

# Introduction à la 5G

EFORT

<http://www.efort.com>

## 1. Introduction

Les générations 2G, 3G puis 4G de téléphonie mobile se sont inscrites dans une continuité technologique : chaque nouvelle génération était une évolution de la précédente.

Avec la 5G, une véritable rupture technologique est née. Cette nouvelle norme de téléphonie mobile exploite par exemple les technologies IP et HTTP2 pour la gestion de la signalisation dans le contrôle du réseau pour une intégration dans le cloud.

La 5G n'est plus seulement un accès mobile très performant, mais un véritable réseau d'intégration comprenant un ensemble de technologies permettant d'offrir des services adaptés aux verticales, comme la télémédecine, la ville intelligente, la voiture connectée, l'accès à Internet à très haut débit pour des applications de réalité virtuelle, etc.

C'est une évolution encouragée par la communauté internationale, via le consortium 3GPP.

En effet, les exigences de ces verticales, telles que la qualification des débits montants et descendants, les seuils de latence acceptables ou encore le niveau de disponibilité des systèmes sont pleinement considérées dès la phase initiale des spécifications de la 5G.

La 5G offrira un débit accru et de très faible latence, proche de celui de la fibre optique, permettant de proposer une connectivité performante là où la fibre est absente, avec le FWA (Fixed Wireless Access). Du côté des opérateurs, la virtualisation du réseau permettra d'adresser tous les marchés avec une infrastructure unique. L'ère des équipements dédiés est révolue : la même architecture matérielle sera utilisée pour différents types de réseaux et services, grâce au logiciel. La virtualisation ouvre ainsi la porte au réseau à la demande, tout en souplesse. Le but de ce tutoriel est d'introduire la 5G.

## 2. Catégories d'usage de la 5G

Trois grandes catégories d'usages, avec leurs exigences respectives et potentiellement incompatibles entre elles, sont en train d'émerger et permettraient de répondre aux besoins métier :

1. mMTC – Massive Machine Type Communications : communications entre une grande quantité d'objets avec des besoins de qualité de service variés. L'objectif de cette catégorie est de répondre à l'augmentation exponentielle de la densité d'objets connectés ;
2. eMBB – Enhanced Mobile Broadband : connexion en ultra haut débit en outdoor et en indoor avec uniformité de la qualité de service, même en bordure de cellule ;
3. uRLLC – Ultra-reliable and Low Latency Communications : communications ultra-fiables pour les besoins critiques avec une très faible latence, pour une réactivité accrue.

Le premier groupe (mMTC) englobe principalement tous les usages liés à l'Internet des objets. Ces services nécessitent une couverture étendue, une faible consommation énergétique et des débits relativement restreints. L'apport annoncé de la 5G par rapport aux technologies actuelles réside dans sa capacité à connecter des objets répartis de manière très dense sur le territoire.

L'ultra haut débit (eMBB) concerne tous les applications et services qui nécessitent une connexion toujours plus rapide, pour permettre par exemple de visionner des vidéos en ultra haute définition (8K) ou de « *streamer* » sans-fil des applications de réalité virtuelle ou augmentée.

Les communications ultra-fiables à très faible latence (uRLLC) regroupent toutes les applications nécessitant une réactivité extrêmement importante ainsi qu'une garantie très

forte de transmission du message. Ces besoins se retrouvent principalement dans les transports (temps de réaction en cas de risque d'accident, par exemple), dans la médecine (téléchirurgie) et, de manière générale, pour la numérisation de l'industrie.

Afin de mettre en oeuvre ces trois types d'usages, huit indicateurs de performance (KPI – Key performance indicators) ont été établis par l'UIT pour préciser, quantifier et mesurer les caractéristiques de systèmes 5G :

- Débit crête par utilisateur (Gbit/s) ;
- Débit moyen perçu par l'utilisateur (Mbit/s) ;
- Efficacité spectrale (bit/Hz) ;
- Vitesse maximale des terminaux (km/h) ;
- Latence (ms) ;
- Nombre d'objets connectés sur une zone (quantité d'objets/km<sup>2</sup>) ;
- Efficacité énergétique du réseau ;
- Débit sur une zone (Mbit/s/m<sup>2</sup>) ;

**Débit crête par utilisateur** : En référence au débit de données maximal réalisable dans des conditions idéales par utilisateur ou appareil en bit/s. Les exigences minimales de 5G pour un débit de données maximal sont de 20 Gbit/s dans le sens descendant (DL, Downlink) et 10 Gbit/s dans le sens montant (UL, Uplink).

