

Séminaire « Soutenabilités »

Contribution - Covid-19 : pour un « après » soutenable

Nom : Lagrange

Prénoms : Mathieu

Institution ou entreprise : Institut de Recherche Technologique b<>com

Axe(s) :

- Numérique : nouveaux usages, nouvelles interrogations

Intitulé de votre contribution : Pour le développement d'infrastructures numériques parcimonieuses, respectueuses des principes de sécurité, d'égalité et de souveraineté

Résumé de votre contribution :

Après une analyse des évolutions sociétales, dont certaines ont été accélérées par la crise du COVID19, ainsi que des usages adoptés par les nouvelles générations, le document s'attache à proposer des axes d'évolution des infrastructures numériques et en particulier des infrastructures réseaux, permettant de soutenir un développement de notre société à la fois harmonieux (pour l'ensemble de la population), parcimonieux (dans un souci d'efficacité), respectueux des principes de sécurité numérique et soucieux de notre souveraineté.

Pour le développement d'infrastructures numériques parcimonieuses, respectueuses des principes de sécurité, d'égalité et de souveraineté

Comment Réduire l'empreinte environnementale du numérique ?

Un impératif sociétal

Au-delà de la 25^e conférence des parties définissant les objectifs globaux de réduction d'émission des gaz à effets de serre afin de pouvoir contenir les impacts liés au réchauffement climatique, au-delà de la conférence mondiale sur la biodiversité d'Aichi en 2010 définissant le plan stratégique pour la biodiversité 2011-2020, ce sont plusieurs urgences associées à des pans entiers d'impératifs de nature environnementale que la crise du COVID-19 a mis en exergue. Ainsi, l'occurrence d'un risque pandémique identifié depuis de nombreuses années aura permis une prise de conscience très forte de l'urgence à agir. Cette prise de conscience est d'autant plus forte qu'elle est accompagnée d'une perception très concrète des impacts qui pourraient résulter de futures crises présentant une nature systémique incluant la fragilisation de systèmes internationaux de gouvernance. La différence avec la crise du COVID-19 étant que les risques tels que ceux associés au changement climatique sont associés à une probabilité d'occurrence toujours plus élevée¹ si rien n'est mis en œuvre qui soit à la hauteur des enjeux et de la nature des crises, que ces dernières soient environnementales, sanitaires, économiques, sociales ou encore sociétales.

Les problèmes qui doivent faire l'objet du développement de solutions sont d'un niveau de complexité extrêmement élevé, à l'image de ce que la crise COVID-19 a aussi mis en exergue c'est-à-dire le degré d'interconnexion et d'interdépendance entre environnement, santé, social, économique, sociétal. Ainsi, la notion de système contraint, de par la limite des ressources disponibles, est-elle aussi apparue évidente. Ainsi, les travaux réalisés par Kate Raworth concernant la théorie du doughnut et publiés en 2017, ont un écho particulier. Ainsi, les notions de souveraineté, de résilience, d'efficacité prennent là encore un sens particulier, celui de l'action.

Le numérique, qui est lui aussi une caractéristique de la globalisation du monde, a considérablement transformé de nombreux secteurs économiques, permettant de développer une efficacité accrue de ces secteurs à travers ubiquité et immédiateté de toute forme d'échange. Cette crise montre à quel point le numérique devient désormais un élément majeur pour adresser les enjeux de souveraineté, de résilience et d'efficacité qui se posent. Enfin, le paysage dressé ne serait pas complet sans aborder la notion de sobriété numérique comme élément majeur d'efficacité, sur le plan comportemental et celui des usages au-delà de toute notion de frugalité énergétique.

¹ World Economic Forum, éditions annuelles du suivi des risques globaux

Enfin, il est nécessaire de repenser la manière dont l'innovation technologique doit être conduite par rapport à ces éléments et dresser les contours de ce que devra être une innovation responsable face à ces enjeux et leur appropriation par les parties prenantes. Une des difficultés qui devra être surmontée est liée à la caractéristique globale des technologies et usages du numérique. Il est très probable que jamais l'enjeu de pouvoir disposer de technologies interopérables et par conséquent standardisées n'ait été aussi fort. En complément du sujet lié aux standards, la question de la souveraineté doit poser la question non seulement du développement de technologies en open-source mais aussi celle de leur exploitation sur le plan économique.

La suite de ce document ne porte pas d'ambition d'exhaustivité. Beaucoup de sujets auraient pu être évoqués. L'Institut de Recherche Technologique b<>com se tient à disposition et sera ravi de contribuer à l'élaboration de tout futur plan concernant les ambitions nationales sur les technologies et usages du numérique.

Un terrain favorable

Les Millennials

La génération des 18-35 ans, souvent appelée « Millennials », se situe à l'avant-garde de la révolution numérique et montre une vision de la vie, des goûts, certaines habitudes, différentes de celles des générations précédentes, qui influent sur sa façon de consommer.

Ainsi, ayant grandi dans un monde de plus en plus contraint qu'il s'agisse d'économie comme d'environnement, les Millennials aiment le bio, le local, l'environnement. Ils sont moins attachés aux grandes marques et surtout préfèrent l'usage à la possession.

Les Millennials, qui représentent aujourd'hui un tiers de la population active mondiale, sont connectés et deviennent même prescripteurs, en multipliant les recommandations à l'aide des outils numériques modernes, étendant ainsi leur capacité d'influence. Ceci crée un terrain favorable à l'émergence de nouveaux services de type collaboratifs focalisés sur l'usage et l'échange.

L'économie collaborative

L'économie collaborative est une économie de pair à pair qui repose sur le partage ou l'échange entre particuliers de biens (voiture, logement, parking, perceuse, etc.), de services (covoiturage, bricolage, etc.), ou de connaissances (cours d'informatique, communautés d'apprentissage, etc.), avec échange monétaire (vente, location, prestation de service) ou sans échange monétaire (dons, troc, volontariat), par l'intermédiaire d'une plateforme numérique de mise en relation.

Aujourd'hui, ce type d'interaction économique se développe dans un grand nombre de secteurs d'activités, comme :

- Le logement : location entre particuliers, échange d'appartement ou maison pour les vacances, colocation, ...
- Les transports : location de véhicules entre particuliers, covoiturage, livraison collaborative, ...

- Culture, enseignement : cours en ligne, soutien scolaire, ...
- Équipements divers : location, don, troc, revente/achat de vêtement et/ou matériel

On peut noter que d'une manière générale, trois facteurs sous-tendent le développement de l'économie collaborative :

- La disponibilité d'outils numériques performants
- Un environnement économique morose (surtout depuis la crise de 2007/2008)
- La volonté des participants d'adopter un style de vie plus respectueux et raisonné

Il s'agit ici, encore une fois de privilégier l'usage à la possession, de donner une seconde vie à des ressources autrement non utilisées, ou de densifier l'usage d'un bien ou d'une infrastructure (par exemple le covoiturage).

