

EUROLINC FRANCE

La Maçonnerie Générale d'Angers

tentée par l'internet

De la difficulté d'utiliser l'Internet quand on prétend écrire en français : histoire des déboires d'une PME et tentative d'explication

Albert Truelle dirige une florissante entreprise de maçonnerie sous l'enseigne "Maçonnerie Générale d'Angers".

Albert est poussé par les démarcheurs de son opérateur de télécom à ouvrir un site internet.

Enfin convaincu, il se fait expliquer les détails pratiques de mise en œuvre. Et comme il s'agit d'abord de trouver un nom de domaine pour son site, Albert propose naturellement "Maçonnerie Générale d'Angers.com".

- "Ah non ! Voyons Monsieur Truelle, pas d'espace, pas de majuscules, pas d'apostrophe, pas de cédille, pas d'accents dans les noms de domaine ..."
- "D'accord, alors je prends le système français : "Maçonnerie Générale d'Angers.fr."
- "Mais non, Monsieur Truelle, c'est exactement la même chose avec le système français .fr."
- "Il ne me reste que "maconnerie-generale-dangers", alors ?"
- "Oui, c'est cela même..."

Pourquoi cette restriction – intolérable – de l'alphabet utilisable ? C'est à cause de l'annuaire utilisé pour transmettre les messages et les données qui, créé et géré par des Anglo-Saxons, ne satisfait entièrement que les besoins des Anglo-Saxons.

Les données qui traversent l'internet sont transmises par morceaux appelés **paquets**. Chaque paquet porte une **adresse** qui permet de l'acheminer au travers des commutateurs du réseau (appelés aussi nœuds ou routeurs). Cette méthode est la réplique électronique des opérations effectuées pour acheminer des cartes postales.

L'adresse placée dans les paquets est une série de 32 chiffres binaires (bits). Pour faciliter leur expression dans un contexte plus humain, ils sont représentés par une succession de quatre nombres décimaux séparés par des points, p.ex. 61.175.132.115. Dans le jargon de l'internet cette adresse est appelée **numéro IP**.

Il existe aussi une autre catégorie d'adresses, appelées selon les locuteurs, adresse email, site web, nom de domaine, URL, ou simplement adresse. La précision du langage n'est pas le point fort du milieu de l'internet. C'est le contexte qui permet de donner un sens aux termes utilisés.

P.ex.: info@isoc.fr, ou <http://www.itu.int>.

Ces **noms**, sensés être plus mnémoniques que les numéros, sont généralement utilisés par les humains dans leur courriel, et imprimés sur cartes de visite, papiers à lettres, brochures, etc.

Il n'en reste pas moins que seuls les numéros IP servent à aiguiller les données dans le réseau. Il est donc nécessaire de convertir (traduire) les adresses-noms en numéros IP avant d'expédier des messages, fichiers, requêtes, etc. Cette opération est réalisée par des logiciels situés dans le PC de l'utilisateur ou chez son fournisseur d'accès (FAI).

Ainsi, l'internet et le téléphone moderne utilisent les mêmes procédés pour établir les communications. En téléphonie, l'utilisateur choisit (sur son mobile ou son poste fixe) le nom d'une personne ou d'une société.. L'appareil effectue la conversion du nom en numéro de téléphone avant de réaliser l'appel. Dans l'internet, l'utilisateur choisit (dans son carnet d'adresses ou dans le message auquel il répond) le nom d'une personne ou d'une société. Son logiciel (mailer ou brouteur) ou celui de son FAI, effectue la conversion en numéro IP avant d'émettre le message.

Le numéro IP est ainsi, dans l'internet, l'équivalent fonctionnel d'un numéro de téléphone.

On retrouve alors la même problématique dans l'internet et le téléphone. Chacun se construit un annuaire personnel des numéros dont il se sert le plus souvent. C'est pratique, mais forcément limité. Il y a donc besoin dans l'internet comme dans le téléphone d'un annuaire pour trouver les numéros que l'on n'a pas mis dans son annuaire personnel.