**Débit aperçu par l'utilisateur** : en référence au débit de données réalisable disponible partout à travers la zone de couverture à un utilisateur ou un appareil mobile en bits par seconde. Ce KPI représente une sorte d'expérience utilisateur minimale dans la zone de couverture. Cette exigence est définie par l'UIT-R et est fixée à 100 Mbit/s en DL et 50 Mbit/s en UL.

**Efficacité spectrale** : Définie comme le débit de données maximal dans des conditions idéales normalisées par la bande passante du canal, en bit/s/Hz. L'objectif fixé par l'UIT-R est de 30bit/s/Hz en DL et 15bit/s/Hz en UL. La combinaison de cet indicateur clé de performance et de débit maximal susmentionné entraîne la nécessité de 1 GHz de spectre pour répondre aux exigences énoncées, le minimum de bande étant de 100 MHz.

**Vitesse maximale des terminaux** : Il s'agit de la vitesse maximum à laquelle les terminaux peuvent se déplacer en garantissant leur mobilité, i.e., sans perte de la communication courante. Quatre classes de mobilité sont considérées :

- Stationnaire : 0 km/h
- A pied : 0km/h à 10km/h
- En véhicule : 10km/h à 120km/h
- En véhicule à haute Vitesse : 120km/h à 500km/h.

**Latence du plan usager** : L'exigence de latence unidirectionnelle de bout en bout (E2E) est fixée à 4 ms pour les services eMBB et à 1 ms pour les services URLLC.

**Nombre d'objets connectés sur une zone** : correspond au nombre total de devices connectés et/ou accessibles par unité de surface. L'UIT-R a spécifié un objectif de 1 million de dispositifs par km<sup>2</sup> pour des communications de type machine.

**Efficacité énergétique du réseau** : se réfère à la quantité de bits d'information transmise aux usagers ou reçue des usagers, par unité de consommation d'énergie du RAN, et du côté des appareils se réfère à la quantité de bits d'information par unité de consommation d'énergie du module de communication. La spécification donnée par l'UIT-R à cet égard est la suivante : les interfaces doivent pouvoir prendre en charge un taux de sommeil élevé et une longue durée de sommeil.

Pour les besoins d'ultra haut débit (eMBB), comme la vidéo 4K, 8K, 3D ou la réalité virtuelle, un certain nombre de performances, comme l'efficacité spectrale, le débit maximal et la capacité globale du réseau, peuvent être atteintes au détriment d'autres, comme la latence ou la densité de connexions simultanées.

A l'inverse, lorsque une connexion massive simultanée d'objets connectés (mMTC) doit être gérée, le réseau concentre ses ressources et utilise les technologies nécessaires à la résolution de cette tâche, mais il n'est pas en mesure, par exemple, d'utiliser aussi efficacement le spectre ou d'assurer une faible latence.

Enfin, lorsque des communications ultra fiables, avec une très faible latence sont nécessaires (uRLLC), le nombre de communications simultanées, le débit ou encore l'efficacité spectrale peuvent être réduits.

L'UIT définit la 4<sup>ème</sup> génération, sous la terminologie IMT-advanced, suivant les 8 KPI susmentionnés avec les valeurs présentées sur le schéma ci-dessous. La 5<sup>ème</sup> génération, appelée IMT-2020, y est également représentée. A cet égard, il convient de noter que la définition officielle de l'UIT de la 4G correspond en réalité à ce que les réseaux commerciaux n'ont déployé que dans un second temps, sous l'appellation 4G+ ou LTE-Advanced.

Selon ces objectifs, la 5G devra pouvoir offrir un débit utilisateur et un débit maximal respectivement 10 et 20 fois supérieur à ce qui est disponible actuellement avec la 4G. La densité maximale de connexions sera multipliée par 10 et la latence divisée par au moins 10 (la latence point à point cible est de 1 ms, contre 30 à 40 ms à ce jour en 4G).