Selon le cabinet PwC, le chiffre d'affaire généré par l'économie collaborative en Europe en 2016 était de 24 milliards d'euros, avec une prédiction à 570 milliards d'euros en 2025 comme indiqué sur la figure ci-dessous :

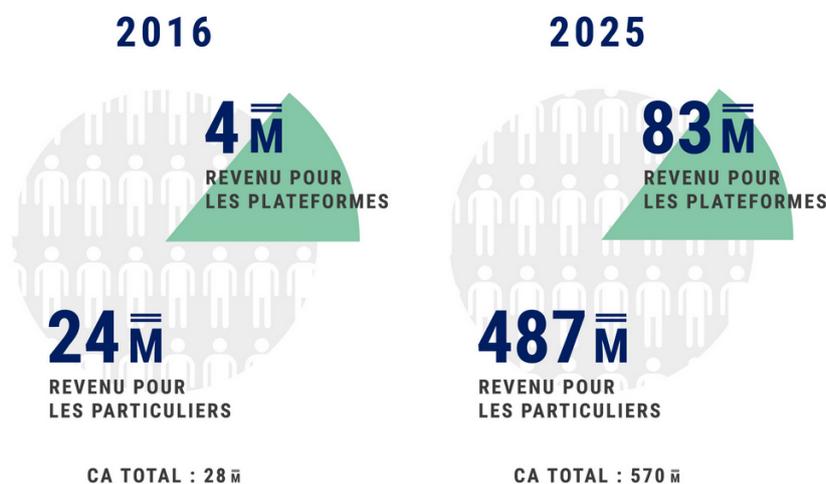


Figure 1: Les chiffres de l'économie collaborative en Europe

Plus significatif, l'ensemble de l'économie européenne enregistre une croissance généralement comprise entre 2% et 3% contre plus de 35 % par an pour l'économie collaborative. Il s'agit d'une tendance de fond représentative de l'évolution des habitudes de consommation des particuliers et des entreprises.

Dans le cadre de la crise du COVID-19, il est intéressant de noter que malgré les risques associés au partage de ressources entre individus, les valeurs essentielles d'entraide ont été préservées et que bon nombre de personnes, d'entreprises, à leur niveau, se mettent au service des communautés, ou de l'intérêt national. Ainsi, il apparaît que la capacité à collaborer, à servir autrui, est largement perçue comme une qualité permettant de surmonter des difficultés que chaque individu ne saurait résoudre sur ses seules compétences ou forces. On découvre que les plus petites tâches, perçues auparavant comme anodines ou sans valeur, sont finalement essentielles. Que l'isolement, ressenti d'autant plus fortement qu'il est imposé, est un point de fragilité, une mise en danger. Il apparaît donc opportun de fonder les solutions technologiques de demain sur ces valeurs, pour en favoriser l'acceptabilité et s'inscrire au plus près des nouveaux usages.

Ces solutions devront permettre de construire des infrastructures répondant aux mêmes valeurs de résilience et de frugalité, au service du bien commun et des usages des générations montantes, capables de créer les conditions nécessaires à l'émergence d'innovations et d'initiatives disruptives mais aussi capables de dégager les conditions financières suffisantes à leur pérennité.

Internet of Everything (IoE) – des challenges

L'IoE relie les personnes, les données, les objets et les processus au sein de réseaux de millions, voire de milliards de connexions. Ces échanges créent de grandes quantités de données qui, lorsqu'elles sont analysées et utilisées intelligemment, permettent de réaliser d'innombrables innovations et gains d'efficacité.

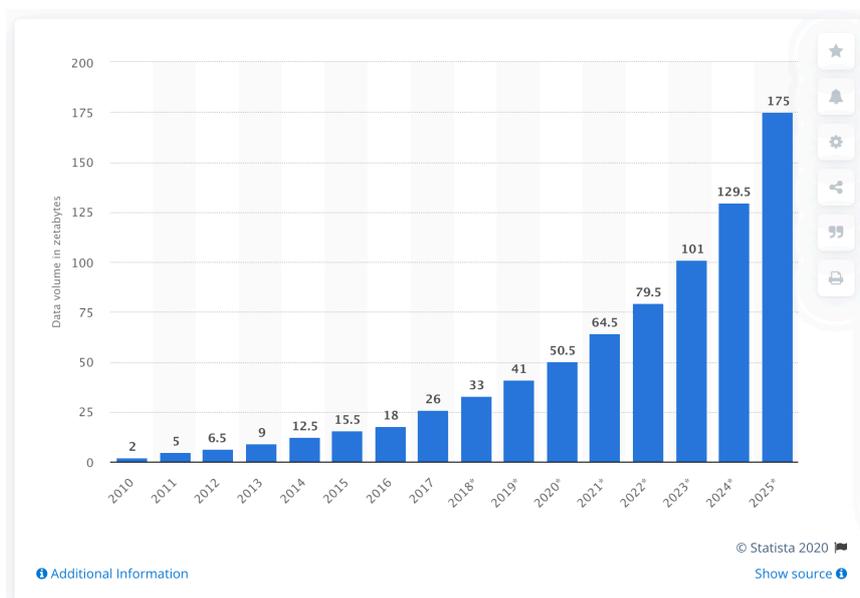


Figure 2: Volume of data/information created worldwide from 2010 to 2025 (in Zetabytes)

Comme souligné sur la figure ci-dessus, la quantité de donnée générée croît exponentiellement. Cela dit, une grande partie de ces données est de nature éphémère, n'est pas destinée à être stockée et entre plutôt en jeu dans le calcul d'autres données, consolidées, plus utiles et exploitables par différents services et applications. Le traitement de cette énorme quantité de données est un défi en soit et devra être localisé au plus près des objets, personnes, systèmes qui les ont générées. Comme on peut d'ailleurs le remarquer sur la figure ci-dessous, ce mouvement vers le traitement des données en temps réel sur des infrastructures de plus en plus distribuées, dites de « Edge » ou « Fog » Computing se reflète dans les investissements réalisés et à venir consentis par les fournisseurs de services numériques.

