

Mon glossaire en informatique pour ce blog

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 7 octobre 2014

<https://www.bortzmeyer.org/mon-glossaire.html>

Ce blog comprend un certain nombre d'articles techniques en français sur l'informatique. J'essaie d'utiliser moins de mots anglais que la moyenne des informaticiens. Résultat, dans ce domaine où le vocabulaire est très anglicisé, des termes français perturbent les francophones, qui ne les comprennent pas car ils ne sont pas largement répandus. D'où une liste partielle des termes les moins classiques.

Mais d'abord, les principes que j'essaie de suivre dans l'écriture de ces articles :

- Quand l'article est en français (une minorité </auto/summary> est en anglais), l'écrire en français correct. Cela veut dire pas trop de fautes de grammaire ou d'orthographe <<https://www.bortzmeyer.org/remerciements.html>> mais également pas de mots en anglais s'il existe un meilleur terme en français. Utiliser un terme en anglais est souvent un signe de paresse (ce qui peut être une bonne idée) ou de non-compréhension (ce qui est plus grave).
- Ne pas refuser par principe tout terme anglais. Le français, comme toutes les langues, récolte de temps en temps des termes étrangers et il serait ridicule de vouloir figer la langue dans un état donné. Dans la technique, le vocabulaire est en général défini par ceux qui inventent et maîtrisent la technique, comme l'a bien montré Pol Corvez <<https://www.bortzmeyer.org/mots-nes-mer.html>> à propos du vocabulaire de la marine.
- S'agissant de l'anglais, une difficulté supplémentaire est que bien des mots anglais viennent du français, ont franchi la Manche avec Guillaume le Conquérant et nous reviennent aujourd'hui à travers l'Atlantique. C'est le cas de termes comme "*challenge*" comme le note Henriette Walter <<https://www.bortzmeyer.org/honni-soit-qui-mal-y-pense.html>> dans son excellent livre.
- Je ne privilégie pas forcément les traductions officielles (celles qui sont publiées au Journal Officiel et qui doivent obligatoirement être employées par l'administration française, dont ce blog ne dépend heureusement pas). La plupart du temps, elles sont élaborées en comité fermé, sans tenir aucun compte des usages existants, et sans même comprendre le concept traduit. Ceux et celles qui cherchent ces termes officiels peuvent consulter FranceTerme <<http://www.culture.fr/franceterme/>>. Ceux qui préfèrent suivre l'usage peuvent suivre Wikipédia mais cette encyclopédie utilise en général davantage de termes français que l'usage courant, surtout dans le langage parlé. (Il y a aussi le Bitoduc <<http://bitoduc.fr/>>.)

Alors, embarquons-nous maintenant pour un pot-pourri de termes divers que j'ai eu besoin de citer dans mes articles. D'abord, un des plus difficiles, "*socket*". Il y a eu des tentatives humoristiques de traduction (chaussette...) mais puisqu'une "*socket*" sert à se brancher sur le réseau, je préfère dire **prise** (c'est l'équivalent d'une prise électrique, mais pour le logiciel). Je n'ai jamais vu personne utiliser le « connecteur logiciel » officiel.

Et le très fréquent "*buffer*" ? Là aussi, j'ai vu des traductions amusantes qui essayaient de coller au mot anglais comme buffet (c'est vrai qu'on stocke des choses dans un buffet...) mais je préfère **tampon** (qui est aussi le terme officiel). Au lieu de mettre les données dans le buffet, je les tamponne.

Et le "*cache*" ? Je ne le traduis pas car je n'ai pas trouvé de bon terme (le terme officiel, antémémoire, est bien trop restrictif, le concept de cache en informatique est bien plus riche que celui d'antémémoire puisqu'on trouve par exemple des serveurs cache pour le Web). Mais comme le mot cache en français a un autre sens, cela peut faire des phrases très bizarres, comme « pour mieux accéder à la ressource, on la cache » ou « on la met dans une cache à proximité »...

Et le très répandu (le mot et la chose...) "*bug*" ? J'utilise **bogue**, le terme officiel, qui évoque les pi-quants de la châtaigne, et qui permet de les chasser avec un débogueur. Mais j'avais un professeur d'informatique qui appelait le défaut une bourde et gdb un débourdeur, termes qui fleuriraient bon le terroir.