L'objectif de la 5G est aussi de fournir :

- un réseau extrêmement fiable, avec des performances plus homogènes, quelle que soit la position de l'utilisateur par rapport à la station de base ;
- une connexion stable même en mobilité (avec de vitesses de l'ordre de 500 km/h) ;
- une augmentation de l'efficacité énergétique (batteries jusqu'à 100 fois moins énergivores).

Le tableau 1 compare la 5G par rapport à LTE-Advanced.

<b>Performances/Génération</b>	<b>4G</b>	<b>5G</b>	
1. Débit maximal (Gbit/s)	1	20	eMBB
2. Débit aperçu par l'utilisateur (Mbit/s)	10	100	eMBB
3. Efficacité spectrale	1x	3x	eMBB
4. Vitesse (km/h)	350	500	uRLLC
5. Latence (ms)	10	1	uRLLC
6. Nombre d'objets connectés sur une zone (quantité d'objets/km <sup>2</sup> )	10 <sup>5</sup>	10 <sup>6</sup>	mMTC
7. Efficacité énergétique du réseau	1x	100x	mMTC
8. Débit sur une zone (Mbit/s/m <sup>2</sup> )	0.1	10	eMBB

Tableau 1 : Comparaison entre les performances de la 4G et de la 5G

### 3. Le spectre pour la 5G

La 5G va utiliser les bandes de fréquence pionnières suivantes :

- Fréquences moyennes ou "Coverage and Capacity Layer" s'appuie sur des fréquences dans la bande de 2 et 6 GHz (e.g., 3400-3800 MHz) pour délivrer le meilleur compromis entre capacité et couverture

- Fréquences hautes ou "Super Data Layer" s'appuie sur des fréquences dans la bande au dessus de 6 GHz (e.g., 24.25-29.5 GHz et 37-43.5 GHz) afin d'adresser les cas d'usage spécifiques exigeant des débits très élevés. Il s'agit des fréquences millimétriques.
- Fréquences basses ou "Coverage Layer" exploite des fréquences dans la bande sous 2 GHz (e.g. 700 MHz et 1400 MHz) fournissant une couverture très importante.

La 5G est la première technologie mobile à fonctionner dans des fréquences à la fois basses, hautes, et surtout très hautes (6 GHz et plus). Ces dernières, vraie rupture technologique même si ces fréquences sont à faible propagation en distance, répondent à l'incessante augmentation des débits et à l'inflation des volumes de données échangés. Pour la transmission et les réseaux de collecte, la fibre est préférée avec la technologie FTTS (Fiber to the site), et les faisceaux hertziens seront toujours utiles.

### 3.1. Bandes millimétriques (mmWave : Millimeter Wave)

Avec la croissance explosive de la demande de trafic de données mobile, la contradiction entre les exigences de capacité et la pénurie de spectre devient de plus en plus importante. Le goulot d'étranglement de la largeur de bande devient un problème clé des réseaux mobiles 5G.

D'autre part, avec une bande passante énorme dans la bande millimétrique (mmWave) de 28 GHz à 300 GHz, les communications à ondes millimétriques (mmWave) ont été proposées comme une partie importante du réseau mobile 5G pour fournir des services de communication multi-gigabit tels que la télévision haute définition (HDTV) et la vidéo ultra haute définition (UHDV).

La plupart des recherches actuelles se concentrent sur la bande de 28 GHz, la bande de 39 GHz, la bande de 60 GHz et la bande E (E-band) (71-76 GHz et 81-86 GHz).

L'utilisation de bandes millimétriques constitue l'une des technologies de rupture de la 5G. Les bandes millimétriques n'ont encore jamais été prises en compte pour le déploiement des réseaux mobiles pour des raisons de maturité technologique et de qualité de propagation. Les bandes millimétriques offrent la réserve de spectre et leur utilisation permettrait d'atteindre les très hauts débits attendus avec la 5G. En contrepartie, leur utilisation impose le développement de toutes les technologies nécessaires, miniaturisées, à bas coût et avec une consommation énergétique compatible avec des terminaux portables (amplificateurs, codeurs, traitement de signal, antennes). En particulier, à cause de la faible qualité de propagation des ondes millimétriques, chaque cellule aura une couverture réduite, ce qui nécessitera la mise en place de techniques de *beamforming*, pour mieux focaliser l'énergie transmise par les antennes.

