Master – Technologie 4G

Terminologie

Terminal

Le terminal s'appelle UE pour User Equipement (smartphone par exemple) Il est muni d'une carte USIM (Universal Subscriber Identity) fournie par l'opérateur, cette carte contient les données d'abonnement.
La puissance d'un UE est généralement de 0,2 W et sa portée de quelques kilomètres.

Stations de base

C'est un ensemble d'émetteurs-récepteurs nommé (eNodeB – Evolved Node Base) munie d'antennes et placé en un endroit stratégique permettant de fournir le service d'accès au réseau.

Réseau d'accès et réseau cœur

Les eNodeB sont reliés au réseau IP de l'opérateur. Les réseaux des opérateurs sont eux-mêmes interconnectés via Internet

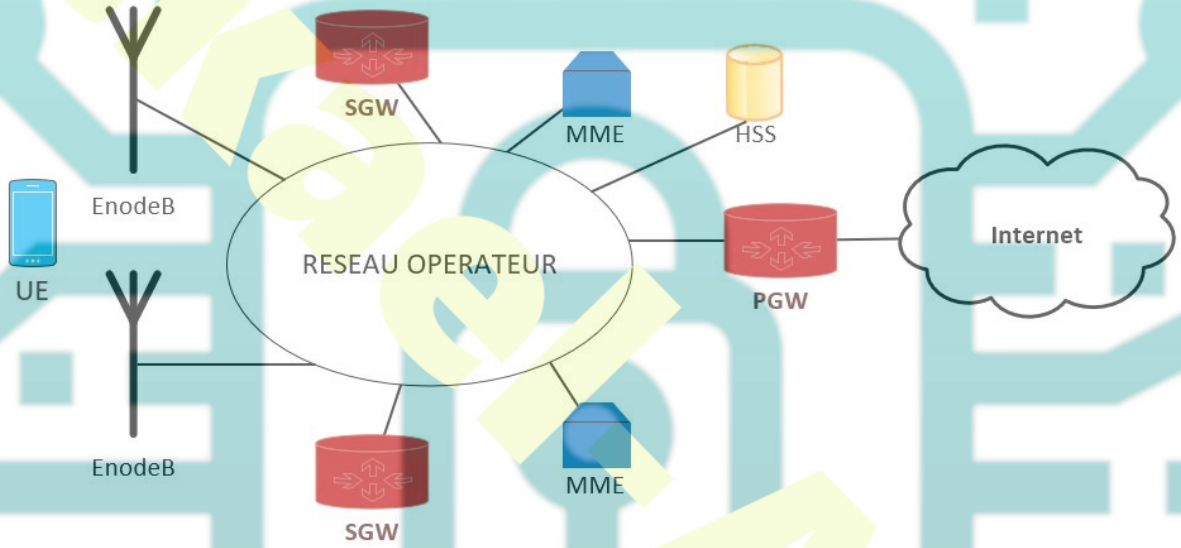
eUTRAN (Evolved Universal Terrestrial Radio Access Network) est un réseau cellulaire qui assure la connexion via des ondes radio UHF entre les terminaux mobiles et le cœur de réseau de l'opérateur mobile.

Réseau cœur, EPC = Evolved Packet Core

Les différents réseaux

Version	Nom	Type accès
2	GSM	TDMA
2.5	GPRS-EDGE	Accès paquet/nouvelle modulation
3	UMTS	CDMA
3.9	HSDPA	CDMA + accès paquet/nouvelle modulation
4	LTE-LTE+	OFDMA

Architecture



Couvrir une zone

Pour arriver à cet objectif, l'opérateur découpe le territoire en cellules via un eNodeB.

Dans les zone à forte densité, les eNodeB sont déployés pour fournir une capacité (en Mbit/s par km2) supérieure au trafic engendré par les clients.

Dans une zone rurale les eNodeB sont déployés assurer une couverture minimale (le trafic étant moindre)

https://enb-analytics.fr/

Voir la répartition en France

Les passerelles (Packet GateWay)

Le **PGW**, assure un lien unique entre le réseau IP de l'opérateur et Internet. Il achemine les données internet vers le terminal et réciproquement.

Les passerelles régionales (Serving Gateway)

Les **SGW** sont des passerelles qui s'occupent d'une zone géographique et qui communiquent avec les PGW et les terminaux via les eNodeB.

Équipements de contrôle dans le réseau cœur

Le **HSS** (Home Subscriber Server) est la base de données des abonnés. Elle comporte les informations de tout abonné autorisé à utiliser le réseau de l'opérateur et sa localisation dans ce réseau. Cet équipement n'échange que des données de signalisation et pas les données transférées par l'utilisateur.

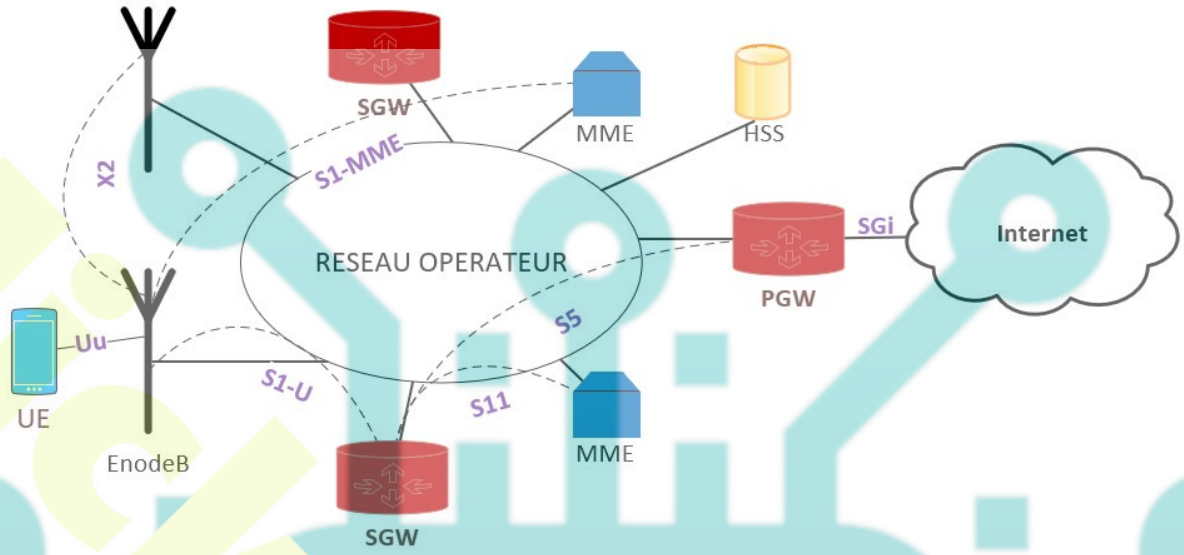
Gestion de la mobilité

Le **MME** (Mobility Management Entity) est chargé de dialoguer avec le HSS pour récupérer les profils et les données de sécurité des terminaux situés dans la région du MME et dialogue également avec les eNodeB.

Il sélectionne le PGW et les SGW qui doivent être utilisés. Il joue le rôle d'un cache pour le HSS.

Les interfaces

Tous les équipements du réseau possèdent la pile IP et dialoguent entre eux, soit parce qu'ils sont reliés directement, soit par l'intermédiaire des routeurs IP.



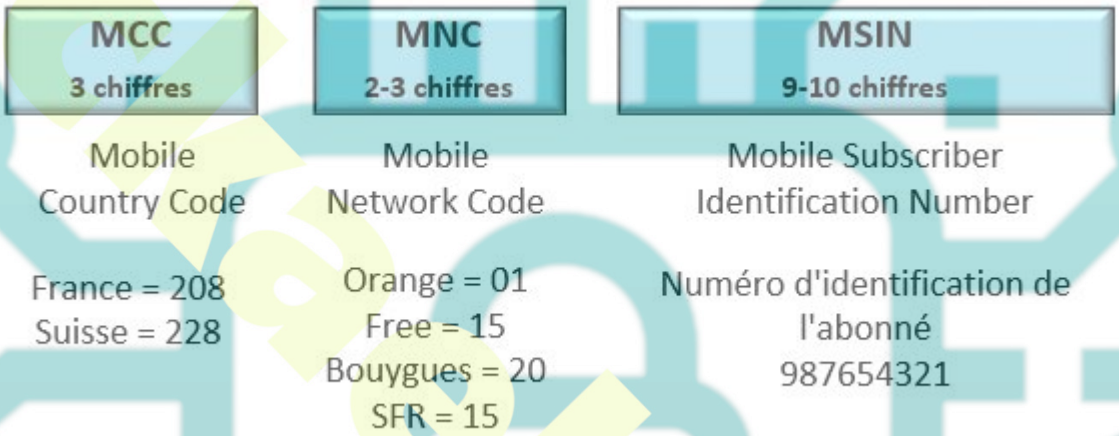
Ces interfaces sont numérotées en fonctions de leurs rôles :

- Interface **SGi** se trouve entre le PGW et le réseau IP externe (Internet)
- Interface **S5** se trouve entre le SGW et le PGW (d'un même réseau) – transport de données et messages de signalisation
- Interface **S11** se trouve entre le SGW et le MME – Transport de messages de signalisation
- Interface S6a se trouve entre le MME et le HSS – Transport de messages de signalisation
- Interface **S1–MME** se trouve entre l'eNodeB et le MME – Transport de messages de signalisation
- Interface **S1-U** se trouve entre l'eNodeB et le SGW – Transport de données utilisateurs
- Interface **X2** se trouve entre 2 eNodeB – Transport des données utilisateurs et de messages de signalisation
- Interface **Uu** ou interface radio se trouve entre le terminal (UE) et l'eNodeB – Transport des données utilisateurs et des messages de signalisation
- Interface **S8** se trouve entre le SGW et le PGW d'un autre réseau
- **EIR** (Equipment Identity Register) : base de données des terminaux (volés) interface **S13** avec le MME
- **PCRF** (Policy and Charging Rules Function) : gestion de la qualité de service Interface **Gx** avec le PGW

Procédure de connexion

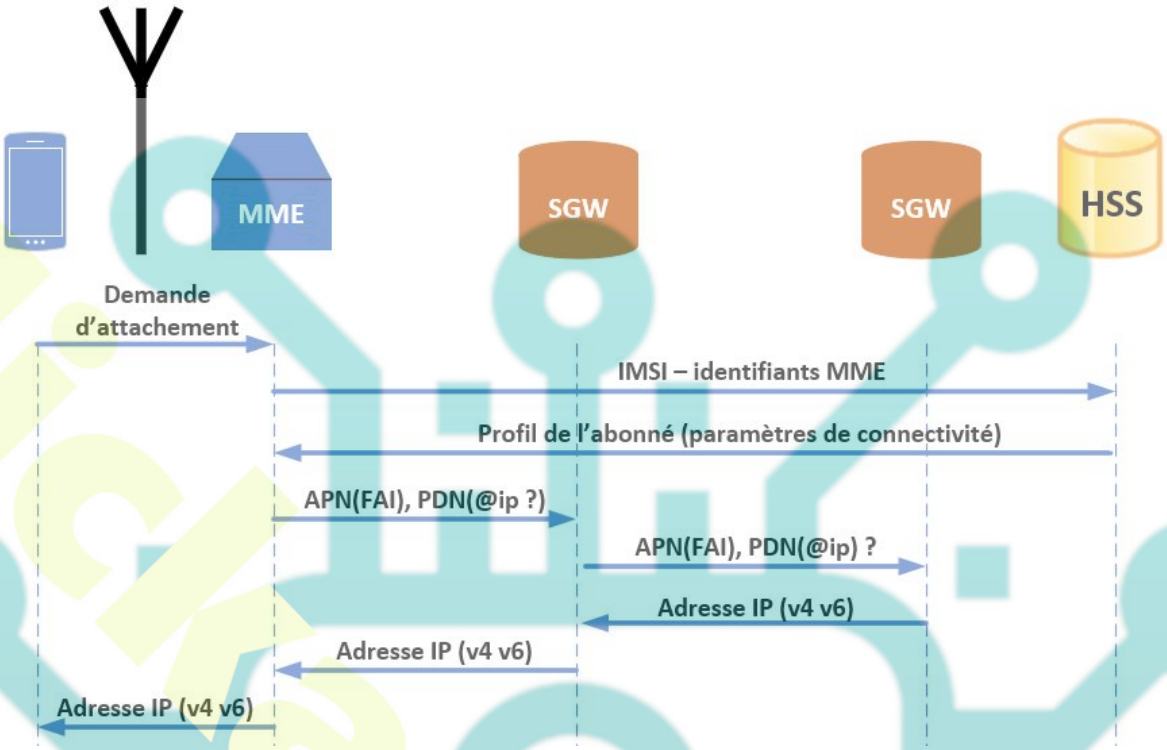
Pour se connecter, l'abonné possède une carte USIM qui lui permet de se rattacher au réseau de l'opérateur. Cette carte possède un identifiant appelé IMSI (International Mobile Subscriber Identity)

L'IMSI est un numéro unique, qui permet à un réseau de téléphonie mobile d'identifier un usager. Il se décompose comme suit :



Lorsque l'on se connecte la première fois, l'UE (terminal) recherche le réseau de l'opérateur en écoutant le **beacon** émit par l'eNodeB.
Le terminal indique également le type de service APN (quel PWG utiliser) et envoie son IMSI à l'eNodeB.