Rien d'exotique à première vue. Il suffirait, pourrait-on penser, d'éditer, comme pour le téléphone, des annuaires par pays donnant les numéros des abonnés. On y trouverait donc **Maçonnerie Générale d'Angers** en France et $\Phi\Lambda\text{I}\Psi$ en Grèce. Ce serait incongru d'écrire PHILIPS en Grèce, alors que c'est un nom d'origine grecque. Les russes n'auraient sans doute pas besoin d'annuaire puisqu'ils ont pris l'habitude de s'en passer.

C'est ici qu'apparaissent des différences substantielles. Les annuaires téléphoniques sont conçus pour une consultation

manuelle et visuelle, alors que dans l'internet on n'a prévu qu'une consultation par des ordinateurs. L'utilisateur n'est pas sensé connaître les numéros IP mais seulement les noms. Ce sont des ordinateurs qui vont chercher les numéros dans l'annuaire au moment où on veut établir une communication.

A première vue cette automaticité n'aurait rien de déplaisant, sauf que l'annuaire de l'internet, réalisé par des techniciens américains, n'a en aucune manière anticipé l'usage d'autres alphabets que le leur, c.a.d. l'ASCII, une espèce limitée d'anglais.

Pour faire simple et de mauvais goût, les seuls caractères autorisés pour exprimer les noms des abonnés sont: A à Z, 0 à 9, et le tiret. (Les minuscules sont autorisées mais sont interprétées comme des majuscules). Espace, accent, cédille, et autres lettres diacritiques sont exclus.

Que cette situation relève d'un nombrilisme atavique ou d'une politique de domination, il en résulte que l'internet force les utilisateurs à dégrader leurs noms en un charabia angloïde.

Un peu d'histoire : comment en est-on arrivé à ce point d'absurdité, ou les étapes de la création de " the " annuaire Internet, et l'importance de la " racine " .

L'annuaire de l'internet n'existait pas à ses débuts (1972). Il a été réalisé vers 1983 par une petite équipe de techniciens américains de la côte ouest. Ce système est connu sous le nom **DNS**.

Le DNS est conçu pour être l'annuaire unique, global, mondial, et naturellement américain, de l'internet. Les informations contenues dans le DNS sont physiquement réparties dans un grand nombre d'ordinateurs, mais elles constituent logiquement une base de données unique sous contrôle américain.

Pour enregistrer un nom, il faut s'adresser d'abord à un **registreur**, (c.a.d. un détaillant), lequel s'adresse à un **registre** (c.a.d. un grossiste), qui gère une tranche de l'annuaire appelée **domaine**. Chaque domaine est repéré par le mot le plus à droite du nom, p.ex. COM, EDU, ORG, FR, UK. Cette fragmentation de l'annuaire a été introduite pour des raisons commerciales: créer une certaine concurrence. L'ensemble des noms appartient au DNS.

Chercher un nom, et son numéro IP associé, dans le DNS, s'effectue selon une procédure imposée par le système. En bref, il faut consulter une cascade de listes donnant finalement l'information recherchée. La première liste consultée contient les adresses des sous/annuaires des domaines (COM, EDU, ORG, etc.). C'est elle qu'on appelle la **racine**.

La racine est la porte d'entrée du DNS. C'est dire que le réseau deviendrait inopérant si elle était hors service. Afin de pallier cette éventualité, la racine est recopiée dans 12 ordinateurs (9 aux Etats Unis, 2 en Europe, 1 au Japon). Ceci permet de répartir la charge de traitement des interrogations, et de diminuer le risque de panne générale.

La racine originale est gérée par la société Verisign, par contrat avec le Département du Commerce USA (DOC). L'organisme californien ICANN, qui assure la gouvernance de l'internet, n'est qu'une façade du DOC. Par ailleurs, Verisign est aussi un registre et un registraire qui loue les noms en COM, soit plus de 70% de l'ensemble des noms. De plus, sa maison mère SAIC est un des principaux fournisseurs du Pentagone. Les vaches à lait de l'internet sont donc bien gardées.