### 3.2. Massive MIMO

Indépendante de la largeur de la bande de fréquence, le **Massive MIMO** (*Multiple Inputs - Multiple Outputs*) est considéré pour la 5G. Cette technologie se caractérise par l'utilisation d'un nombre élevé de micro antennes « intelligentes », situées sur le même panneau (de 8 à 128 actuellement, mais le nombre augmentera avec l'utilisation de fréquences supérieures à 30 GHz, mmWave). L'attrait de l'utilisation du massive MIMO est double : d'une part, cette technologie permet d'augmenter les débits, grâce au multiplexage spatio-temporel ; d'autre part, elle permet de focaliser l'énergie sur un terminal, pour améliorer son bilan de liaison, grâce à la formation de faisceau, ou *beamforming*.

## 4. Système 5G

Le système 5G doit supporter toutes les capacités EPS (4G) avec les exceptions suivantes :

- SR-VCC et/ou CS-Fallback : Aucune mobilité ne peut être assurée entre le domaine 5G avec le domaine circuit

- handover entre 5G-RAN et GERAN (2G) ou entre 5G-RAN et UTRAN (3G) : Les seules mobilités possibles sont entre les radio 4G, 5G et WiFi.
- Accès au réseau cœur 5G via GERAN ou UTRAN : Le réseau cœur 5G n'offre d'interface qu'aux accès 4G, 5G et WiFi.

Le système 5G ne supporte donc que trois types d'accès : E-UTRAN, WLAN et NR (New Radio)

La voix est mise à œuvre via l'IMS et la mobilité ne peut être que paquet-paquet.

La figure 1 montre l'architecture de haut niveau qui peut être utilisée comme modèle de référence. La figure montre les éléments NextGen UE, NextGen RAN (Réseau d'accès), NextGen Core (Réseau cœur) et leurs points de référence.

N2: Point de référence pour le plan de contrôle entre NextGen (R)AN et NextGen Core (NGCN)

N3: Point de référence pour le plan usager entre NextGen (R)AN et NextGen Core.

N1: Point de référence pour le plan de contrôle entre NextGen UE et NextGen Core.

N6: Il s'agit du point de référence entre NextGen Core et le réseau de données (PDN, Packet data Network). Le réseau de données peut être un réseau de données public ou privé externe d'opérateur ou un réseau de données intra-opérateur, e.g., pour la fourniture de services IMS. Ce point de référence correspond à SGi dans le contexte 4G.

Les principes d'architecture sont les suivants :

- L'UE peut être rattaché au réseau sans avoir de session établie pour la transmission de données (important notamment pour les devices IoT).
- L'UE ne peut dialoguer avec le 5GC que si l'UE supporte le protocole NAS (Non-Access Stratum) N1.
- Les RANs (Radio Access Networks) ne peuvent dialoguer avec le réseau cœur 5G appelé NGCN ou 5GC que s'il supportent les interfaces N2 (Plan contrôle) et N3 (Plan usager). Les seules radio autorisées à s'interfacer au 5GC sont LTE et ses évolutions (LTE-Advanced, LTE-Advanced Pro), New Radio basée sur les évolutions futures LTE et mmWave et enfin WiFi (aussi bien WiFi trusted que WiFi untrusted). Dans le cas de l'accès WiFi untrusted, un élément d'interfonctionnement a été spécifié appelé N3IWF (Non-3GPP Interworking Function).
- Le réseau 5GC donne accès à des réseaux externes IP (e.g., Internet, Intranet, IMS) appelées Data Networks.