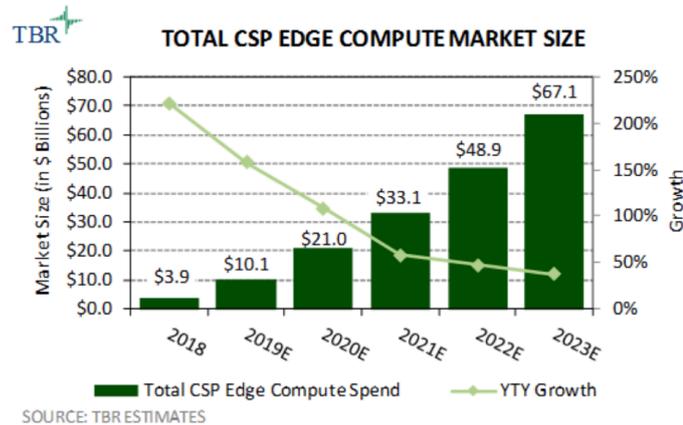


Figure 3: Croissance des investissements en infrastructures de Edge Computing

Il est à noter que cette tendance à l'explosion des données et à la distribution des traitements s'est trouvée temporairement renforcée par les mesures de confinement prises dans le cadre de la crise du COVID-19, mais pourrait également perdurer par l'évolution des mentalités en matière de télétravail, et la prise de conscience que celui-ci, encadré et pratiqué de manière éclairée, peut permettre à la fois de dégager des gains en productivité et des économies pour l'entreprise. En cela, il est probable que les besoins en ressources de calcul distribué vont augmenter significativement dans les années à venir.

Les infrastructures numériques devront donc être à même de permettre, faciliter ces transformations (qui ont aussi des vertus de soutenabilité) de la manière la plus efficace possible, en mettant en œuvre un cycle de vie des équipements et des services qui soit respectueux de l'environnement.

Évolution des infrastructures numériques

Nous traiterons dans ce chapitre des infrastructures numériques au sens large, c'est-à-dire des moyens qui sont mis en œuvre dans le cadre de la production puis du transport jusqu'à l'utilisateur de contenus numériques. Il apparaît d'ailleurs pertinent d'examiner l'évolution de celles-ci en prenant en compte les éléments évoqués dans les sections précédentes. Ainsi, nous assistons à deux phénomènes, déjà présents et sans doute renforcés par la crise actuelle du COVID-19, pouvant être perçus comme antagonistes :

- A l'éveil d'une conscience écologique, en particulier pour les nouvelles générations. Une certaine conscience des enjeux environnementaux qui se déploie à l'échelle planétaire et, de plus en plus, conduit à une évolution des modes de pensée, mais aussi des systèmes de valeurs qui régissent le fonctionnement actuel de nos sociétés. Cette prise de conscience est d'ailleurs renforcée par la crise actuelle liée au COVID-19, ressentie par beaucoup comme une conséquence de la disparition ou la modification des écosystèmes naturels de certaines espèces (dont les chauve-souris par exemple), qui a terme favoriserait, par la proximité forcée de celles-ci avec l'homme ou avec des animaux domestiques, la contamination inter-espèce et ensuite l'émergence de nouvelles maladies.

- Une évolution des usages, notamment professionnels, qui favorise le traitement des données en temps réel en périphérie de réseau, et stimule le besoin en infrastructures distribuées de calcul et de communication, présentant des capacités de résilience et d'adaptation et permettant nativement le maintien des services numériques essentiels en situation de crise.

Les infrastructures réseau

Un angle intéressant est celui des infrastructures cellulaires tant elles sont devenues familières, intégrées dans nos usages personnels et professionnels et tant les investissements à consentir pour leur déploiement sont massifs. En particulier, les réseaux 5G sont des infrastructures fondatrices, sur lesquelles les attentes sont très fortes mais qui présentent un certain nombre d'écueils si le sujet n'est pas abordé rationnellement.

Aujourd'hui, par la volonté des grands opérateurs télécom internationaux, la très vaste majorité des déploiements 5G, sont réalisés selon l'option 3GPP dite « Non-StandAlone » par laquelle les smartphones compatibles 5G se connectant aux fréquences 5G, utilisent toujours préalablement et systématiquement le réseau 4G pour toutes les interactions liées à la gestion des sessions utilisateur et à la sécurité des communications. En cela, l'accès radio LTE et le réseau central (EPC) existants servent d'ancrage à la nouvelle radio 5G (5G-NR) en utilisant la fonction 3GPP dite de double connectivité.

En conséquence, ces déploiements sont donc réalisés selon une logique d'ajout de capacité plus que de mise en œuvre d'une connectivité au service de nouveaux usages. Le modèle de déploiement n'est pas nouveau, il s'agit, en plus des antennes 4G déjà présentes, d'ajouter des antennes 5G supplémentaires sur tout le territoire, le tout se connectant sur le cœur de réseau mobile existant. Les opérateurs de réseaux mobiles s'attachent donc à fournir, comme auparavant, les services mobiles à large bande (eMBB) permettant de consommer des données et des contenus multimédias dans des conditions proches de ce que la 4G a pu proposer jusqu'à présent. Il faut d'ailleurs noter que la fourniture de ce type de services de communication, dont le prix est aujourd'hui fonction de la quantité de données consommées, ne les aidera pas à améliorer leur revenu sur des marchés où la concurrence est déjà très vive et pourrait donc compromettre leur capacité à monétiser leur toute nouvelle infrastructure 5G-NR. Par ailleurs, la nature ayant horreur du vide et la croissance des besoins étant exponentielle, il est certain que ce surcroît de capacité ne saurait que temporairement satisfaire la demande et ne constitue en soi que la poursuite d'une logique de consommation par laquelle les nouvelles technologies de communication, génération après génération, sont proposées au grand public avant même que les usages ne soient clairement identifiés.

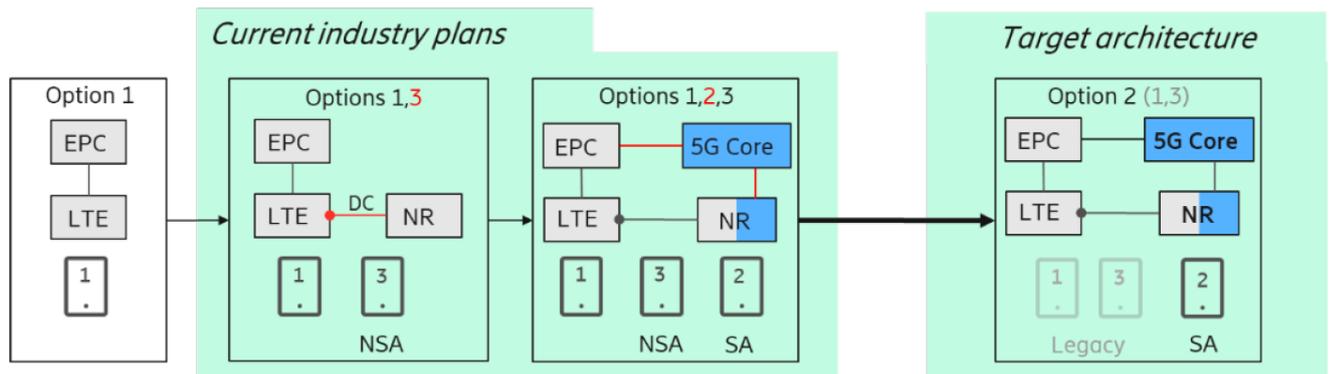


Figure 4: Plan de migration 5G de l'industrie

A l'opposé, dans un scénario de déploiement présentant des bénéfices forts en matière de souplesse, de résilience et de performances, appelé par le 3GPP, l'option "Standalone", la 5G-NR (5G-New Radio) est contrôlée par un cœur de réseau 5G (5G NGC) construit et pensé en fonction des technologies utilisées dans le monde du cloud computing. Ce cœur de réseau s'éloigne résolument de l'approche traditionnelle faisant intervenir des fonctions réseaux monolithiques communiquant entre elles par l'intermédiaire de protocoles complexes, spécifiques et éminemment statiques. En effet, il a été conçu selon une architecture modulaire dont les composants fonctionnels exposent des interfaces asynchrones dites « Service Based », conformes aux API REST des applications Internet.

