
Introduction

Les réseaux en sont à leur troisième révolution. La première a été le passage du mode circuit au mode paquet et la deuxième du mode terrestre au mode hertzien. La troisième révolution que nous décrivons dans ce livre concerne le passage du mode matériel au mode logiciel. Regardons rapidement ces trois révolutions avant de s'attacher à la troisième, qui sera étudiée en détail dans ce livre.

1.1. Les deux premières révolutions

Le circuit correspond à un ensemble de matériels et de logiciels alloués à deux utilisateurs, un à chaque extrémité du circuit. Les ressources du circuit n'appartiennent qu'aux deux utilisateurs, personne d'autre ne peut les utiliser. Ce mode a été particulièrement utilisé dans le cadre du réseau téléphonique commuté. En effet, la parole téléphonique est une application continue qui s'adapte très bien au circuit.

Une évolution forte a entraîné la première grande révolution du monde des réseaux. Elle concerne les applications asynchrones et irrégulières. Les données transportées pour ces applications ne remplissent que très mal les circuits mais s'adaptent bien au mode paquet. Lorsqu'un message doit être transmis d'un émetteur à un récepteur, les données à envoyer sont mises dans un ou plusieurs paquets en fonction de la taille totale du message. Pour un petit message, un paquet est suffisant ; en revanche, pour un long message, plusieurs paquets sont nécessaires. Les paquets passent ensuite par des nœuds de transfert intermédiaires entre l'émetteur et le récepteur permettant aux paquets de transiter jusqu'au destinataire. Les ressources nécessaires, pour prendre en charge les paquets, sont composées de mémoires, de liaisons entre les nœuds et d'émetteurs-récepteurs. Ces ressources sont partagées entre tous les utilisateurs. Le mode paquet demande une architecture et des protocoles, c'est-à-dire des règles, pour réaliser la communication d'une extrémité à l'autre. Beaucoup de propositions

d'architectures ont été faites en utilisant des couches protocolaires et des algorithmes associés. Au départ chaque équipementier avait sa propre architecture (SNA, DNA, DecNet, etc.). Puis le modèle OSI (*Open System Interconnection*) a essayé de rendre compatibles toutes ces architectures. L'échec de la compatibilité entre équipementiers, même avec un modèle commun, a poussé à reprendre une des toutes premières architectures introduites pour le mode paquet : TCP/IP (*Transport Control Protocol/Internet Protocol*).

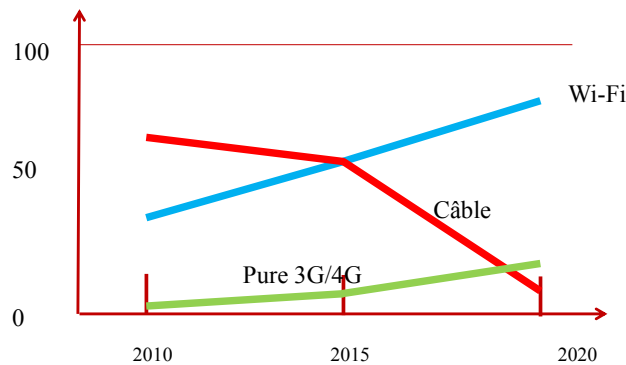


Figure 1.1. La connexion terminale en 2020

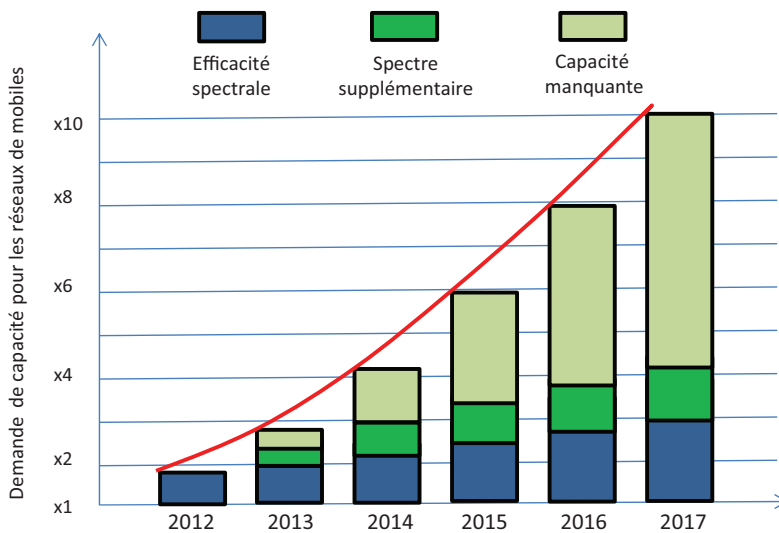


Figure 1.2. Le gap entre les progrès technologiques et la demande des utilisateurs