Terme plus rare mais qui était aussi difficile à traduire, le "*nonce*" (au fait, l'article du Wikipédia anglophone le réduit à un usage dans la cryptographie alors qu'il est courant dans les protocoles réseau, même quand il n'y a pas de cryptographie du tout). Utiliser le terme anglais entraîne des risques de confusion avec l'ambassadeur du pape. Il n'y a pas de traduction officielle. Cela a été un long débat, documenté dans un autre article <<https://www.bortzmeyer.org/nonce.html>>, avant que je ne choisisse **numnique**.

Passons maintenant au "*streaming*". La quantité énorme d'octets qui sont poussés chaque jour vers les temps de cerveau disponibles, grâce à cette technique, nécessite un terme. C'est la diffusion en flux pour le Journal Officiel, terme qui me semble bien lourd. Ou bien parler de « flux temps réel » ? Comme le terme anglais est construit pour évoquer la rivière, on pourrait parler de rivièrage ? De ruissellement (« Tu regardes YouTube sur ta Freebox en ruissellement ») ? Ou, toujours en filant la métaphore aquatique, de diffusion en flot ? en courant ? Je ne me suis pas vraiment décidé.

Le nom "*stream*" pose d'ailleurs les mêmes problèmes. La norme HTTP 2 (RFC 9113¹) utilise beaucoup "*stream*", un concept très important pour ce protocole. Je l'ai traduit par « ruisseau ».

Et le "*flow*", concept très utilisé dans les réseaux (rien à voir avec le "*streaming*", traité plus haut), par exemple dans les protocoles comme IPFIX (RFC 3917) ? J'utilise en général **flot** dans le contexte de l'analyse du trafic réseau (cas d'IPFIX). Mais flux pourrait également convenir. Sauf que je l'utilise pour le terme suivant.

Toujours dans les choses qui coulent d'un point à un autre, les "*feeds*" de syndication. Je dis **flux**, qui est également le terme officiel, et apparemment le plus utilisé en pratique.

1. Pour voir le RFC de numéro NNN, <https://www.ietf.org/rfc/rfcNNN.txt>, par exemple <https://www.ietf.org/rfc/rfc9113.txt>

Passons au monde du routage maintenant. L'anglais a deux termes, "*routing*", qui désigne la construction des tables de routage par les routeurs (on dit aussi "*control*"), et "*forwarding*", qui est le passage d'un paquet d'un routeur à un autre, en suivant les tables de routage. En français, "*routing*" se traduit bien par routage mais "*forwarding*" n'a pas de traduction équivalente. Je propose simplement **transmission**.

Une des raisons pour lesquelles l'usage abusif de l'anglais me hérise, c'est que c'est souvent le résultat d'une incompréhension. La proximité même des deux langues fait que les faux amis sont nombreux, et on voit souvent des gens utiliser un terme anglais (pourquoi pas) à contre-sens (c'est déjà plus embêtant). Un exemple fameux est le logiciel "*malveillant*" stupidement traduit par « logiciel malicieux », ce qui n'a rien à voir. Un autre exemple est "*regular expression*" traduit par « expression régulière » (elles n'ont rien de régulier, ces expressions rationnelles). Attention donc à la paresse, elle peut cacher une absence de compréhension du concept.

Passons maintenant à un autre terme difficile : hash (et "*hashing*"), à savoir le résultat d'une fonction qui va produire une chaîne de bits, souvent de longueur constante mais surtout de taille réduite, donc plus facile à manipuler (notamment à signer), à partir d'un texte de longueur quelconque. Pour une fois, la traduction officielle et l'usage coïncident pour parler de hachage (pour traduire "*hashing*" ; je ne trouve pas de traduction officielle pour le "*hash*", le résultat du hachage). Je n'aime guère ce terme qui donne l'impression que le résultat est une purée quelconque, ce qui fait perdre la principale caractéristique des fonctions de « hachage », leur reproductibilité. Je préfère donc parler de **condensation**, ce qui a l'avantage de donner des termes dérivés comme **condensat** pour "*hash*". À noter qu'il s'agit de condensation au sens de réduction, pas au sens de passage à l'état solide (passage qui, lui, est réversible)... (Je ne suis pas toujours cohérent dans mes articles : j'ai parfois utilisé le terme « résumé » à la place de « condensat ».)