La racine est un instrument de contrôle, de gestion, et de surveillance, mais pas d'organisation

On a vu précédemment que la racine contient une liste d'adresses de sous/annuaires. Lorsque je veux envoyer un mël à truelle@maconnerie-generale-dangers.fr, mon logiciel demande à la racine l'adresse de maconnerie-generale-dangers.fr. Comme la racine n'en sait rien, elle renvoie une liste des sous/annuaires du domaine FR, que mon logiciel va de suite consulter pour obtenir le numéro IP de maconnerie-generale-dangers. Tout va bien.

Il se pourrait, toutefois, que suite à une fâcherie sur les modalités de chasse aux dahus de destruction massive, le Pentagone veuille museler des trouble-fêtes de l'Hexagone. Un clic suffit pour supprimer FR de la racine. Aucun mël n'ira plus vers un abonné en FR. Le logiciel de la racine peut aussi bien tester l'origine des interrogations, et répondre qu'il y a erreur lorsque l'origine est en FR.

Si ce scénario paraît improbable à certains, il en est d'autres moins bruyants. Au lieu de la liste des sous/annuaires FR, la racine peut fort bien renvoyer le numéro IP d'un service d'écoute. Le mël émis vers des utilisateurs est alors transmis d'abord au service d'écoute, qui le fait suivre ensuite aux destinataires. C'est l'équivalent d'une bretelle sur une ligne de téléphone. Avec en prime la possibilité de modifier le contenu du mël. Au minimum la racine permet de collecter des statistiques sur qui communique avec qui, ou visite quel site web.

Bof, diront les fatalistes. Il y a déjà Echelon et Carnivore. Que l'internet peut-il apporter de plus ? Eh bien, est-ce que l'internet ne serait pas justement une composante essentielle de ces grandes oreilles ?

Aucun autre système de communication à l'échelle planétaire ne fonctionne ainsi ; il y aurait un moyen de mettre de la cohérence dans le système : la multiplicité des annuaires Internet.

L'architecture des grandes bases de données distribuées est un champ de recherche ouvert depuis le milieu des années 1970. Il y a belle lurette que les systèmes coopérants et décentralisés sont utilisés de préférence aux systèmes centralisés et contrôlés par un seul pays, comme le DNS. Prenons quelques exemples.

- Le système postal.

Il est ancien. L'écriture des noms et adresses (postaux) n'a jamais été fixée au plan mondial par aucun pays. Les contraintes se résument à deux:

- 1) le nom du pays destinataire doit être lisible par les agents postaux qui assurent l'acheminement du courrier,
- 2) noms et adresses doivent être lisibles dans le pays de destination.

La première contrainte conduit à l'adoption d'un alphabet international pour désigner les pays. Le reste est affaire de chacun des pays. Je peux donc poster à Paris une lettre dont l'adresse est en hindi, pourvu qu'il y ait INDE ou INDIA sur l'enveloppe.

- Le réseau téléphonique.

Sous sa forme traditionnelle, il n'y a pas de noms, seulement des numéros. Mais il n'y a pas de format unique de numéro. Le nombre de chiffres varie selon les pays, ou selon les zones, ou selon l'abonné (standard ou sélection directe du destinataire). La seule contrainte internationale est le préfixe (un ou plusieurs chiffres) indiquant le pays de l'appelé.. Le reste est affaire de chacun des pays.

- Le réseau de mobiles GSM.

Le numéro d'appel d'un mobile n'est autre chose qu'un nom. Il ne désigne pas une adresse physique comme le numéro IP, puisque l'on ne sait pas où sera l'abonné. Logiquement, c'est un nom comme dujardin@wanadoo.fr.

Pour localiser l'abonné appelé, le réseau interroge l'annuaire où est enregistré l'abonné (HLR). L'annuaire renvoie l'adresse d'un autre annuaire, celui des visiteurs de la zone où l'appelé s'est manifesté (VLR). On peut alors le sonner.

Il n'y a pas d'annuaire unique ni de racine dans le GSM. Le même préfixe que pour le téléphone traditionnel permet d'identifier le pays d'enregistrement de l'appelé, donc son annuaire (HLR).