La deuxième révolution est le passage du mode terrestre à un mode hertzien. A la figure 1.1, on peut voir qu'en 2020, la connexion terminale sera essentiellement hertzienne et réalisée soit avec la technologie Wi-Fi soit la technologie 3G/4G/5G. En fait, les deux techniques s'interpénètrent de plus en plus, car elles sont plus complémentaires que concurrentes. De plus, lorsque l'on regarde la courbe, présentée à la figure 1.2, de la demande des utilisateurs mondiaux et la croissance de ce que peut apporter la technologie 3G/4G/5G le gap est tellement important que seule la technologie Wi-Fi est capable de le prendre en charge. Nous reviendrons sur les architectures hertziennes, car la dernière révolution a également un impact important sur cette transition vers les technologies radio.

1.2. La troisième révolution

La troisième révolution qui nous intéresse dans ce livre concerne le passage du mode matériel vers le mode logiciel. Cette transition s'effectue par la virtualisation qui permet de transformer les équipements de réseau matériels en équipements de réseau logiciels.

Regardons les différents éléments qui militent pour une nouvelle génération de réseau. En premier lieu on peut citer le Cloud. Le Cloud est un ensemble de ressources que l'on déplace de l'entreprise ou du particulier vers l'Internet. On délocalise les ressources pour les rassembler dans des centres de ressources que l'on appelle des centres de données ou datacenters.

Les raisons de la naissance du Cloud proviennent de la mauvaise utilisation des serveurs dans le monde : dix pour cent seulement. Cette faible valeur provient de l'utilisation quasiment nulle des serveurs durant la nuit et du peu d'utilisation en dehors des heures de pointe qui ne représentent pas plus de 4 à 5 heures par jour. De plus, le coût relativement bas du matériel implique en général des serveurs fortement surdimensionnés. Un autre critère à prendre en compte sérieusement provient du coût croissant du personnel de gestion et de contrôle des ressources. Pour optimiser à la fois le coût des ressources et celui du personnel spécialisé, il faut les partager. Les Clouds sont là pour réaliser ce partage de façon efficace.

La figure 1.3 montre l'augmentation du marché du Cloud public. Elle est importante certes, mais finalement relativement faible par rapport à ce que cela aurait pu être s'il n'y avait pas de problème de sécurité. En effet, la sécurité des données étant assez mal assurée, les Clouds privés se sont démultipliés en prenant la place des Cloud publics. Nous examinerons au chapitre 7, les avancées en matière de sécurité avec l'apparition des Clouds de sécurité.

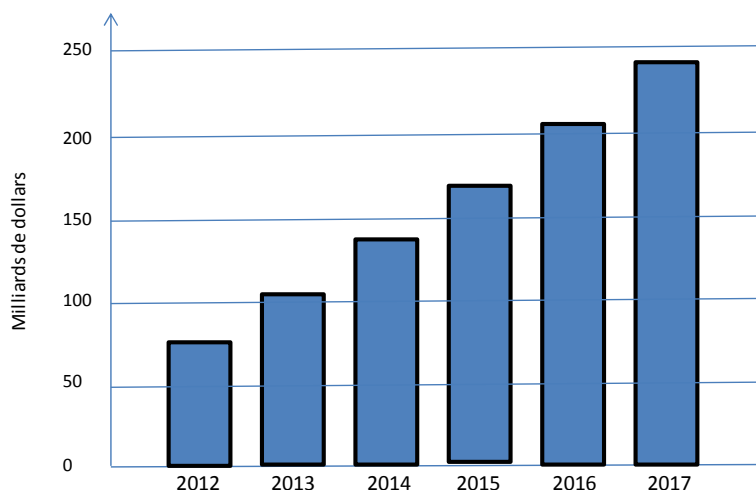


Figure 1.3. Marché du Cloud public et augmentation annuelle

La virtualisation est également un facteur-clé comme nous l'avons déjà indiqué au début de ce chapitre. L'augmentation du nombre des machines virtuelles est indéniable et en 2015 plus des deux tiers des serveurs disponibles dans le monde sont des machines virtuelles. Les machines physiques supportent de plus en plus de machines virtuelles et cette augmentation est indiquée à la figure 1.4 avec en 2015 approximativement huit machines virtuelles par serveur physique.