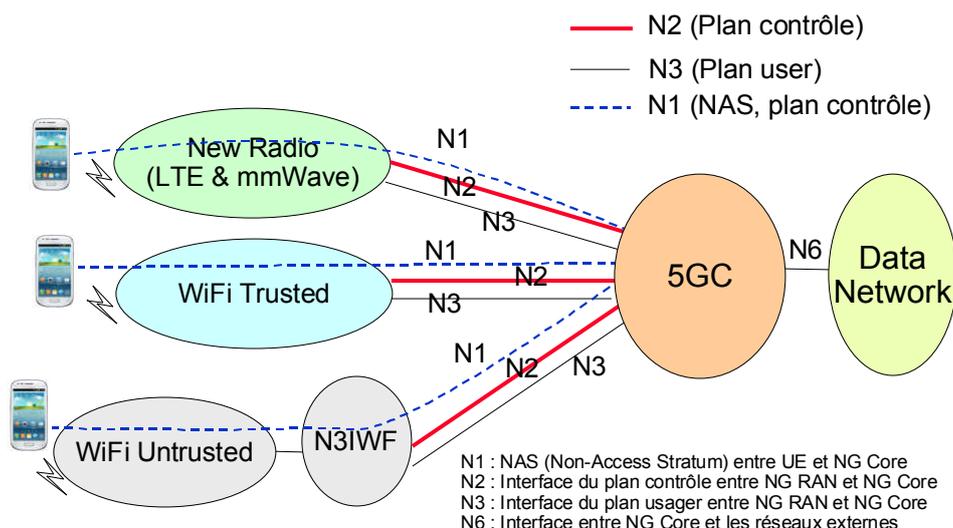


Figure 1 : Architecture 5GC

Le nouveau RAN (NG-RAN) est constitué des noeuds logiques suivants:

- gNBs comme nouveaux NodeBs qui implémentent la New Radio.

- ng-eNB est une évolution d'eNB qui prend en charge la connectivité à EPC et 5GC mais qui n'implémente que la radio LTE.

gNB fournit les terminaisons de protocole des plans Usager et Contrôle New Radio (NR) vers l'UE;

L'interface du plan de contrôle entre gNB ou ng-eNB et le 5GC est appelée N2 et est basée sur NG-AP qui sur un transport SCTP sur IP.

L'interface du plan utilisateur entre gNB ou ng-eNB et le 5GC est appelée N3 et est basée sur GTP-U sur un transport UDP sur IP.

De même les accès WiFi trusted et untrusted doivent montrer une interface N2 pour le plan contrôle et une interface N3 pour le plan usager afin de permettre la connexion de WiFi à 5GC et offrir un service comme la VoWiFi.

L'interface Xn permet la mobilité en mode actif entre les gNBs et les gNBs et les ng-eNBs. Xn utilise le protocole Xn-AP sur un transport SCTP/IP pour le plan contrôle et le protocole GTP-U sur un transport UDP/IP pour le plan usager comme indiqué à la figure.