Schéma de la connexion simplifié



1. L'eNodeB envoie cette demande au MME.
2. Le MME regarde s'il n'a pas l'information en cache, dans le cas où il ne l'a pas, il transmet la demande au HSS (base de données des profils)
3. Le HSS transmet le profil au MME qui le stocke
4. Le MME envoie la demande d'adresse IP au SWG qui la transmet au PWG (seul élément ayant autorité pour distribuer une adresse IP dynamique)
5. Le PWG envoie l'information au SWG qui la transmet au MME
6. Le MME la transmet à l'UE via l'eNodeB

Les principaux mécanismes de sécurité

Authentification

Pour se protéger de l'utilisation non autorisée du réseau, le terminal et le HSS vont négocier une authentification mutuelle par l'intermédiaire d'un défi.
Ceci étant, pour des raisons de mise à l'échelle, l'authentification ne se fait pas directement entre l'UE et le HSS mais entre l'UE et le MME. En effet, le HSS délègue cette négociation au MME pour répartir la charge.

Chiffrement

Pour se protéger de l'écoute, les opérateurs utilisent les algorithmes de chiffrement standardisés (Snow 3G, AES)
Le chiffrement va permettre de protéger les échanges entre l'UE, l'eNodeB et le MME.

Intégrité

Pour vérifier qu'un message n'est pas modifié, on utilise un système de hachage (SHA2)

Identité temporaire

Pour permettre une authentification moins lourde et pour empêcher un pirate de suivre nos déplacements.

Procédure d'authentification

L'UE envoie son IMSI au MME qui le transfère au HSS.

Grâce à la clé partagée (K) entre lui et l'UE, le HSS (en s'appuyant sur l'algorithme AES) calcule un chiffre aléatoire (RAND) + un défi (XRES) + un jeton d'authentification (AUTN)+ clé K*asme*

K = clé secrète partagée en le HSS et la carte USIM
XRES = est le résultat du chiffrement de la variable aléatoire avec la clé secrète contenue sur la carte USIM
AUTN = obtenu avec la clé partagée K+RAND+SQN
SQN = chiffre incrémenté à chaque nouvelle connexion pour éviter le rejeu
Kasme = la clé Kasme est créée à partir de RAND des identifiants de l'opérateur et du SQN.

- Le HSS envoie la séquence au MME pour qu'il s'occupe de la suite des échanges et notamment la création des clés de chiffrement.
- Le MME transmet cette séquence à l'UE.
- L'UE de son côté va calculer le défi et créer son RES et le AUTN avec sa clé K, et créer les différentes clés de chiffrement pour la suite des communications.
- L'UE vérifie si le AUTN est le même des 2 cotés. Dans le cas où le résultat est positif, le client a authentifié l'opérateur.

- L'UE transmet son RES avec la réponse d'authentification au MME qui le compare au XRES. Si les deux résultats sont équivalents, l'UE est authentifié.

Procédure de chiffrement

La carte USIM de l'UE possède une clé K partagée avec le HSS. Seules ces deux entités possèdent la clé. Cependant, ce n'est pas cette clé qui va être utilisée pour la génération des autres clés, c'est une nouvelle clé (Kasme)

- Le HSS envoie la clé Kasme au MME pour qu'il gère ensuite la distribution des différentes clés.
- De son côté, l'UE calcule à son tour la clé Kasme. La clé Kasme contient aussi les informations (MCC + MNC) de l'opérateur ce qui permet de garantir qu'elle n'est utilisable que dans le réseau de l'opérateur.
- Ensuite, grâce à la clé *Kasme*, on fabrique une clé KEnobeB sur le MME et sur l'UE. Cette clé n'est pas utilisée pour le chiffrement des échanges mais pour créer les autres clés.
- On crée les clés de chiffrement et d'intégrité pour la communication entre l'UE et le MME (KNASEnc et KNASInt)
- On fabrique les clés de chiffrement et d'intégrité pour la signalisation entre l'UE et l'EnodeB (KARCEnc et KARCEInt) + la clé utilisateur (KUPEnc) pour les données.

Utilisation de l'identité temporaire

Pour éviter qu'un pirate ne puisse localiser un utilisateur en décryptant le IMSI, le réseau alloue à chaque UE un numéro temporaire appelé TMSI.

Ce TMSI (Temporal IMSI) est créé après une première authentification classique utilisant l'IMSI.

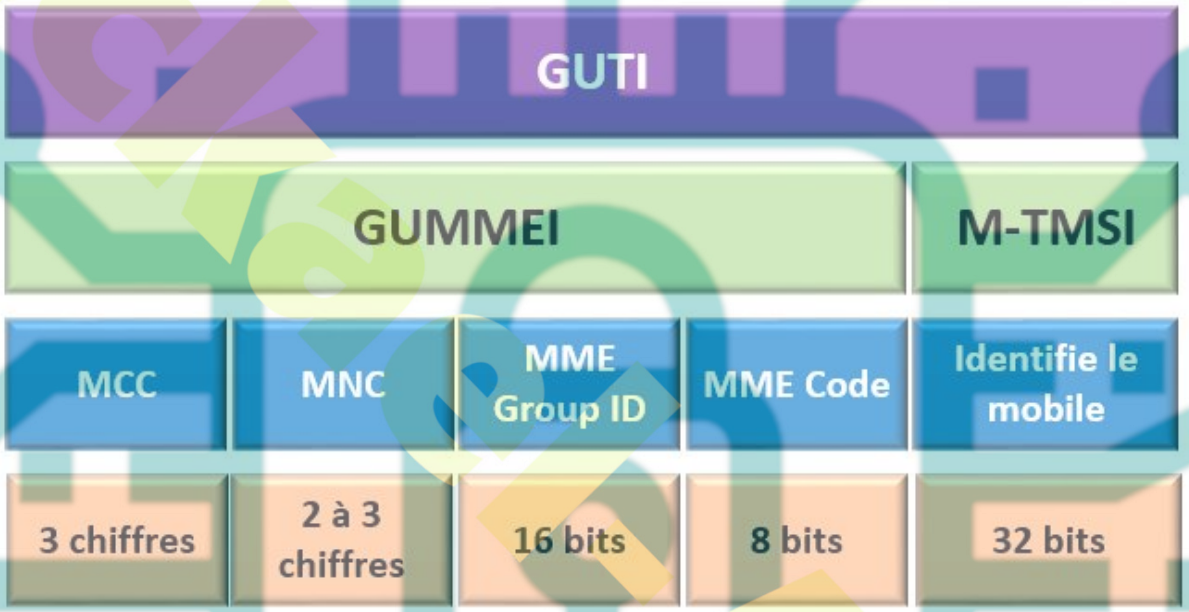
Une fois la première authentification effectuée, on dispose des différentes clés de chiffrement et les échanges qui suivent sont sécurisés et notamment l'envoie du TMSI.

A partir de cette étape, le HSS n'est plus jamais sollicité.

Ce TMSI est fréquemment changé selon la politique de l'opérateur, en cas de changement de MME ou dans le cas où l'on change de carte USIM.

Cependant, le TMSI est trop petit (32 bits) et il existe le risque que 2 utilisateurs obtiennent le même TMSI. Pour éviter cela, on utilise le GUTI (Globally Unique Temporary Identifier)

L'objectif du GUTI est de fournir une identité unique à l'UE sans dévoiler l'IMSI. Le GUTI est composé du GUMMEI (Globally Unique Mobility Management Entity Identifier) pour identifier le MME de manière unique et du TMSI pour identifier de manière unique l'UE.



Si l'UE change de MME et envoie le GUTI, le nouveau MME, grâce aux informations situées dans le GUMMEI, peut découvrir l'ancien MME et lui relayer l'information. Il obtient en retour l'IMEI et les informations d'authentification.

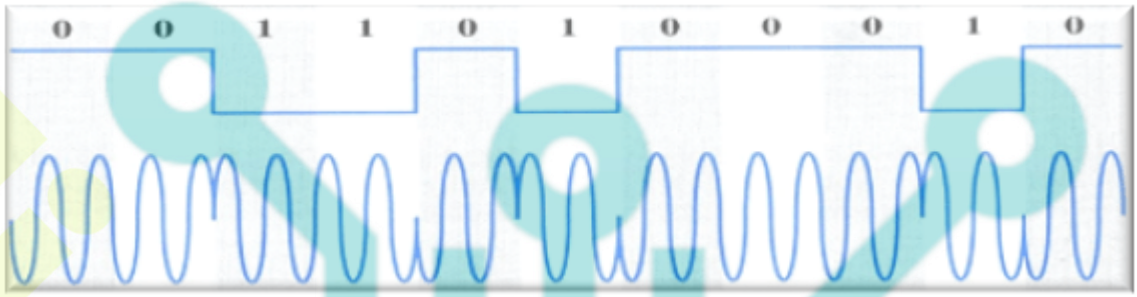
Lorsque le MME à besoin de solliciter un UE en mode veille, il envoie un message de PAGING aux eNodeB qu'il gère et ceux-ci répondent le TMSI de l'abonné, ce qui permet au MME de le contacter.

Le S-TMSI est utilisé dans le cadre de la procédure de Paging. Lorsque l'UE est enregistré, le S-TMSI est déduit du GUTI à partir des 40 derniers bits du GUTI.

Transmission des ondes radio

Le principe du PSK (modulation par changement de phase)
Utiliser un signal de 0 degré pour transporter un zéro et un signal de 180° pour

transporter un 1.

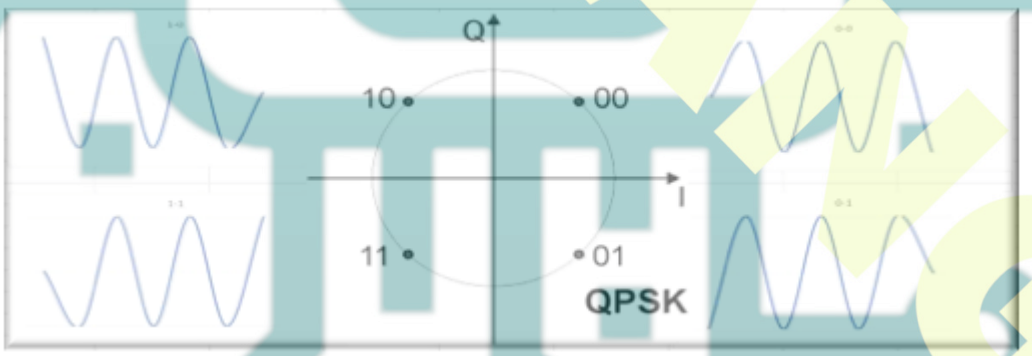


Le principe du BPSK Le BPSK est la forme la plus simple du PSK. Il utilise deux phases qui sont séparées de 180°.