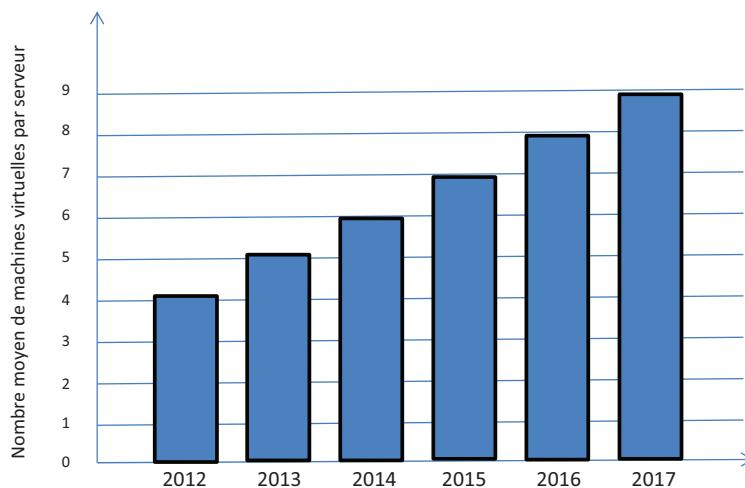


Figure 1.4. Nombre de machines virtuelles par serveur physique

L'utilisation du Cloud implique une forte augmentation des débits des réseaux. En effet, les traitements sont maintenant réalisés en central et les données ainsi que la signalisation doivent aller vers le Cloud et retourner après traitement. On peut voir ce besoin de débit en examinant le marché des cartes Ethernet dans les accès aux datacenters. La figure 1.5 montre la vente de cartes Ethernet à 1 Gbit/s en comparaison des cartes 10 Gbit/s. On remarque que les cartes Ethernet à 1 Gbit/s, pourtant déjà assez rapides, sont remplacées par des cartes dix fois plus puissantes.

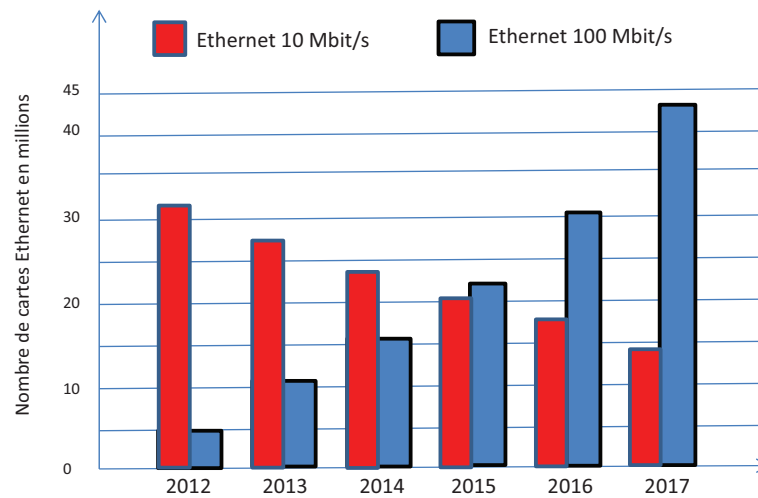


Figure 1.5. La montée en puissance des accès

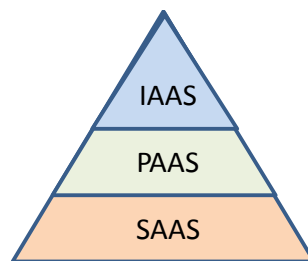


Figure 1.6. Les trois principaux types de Cloud

Le monde du Cloud est en fait assez diversifié en regardant le nombre de fonctions qu'il peut mettre en œuvre. De nombreux types de Clouds sont répertoriés mais trois catégories, indiquées à la figure 1.6, suffisent pour bien les différencier. La catégorie

offrant le plus de potentiel concerne le Cloud SaaS (*Software as a Service*). Celui-ci offre tous les services disponibles à l'utilisateur que ce soit du calcul, du stockage ou du réseau. Dans cette solution l'entreprise demande à son fournisseur de Cloud de lui fournir l'ensemble des applications dont elle a besoin. En fait, l'entreprise soustrait son système d'information au fournisseur de Cloud. La deuxième solution ou PaaS (*Platform as a Service*) laisse les applications à l'initiative de la société. Le fournisseur de Cloud propose une plate-forme complète et ne laisse que la gestion des applications à l'entreprise. Enfin, la troisième solution IaaS (*Infrastructure as a Service*) laisse nettement plus d'initiatives à la société. Le fournisseur offre toujours le calcul, le stockage et le réseau mais laisse à l'entreprise les applications et les environnements nécessaires aux applications comme les systèmes d'exploitation et les bases de données.