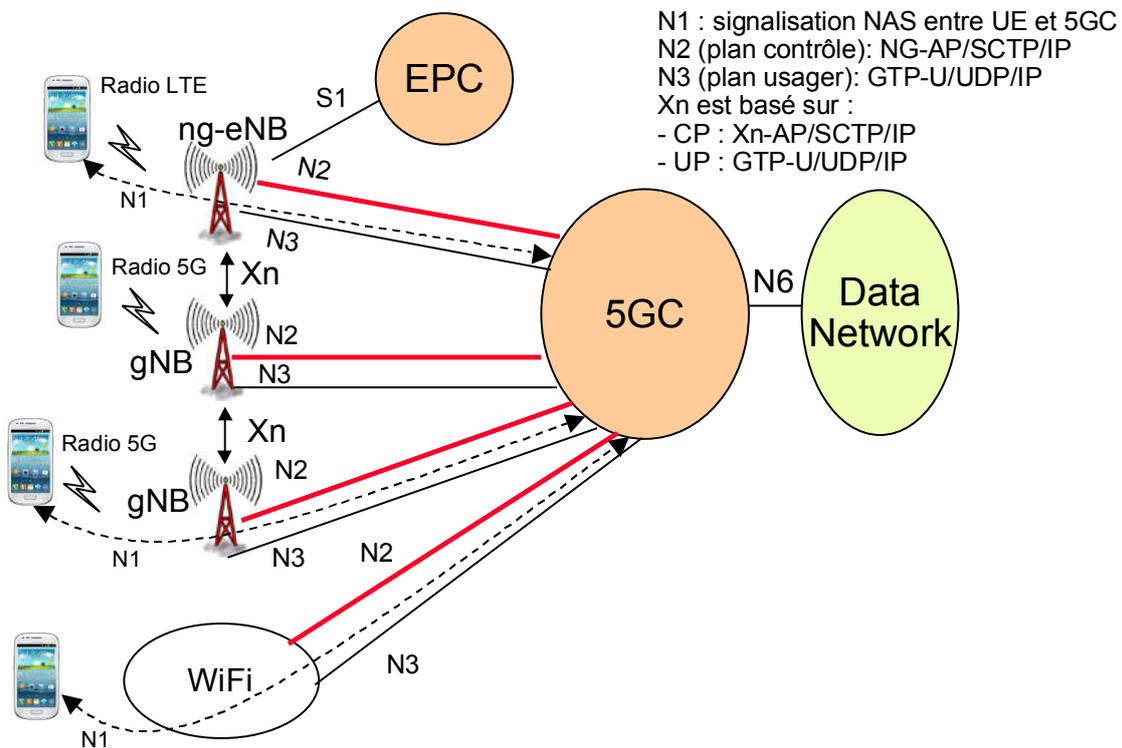


Figure 2 : Interfonctionnement entre New Radio (NR) et réseau cœur 5GC

#### 4.1. Déploiement de la 5G en Mode NSA puis en mode SA

Comme certains opérateurs ont un calendrier de déploiement de la New Radio agressif, ils cherchent à la déployer commercialement avant la disponibilité du cœur de réseau 5G qui suivra plus tard. Il existe deux manières de déployer la radio 5G : mode non-standalone (NSA) et mode standalone (SA) comme montré à la figure 3.

En mode NSA, la New Radio (connue comme gNodeB (gNB)) est colocalisée avec l'eNodeB (eNB) et se connecte au réseau cœur 4G (ePC) via l'interface S1-U pour le trafic du plan usager. Les communications du plan contrôle (e.g., pour EMM et ESM) entre l'UE et le réseau restent sur la radio LTE, et donc le réseau cœur 4G. Dans ce modèle, la radio 5G agit comme une radio secondaire donc le seul but est de booster le débit et la capacité. Cela ne requiert pas de réseau cœur 5G et représente une solution attractive pour certains opérateurs.

Pour déployer la 5G en mode SA, un réseau cœur 5G est requis. Certains opérateurs choisiront ce mode pour leur déploiement 5G initial notamment pour déployer des services

uniquement 5G dans une aire géographiquement limitée sans interfonctionnement avec la 4G. A long terme, il y a un besoin d'intégration avec l'ePC afin de permettre la mobilité entre les accès 4G et 5G et l'intégration avec le RAN LTE évolué qui se connectera directement au réseau cœur 5G.

Sur le long terme, il est attendu que le réseau cœur 5G soit le réseau cœur commun pour tous les types d'accès. Ainsi l'investissement sera transféré de l'ePC au 5GC en ligne avec le taux de migration de la base de clients. Par ailleurs, afin de fournir des nouveaux services avancés, tels que ceux qui requièrent une mobilité élevée, une latence très courte et une indépendance par rapport à l'accès, l'opérateur devra déployer le 5GC. Un tutoriel EFORT est consacré au cœur de réseau 5G : <http://efort.com/index.php?PageID=5&l=fr>  
 Parmi les différentes options NSA, l'option 3X est celle retenue par les opérateurs pour leur déploiement initial de la 5G.