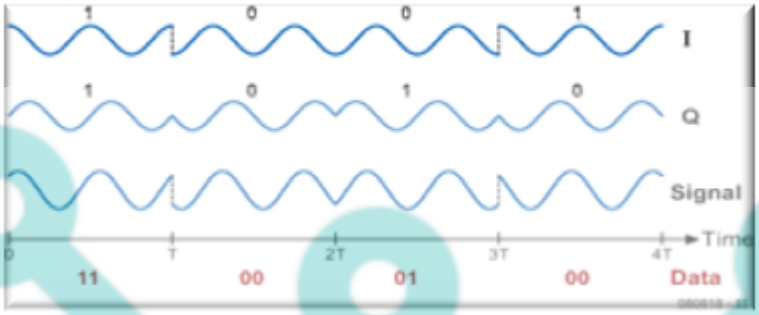
Si l'on veut transporter une porteuse à 1Ghz (10^9 hertz) avec un signal 1Mb/s (10^6 bits) on aura 10^9/10^6=1000 cycles de porteuse pour transporter un bit.

Le problème est que l'on ne transporte qu'un bit par symbole soit un 1 ou un 0

Le principe du QAM (modulation d'amplitude en quadrature)
Cette modulation joue à la fois sur la phase et l'amplitude ce qui permet de transporter plus d'un bit par symbole.
16-QAM (4 bits) 64-QAM (6bits) 256-QAM (8 bits)



On peut considérer le QPSK comme deux modulations indépendantes. Avec cette interprétation, les bits pairs (ou impairs) sont utilisés pour moduler la composante In-Phase (I) donc le cosinus, tandis que les bits impairs (ou pairs) sont utilisés pour la Quadrature-phase (Q) donc le sinus.



Bits impairs correspondant à la composante I (Phase) **1** 1 **0** 0 **0** 1 **1** 0
Bits pairs correspondant à la composante Q (quadrature) 1 **1** 0 **0** 0 **1** 1 **0**

Cependant, lorsque l'on utilise QPSK on va 2 fois plus vite mais les états sont plus proches les uns des autres ce qui peut entrainer des recouvrements et donc des erreurs.

La correction d'erreurs

Pour corriger les erreurs on utilise des messages redondants pour que le récepteur puisse retrouver le message original.

Cette correction est appelée FEC (Forward Error Correction). Le taux de codage est calculé comme suit : information utile / information transmise (1/3 en forte, 1/1 en faible)

Par exemple, on peut envoyer 3 messages de redondance en cas de forte perturbation.

BONJOUR —BANJOUR – RONJAUR – BOJBOUH — BONJOUR

Le récepteur va pourvoir facilement reconstruire le message d'origine.

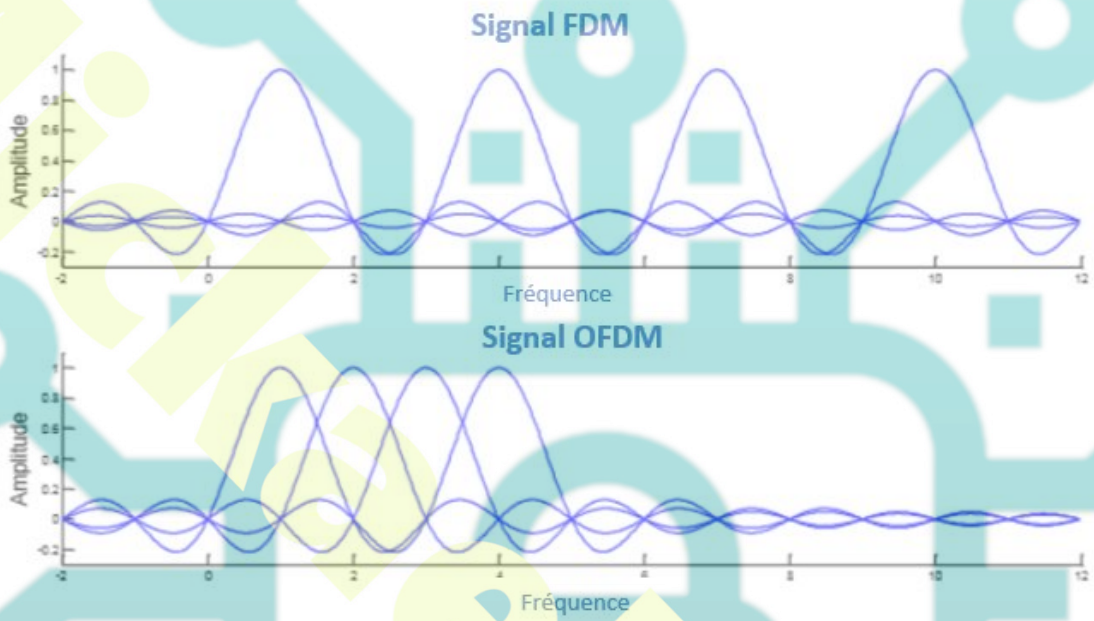
LTE est capable d'adapter son débit en fonction des conditions de propagation en temps réel et indépendamment pour chaque utilisateur.

L'association de la modulation et du taux de codage définit le MCS (Module Coding Scheme). Il existe 29 MCS utilisables (de QPSK à QAM).

LTE utilise, comme dans le cadre du WIFI, l'OFDM qui permet notamment le multi-trajets.

L'**OFDM** est un procédé de codage de signaux par répartition en fréquences orthogonales sous forme de multiples sous porteuses. Cette technique permet de

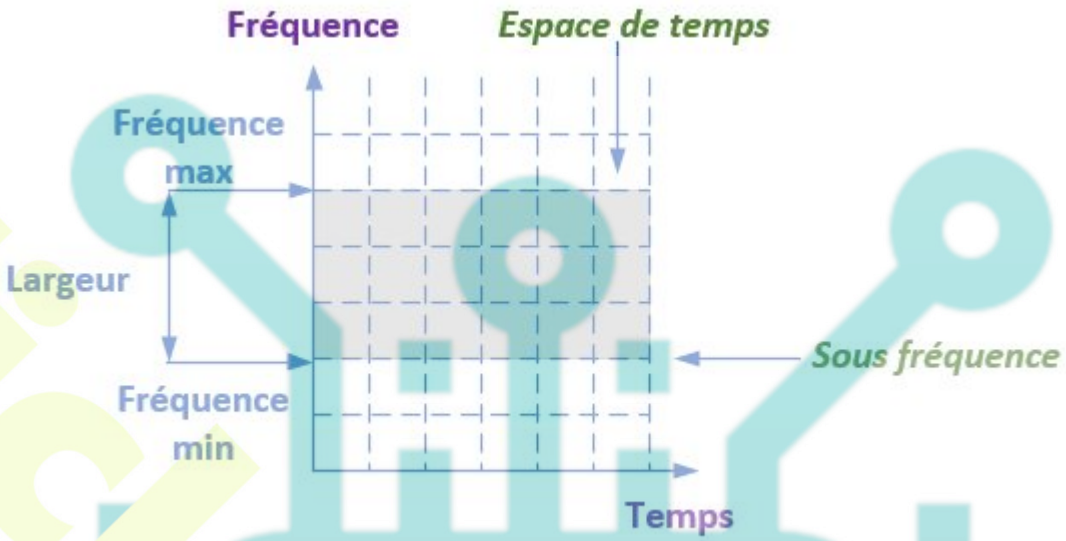
transmettre plus de signaux et de gagner de la bande passante. Cependant, pour éviter la confusion liée au chevauchement des fréquences, on s'intéresse aux crêtes et on annule les sous porteuses avoisinantes.



Allocations de ressources

Le partage des ressources s'effectue entre les différents utilisateurs et les différents usages (signaux/données). Ce partage doit être dynamique et s'adapter dans le temps.

LTE utilise différentes fréquences (700Mhz, 1.8Ghz, 2.6Ghz…) et différentes largeurs de bandes de 1.4 Mhz et 20Mhz. Ensuite, ces fréquences sont divisées en sous fréquences et chaque temps est alloué aux besoins des utilisateurs.

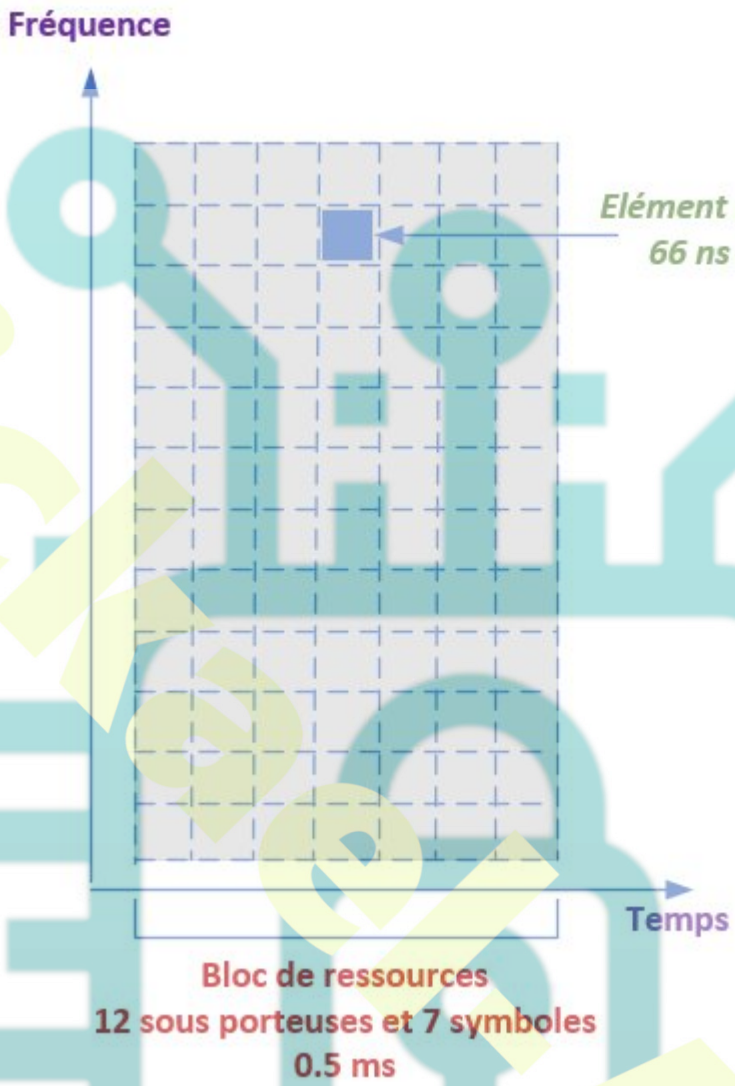


Un resource block (RB)

Pour partager les ressources, LTE utilise 12 sous porteuses de 15Khz (180 Khz) et 7 symboles de 66 ns (0.5 ms)

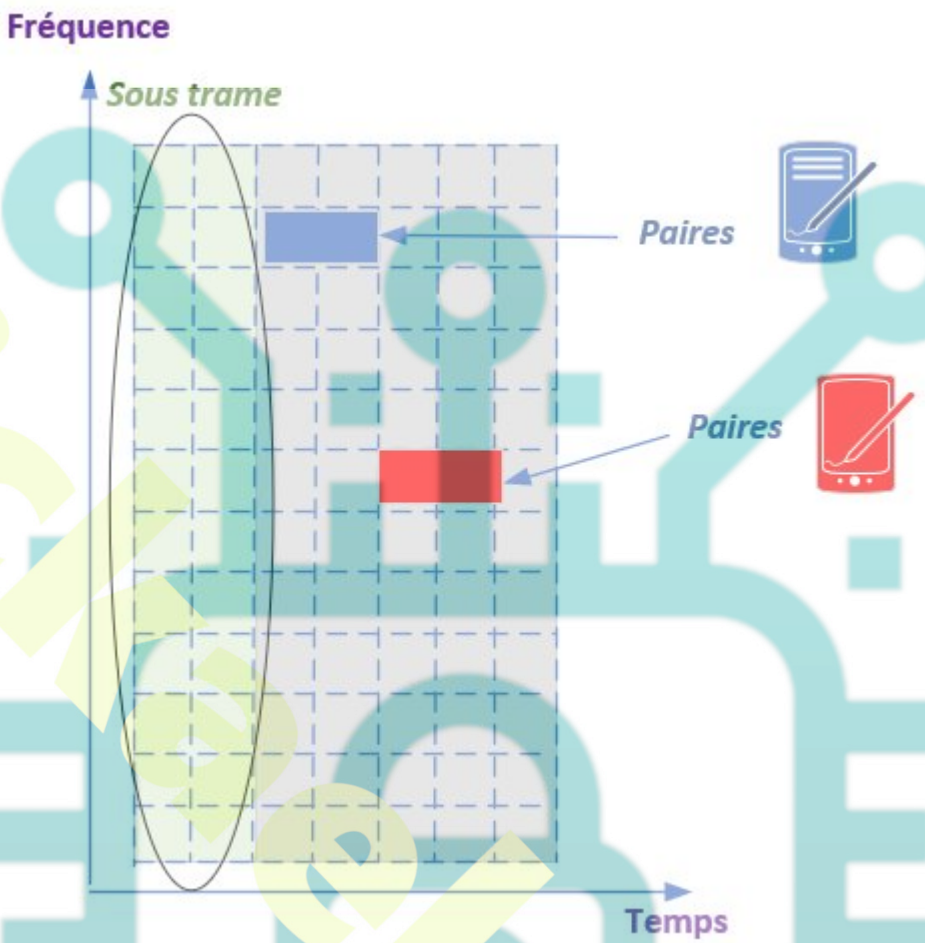
La plus petite valeur est l'élément de ressources (RE). Ces éléments sont regroupés en bloc de ressources (RB). Un bloc de ressources contient 84 éléments de ressources (7*12).