D'une façon un peu plus précise, nous pouvons définir les trois architectures de Cloud dans les quelques lignes suivantes.

IaaS (*Infrastructure as a Service*) : toute première approche, avec une partie de la virtualisation prise en charge par le Cloud, telle que serveurs réseau, serveurs de stockage, réseau lui-même. On déporte dans le réseau Internet les machines de type PABX, pare-feu ou serveurs de stockage, et plus généralement les serveurs liés à l'infrastructure réseau.

PaaS (*Platform as a Service*) : second modèle de Cloud, avec, en plus de l'infrastructure, le logiciel intermédiaire correspondant à la plate-forme Internet. Les serveurs de l'entreprise ne prennent en compte que les applications.

SaaS (*Software as a Service*) : permet au fournisseur de Cloud de proposer, en plus de l'infrastructure et de la plate-forme, les applications elles-mêmes. Globalement, il ne reste plus rien dans l'entreprise, si ce n'est des terminaux d'accès à Internet. Cette solution, également appelée Cloud computing, extériorise quasiment l'ensemble de l'informatique et des réseaux de l'entreprise.

La figure 1.7 montre les fonctions des différents types de Cloud en comparaison du modèle classique d'aujourd'hui.

La peur principale du Cloud auprès des entreprises provient de la sécurité. Effectivement rien n'empêche le fournisseur d'aller jeter un coup d'œil aux données ou encore beaucoup plus classiquement que les données soient réquisitionnées par les Etats où elles sont placées ; les fournisseurs doivent s'exécuter. La montée des Clouds souverains est également à noter : les données ne doivent pas sortir des frontières. La plupart des Etats pour leurs propres données exigent cette propriété.

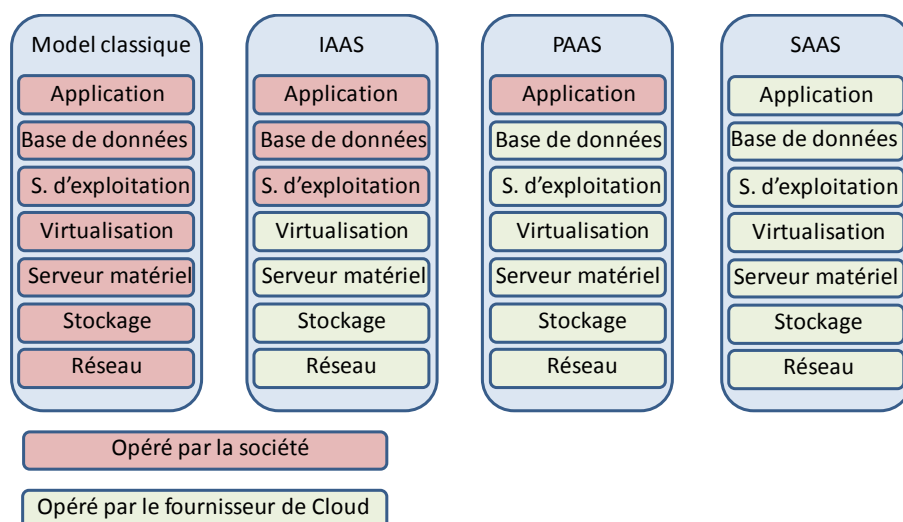


Figure 1.7. Les différents types de Clouds

L'avantage du Cloud provient de la puissance des datacenters qui permettent de prendre en charge beaucoup de machines virtuelles et de leur donner la puissance nécessaire à leur exécution. Le multiplexage entre un grand nombre d'utilisateurs abaisse les coûts de façon forte. Les datacenters vont également être à l'origine des réseaux logiciels et le support des machines virtuelles pour les créer. C'est la raison pour laquelle de nombreux opérateurs de télécommunications ont développé des sociétés qui fournissent des services Cloud pour eux-mêmes et également pour leurs clients.

Dans les techniques que l'on examinera en détail dans la suite, on retrouvera le SDN (*Software-Defined Networking*) qui permet de définir de multiples tables de transfert et seule la puissance des datacenters sera capable de faire tous les calculs nécessaires. Une des problématiques est de déterminer la taille des datacenters et leur positionnement. Très grossièrement l'échelle va de très gros datacenters avec un million de serveurs jusqu'à des femto-datacenters avec l'équivalent de quelques serveurs en passant par toutes les tailles intermédiaires.