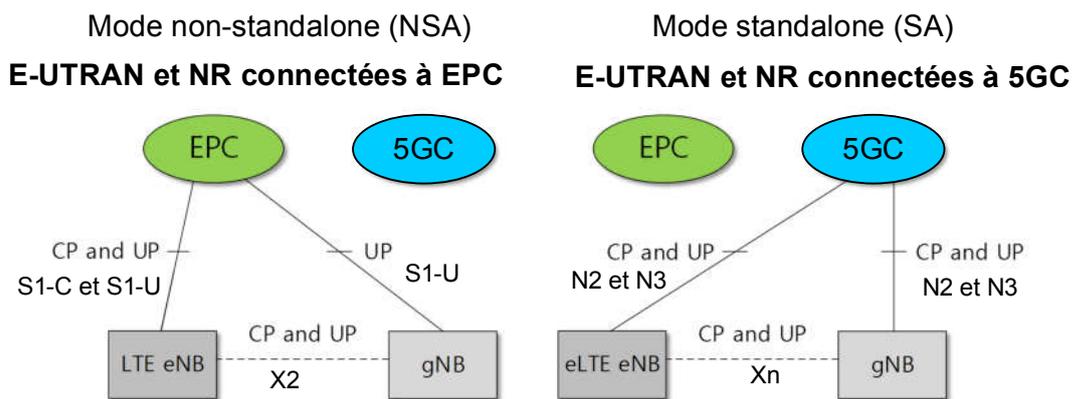


Figure 3 : Mode NSA versus Mode SA

## 5. Du RAN au Cloud RAN (C-RAN)

Un Node B (3G), eNodeB (4G) ou gNodeB (5G) est composé du BBU (Baseband Unit) et du RRH (Remote Radio Head) qui sont respectivement des équipements Indoor et Outdoor. Le RRH est aussi appelé RRU (Remote Radio Unit). Le RRH est une unité radio et le BBU est une unité de traitement de signal bande de base. Le lien entre les deux est en fibre optique ou câble Ethernet via l'interface CPRI (Common Public Radio Interface).

Pour réaliser une architecture C-RAN (Cloud RAN) conforme au 3GPP, il est possible de couper en deux la station de base actuelle, en gardant l'architecture du RRH et la partie processing éloignées l'une de l'autre (Figure 4) ; la partie de la station de base qui était en bas du pylône peut être mise dans un cluster quelque part dans la ville et son raccordement au pylône peut être fait à l'aide d'un câble Ethernet ou d'une fibre optique. En C-RAN (Cloud RAN), afin d'optimiser l'utilisation des BBU fortement et faiblement chargés, les BBU sont centralisés en une seule entité appelée un pool BBU.

L'objectif des opérateurs doit être d'augmenter leurs capacités, d'optimiser l'utilisation du spectre qu'ils ont payé très cher, de réduire leurs coûts d'installation avec des solutions pérennes, de faciliter le déploiement, de réduire la consommation électrique et donc les coûts de fonctionnement de l'architecture choisie, de prendre en compte l'aspect environnemental et surtout de gérer les interférences. La solution est de centraliser.

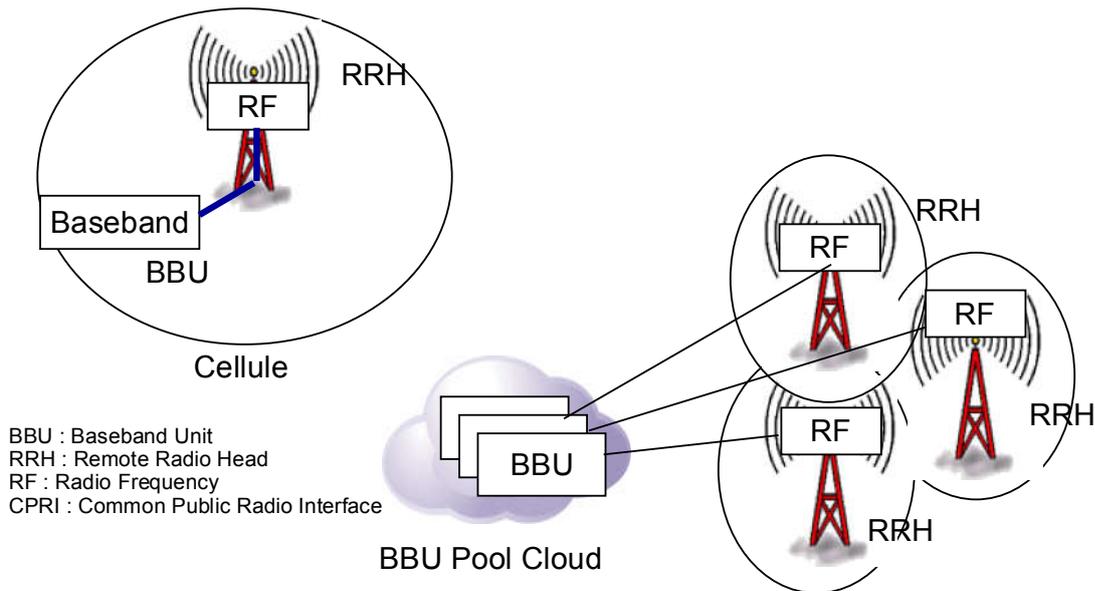


Figure 4 : De RAN à C-RAN

### 5.1. De RAN Backhaul à RAN Fronthaul

Dans les réseaux existants, le segment du réseau qui interconnecte les eNodeB à l'ePC s'appelle le backhaul. La fibre optique ou FH (Fréquence Hertzienne) paquet joue ce rôle avec un backbone IP/Gigabit Ethernet.