Étant donné que LTE peut fonctionner sur plusieurs bandes de fréquence, il utilisera 6 blocs de ressources en 1.4 Mhz et 100 blocs de ressources en 20Mhz.



LTE affecte les blocs de ressources par paires pour les utilisateurs, et cette paire dure donc 1 ms (2*0.5).

Le motif formé toutes les millisecondes par l'ensemble des blocs de ressources forme une sous trame qui est l'unité de base pour l'allocation.



Au sein d'un bloc de ressources, certains des éléments de ressources (RE) sont utilisés pour la synchronisation, l'estimation du canal et les fonctions de contrôle (ACK, allocations). Un ensemble de ressources réservées est un canal physique.

Bloc de transport

C'est un bloc de données à transmettre sur une seule sous trame, sur un ou plusieurs blocs de ressources. C'est l'eNodeB qui alloue un volume de données à chaque terminal toutes les 1 ms.

La taille des blocs est variable en fonction de la modulation, du codage. Ce couple modulation/codage est appelé le MCS (Modulation and Coding Scheme)

Tableau index valeurs MCS

Nombre de paires de blocs de ressources										
Index MCS		1	2	3	4	5	6	...25	...50	...100
0	QPSK	16	32	56	88	120	152	680	1384	2792
1	QPSK	24	56	88	144	176	208	904	1800	3624
2	QPSK	32	72	144	176	208	256	1096	2216	4584
3	QPSK	40	104	176	208	256	328	1416	2856	5736
4	QPSK	56	120	208	256	328	408	1800	3624	7224
5	QPSK	72	144	224	328	424	504	2216	4392	8760
6	QPSK	328	176	256	392	504	600	2600	5160	10296
7	16 QAM	104	224	328	472	584	712	3112	6200	12216
8	16 QAM	120	256	392	536	680	808	3496	6968	14112
9 et 10	16 QAM	136	296	456	616	776	936	4008	7992	15840
16,17	64 QAM	280	600	904	1224	1544	1800	7736	15364	30576
23	64 QAM	488	1000	1480	1992	2472	2984	12576	25456	51024
24	64 QAM	520	1064	1608	2152	2664	3240	13536	27376	55056
25	64 QAM	552	1128	1736	2280	2856	3496	14112	28336	57336
26	64 QAM	584	1192	1800	2408	2984	3624	15264	30576	61664
27	64 QAM	616	1256	1864	2536	3112	3752	15840	31704	63776
28	64 QAM	712	1480	2216	2984	3752	4392	18336	36696	75376

- L'intersection ligne/colonne donne la taille en bit. On peut donc calculer le débit obtenu.

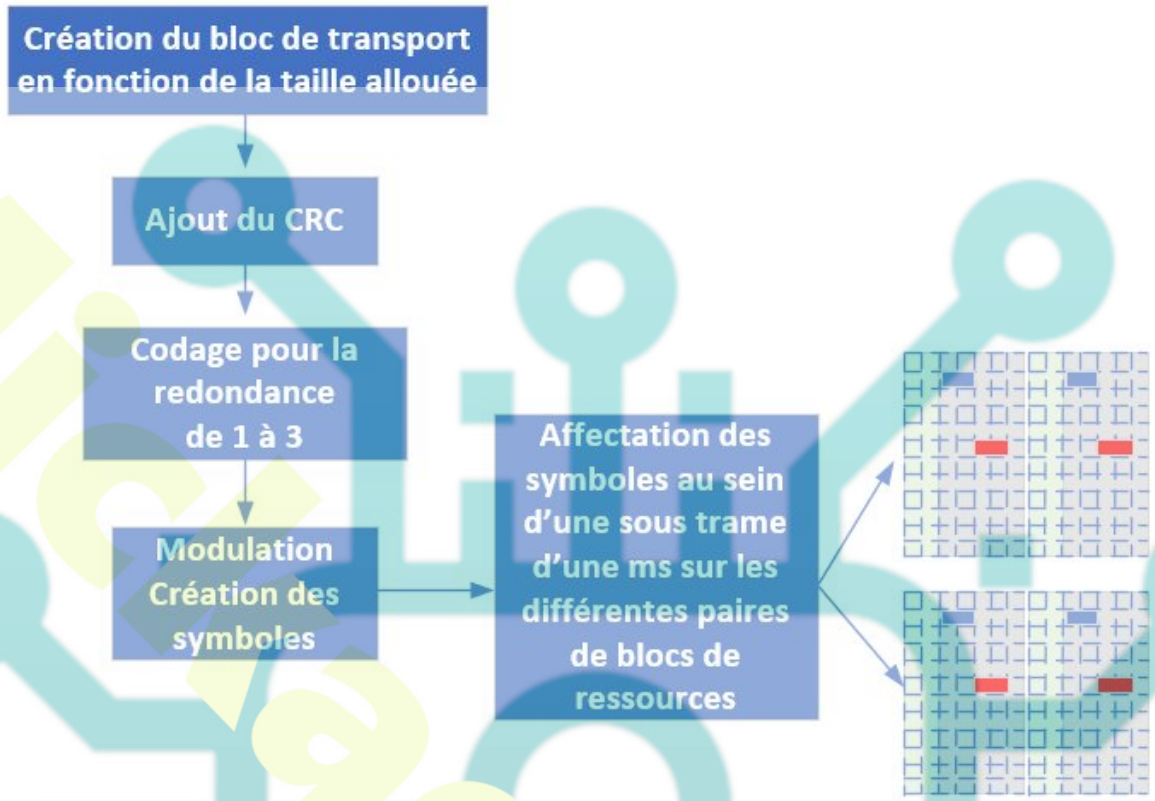
Par exemple, si l'on prend l'intersection 0/1 on obtient la valeur 16 bits pour 1 milliseconde, on multiplie 16 par 1000 (1000 ms = 1 seconde) et on obtient 16 Kb/s. Pour l'intersection 28/100 avec le même calcul on obtient le débit max de 75 Mb/s (version de base de LTE)

On peut également observer que certaines valeurs reviennent plusieurs fois. Cette redondance peut être utile en cas de dégradation du lien et de changement de modulation.

En effet, imaginons que l'on est constitué un bloc de 256 bits avec un MCS de 8 et 2 paires de blocs, si les conditions se dégradent, on va pouvoir se rabattre sur un MCS de 4 utilisant 4 blocs de ressources pour transférer ce bloc de données.

Chaine de transmission

Le processus suivant se répète toutes les 1 ms.



Allocations des paquets

Pour l'allocation des ressources on utilise un ordonnanceur et on alloue les ressources que si le besoin de transmission est effectif. Cet ordonnanceur est géré par l'eNodeB pour la voie montante et descendante.

L'eNodeB informe les terminaux de l'allocation des ressources via une table d'allocation qui est publiée à chaque sous trames.

Comme en radio tout le monde entend ce qui se passe, il est nécessaire de pouvoir adresser les données au bon élément.

Pour allouer les paquets, l'eNodeB utilise une adresse spécifique pour repérer les mobiles – le RNTI (Radio Network Temporaly Identifier)

Cette adresse est codée sur 16 bits, c'est elle qui utilisée pendant toute la communication entre le mobile et l'eNodeB et est limitée à la cellule où se trouve le mobile.

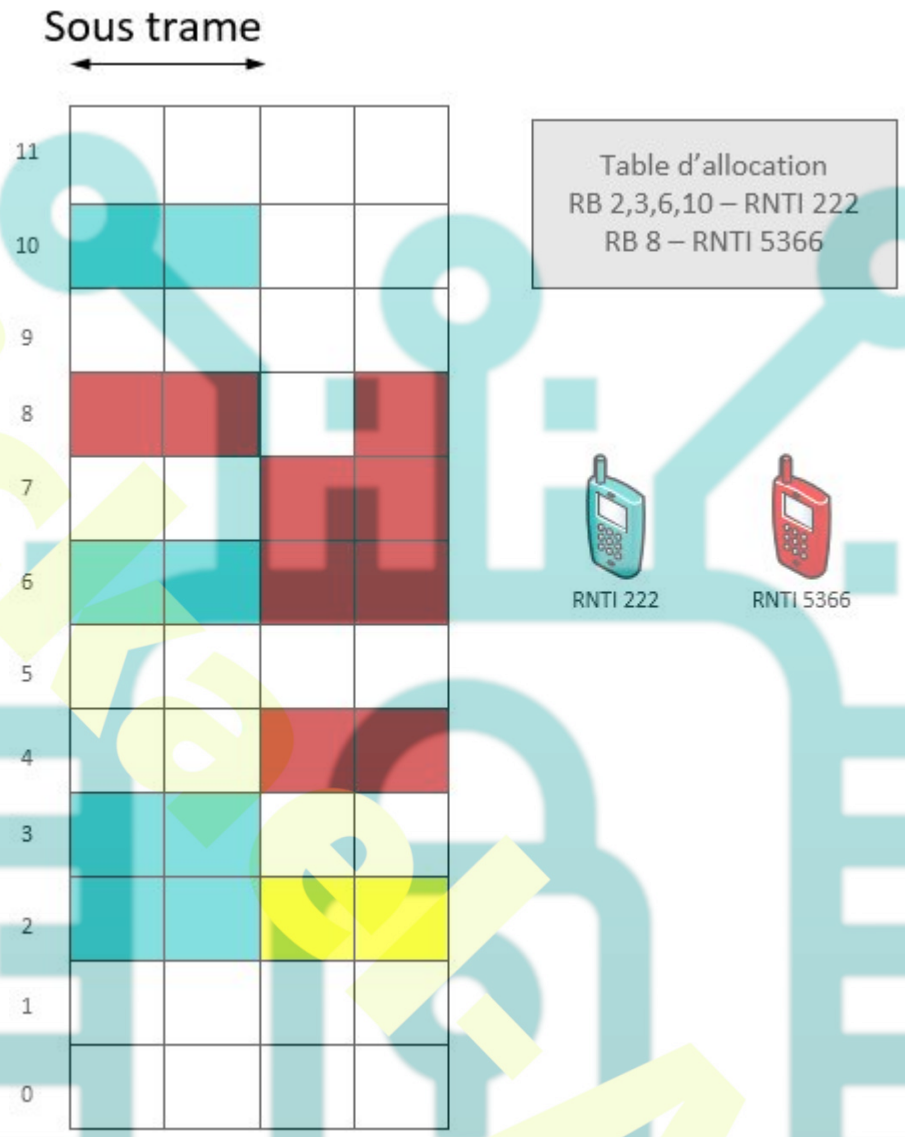
On utilise cette adresse et non celles vues précédemment car les adresses précédentes doivent identifier le mobile dans tout le réseau et sont donc trop longues pour les besoins

de communication avec l'eNodeB. C'est pour cela que l'on utilise une adresse plus courte codée sur 16 bits et valable uniquement dans une cellule.