1.3. La cloudification des réseaux

L'arrivée de cette nouvelle génération de réseaux fondés sur les datacenters est également due à la consommation d'énergie du monde des TIC. Cette consommation est estimée entre 3 % et 5 % de l'empreinte carbone en fonction des études. Mais cette

fraction s'accroît extrêmement rapidement avec le déploiement rapide des datacenters et des antennes pour les réseaux de mobiles. Pour donner des exemples, un datacenter d'un million de serveurs consomme approximativement 100 MW. Un fournisseur de Cloud qui aurait une dizaine de tels datacenters consommerait 1 GW, c'est-à-dire l'équivalent d'une tranche de centrale nucléaire. Ce nombre total de serveurs est déjà atteint ou dépassé par une dizaine de grandes sociétés bien connues. De même le nombre d'antennes 2G/3G/4G dépasse les 10 millions dans le monde. La consommation étant en moyenne de 1 500 W par antenne (2 000 W pour les antennes 3G/4G mais nettement moins pour les antennes 2G), cela représente 15 GW dans le monde.

En continuant sur notre lancée, l'empreinte carbone de la consommation énergétique devrait arriver à 20 % en 2025 pour le monde des TIC. Il est de ce fait essentiel de trouver des solutions pour amortir cette montée. Nous y reviendrons au chapitre 10 de ce livre mais des solutions existent et commencent à être employées. La virtualisation représente une bonne solution en rassemblant les machines virtuelles sur des machines physiques communes et en mettant en veille un grand nombre de serveurs. Les processeurs devraient également avoir la possibilité de descendre à de très basses vitesses dès que nécessaire. En effet, la consommation électrique est fortement proportionnelle à la vitesse du processeur. Lorsque le processeur n'a rien à faire, il s'arrête presque et il accélère en fonction du travail à effectuer.

La mobilité représente également un autre argument pour aller vers une nouvelle architecture de réseau. On montrera qu'en 2020, 95 % des machines terminales seront connectées sur la partie terminale par une solution hertzienne. Il faut de ce fait gérer les problèmes de mobilité et donc en premier lieu la gestion du *multi-homing*, c'est-à-dire être capable de se connecter à plusieurs réseaux simultanément. Le mot *multi-homing* vient du fait que le terminal reçoit plusieurs adresses IP provenant des différents réseaux connectés. La gestion de ces multi-adresses est complexe et demande des caractéristiques particulières. La mobilité doit également gérer la connexion simultanée à plusieurs réseaux. Les paquets en fonction de critères à déterminer peuvent se séparer pour passer par des réseaux différents. Ils doivent donc être remis en ordre à l'arrivée ce qui peut poser de nombreux problèmes. La mobilité pose également des problèmes d'adressage et d'identité. Si l'on utilise l'adresse IP, elle peut être interprétée de deux manières : c'est l'identification de l'utilisateur qui permet de déterminer qui est l'utilisateur mais c'est également une adresse qui permet la localisation de l'utilisateur. La difficulté est de gérer ces deux notions simultanément. Pour cela lorsqu'un client se déplace suffisamment pour sortir du sous-réseau dans lequel il est répertorié, il faut lui changer son adresse IP. Et ceci est complexe du point de vue identité. On verra qu'une des solutions est de donner deux adresses IP à un même utilisateur : une adresse d'identification et une adresse de localisation.

Une autre révolution en cours concerne l'Internet des choses (*Internet of Things*) : des milliards de choses seront connectées dans quelques années. On parle de 50 milliards en 2020. En d'autres termes, le nombre de connexions devrait être multiplié par dix en quelques années. Les choses proviennent de différents horizons :

- 1) le domicile avec les matériels électroménagers, la médecine à domicile, la gestion du domicile, etc. ;
- 2) la médecine avec des capteurs de toutes sortes sur le corps et dans le corps pour mesurer, analyser et réaliser des actions ;
- 3) les entreprises avec des capteurs de luminosité, de température, de sécurité, etc.

De nombreux problèmes se posent dans ce nouvel univers comme la gestion de l'identification et la sécurité des communications avec les capteurs. L'identification est souvent évaluée à 40 \$ par objet, ce qui est absolument incompatible avec le coût d'un capteur souvent inférieur à 1 \$. La sécurité est également un facteur complexe puisque le capteur a très peu de puissance et ne peut pas faire de chiffrement suffisamment sophistiqué pour assurer la confidentialité des transmissions.