Cette nouvelle architecture présente par nature de fortes affinités avec les principes dits « Cloud Native » et a aussi été spécifiée en considérant les concepts de virtualisation des fonctions réseau (NFV – Network Function Virtualisation) et de réseau programmable (SDN – Software Defined Networks), établissant une base technologique solide pour un meilleur usage des infrastructures numériques sous-jacentes, celles-ci étant partagées entre plusieurs usagers, et optimisées en fonction des besoins. Il s'agit donc d'un changement de paradigme majeur, où les équipements matériels statiques déployés dans le cadre des réseaux 4G sont virtualisés, remplacés par des composants logiciels distribués dont le cycle de vie est orchestré par les outils de gestion réseau des opérateurs.

De plus, l'approche "5G Standalone" va permettre aux opérateurs et industriels possédant leurs propres réseaux privés d'orchestrer des réseaux virtuels, instanciés sur la même infrastructure matérielle sous-jacente (des serveurs informatiques, des routeurs ou commutateurs réseau banalisés), chaque réseau logique étant dimensionné pour utiliser une quantité définie de ressources informatiques pendant un certain temps et selon un degré d'isolation défini à l'avance. En 5G, ces réseaux virtuels sont appelés "network slices", ils sont constitués d'un ensemble de services réseau, eux-mêmes construits sur la base d'un graphe de fonctions réseau logicielles déployées dans des machines virtuelles ou des containers.

Nous pensons que cette nouvelle manière d'opérer les réseaux permettra d'ailleurs aux opérateurs d'offrir à la demande des services de connectivité « à la carte » à leurs clients industriels, sur des zones géographiques précises, et d'en dégager des profits substantiels. Il s'agit d'une révolution du même ordre que celle vécue par l'industrie de l'informatique avec le « Cloud Computing » qui ouvre la porte à la consommation à la demande et au plus près des ressources

physiques associées aux services de communication. C'est en cela, et par la mutualisation des infrastructures de calcul et de communication que la 5G pourra amener des gains substantiels en matière d'efficacité énergétique.

Enfin, il est important de noter que la capacité à déployer des réseaux virtuels (slices) n'est pas qu'un enjeu technologique mais aussi le moteur permettant d'adopter une approche de réseau-en-tant-que-service (Network-as-a-Service) par laquelle un opérateur déploie une offre de réseau virtuel managée à d'autres fournisseurs de services numériques (voir à des opérateurs virtuels). Par exemple, dans un contexte de contribution vidéo, un diffuseur média ou une chaîne de télévision désirent couvrir un événement particulier à un endroit spécifique pourrait louer un réseau virtuel sur une base temporelle et selon des contraintes de qualité de service définie à l'avance à un prix contractuel dépendant des capacités et des performances demandées.

Les dernières évolutions en la matière, spécifiées par le 3GPP, ouvrent de plus la porte à de nouveaux acteurs louant des services plus élémentaires qu'un réseau de communication complet (fut-il virtuel). Ces services élémentaires, pourront être enrôlés de manière automatique dans la construction d'un service de plus haut niveau et permettront l'émergence d'un nouvel écosystème sur lequel seront construites les infrastructures numériques de demain. Ils sont par nature très variés, mais on peut notamment citer :

- Une infrastructure radio : le propriétaire d'un bâtiment, stade ou autre, équipé avec des têtes radio 4G/5G pourra vouloir mettre cette infrastructure physique à disposition d'un opérateur lors d'un événement médiatique ponctuel ayant lieu sur son site.
- Une bande de fréquence : Le possesseur d'une licence d'accès radio dans une bande de fréquence donnée pourra offrir une partie de cette capacité radio sur une base temporelle définie à l'avance à un fournisseur de services de communication
- Des capacités de calcul : un industriel possédant d'importantes capacités de calcul non-utilisées et orchestrées via un Gestionnaire d'Infrastructure Virtualisée (VIM) pourra exposer ses APIs de VIM à l'orchestrateur de services réseau d'un fournisseur de services de communication
- Un réseau fibre : un opérateur de boucle locale ayant déployé un réseau de fibres optiques programmable à l'aide d'un contrôleur SDN (Software Defined Network) peut proposer un service de connectivité entre des têtes radio 4G/5G et des fonctions RAN virtualisées instanciées dans un datacenter proche.

Dans une approche encore plus avancée dite approche micro-service, les fonctions réseau logicielles sont fragmentées en composants fonctionnels élémentaires, plus agiles, et sans états, pensés pour être exécutés dans des environnements informatiques distribués, hétérogènes et non fiabilisés. Ceci composant au total une architecture de réseau plus résiliente, dont les microéléments sont distribués aux endroits les plus appropriés pour respecter les contraintes de qualité de service (latence, sécurité, fiabilité, ...) et peuvent être redémarrés de manière transparente pour l'utilisateur en cas de panne ou de dysfonctionnement.

C'est le paradigme sur lequel l'Institut de Recherche Technologique b<>com travaille activement depuis plusieurs années et qui permettra à terme de jeter les bases des réseaux du futur en répondant aux besoins de densification, de résilience et d'efficacité énergétique. Les concepts avancés sont d'ailleurs par définition indépendants de la technologie d'accès radio et s'appliquent

aussi bien à la 5G, à la 4G, qu'aux réseaux WLAN ou IoT. Il apparaît donc pertinent de travailler en priorité sur la mise en place d'une infrastructure cœur flexible, programmable et convergente plutôt que d'investir dans l'ajout de capacité radio pour des usages malheureusement encore mal définis aujourd'hui.

Les infrastructures de calcul

On l'a vu dans la section précédente, les infrastructures de calcul, autrement dit les datacenters, constitueront l'épine dorsale des réseaux du futur, l'infrastructure matérielle sur laquelle leurs composants logiciels, virtualisés, seront déployés et exécutés. La mutualisation de ces moyens de calcul pour le traitement de données en Edge est souhaitable et relève de la même logique de distribution, implacable et liée à l'augmentation exponentielle de la production et de la consommation des données. Mais, cela soulève également la question centrale, posée aujourd'hui par les acteurs industriels, de sa rentabilité en l'absence de facteur échelle tel que celui réalisé sur les grosses structures centralisées.