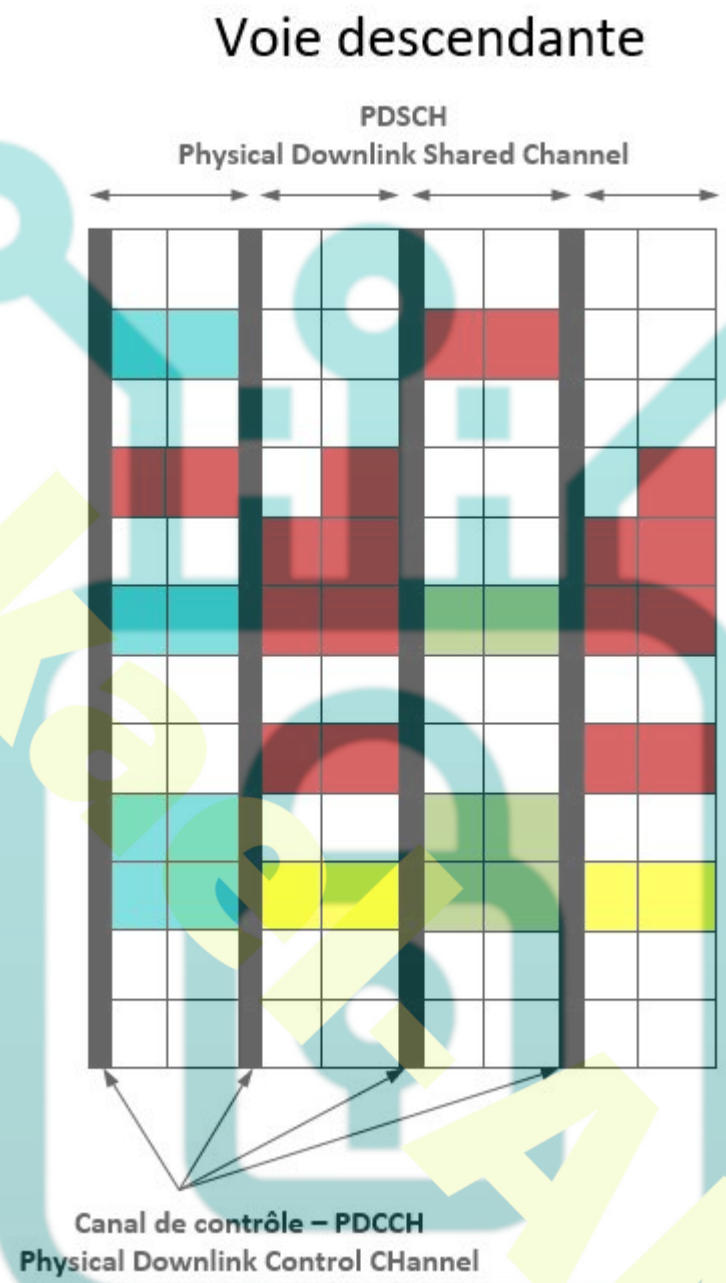
Sachant qu'il y a moins de terminaux dans une cellule que dans le réseau complet, le codage sur 16 bits permet de gérer jusqu'à 65462 (de 61 à 65523) terminaux au sein d'une cellule, les 74 adresses restantes sont réservées au broadcast, paging…

Allocation sur la voie descendante

A chaque sous trame, l'eNodeB détermine l'allocation des ressources à affecter à chaque mobile. Pour que les mobiles sachent qu'il y a des données pour eux, l'eNodeB publie dans chaque sous trames une table d'allocation des ressources qu'il positionne dans un canal de contrôle.



L'avantage de cette technique est d'éviter qu'un terminal ne consomme des ressources inutilement. Le terminal sait lorsqu'il doit travailler ou se remettre en état de veille entre chaque sous trame. Au final le mobile ne consomme de l'énergie que pour lire des RB.

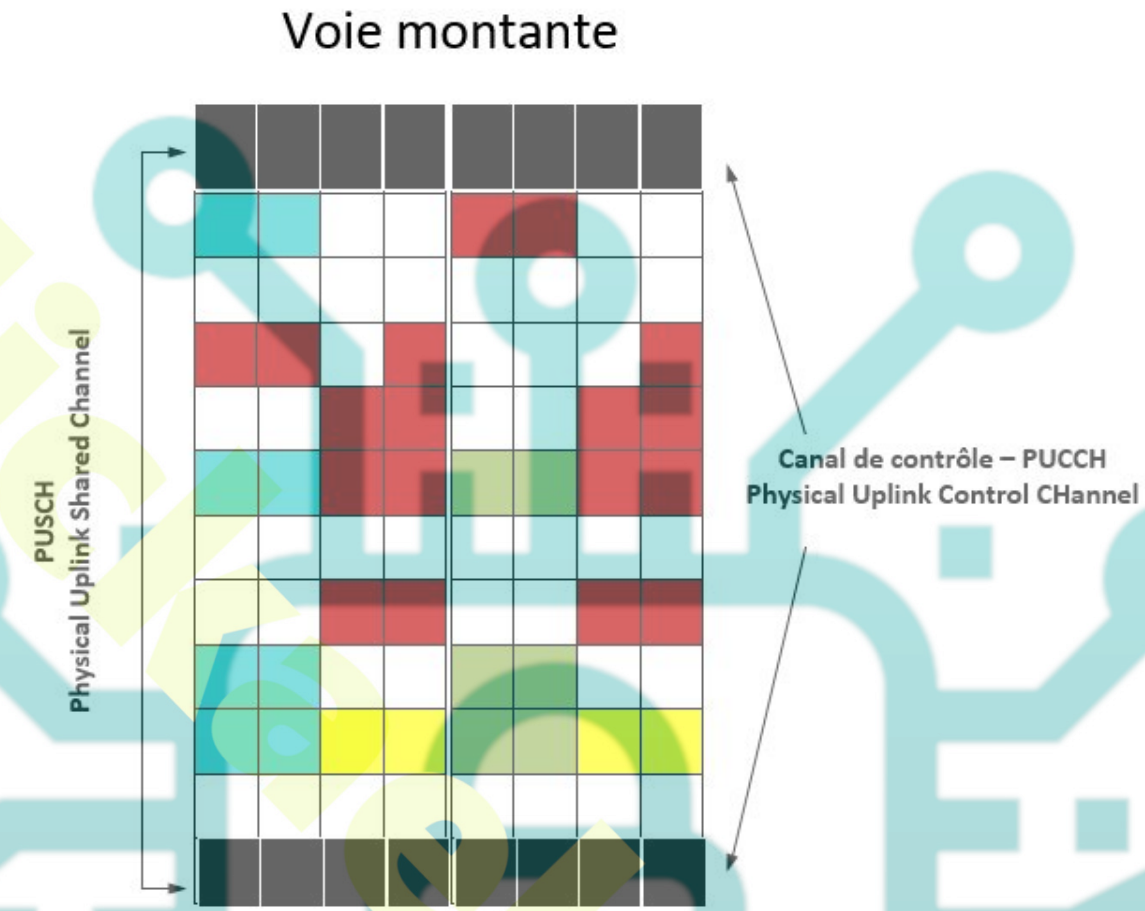


Allocation sur la voie montante

C'est également l'eNodeB qui alloue les ressources, il utilise la table des ressources qu'il a publié dans la voie descendante. Cependant, la grande différence entre la voie descendante et montante, c'est que l'eNodeB ne sait pas à priori si un mobile a besoin d'envoyer des données.

Le mobile doit dans un premier temps, formuler une demande auprès de l'eNodeB via un canal de contrôle, puis les ressources lui sont allouées.

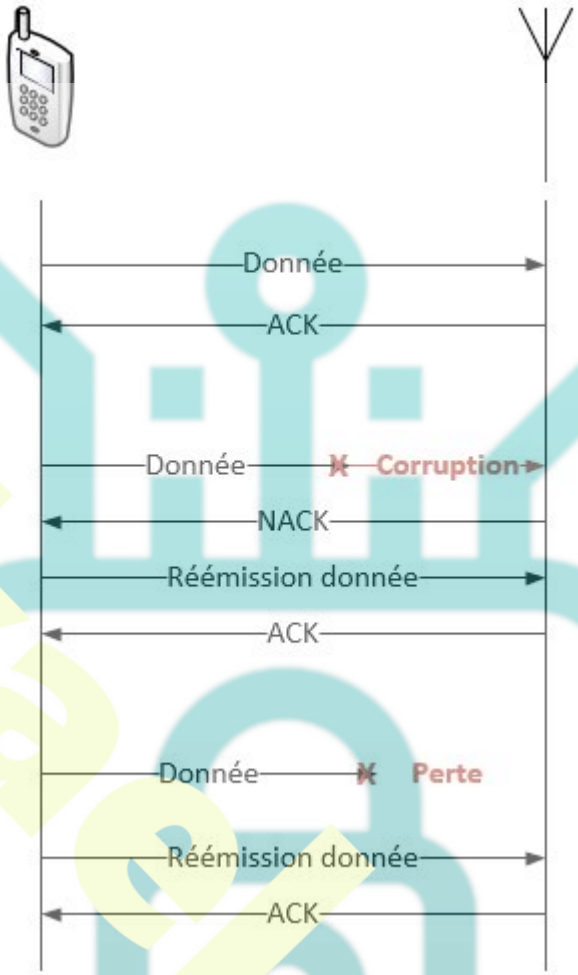
Pour contourner le fait que le mobile ne soit pas prêt à émettre lorsqu'il reçoit l'allocation, on décale son droit d'émettre de 4 sous trames (4 ms plus tard)



Détection des erreurs – couche MAC

La couche physique détecte une partie des erreurs, mais cela ne suffit pas. La couche MAC va utiliser d'une part un CRC et d'autre part un ACK via ARQ (Automatic Repeat Request).

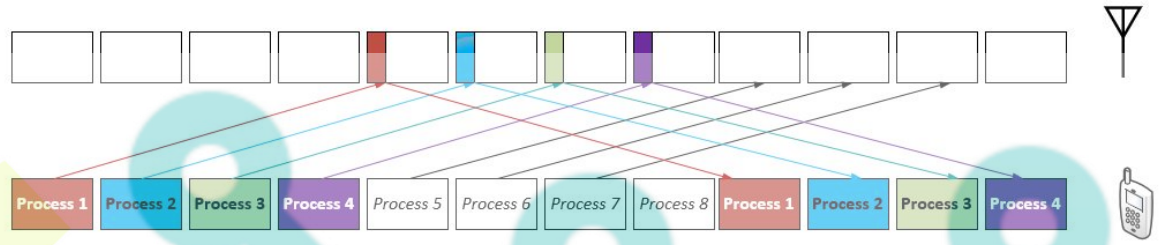
Processus de SEND and WAIT ARQ



Au bout d'un certain nombre de réémission, on abandonne aux couches supérieures le problème.

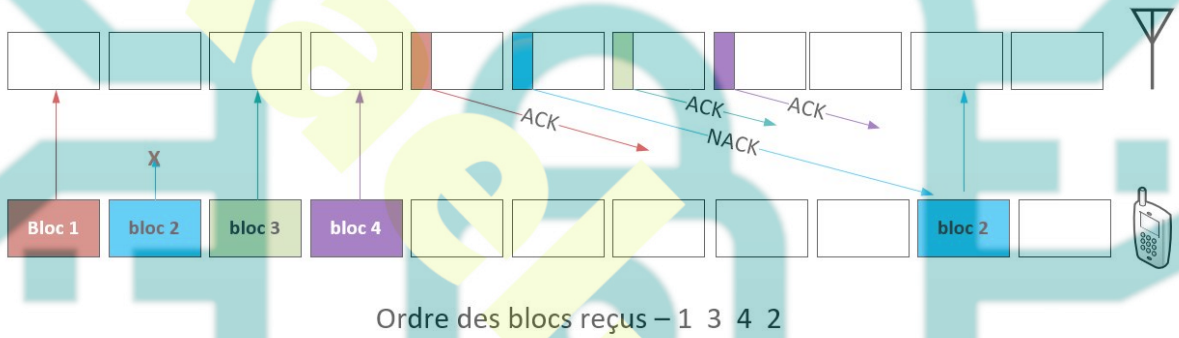
Lors du processus acquittement, le récepteur à 4 sous trames pour réagir et c'est aussi le cas pour l'émetteur lorsqu'il reçoit un ACK, NACK ou rien. Au total le processus dure 8 sous trames soit 8 ms.