Maintenant, prenons un terme plus rare et bien moins utilisé, "*scheme*", au sens qu'il a dans les URI (RFC 3986, section 3.1). Par exemple, dans `<http://prsums-terroristes.fr/>`, le plan est `http:`. L'usage est souvent de le traduire par schéma mais je préfère personnellement **plan** (en français, un schéma est plus approximatif qu'un plan, ce second terme convient donc mieux au caractère très formalisé des plans d'URI). En tout cas, ce n'est pas un « protocole » et il ne faut pas utiliser ce terme qui montre qu'on n'a rien compris aux URI (un URI est un identificateur, pas un localisateur, il n'y a pas forcément de protocole utilisable dedans, même quand le plan est `http:`).

Bon, sauf si on travaille au W3C, on n'utilise pas ce terme de « plan » tous les jours. Mais il y a aussi des termes qui sont très importants en informatique, des concepts essentiels, mais pour lesquels la traduction frustre toujours les terminologues les plus avertis. C'est le cas de la "*scalability*", la capacité d'un système à garder ses propriétés essentielles lorsqu'on change sérieusement d'échelle (par exemple quand on passe de mille à un million d'utilisateurs). Le terme revient tout le temps dans les discussions à l'IETF (« "*The scalability of ARP is very poor since every machine on the link must process every ARP request*" ») et le traduire est difficile. La traduction officielle est « extensibilité », ce qui est à la limite du contre-sens. L'extensibilité est une autre propriété des systèmes, la capacité à recevoir de nouvelles fonctions. Ici, il ne s'agit pas d'étendre le protocole, par exemple en lui ajoutant des fonctions, mais d'étendre le nombre d'utilisateurs. Je parle donc de **passage à l'échelle** en sachant bien que c'est nettement moins percutant que le terme anglais.

Passons ensuite à la génération de textes (par exemple une page HTML) à partir d'un "*template*". Mais comment on dit "*template*" en français? La traduction officielle est « modèle », ce qui me semble curieux (un modèle est plutôt conçu pour être copié tel quel, alors que l'expansion d'un "*template*" va donner des résultats différents selon la valeur des variables utilisées). Wikipédia ne traduit pas, mais propose éventuellement « patron », ce qui est joli (ce terme vient de la couture) mais ne me semble pas meilleur (un patron est également appliqué tel quel, il n'y a pas l'idée qui est essentielle avec les

"*templates*", d'incarnation d'une entité spécifique, en combinant "*template*" et variables). Je préfère donc de loin le terme de **gabarit** au sens où il est utilisé par les maquettistes.

Bon, et puisque j'écris ce paragraphe au moment des manifestations de Hong-Kong, c'est le moment de traduire le terme de réseau "*mesh*". Il n'y a, semble-t-il, pas de traduction officielle (peut-être que l'État français n'aime pas cette idée de citoyens se connectant directement entre eux?) Réseau maillé semble le plus fréquent mais c'est un peu idiot, tous les réseaux sont maillés. Cette traduction ne rend pas compte de la propriété importante des réseaux "*mesh*", leur auto-organisation, qui permet la résistance à la censure <http://www.lemonde.fr/pixels/article/2014/05/23/thailande-taiwan-les-reseaux-mesh-4420312_4408996.html>.

Et le "*log*" des connexions, comment l'appeler? Je ne trouve pas de traduction officielle mais Wikipédia propose de parler d'historique, ce qui est raisonnable. Pour une fois, je préfère une traduction littérale : le "*log*" en anglais vient de la marine (pour parler du journal de bord) et je propose donc de parler tout simplement de **journal** (et, donc, "*loguer*" une information est l'écrire dans le journal ou la journaliser, quoique ce dernier terme entre en collision avec certains systèmes de fichiers).

Enfin, certains termes sont problématiques, non pas parce qu'ils viennent de l'anglais, mais tout simplement parce qu'ils sont techniquement incorrects. C'est ainsi que je ne parle pas de propagation DNS <<https://www.bortzmeyer.org/dns-propagation.html>>, ni de cryptage <<https://www.bortzmeyer.org/cryptage-n-existe-pas.html>>.

Mille mercis à tous ceux et toutes celles qui ont discuté avec moi de terminologie, notamment sur Twitter (pardon, Gazouilleur), à chaque fois que je demande « et vous, comment vous traduiriez ce terme? ».

Et il y a plein d'autres termes pas évidents à traduire et sur lesquels je vous laisse réfléchir (merci à Bertrand Petit pour les suggestions) : "*closure*" (le terme venant des maths, il faut probablement chercher une traduction dans ce secteur), "*mixin*", "*bytecode*", "*tail-recursivity*" (récursivité terminale semble répandu?), "*mapping*" (je suggère « correspondance »), "*garbage collector*"...