En effet, dans une logique de distribution, l'emplacement de ces infrastructures est dorénavant contraint par la nécessaire proximité avec les usagers et leurs données, et ne permet pas de se reposer sur un environnement favorablement efficace en matière de climatisation et donc de consommation électrique. Dans un contexte où le secteur des nouvelles technologies représente à lui seul entre 6 et 10% de la consommation mondiale d'électricité, selon les estimations, dont 30% directement imputable aux datacenters, une pression croissante incite les propriétaires de ces infrastructures numériques à en réduire la consommation énergétique.

Ces chiffres de consommation sont en croissance constante, de 5 à 7% d'augmentation par an, stimulés par la consommation exponentielle de services numériques, ainsi que par le dimensionnement des infrastructures qui doivent être capables de faire face à tout moment à des pics de consommation n'arrivant en réalité que quelques heures par jour. Tout cela milite clairement en faveur d'une meilleure efficacité opérationnelle.

Sur le sujet de l'efficacité énergétique qui en est un volet important, trois variables clés ont une incidence majeure sur la consommation électrique des datacenters : leur emplacement, leur charge informatique et leur efficacité électrique. Mais une analyse complète doit également inclure la contribution en carbone embarqué de toutes les phases du cycle de vie produit, ceci incluant la fabrication des serveurs, leur emballage, leur transport jusqu'au datacenter, leur installation et enfin l'éventuel recyclage des matériaux lors du remplacement des équipements.

L'efficacité électrique est une variable dépendant principalement de la structure du bâtiment, de la conception des systèmes de refroidissement, de l'emplacement du datacenter et de la nature de l'énergie électrique délivrée par les fournisseurs d'électricité, et se situe donc sous le contrôle des propriétaires de datacenters. Il existe cependant des opportunités de réduction de l'empreinte carbone globale du datacenter reposant sur une meilleure optimisation de la charge informatique des serveurs et sur le rallongement de leur cycle de vie, autant de variables dont peuvent se saisir les hébergeurs et fournisseurs de service numériques. En effet, ceux-ci peuvent réduire leurs impacts écologiques en adoptant de nouveaux équipements plus économes, mais si la démarche se fait dans la précipitation, elle aura l'effet inverse. Si tous les serveurs sont remplacés sur une courte période, l'empreinte carbone, liée à la destruction et au recyclage des matériaux et des métaux utilisés dans

les équipements, sera très élevée. À cela vient s'ajouter l'empreinte liée à la conception et à l'acheminement du nouveau matériel. Par conséquent, au lieu de faire exploser la demande – et donc l'offre – de nouveaux serveurs moins énergivores, il est sans doute préférable d'adopter une stratégie parcimonieuse et d'allonger le cycle de vie des serveurs existants, lorsque cela fait sens, pour réduire la demande, créant ainsi un cercle vertueux en termes d'écologie. La demande en ressource IT étant soutenue et croissante, il importe donc dans une démarche de développement durable et équitable de pouvoir faire plus avec moins de moyens.

Les enjeux technologiques sur les nouvelles infrastructures matérielles pourraient donc se résumer dans les points suivants :

- Efficacité énergétique globale dans un contexte d'augmentation massive des besoins de traitement,
- Rationalisation de l'usage des infrastructures à toutes les échelles,
- Optimisation des coûts opérationnels et d'investissement dans un contexte de distribution des infrastructures matérielles,
- Architectures résilientes et souveraines au service des communautés

Une solution technique, répondant à ces enjeux et permettant à un ensemble d'acteurs de collaborer pour augmenter le taux d'usage de leurs infrastructures est par exemple en cours de développement chez b<>com. Elle consiste à créer un cloud collaboratif qui regroupera et exposera les ressources informatiques non utilisées de multiples entreprises et/ou administrations publiques. L'ambition est ainsi de réaliser des passerelles entre cloud privés, de la même manière que le cloud hybride établit une passerelle entre le cloud privé et le cloud public.

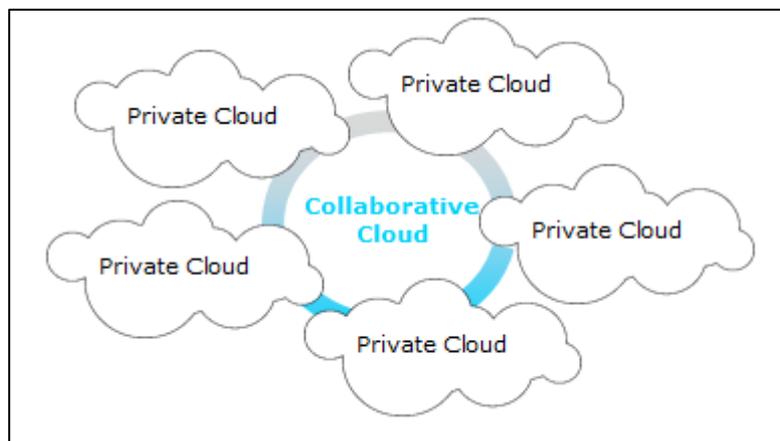


Figure 5: le cloud collaboratif

Concrètement, il s'agit de construire un cloud collaboratif en interconnectant entre eux différents datacenters privés ayant chacun un profil de consommation spécifique, puis d'exposer au travers d'APIs standards les ressources hétérogènes dormantes de cette communauté de datacenters à des consommateurs moyennant un revenu de location.

L'intégration dans la solution de mécanismes d'apprentissage automatique, d'algorithmes de prédiction de la disponibilité des ressources informatiques dormantes permet de maximiser l'usage de celles-ci tout en répondant de manière optimisée aux besoins des applications. L'ambition est de pouvoir garantir la disponibilité des ressources sur plusieurs heures (éventuellement jusqu'à 24h).

De plus, un mécanisme réactif permettra de libérer automatiquement et immédiatement les ressources allouées à la communauté au niveau d'un datacenter si des besoins imprévisibles apparaissent.

Les grands différentiateurs cette offre de cloud collaboratif sont donc les suivantes :

- **Augmentation des revenus** : La génération d'une nouvelle source de revenus au travers de la valorisation des ressources dormantes des infrastructures privées
- **Rationalisation de l'usage du matériel** : La mise en place d'un écosystème collaboratif vertueux permettant de densifier l'usage des infrastructures existantes et de limiter les investissements en matériel supplémentaire
- **Flexibilité des politiques de cloudification** : L'accompagnement d'une entreprise à la cloudification de certaines de ses applications. Les outils développés par b<>com permettront aux utilisateurs de créer leurs propres zones d'affinité et donc d'externaliser leurs données et applications uniquement sur les cloud privés qu'ils auront choisis. A l'extrême, un utilisateur peut commencer à placer ses charges de calcul sur sa propre infrastructure matérielle. De plus la solution devra reposer sur des APIs largement répandues et utilisées dans l'industrie, ce qui facilitera encore la migration.
- **Coût des ressources réduit** : Le coût des ressources, attendu à moins d'un tiers du prix pratiqué par les opérateurs hyperscale, générera des opportunités de gain importantes pour l'opérateur d'une communauté de clouds collaboratifs et un prix d'instance très bas pour l'utilisateur final.