Pour augmenter l'efficacité, on va donc envoyer des données à un autre processus avant le ACK du premier processus pour éviter d'attendre 4 sous trames pour rien.



Le dé-séquencement

Il se produit lorsqu'un bloc se perd et cela nécessite la réémission du bloc. L'ordre d'envoi devient différent de l'ordre de réception.



Ordre des blocs reçus – 1 3 4 2

La couche mac ne peut pas gérer ce problème et doit laisser aux couches supérieures (RLC) la gestion du problème.

La méthode HARQ

LTE utilise une version améliorée d'ARQ en ajoutant la conservation des données erronées par le récepteur, c'est la méthode HARQ (Hybrid ARQ).

Cette méthode permet à la couche MAC de collaborer avec la couche physique en lui indiquant s'il s'agit d'un nouveau message ou un message de répétition

(BONJOUR — B**A**NJOUR – **R**ONJ**A**UR – BO**JB**OU**H** — BONJOUR)

La couche physique sait alors quel modèle de correction affecter

NB. Lors de l'émission de répétition, l'émetteur change la séquence de codage.

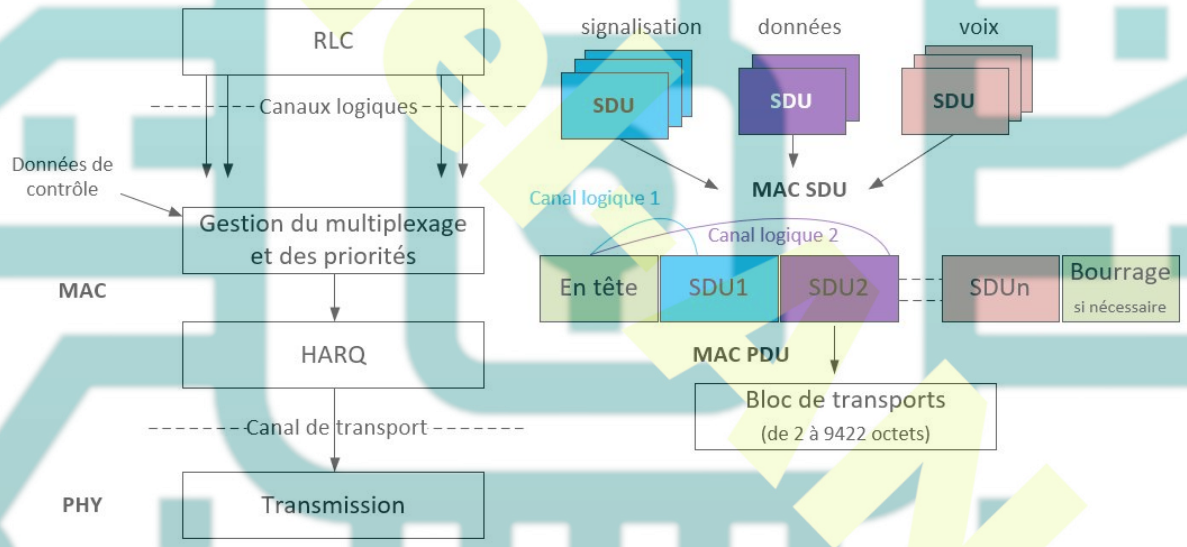
Multiplexage couche MAC

Les données reçues par la couche MAC viennent de la couche RLC (MAC SDU – données reçues par la couche MAC) au travers de canaux logiques (files d'attente) qui correspondent à différents niveaux de service (voix, données, signalisation…)

Le rôle de la couche MAC est de former des blocs de transport le MAC PDU (données envoyées par la couche MAC).

La couche MAC va prendre dans la file d'attente les SDU pour former un bloc de transport. Elle réalise donc le multiplexage en prenant plusieurs voies en entrée mais une seule en sortie.

Pour que le récepteur puisse démultiplexer, la couche MAC ajoute un en-tête en indiquant ou commence chaque message et à quel canal logique il appartient.



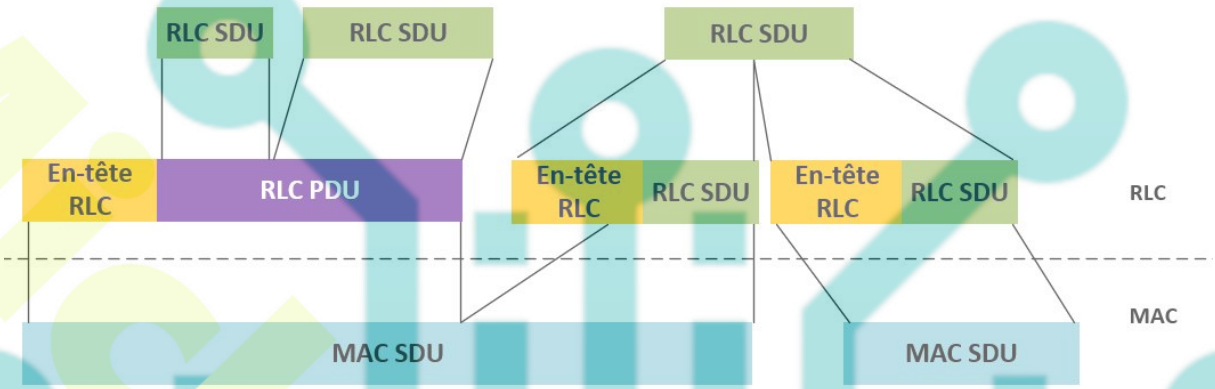
La couche RLC et le QOS

Le QOS réseau doit fournir un niveau de service correspondant aux besoins des applications selon des critères comme le délai, le débit ou le taux d'erreur.

La couche MAC possède des limites comme l'arrêt des transmissions au bout d'un certain temps, le dé-séquencement des messages et l'impossibilité de segmenter des messages.

La couche RLC (Radio Link Control) propose des services pour augmenter la fiabilité des échanges, pour gérer le re-séquencement et la segmentation. Cependant, ces options

entrainent une latence plus importante.



On peut remarquer que la couche MAC peut regrouper dans un même bloc de transport des SDU provenant de différentes instances de RLC.

RLC propose 3 modes de fonctionnement :

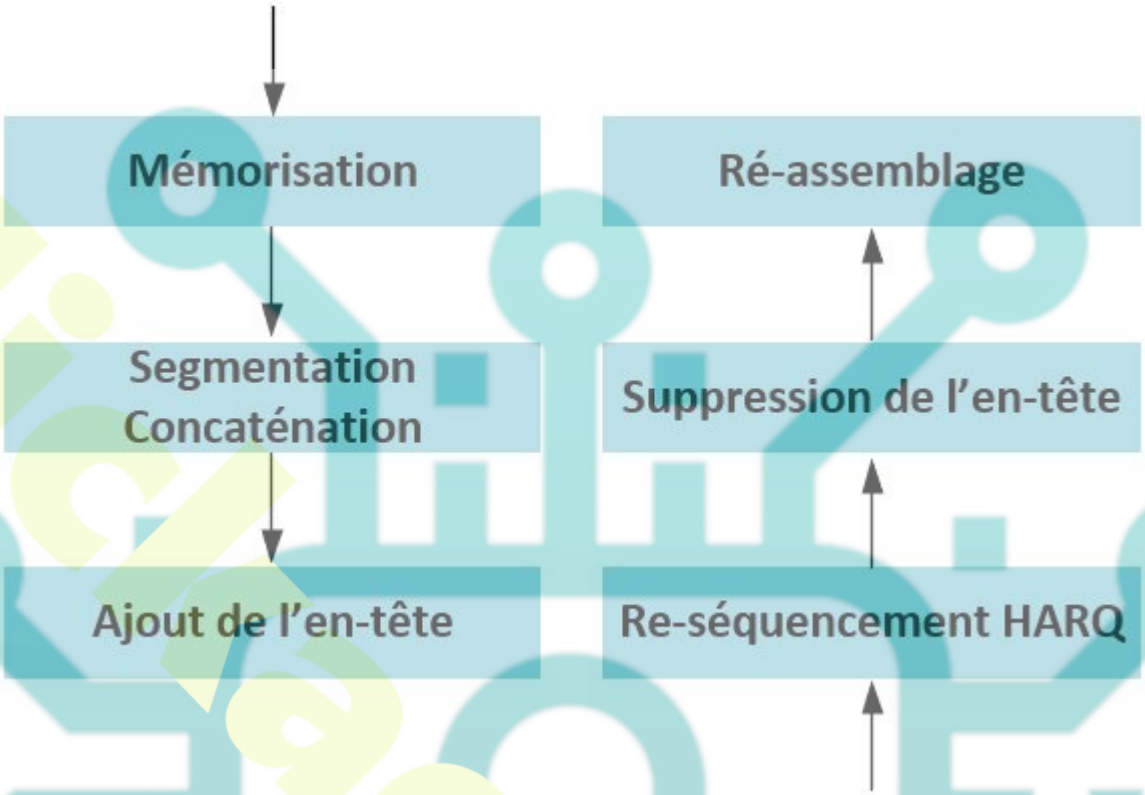
Le mode transparent (TM – Transparent Mode)
RLC ne fait rien, il est utilisé pour les messages courts sans segmentation comme la signalisation par exemple.

Le mode non-acquitté (UM – Unacknowlegde Mode)
Permet le re-séquencement des messages, la segmentation et la concaténation. Il a une faible latence mais peu de fiabilité. Il convient aux données temps réel comme la Voix, la vidéo.

Lorsqu'elle reçoit un paquet de la couche supérieure, la couche RLC le mémorise dans une mémoire tampon en attendant que la couche MAC lui demande un MAC SDU. Lorsque la couche MAC demande un paquet, elle précise la taille du MAC SDU qu'elle attend.

RLC assemble les paquets qu'il a en mémoire pour former un RLC PDU de la taille demandée. Pour cela, il peut concaténer plusieurs messages et couper les messages pour obtenir la taille spécifiée par la couche MAC.

RLC ajoute son en-tête pour préciser l'emplacement des fragments et les numéros de séquence.



De plus, comme il y a re-séquencement le mode UM pourrait rester bloqué si un segment n'arrivait jamais. Pour éviter cela, RLC va estimer qu'au bout d'un certain temps le paquet est perdu et qu'il faut passer au suivant.

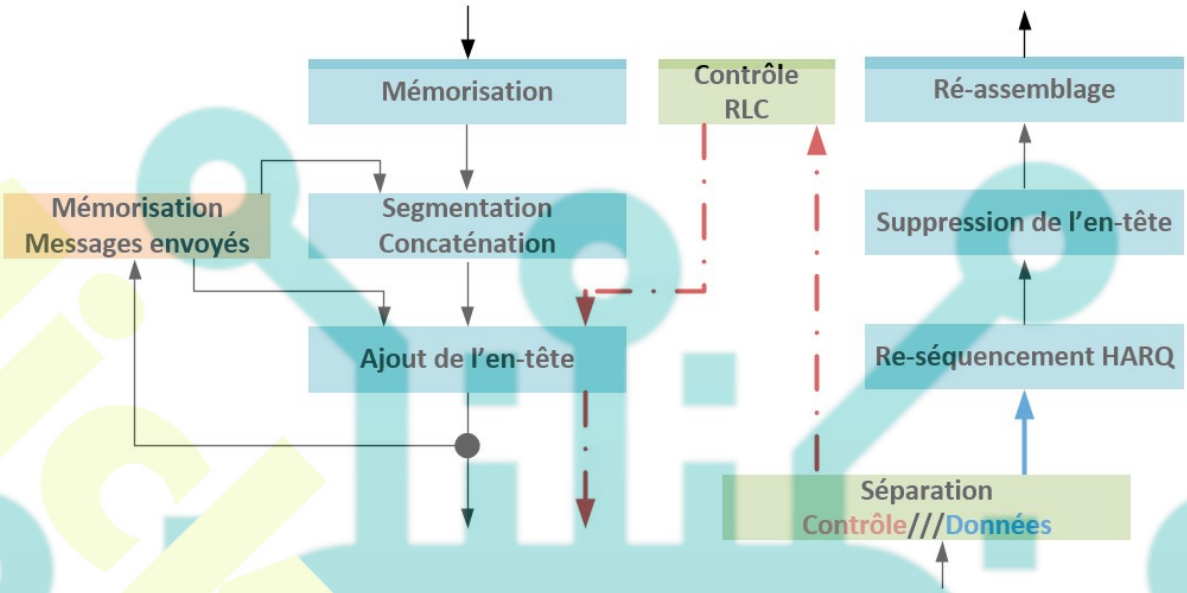
Le mode acquitté (AM – Acknowlegde Mode)
Il propose la réémission des paquets perdus et convient aux applications de type web, transfert de fichiers et messagerie. Il une bonne fiabilité mais au détriment de la latence.