Tous ces systèmes marchent. Ils reposent sur une coordination multi-latérale entre états ou opérateurs responsables, pour fixer des normes d'échanges inter-réseaux. La gestion des noms ou numéros à l'intérieur de chaque réseau relève de l'autorité de chacun.

Revenons à l'internet. Les noms se terminent par une étiquette indiquant le domaine d'enregistrement. C'est l'équivalent d'un code pays dans le téléphone, et c'est bien ainsi que le traite la racine en renvoyant une liste des annuaires du domaine indiqué. Comme la liste des pays ou domaines change rarement, il n'y a nullement besoin d'une racine unique. L'internet pourrait aussi bien fonctionner sur le modèle d'annuaires du GSM.

Alors, pourquoi une racine dans l'internet ?

Le DNS, dont la conception remonte à 1982, est un cas pathologique dans un univers technique en constante évolution. Son architecture centralisée n'a pas évolué, alors qu'elle est intrinsèquement fragile aux attaques. De plus, toute proposition d'évolution se heurte à des tabous de préservation de l'existant réel ou imaginaire.

Ce genre de situation techniquement anormale s'explique fort bien si l'on aborde le sujet sous un angle politique. Le contrôle américain de l'internet repose sur le mythe de la nécessité du DNS. Ce mythe est entretenu par l'IETF, un organisme de standards contrôlé par les américains, qui s'oppose donc à toute évolution susceptible de remettre en cause l'architecture actuelle centralisée.

Annuaire pluriels et autonomes, seule possibilité d'introduire un "vrai" multilinguisme dans l'Internet : La Maçonnerie Générale d'Angers enfin reconnue

L'autonomie des annuaires permet l'utilisation naturelle d'alphabets éventuellement différents dans différents annuaires (analogie postale). Prenons un exemple.

La Zigolie est un pays peu connu, et à part ses habitants, les zigols, seuls quelques lettrés ou expatriés comprennent la langue. L'alphabet est assez singulier. Son code de pays est ZG.

Un mèl destiné à un correspondant zigol pourrait s'écrire ainsi:

De: Musée ethnographique @ Alençon.fr

À: ☉ ☽ ↶ ↷ ↸ .zg

.. le reste en français ou zigol, ou toute autre langue connue des correspondants.

Il est clair que les caractères des noms zigols devront être représentés dans un mèl par des suites de chiffres binaires (bits). Le codage nécessaire est exclusivement du ressort d'un organisme de normalisation de Zigolie. Le correspondant français ne fait que coller dans son mèl les champs de noms zigols copiés dans un mèl reçu d'un correspondant zigol. Ces caractères peuvent ne pas être visibles dans leur forme zigole sur un PC français, mais ceci n'est qu'une question de polices disponibles localement, et n'a pas d'incidence sur le transport du mèl.

Le seul élément que l'internet doit pouvoir reconnaître et interpréter est le code du pays, **ZG**

Tout mèl destiné à un correspondant dont le nom de domaine est terminé par ZG sera transmis à un annuaire de Zigolie, qui fera suivre le mèl au site zigol spécifié dans l'en-tête du mèl. Une alternative est d'interroger un annuaire zigol pour obtenir le n° IP du destinataire, auquel le mèl sera ensuite transmis directement.

Il appartient à un organisme de Zigolie de définir une politique d'enregistrement de ses noms nationaux. L'essentiel, vu de France, est qu'un annuaire zigol assure l'acheminement des mèls envoyés en Zigolie ou la conversion de ses noms en n° IP.

Noter que le nom zigol de l'exemple ne contient pas de @. Ce caractère n'existe pas en zigol. La représentation d'un nom et sa structure, hors le code pays, ne concernent que la Zigolie. Autrement dit, l'organisation du système d'acheminement zigol est autonome et indépendant de celui imposé dans l'internet.

Ce qui est applicable au zigol l'est aussi bien à tout alphabet, notamment le français. Les principes énoncés limitent au strict nécessaire les contraintes d'un nommage multi-lingue. Des contraintes supplémentaires peuvent être souhaitables, pour faciliter les échanges avec ses voisins. Mais ceci relève d'accords négociables multi-latéraux.

