation militent pour que l'on considère le réseau télécoms comme un monopole naturel sur la zone géographique où la matrice de

trafic forme un bloc – c’est-à-dire, dans notre cas, sur le territoire français.

Le monopole naturel est une notion *physique* liée à la fonction de coût du réseau, ce n’est pas une notion juridique. Rien ne s’oppose donc en principe à ce que le réseau soit construit et exploité par des entreprises juridiquement distinctes, à *condition* que leurs choix techniques et le cadencement de leurs investissements soient coordonnés par le régulateur et que le fonctionnement d’ensemble du réseau soit supervisé.

Du point de vue physique qui est celui de la fonction de coût, tout se passerait alors comme si le régulateur était une direction générale et les entreprises des établissements déconcentrés dont les relations relèveraient d’un partenariat et non de la concurrence.

Soulignons enfin deux faits cruciaux :

- 1) le réseau télécoms qui équipe un pays constitue *un* réseau et non une collection de plusieurs réseaux juxtaposés, car il est soumis à une forte contrainte de cohérence ;
- 2) l’automate qu’exploite l’opérateur historique est l’une des ressources informatiques importantes du pays, voire la plus importante ; cet opérateur devrait donc, en principe, être un acteur important du processus d’*informatisation* de la Nation.

5.2 Stratégie extensive ou intensive

Le réseau doit croître pour survivre. Cela peut se faire de deux façons. Soit il diversifie l’offre de services sur une même plate-forme géographique (croissance intensive), soit il élargit la plate-forme pour offrir le même service de base à une clientèle plus nombreuse (croissance extensive).

La croissance intensive est difficile à concevoir et à organiser, elle suppose l’aptitude à former et faire vivre des partena-

riats, mais elle est relativement peu coûteuse car elle tire parti des infrastructures existantes. La croissance extensive n'exige pas de gros effort intellectuel mais elle suppose des investissements coûteux et à la rentabilité incertaine, car le service de base sera vigoureusement concurrencé.

Il est naturel que la culture d'entreprise, peu consciente des risques futurs, préfère la croissance extensive qui permet de continuer à faire ce que l'on sait déjà bien faire. La croissance intensive obligerait à diversifier les activités, à écouter des partenaires, à faire entrer dans les équipes dirigeantes des personnes possédant de nouvelles compétences et dont il faudrait comprendre le langage. C'est difficile pour une corporation sûre de ses valeurs.

Cette description bipolaire de la stratégie peut sembler simpliste : elle transcrit cependant le choix essentiel pour tout exploitant de réseau, qu'il s'agisse des télécoms, du transport aérien, du transport routier, de l'électricité etc. Le service de base, que l'exploitant maîtrise, est concurrencé et donc peu rentable à terme. La différenciation, difficile à réaliser, permet par contre de mieux rentabiliser la plate-forme.

La difficulté n'est ni technique (les réseaux exploitent une plate-forme informatique puissante), ni économique (l'investissement initial nécessaire pour monter un service à valeur ajoutée est réduit), elle est sociologique – et donc peut-être insurmontable.

Le fait est que les chercheurs de France Telecom ont conçu, dans les années 1990, l'équivalent de Google, de Wikipédia, d'Amazon, des intermédiations financières, des services de télé-médecine et de télé-enseignement, de l'impression 3D. Des business plan ont été préparés sous la direction de Jacques Lasisz mais aucun d'entre eux n'a trouvé, dans la haute hiérarchie, le sponsor qui aurait permis de lancer l'une ou l'autre de

ces affaires, qui n'ont pas intéressé les filiales de capital-risque du groupe.

Le montant des investissements nécessaires pour monter une de ces affaires était trop modeste pour motiver des dirigeants dont le pouvoir s'évaluait selon l'ampleur du budget qui leur était confié – il aurait fallu qu'il fût de quelques centaines de milliers d'euros pour sembler *sérieux*.

Les compétences nécessaires pour discuter d'égal à égal avec des partenaires (avec par exemple les hôpitaux et le SAMU pour une télésurveillance des grossesses à risque) n'étaient pas disponibles dans une grande entreprise habituée à des relations de domination de type impérial.

Ces services à valeur ajoutée procuraient enfin peu de trafic au réseau (quelques octets au plus pour l'exemple cité), alors que la stratégie du « delta minute » mettait l'entreprise sous pression pour faire croître le trafic, approche purement quantitative qui s'écartait de la stratégie intensive.

* *