Dans ce mode, RLC mémorise les messages qu'il envoie à la couche MAC (MAC SDU) et il demande régulièrement au destinataire un état des messages reçus par l'intermédiaire du bit polling situé dans l'en-tête.

Le destinataire répond par un message de signalisation RLC envoyé au milieu des données.

Contrairement au début du cours, on n'utilise pas les canaux dédiés pour la signalisation.

A la réception, il faut donc séparer ce message des données et l'envoyer au contrôleur RLC. Ce dernier peut alors vider son buffer et passer aux messages suivants.



Les méthodes d'accès

LTE n'utilise pas comme le WIFI une méthode en contention, il utilise une méthode de réservation. La ressource est divisée dans le temps et en fréquence et est allouée dynamiquement.

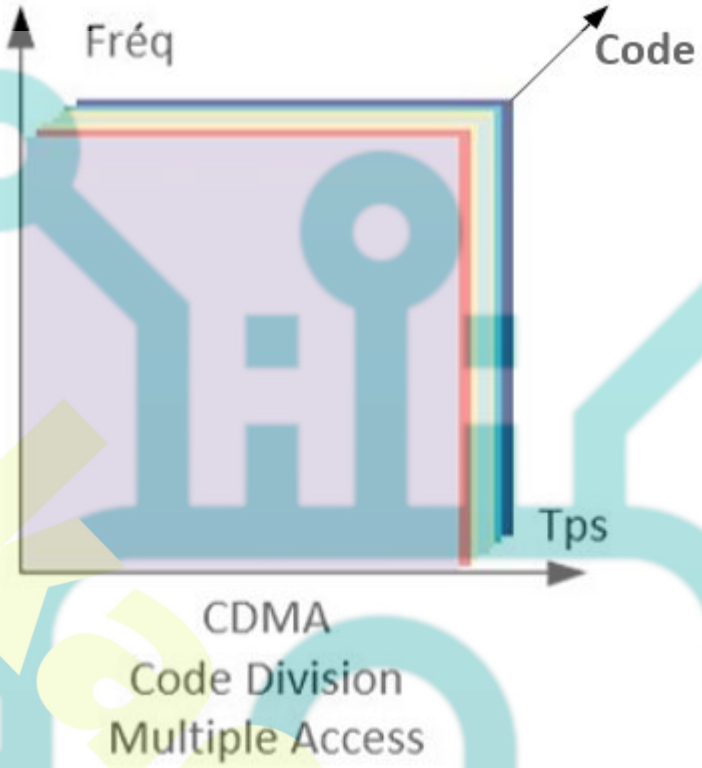
Dans le sens descendant, il n'y a pas de problème puisque c'est l'eNodeB qui gère les émissions.

Dans le sens montant, le mobile doit demander, via le canal de contrôle de la voie montante, l'allocation des ressources. Pour cela, les mobiles doivent s'inscrire au préalable.

Le modèle LTE propose d'insérer une dose de contention au sein de la voie montante pour permettre aux nouveaux arrivants de se signaler et de s'inscrire.

Un groupe de ressources est réservé pour permettre cette inscription. Ce groupe est un canal physique (**PRACH** – Physical Random Access Channel), ce canal est constitué de 6 paires de blocs de ressources contigus qui revient toutes les 1 à 20 ms.

L'accès utilise le CDMA qui utilise 64 séquences différentes. Le CDMA consiste à utiliser une technique d 'étalement par codes qui permet la transmission simultanée de plusieurs canaux, chacun étant étalé en temps et en fréquence. On attribue à chaque utilisateur un code d'étalement pour moduler son signal d'information, Les utilisateurs occupent la même bande au même instant.



Dans les 6 paires de blocs, on insère juste un symbole pour que l'eNodeB s'aperçoive qu'un nouveau terminal est présent. Comme il existe 64 séquences, cela veut dire que 64 nouveaux terminaux pourront faire leur demande simultanément.

Traitement des demandes d'accès

Lorsqu'un nouveau terminal arrive dans une cellule, il écoute les émissions de l'eNodeB pour connaitre à quel moment est programmé le PRACH.

Le terminal choisit une séquence parmi les 64 puis émet sur le canal à accès aléatoire (PRACH).

L'EnodeB détecte cette séquence et lui attribue un RNTI (Radio Network Temporaly Identifier). Cependant, le RNTI n'est connu que de l'eNodeB et ne peut pas servir d'adresse pour contacter le mobile. Alors, l'eNodeB créé un RA-RNTI (Randow Access RNTI) qu'il calcule par rapport au numéro de séquence qui a été utilisé par le terminal.

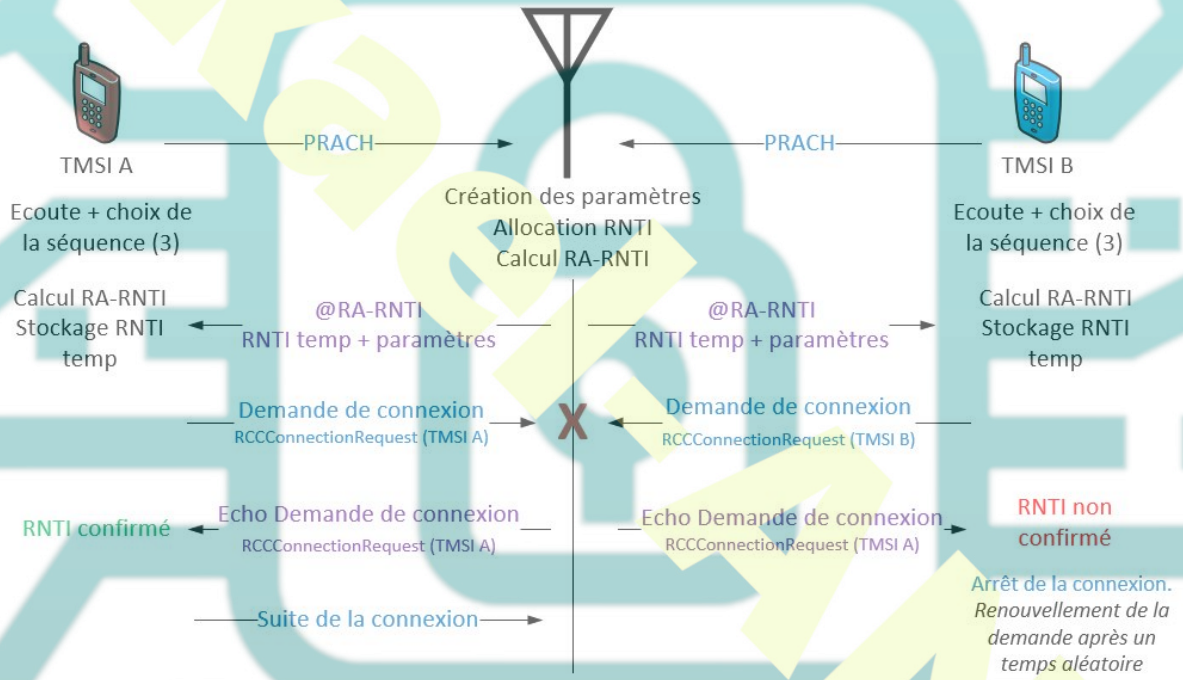
Une fois le RA-RNTI créé, l'eNodeB envoie un message qui contient le RNTI et d'autres paramètres notamment une réservation de ressources pour que le terminal puisse

s'identifier.

Ce RA-RNTI peut être alors détecté par le mobile avec le même calcul. Le mobile peut alors récupérer le message contenant son RNTI et commencer sa procédure de connexion vue plus haut dans le cours.

Pour éviter les conflits entre 2 appareils L'eNodeB envoie un ECHO de vérification de la demande du client.

RRC : Hérité de la 3G, le protocole RRC permet à l'UE et à l'eNodeb d'échanger de la signalisation (messages RRC).



Interconnexion avec le réseau

Le protocole PDCP (Packet Data Convergence Protocol) assure l'interface avec le protocole de réseau c'est à dire IP pour les données utiles et RCP pour les messages de signalisation.

Le protocole PDCP assure 3 types de services :

La compression d'en-tête

Pour limiter la consommation inutile de bande passante. En effet, sur la VOIP les paquets font 30 octets puis ils sont encapsulés dans RTP qui ajoute 12 octets puis de nouveau

encapsulés dans UDP qui rajoute 8 octets et enfin IP qui rajoute 40 octets dans la version 6.

On s'aperçoit alors que les en-têtes font 60 octets, soit le double des données utiles. Comme ce sont des données qui évoluent peu, il est possible de les compresser en utilisant un protocole standard de l'internet ROHC (RObust Header Compression).

Prévenir les pertes dues aux Handhovers

PDCP pallie le fait que la retransmission des données via RLC ne peut se faire qu'au sein d'une même cellule. Pour cela, il permet à l'ancien eNodeB de prévenir le nouvel eNodeB de l'état des transmissions au moment du changement de cellule et lui faire suivre les données en attente d'émission ou de réception.

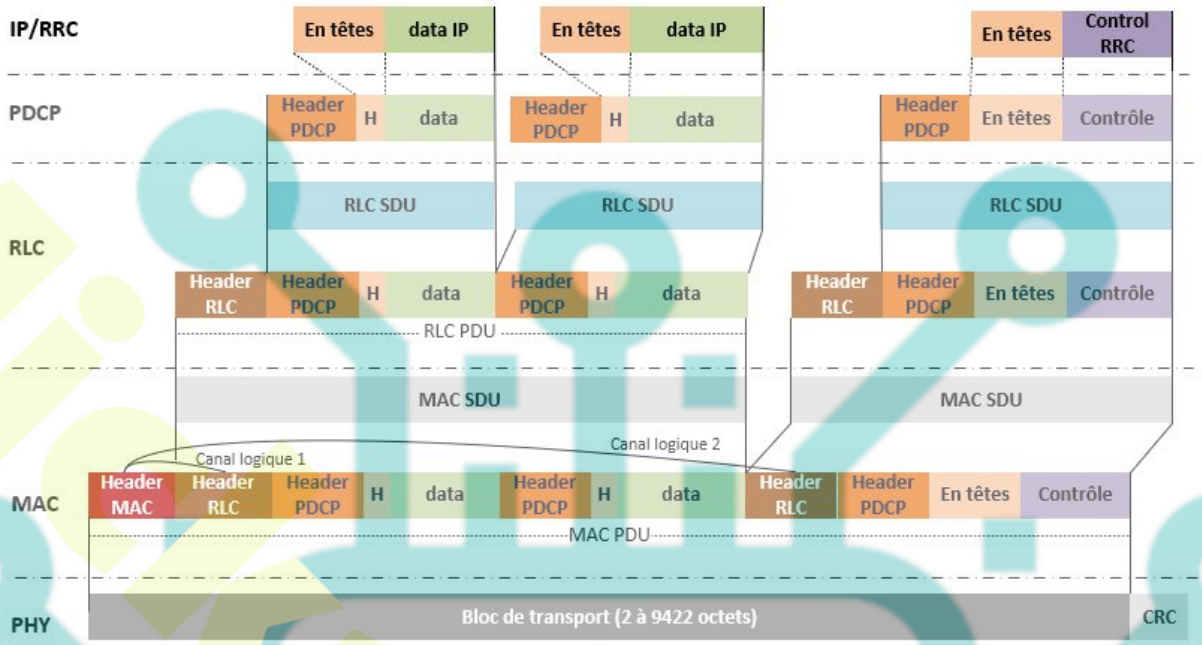
La sécurité

PDCP permet de chiffrer les données et propose un contrôle d'intégrité pour les messages de signalisation.