Comment être moins vulnérables aux cyber-attaques ?

Il est utile ici de se référer au référentiel NIS adopté par l'union européenne en 2016 et notamment aux travaux du « NIS Cooperation Group », groupe par lequel les états de l'Union collaborent, échangent des informations et s'assurent que le référentiel NIS est implémenté de manière cohérente. Ce groupe a notamment évalué le risque cyber associé aux réseaux 5G dans un ensemble de documents mettant en lumière le fait que les concepts technologiques introduits par cette technologie, tendent à augmenter la surface d'attaque globale et le nombre de points d'entrée pour les attaquants potentiels :

- Des fonctionnalités accrues en périphérie de réseau (Edge) en comparaison des générations précédentes de réseaux mobiles ; ce qui signifie que certaines fonctions du réseau cœur pourront être distribuées sur plus d'équipements les rendant plus sensibles au risque cyber ;
- Une part accrue des logiciels dans les équipements 5G entraîne une augmentation des risques liés aux processus de développement et de mise à jour de ceux-ci, ainsi qu'aux erreurs de configuration, et donne un rôle plus important aux choix faits par chaque opérateur de réseau mobile dans les phases de déploiement de leur réseau

De plus, comme évoqué précédemment, ces nouveaux concepts, favorisent l'émergence d'un nouvel écosystème avec de nouveaux acteurs, créant un contexte d'exposition accrue aux attaques facilitées par des fournisseurs tiers dont le profil de risque individuel deviendra particulièrement important, et devra être évalué par les opérateurs de réseau.

Dans ce contexte, être moins vulnérable aux attaques suppose une capacité des réseaux 5G à s'adapter à des contextes de sécurité qui peuvent évoluer avec le temps. Ainsi, il est vain de penser qu'une politique de sécurité établie statiquement puisse protéger efficacement un réseau à terme et il importe de pouvoir prendre en temps réel et de manière automatique les actions nécessaires pour le maintien d'un niveau de protection optimal.

Pour cela, un modèle opérationnel intéressant, général et applicable aux réseaux télécom peut se décomposer en cinq fonctions fondatrices (IPDRR) telles que décrites ci-dessous :

- « **Identifier** » : Développer la compréhension de l'opérateur des risques en matière de cybersécurité sur les systèmes, les actifs, les données et les capacités associées aux déploiements des services de communication.
- « **Protéger** » : En regard de l'analyse de risques effectuée à l'étape précédente, et des contraintes de sécurité associées aux services de communication de l'opérateur, il s'agit de mettre en œuvre les garanties appropriées pour assurer la prestation de ces services. En pratique, cette étape permet de définir précisément les contraintes fonctionnelles (choix des fonctions réseau virtualisées et des composants de sécurité) et non-fonctionnelles (degré d'isolation entre les « slices » de réseau, affinités géographiques et/ou matérielles, surveillance et rapports, ...) associées au déploiement du service.
- « **Détecter** » : Développer et mettre en œuvre les activités appropriées pour identifier l'apparition d'un événement de sécurité. Il s'agira de mettre en œuvre les composants de détection, possiblement dotés de capacités d'intelligence artificielle, capable de détecter une déviation comportementale sur un ou plusieurs éléments de réseau sur la base des rapports de surveillance définis à l'étape précédente.
- « **Répondre** » : Développer et mettre en œuvre les activités appropriées concernant un événement détecté à l'étape précédente. Ces activités seront variées et dépendent du type d'évènement, il pourra s'agir de la mise en œuvre de nouvelles règles dans un pare-feu, de l'isolation d'une fonction réseau virtualisée ou de son déplacement sur un serveur différent, de la redirection d'un flux de données ou même, si nécessaire, du décommissionnement complet du service de communication.
- « **Récupérer** » : Développer et mettre en œuvre les activités appropriées pour restaurer les capacités ou les services et performances associées qui ont été altérées en raison d'un événement de cybersécurité. Il s'agira par exemple de redéployer une fonction réseau compromise, voir un service de communication complet.

Dans les sections précédentes, nous avons vu que les réseaux cellulaires 5G peuvent mettre en œuvre des instances du réseau (les slices) qui sont définies comme un réseau logique utilisant les ressources d'une infrastructure partagée et offrant des services de bout en bout aux utilisateurs avec des exigences fonctionnelles et non-fonctionnelles. L'architecture logicielle permettant cette mise en œuvre repose sur des composants théorisés par l'ETSI sur ses travaux liés à l'orchestration de réseaux (MANO – MANagement and Orchestration) et comprenant principalement trois niveaux de responsabilités différents :

- Le gestionnaire des ressources virtualisées d'infrastructure (VIM) est responsable du contrôle et de la gestion des ressources de calcul, de stockage et de réseau de

l'infrastructure NFV (NFVI), et il offre des services de machines virtuelles et la connectivité réseau (as a service). Le VIM offre des interfaces de programmation permettant l'inventaire et la gestion du cycle de vie des ressources virtualisées, ainsi que le monitoring et la gestion des évènements.

- L'orchestrateur NFV (NFVO) est responsable de l'intégration des nouveaux services réseau (NS – Network Service) et des packages de fonctions réseau virtualisées (VNF) permettant de prendre en compte les besoins fonctionnels du service, de la gestion du cycle de vie des services réseau et des fonctions réseau virtualisées.
- Le gestionnaire de slice réseau (NSM – Network Slice Manager) fait partie du BSS/OSS d'un réseau et comprend les principales fonctions suivantes : la conception des services clients qui pourraient être offerts par un slice réseau sur la base des services réseau plus élémentaires que le NFVO rend disponibles à travers son catalogue, la gestion du cycle de vie des instances de slice réseau (NSI – Network Slice Instances).