Passer d'un annuaire unique contrôlé par les américains à des annuaires pluriels et autonomes soulève des questions d'ordre technique, et d'ordre politique. L'évolution sera laborieuse, et inévitable.

Une première étape serait d'éliminer l'usage de la racine américaine. Ce serait d'ailleurs un impératif national si la sécurité des communications était réellement prise au sérieux. La réalisation est techniquement assez limitée. Elle consiste à fixer les paramètres de fonctionnement du serveur de noms que l'on utilise, de manière qu'il conserve en permanence une copie des adresses des domaines de l'internet (COM, EDU, ORG, etc. et domaines par pays). Le volume de ces informations est de l'ordre de 100K octets.

Une autre solution (non exclusive) serait de mettre à disposition des utilisateurs un annuaire national, dupliqué et sécurisé.

C'est alors qu'apparaissent les obstacles de nature politique. Le choix de consulter la racine pour établir une communication est une décision locale, prise par le logiciel du FAI, ou d'un ordinateur de l'entreprise, ou même d'un PC individuel. La mise à jour de tous ces systèmes requiert du temps et un bon niveau de consensus de la part des utilisateurs. C'est néanmoins faisable. Tous les systèmes informatiques ont besoin de fréquentes mises à jour, quoiqu'il arrive. L'exemple de la fameuse bogue de l'an 2000 a démontré que même un fantôme peut être pris pour une nécessité s'il est bien orchestré.

A supposé que l'on se donne la volonté de changement au plan national, la question ne serait pas réglée pour autant. Il faudrait qu'un nombre significatif de pays adoptent une politique similaire, sans nécessairement choisir les mêmes solutions techniques. Ce niveau de consensus multinational n'est certainement pas un objectif facile à concrétiser. Comme bien d'autres d'ailleurs. Il faut du temps. Il est quand même incohérent que l'Europe se démène pour mettre en place Galileo, tout en abandonnant l'internet aux américains. Un système d'annuaires européens ne serait-il pas tout aussi justifié, et bien moins coûteux ?

Cela dit, il n'est pas techniquement nécessaire que tous les systèmes suivent la même politique d'annuaire. L'usage de la racine américaine restera toujours possible, sauf action contraire des américains. On pourrait donc se contenter d'une migration partielle, étalée dans le temps.

Il va de soi que les opposants à cette évolution ne manqueront pas d'annoncer les catastrophes inévitables qui en résulteraient. L'exemple de la Chine est le meilleur démenti. C'est tout simplement le modèle zigol. L'internet chinois utilise des noms et des annuaires en chinois. Il est autonome. C'est en quelque sorte un intranet chinois. Les échanges avec le reste de l'internet, autrement dit l'intranet américain, sont strictement contrôlés. Ce n'était pas techniquement nécessaire, mais sans doute politiquement plus correct.

Le cas de la Chine est encore unique, mais il serait naïf de penser qu'il le restera. Que ce soit pour des motifs d'indépendance politique et économique, ou de protection linguistique et culturelle, ce modèle ne manquera pas de faire des émules.

Et pour conclure, revenons à notre ami Albert Truelle. Car il pourra alors vivre enfin le bonheur que "Maçonnerie Générale d'Angers.fr" soit devenu un nom de domaine parfaitement recevable.

-+--+--+-

Sigles:

ASCII: American Standard Code for Information Interchange
DNS: Domain Name System
DOC: Department Of Commerce
FAI: Fournisseur d'Accès Internet
GSM: Global System for Mobiles
HLR: Home Location Register
ICANN: Internet Corporation for Assigned Names and Numbers
IETF: Internet Engineering Task Force
IP: Internet Protocol
PC: Personal Computer
SAIC: Science Applications International Corporation
URL: Uniform Resource Locator
USA: United States of America
VLR: Visitor Location Register

version 5 - 29-6-03

version 4 - 28-5-03

version 3 - 23-5-03

version 2 - 20-5-03