Enfin, une dernière raison explique la migration vers un nouveau réseau : la sécurité. La sécurité requiert une vision et une compréhension précise des problèmes qui vont de la sécurité physique jusqu'à la sécurité logique en passant par des attaques parfois totalement imprévisibles. Le monde de l'Internet est aujourd'hui comme un pneumatique qui n'a plus que des rustines et à chaque nouvelle attaque réussie une nouvelle rustine est ajoutée. Cet ensemble tient encore bien la route aujourd'hui mais l'ensemble risque d'éclater si rien de nouveau n'est envisagé dans les années qui viennent. Nous regarderons à la fin de ce livre, dans le chapitre 8, le Cloud de sécurité qui permet de rassembler dans un datacenter tout un ensemble de solutions autour de machines virtuelles spécialisées pour apporter des éléments nouveaux ayant pour objectif de sécuriser les applications et les réseaux.

La sécurité pour être très forte doit avoir un élément matériel : un coffre-fort pour y mettre les éléments importants de l'arsenal nécessaire pour assurer la confidentialité, l'authentification, etc. La sécurité logicielle existe et elle peut largement être suffisante pour de nombreuses applications. Cependant, les éléments de sécurité peuvent toujours être cambriolés lorsque l'ensemble des défenses s'effectue en logiciel. Cela implique, pour les nouvelles générations, un élément matériel local ou situé à distance. Cet élément matériel est un microprocesseur sécurisé que l'on appelle un élément sécurité (*secure element*). Un exemple classique de ce type de matériel provient de la carte à puce particulièrement utilisée par les opérateurs de télécommunications et les banques.

En fonction de son emprise sur le monde des entreprises ou du grand public, l'élément sécurisé se trouve dans le terminal, à proximité du terminal ou loin du terminal. Nous examinerons les différentes solutions dans la suite de ce livre.

La virtualisation a également un impact sur la sécurité : la puissance du Cloud, avec des machines virtuelles spécialisées, permet aux attaquants de disposer d'une puissance de frappe remarquable. En quelques années, les hackers ont gagné un facteur de cinq à six pour casser des algorithmes de chiffrement.

Un autre point important qui doit absolument être introduit dans les réseaux concerne l'intelligence. Il y a eu l'époque des réseaux intelligents mais l'intelligence n'était pas vraiment celle à laquelle on pense mais un ensemble de mécanismes automatiques pour répondre à des problèmes parfaitement déterminés à l'avance. L'intelligence est ici liée à des mécanismes d'apprentissage et de décisions intelligentes en fonction de l'état du réseau et des demandes utilisateur. Le réseau doit devenir un système intelligent capable de prendre des décisions de façon autonome. Une solution pour aller dans ce sens a été introduite par IBM au début des années 2000, il s'agit de l'autonomic. Autonomic signifie autonome et spontané. Autonome dans le sens où chaque équipement du réseau doit être capable de prendre tout seul des décisions avec une connaissance de sa vue située, c'est-à-dire des nœuds autour de lui à moins d'un certain nombre de sauts. Les solutions apportées pour accroître l'intelligence des réseaux sont influencées par la Cloud. Nous les détaillerons dans le chapitre 4 sur *les smart edge*.

Enfin, un dernier point qui pourrait être vu comme la quatrième révolution concerne la concrétisation, c'est-à-dire l'inverse de la virtualisation. En effet, le problème apporté par la virtualisation est une forte baisse des performances provenant du remplacement du matériel par du logiciel. Plusieurs solutions se dessinent pour regagner en performance : les accélérateurs logiciels et surtout le remplacement du logiciel par du matériel par une phase de concrétisation. On remplace le logiciel par du matériel reconfigurable qui se transforme en fonction du logiciel à traiter. Cette phase devrait amener aux réseaux morphiques (*morphware networks*). Cette génération sera décrite au chapitre 9.

1.4. Conclusion

En conclusion, le monde des réseaux change fortement pour les raisons évoquées ci-dessus. Il change plus vite qu'on aurait pu le supposer il y a quelques années. Une proposition de départ a été faite mais elle a échoué : repartir de zéro. C'est la *Clean Slate Approach* : on efface tout pour redémarrer de rien. Malheureusement, aucune proposition concrète n'a pu s'imposer et le transfert de paquets IP continue d'être la solution pour le transport des données. Mais dans les nombreuses propositions la virtualisation et le Cloud forment les deux axes principaux qui sont aujourd'hui fortement exploités et qui sont à la base de ce livre.