Ces trois piliers forment la base architecturale permettant la programmabilité et l'automatisation des réseaux. A ce système, b<>com ajoute les services supplémentaires « Surveillance », « Analyse », « Décision » qui agissent comme une boucle de rétroaction pour renforcer cette automatisation dans le domaine de la sécurité :

- Le service Surveillance rassemble des données statistiques qui peuvent être générées par l'infrastructure virtualisée NFVI, ou par une VNF et les rend accessible via une base de données par exemple
- Le service Analyse exploite les informations mises à disposition par le service Surveillance pour y extraire des évènements
- Le service Décision met en œuvre des actions appropriées concernant un évènement détecté. Les actions peuvent être des opérations invoquées sur le gestionnaire d'infrastructure virtualisé (par exemple pour replacer une VNF sur un autre serveur du même datacenter), sur le gestionnaire des fonctions réseau virtualisées, sur l'orchestrateur lui-même (si l'action a un impact sur un service réseau), ou même le gestionnaire de slices réseau (si le fonctionnement du réseau virtuel complet n'est plus conforme aux SLA demandés).

Les travaux en matière de cybersécurité des réseaux poursuivis chez b<>com visent donc à réutiliser ce cadre théorique ainsi que le modèle opérationnel IPDRR de manière à définir, déployer et maintenir des composants non-fonctionnels de sécurité dans un réseau, et lui conférer en cela des capacités de résilience aux attaques, d'autodiagnostic et de réparation, comme illustré dans le tableau suivant :

| NIST | | Impacts on management, orchestration and PaaS services | | | | | |
|----------------|--|--|------|-----|------------|-----------|----------|
| Core Functions | Description | NSM | NFVO | VIM | Monitoring | Analytics | Décision |
| Identify | Phase de conception des modèles de slices réseau | | | | | | |
| Protect | Préparation et déploiement d'une instance de slice réseau Incluant un service de surveillance et les fonctions cryptographiques adaptées aux données transportées et à certaines exigences non-fonctionnelles comme la fiabilité ou la haute disponibilité. | X | X | X | X | | |
| Detect | Analyse des données collectées et conversion au format adapté au système de détection. En fonction de la politique de sécurité définie pour une instance de slice réseau, cette analyse, si elle permet de détecter un événement anormal peut déclencher une action de contre-mesure. | | | | X | X | |
| Respond | Le service de Décision déclenche une série d'actions selon une logique prédéterminée afin de limiter les effets de l'incident de sécurité. Un exemple d'action possible est l'insertion de composants de sécurité supplémentaires dans le réseau (Pare-feu, ...) ou la mise en quarantaine des machines virtuelles ou physiques affectées. | X | X | X | | | X |
| Recover | Afin de respecter les engagements de performance de l'opérateur, le service de Décision doit restaurer les capacités initiales du réseau, par exemple redéployer les composants affectés et mis en quarantaine. | X | X | X | | | X |

Figure 6: Automatisation de la sécurité des réseaux

Réinventer et démocratiser les usages et numériser la démocratie ?

Numériser la démocratie / les services de l'État

Il ne fait aucun doute que le Cloud Computing permet aux administrations de transformer leurs opérations tout en réduisant leurs coûts. Certaines d'entre elles l'ont déjà compris. Aux États-Unis par exemple, Vivek Kundra, directeur fédéral des systèmes d'information au sein de l'administration Obama, a lancé la stratégie de Cloud Computing du gouvernement en février 2011. L'administration du président Obama a adopté ce qu'elle a appelé l'approche « Cloud First », qui vise à privilégier le modèle du Cloud Computing plutôt que le modèle classique, et depuis 2012, les agences fédérales doivent adopter des solutions de Cloud Computing par défaut.

Dans les administrations, cette approche du Cloud Computing à la demande peut améliorer la qualité des services fournis grâce à des données centralisées et des plates-formes partagées. Ainsi dans les pays nordiques, un certain nombre de DSI du secteur public ont étudié la faisabilité d'un « cloud nordique » grâce auquel des données publiques pourraient ainsi être mises à la disposition de divers pouvoirs publics pour qu'ils puissent les exploiter de l'intérieur de leurs frontières organisationnelles, plutôt que de travailler à partir de leurs propres bases de données en silo. Cette nouvelle stratégie diffère notablement du modèle traditionnel « inside-out » qui se contente de supporter par des moyens externes les services gouvernementaux internes.

Le rejet et le retard ne sont pas une bonne stratégie en ce qui concerne le Cloud. Les administrations qui implémentent une solution de Cloud Computing rapidement réalisent des économies qui pourraient représenter jusqu'à 25% des coûts de fonctionnement dans certains secteurs. Dans les ministères et organismes qui essaient de s'adapter aux coupes budgétaires, ce type d'avantage financier ne peut pas être négligé.

En effet, dans la période actuelle, les administrations souhaitent obtenir plus de résultats avec moins de moyens. De plus en plus elles comprennent que le « plus » implique de répondre aux attentes des citoyens en leur offrant des services en ligne plus sophistiqués (la technologie de Cloud

Computing ayant un rôle crucial à jouer), et que le « moins » fait référence à la réduction du coût des systèmes informatiques existants.

S'inscrivant en plein dans cette tendance, l'administration française devra fonctionner sans échange de papier d'ici la fin du quinquennat, a promis en 2017 Gérald Darmanin, alors ministre de l'Action et des Comptes publics. Il s'agit de miser notamment sur la numérisation, la dématérialisation et la simplification afin de faire gagner du temps et de l'argent. Ici encore, les technologies du Cloud joueront un rôle essentiel dans la réalisation de cet objectif.

Au regard des considérations évoquées ci-dessus, il est donc légitime de se poser la question de la pertinence de la mutualisation des moyens au niveau européen pour la mise en place d'une infrastructure de cloud computing efficace, moderne, et respectueuse des souhaits de souveraineté des différentes nations européennes. Ce sujet de la souveraineté est d'ailleurs central et doit être adressé prioritairement dès lors que l'on veut numériser les services de l'état en s'appuyant sur des technologies de l'état de l'art, disponibles sur le marché et satisfaisant des critères de coûts, de maintenabilité et de pérennité. Malgré l'absence de réelle alternative aux GAFAs sur ces critères, il faut cependant noter une réelle volonté du gouvernement français d'assurer la souveraineté des données produites et hébergées sur le territoire national. Cette volonté de souveraineté est d'ailleurs exacerbée par la promulgation du « Cloud Act » (Clarifyng Lawful Overseas Use of Data Act) en Mars 2018 suite à l'arrivée de Donald Trump au pouvoir aux États-Unis. Pour mémoire, le « Cloud Act » peut être vu comme une extension du « Patriot Act » (promulgué par George W. Bush après les attentats de 2001) et vise à permettre aux autorités judiciaires américaines l'accès aux données électroniques stockées à l'étranger. Son but est de permettre un accès plus rapide à celles-ci en s'adressant directement aux fournisseurs de cloud plutôt que de passer par le biais d'une demande d'entraide judiciaire internationale. Concrètement, le « Cloud Act » stipule que toute société de droit américain doit communiquer, sur demande des autorités américaines bénéficiant d'un mandat dans le cadre d'une enquête pénale, les données placées sous son contrôle, sans considération du lieu où elles sont localisées.