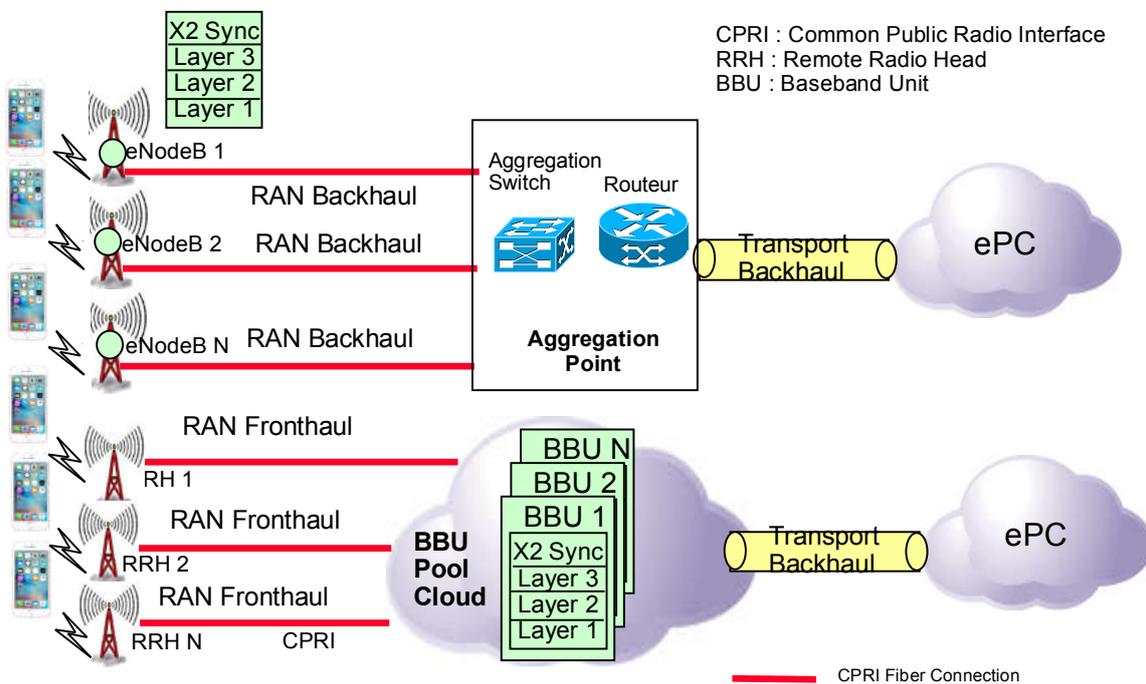


Figure 5 : RAN Fronthaul versus RAN Backhaul

Mais, comme indiqué ci-dessus, la technologie mobile évolue vers les technologies de virtualisation où les ressources radio sont allouée à partir d'un pool commun à différents fournisseurs telco, et leur gestion est centralisée. Cette nouvelle ère reflète en fait des parallèles dans le monde du cloud computing, avec l'apparition des services cloud.

Le paradigme de séparation RRH et BBU change la perception du transport de la communication dans le réseau, du backhauling à l'intégration à la fois d'un segment backhaul et fronthaul.

Dans ce contexte, le backhaul définit la connectivité entre le réseau central C-RAN et le réseau cœur alors que le fronthaul se réfère au segment de connectivité entre le réseau central C-RAN et la petite cellule (Figure 5).

Le module RRH utilise l'interface CPRI (Common Public Radio Interface) pour communiquer avec le BBU. Il s'agit d'une interface numérique pour encapsuler les échantillons radio entre RRH et BBU et qui fonctionne sur fibre. Il ne s'agit pas d'une interface orientée paquet, les signaux sont multiplexés dans un style de timeslot. CPRI offre une latence très courte, une jigue proche de 0, et un taux d'erreur quasi nul. La distance maximum entre RRH et BBU est 17 km aujourd'hui.

La formation EFORT « Introduction à la 5G » décrit le réseau d'accès, le réseau cœur, les services et l'état de la normalisation liés à la 5G.

[http://efort.com/index.php?PageID=21&l=fr&f\\_id=139&imageField.x=2&imageField.y=2](http://efort.com/index.php?PageID=21&l=fr&f_id=139&imageField.x=2&imageField.y=2)