Ainsi, il est possible de se rappeler une initiative telle que le projet Andromède, lancée en 2011 visant à développer une alternative française et européenne aux sociétés américaines, autour de trois grands acteurs de l'informatique et des télécommunications : Orange, Thalès et Dassault Systèmes. Le financement de ce projet (150 millions d'euros) a été réalisé grâce au grand emprunt via la Caisse des Dépôts en tant qu'actionnaire majoritaire de la nouvelle structure. A la suite d'un désaccord entre Orange et Dassault Systèmes, le projet initial a été abandonné et a donné naissance à deux nouveaux projets, soumis en 2012 à l'appel à projet du gouvernement français :

- Cloudwatt initialement porté par Orange et Thalès, racheté intégralement par Orange en 2015 et arrêté en Février 2020,
- Numergy initialement porté par SFR et Bull, Dassault Systèmes étant sorti du projet et racheté intégralement par SFR en 2016, puis définitivement enterré en 2018.

Il faut donc se poser la question de la résilience des technologies et des structures sur lesquelles faire se reposer les services de l'état et questionner la pertinence d'en confier le développement et l'administration à quelques grandes entreprises, dont l'État ne maîtrise pas nécessairement la gouvernance ni la stratégie à moyen terme, et qui souvent se contentent d'intégrer les technologies américaines ou chinoises (serveurs DELL ou CISCO, virtualisation VMware, ...).

Ainsi, une approche orientée filière, en soutien dès le début du processus d'innovation, ancrant le développement des technologies d'infrastructures numériques dans le tissu industriel français, en particulier stimulant l'appropriation de celles-ci par les PME françaises ferait parfaitement sens.

Le sujet de la souveraineté doit également couvrir les infrastructures réseaux qui sont impliquées dans le transport de données potentiellement sensibles ou critiques. Nous pouvons bien sûr compter sur deux champions européens qui cependant présentent un retard technologique et marché indéniable notamment par rapport aux propositions asiatiques poussées par Huawei, Samsung, ZTE, Datang, ...

A l'échelle de notre pays, il semble que nous n'avons malheureusement pas (plus) la possibilité de recréer un champion Français de cette envergure, capable d'adresser les besoins des opérateurs nationaux. Petit à petit, notre industrie, et les formations académiques qui la servent, évoluent vers d'autres horizons.

Mais par ailleurs, avec une approche cloud-native plus innovante sur le plan technologique et des ambitions calibrées sur des marchés plus ciblés, des équipementiers télécom pour réseaux privés se développent dans d'autres pays :

- États-Unis : Radisys, Mavenir, Flexcore, Airspan, Casa Systems, Affirmed Networks, ...
- Royaume-Uni : Quortus
- Italie : Athonet (choisi en 2019 par le ministère de l'intérieur Français pour fourniture de bulles tactiques - appel à projet PCSTORM - lot 1 - Réseaux tactiques et terminaux)
- Allemagne : Blackned

Il ne semble malheureusement pas non plus y avoir de proposition industrielle française avec un positionnement similaire, à tel point qu'aucune entreprise française n'a été retenue par le ministère de l'intérieur pour la fourniture de solutions réseaux critiques.

L'urgence est donc de dynamiser l'écosystème télécom français et favoriser l'émergence de petites structures agiles qui pourront proposer des produits innovants et s'adapter facilement au marché professionnel. Un outil intéressant pour cela consisterait à libérer rapidement du spectre à usage privé sur un modèle non contraignant proche du modèle américain qui permettrait la mise en place rapide de cas d'usages variés. **Dans ce contexte, b<>com pourra accompagner technologiquement les entreprises sur la base d'une offre de formation et de briques technologiques à intégrer dans leurs produits.**

A moyen terme, dans le cadre de la généralisation des solutions de virtualisation et d'orchestration réseaux, cette offre française d'infrastructure cellulaire privée pourrait être déployée par les opérateurs sur des cas d'usage spécifiques, dans des slices réseau spécialisées sur des marchés professionnels du type :

- Mission Critical
- Évènementiel
- Industrie 4.0
- Transport / Infrastructures routières

Démocratiser l'accès aux services de l'État

Par ailleurs, en ce qui concerne l'accès aux applications et services numériques, on peut rappeler que le plan national Français Très Haut Débit (THD) consiste à déployer plus de 30 millions de lignes à très haut débit (principalement en FTTH) d'ici 2022, ce qui représente plus de 20 milliards d'euros d'investissements. Les opérateurs déploient leurs réseaux sur une base commerciale (ou ont manifesté leur intérêt à le faire) dans des zones qui couvrent environ 57 % de la population, soit un investissement de 6 à 7 milliards d'euros. Les autorités locales déploient des réseaux d'initiative publique dans le reste du pays, y compris dans les zones à faible densité de population. Celles-ci offrent un accès aux opérateurs commerciaux à des prix de gros maintenus au même niveau que ceux des zones plus densément peuplées par des subventions publiques (qui représentent environ la moitié des 13-14 milliards d'euros restant à investir). A titre d'exemple, un opérateur de réseau fibre optique a été créé en Bretagne (délégation de service public confiée à Orange).

Au-delà du déploiement strict d'un nouveau réseau fibré haute performance, il s'agit de construire l'épine dorsale permettant d'interconnecter des centres de données ou plus généralement des infrastructures informatiques (potentiellement elles-mêmes support de nos réseaux cellulaires futurs) publiques ou privées. Et donc en ce qui concerne les entreprises, cet Internet très haut débit est une opportunité de développer leur activité et leur compétitivité : indépendants, startups, PME, grands groupes, dans les grandes villes comme dans les plus petites communes. Le gain de productivité est indéniable :

- avec l'échange rapide de données à grande échelle,
- le stockage externe de ces données, « le Cloud Computing »,
- la fluidité des interfaces audio et vidéo pour le travail collaboratif, le télétravail ou les visioconférences.

L'Internet très haut débit est source de nouvelles opportunités pour les entrepreneurs. De nouveaux produits, services et usages émergent, les poussant à partir à la conquête de nouveaux marchés.

Sur le plan de la démocratisation de l'accès aux services numériques, il reste donc essentiel de continuer et renforcer ces déploiements malgré les difficultés auxquelles les fournisseurs de services et leur sous-traitant sont confrontés. En complément, notamment dans les zones rurales, le déploiement de l'accès sans fil haut-débit reste un enjeu important qui pourra être adressé en utilisant par exemple la bande des 700MHz, présentant des qualités de propagation adaptées, et récemment transférée au secteur des télécommunications suite au déploiement de la TNT. A ces fréquences, les largeurs de bande étant limitées, il importe de privilégier l'usage de technologies éprouvées et peu onéreuses sur le segment radio, même si celles-ci peuvent toujours être couplées à un cœur de réseau répondant aux principes énoncés dans la section « Évolution des infrastructures réseaux ».