PDCP ne prend pas en compte certaines options en fonction de l'utilisation.

	RRC		IP	
	UM	AM	UM	AM
Compression ROHC	NON	NON	**OUI**	**OUI**
Pertes Handover	NON	**OUI**	NON	**OUI**
Chiffrement	**OUI**	**OUI**	**OUI**	**OUI**
Intégrité	**OUI**	**OUI**	NON	NON

Récapitulatif des couches



Les tunnels

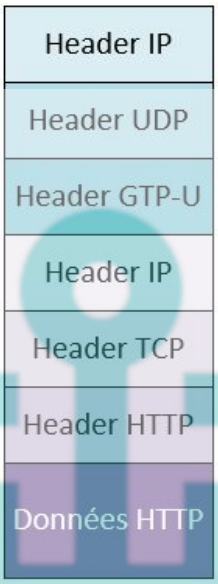
Pour transporter les paquets IP à destination de l'UE, on créé un premier tunnel entre le PGW et le SGW et un deuxième tunnel entre le SGW et l'eNodeB. Enfin, on place le paquet dans un bloc de transport en utilisant le RNTI pour joindre l'UE.

Encapsulation GTP

Le protocole GTP (GPRS Tunnelling Protocol) est la couche qui permet de gérer les tunnels au-dessus d'UDP (plus simple que TCP).

Pour indiquer qu'il s'agit de données utilisateur (vidéo…), on indique qu'il s'agit de GTP-U (user plane). GTP-U est utilisé sur la liaison S1-U (entre eNodeB et le SGW) et sur l'interface S5/S8 (entre SGW et le PGW)

Une encapsulation des données HTTP venant d'internet et à destination d'un UE donnerait l'encapsulation suivante :



La gestion de tunnels multiples

Dans un réseau 4G, il y a des millions d'utilisateurs rattachés à différents eNodeB eux-mêmes connectés à différents SGW attachés à quelques PGW.

Pour gérer très rapidement cette multiplicité de tunnels et de traitement, on utilise un principe de numérotation unique pour chaque extrémité de tunnel sur le SGW.

Cette numérotation s'appelle **TEID** (Tunnel Endpoint Identifier – identifiant codé sur 32 bits).

Chaque nœud concerné par un tunnel attribue un identificateur, ce qui veux dire qu'il existe 2 identifiants par tunnel. Un même numéro peut être utilisé dans le réseau, mais pas au sein d'un même équipement.

Principe de création du tunnel

Un équipement fait une demande de tunnel avec un TEID en utilisant le protocole GTP-C pour les messages de contrôle. L'autre extrémité répond en envoyant à son tour son TEID. A partir de ce moment les deux extrémités connaissent les 2 TEID. Ces informations sont stockées dans une table de correspondance sur les deux entités.

Table de correspondance

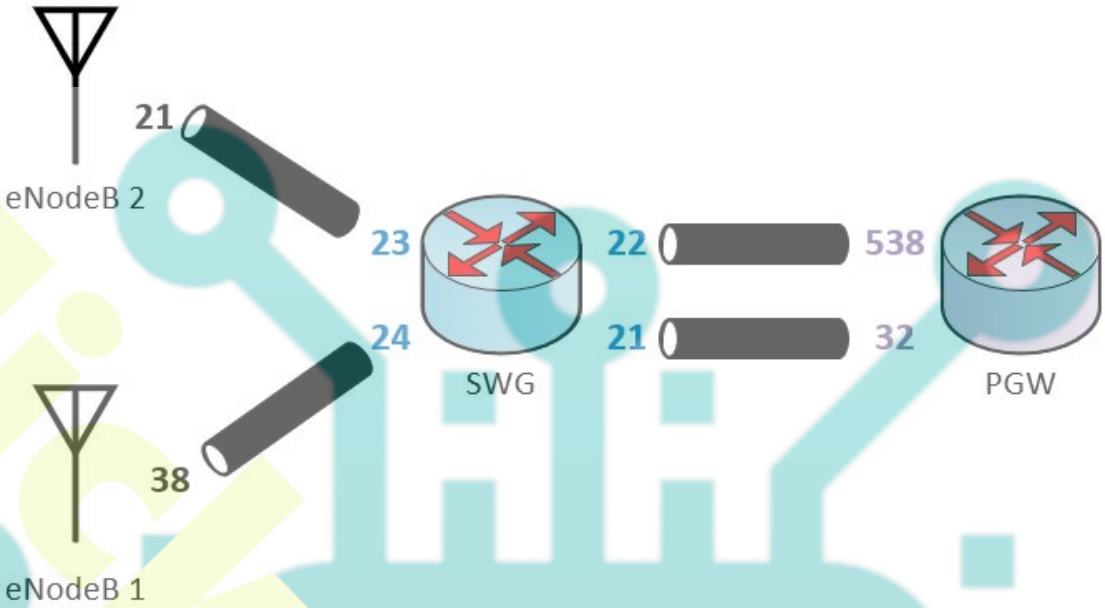


Table de correspondance du SGW

1 - Création des tunnels entre SGW et PGW

TEID	Action	Détails	Peer Entity	
			@IP	TEID
21			@PGW	32
22			@PGW	538

2 - Création des tunnels entre eNodeB et PGW

TEID	Action	Détails	Peer Entity	
			@IP	TEID
21			@PGW	32
22			@PGW	538
23			@eNodeB 2	21
24			@eNodeB 1	38

3 - Création des informations de redirection entre eNodeB et PGW

Règle – ce qui vient de 23/21 (eNodeB-2) est redirigé vers 22/538
ce qui vient de 38/24 (eNodeB-1) est redirigé vers 21/32

TEID	Action	Détails	Peer Entity	
			@IP	TEID
21	Forward	TEID = 24	@PGW	32
22	Forward	TEID = 23	@PGW	538
23	Forward	TEID = 22	@eNodeB 2	21
24	Forward	TEID = 21	@eNodeB 1	38

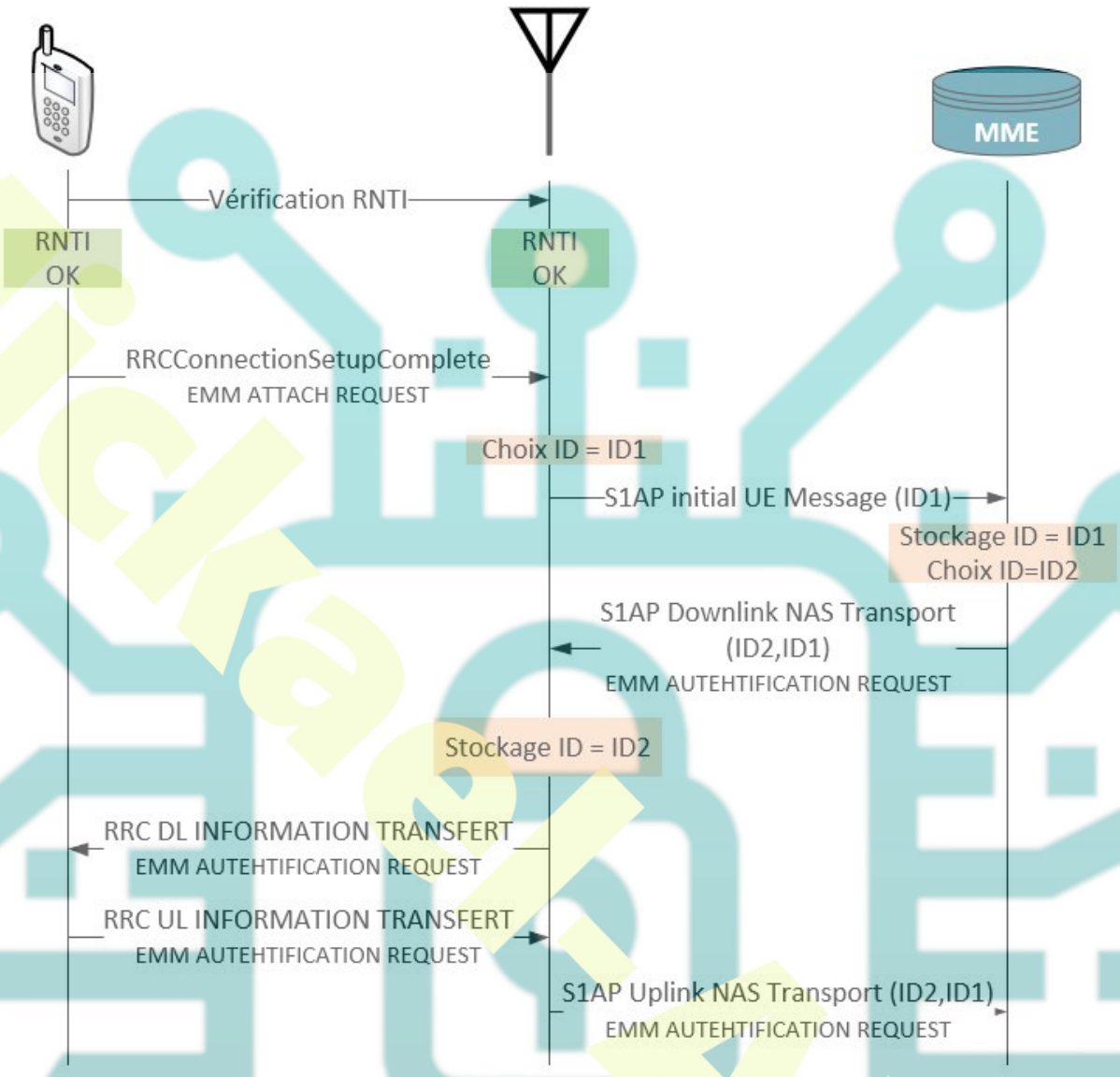
Dialogue entre le MME et l'UE

Pour faciliter le traitement entre le MME et les millions de terminaux qu'il gère, on utilise un protocole appelé **S1-AP** (Application Protocol On S1). Ce protocole est orienté connexion ce qui veut dire qu'il y a autant de connexion que d'utilisateur.

Ce protocole doit permettre d'identifier les communications de l'UE entre l'eNodeB et le MME.

Le protocole S1-AP permet au MME et à l'eNodeB d'échanger des messages de configuration générale et d'activer certaines fonctions en lien avec la connexion d'un terminal. Il permet également à l'eNodeB de signaler au MME des changements d'état d'un terminal

Pour chaque message lié à un terminal particulier, on met dans l'en tête la paire d'identités de connexion (sauf pour le tout premier message qui n'a qu'une identité de connexion)
Les identités de connexion sont codées sur 3 ou 4 octets.



Séparation des rôles

Le plan de contrôle est utilisé pour tous les échanges de signalisation, il n'est pas là pour transporter des données. Le dialogue entre l'UE et le MME ne se place que dans le plan de contrôle, c'est à dire que le MME ne voit pas passer les données utilisateurs.

Le **contrôle** est réservé à la gestion de la mobilité, de la sécurité (MME-UE) ou à l'allocation des ressources radio et l'établissement des connexions (eNodeb – UE). Ce sont donc des messages et protocoles liés à la technologie radio.

Le plan **utilisateur** est utilisé lui pour le transport des données utilisateurs.

Les échanges sont donc séparés entre le **User Plane** et le **Control Plane**.

Par exemple, les échanges S1-AP utilisent des identifiants différents entre le plan de contrôle et le plan utilisateur.

NAS et AS

Les messages **NAS** (Non Access Stratum) sont relayés par l'eNodeB pour la communication en UE et MME sans analyse de contenu.

Les messages **AS** (Access Stratum) sont échangés entre le terminal et l'eNodeB.

Le mode NAS utilise deux protocoles entre l'UE et le MME, ESM (Envolved Session Manager) et EMM (Envolved Mobility Manager).

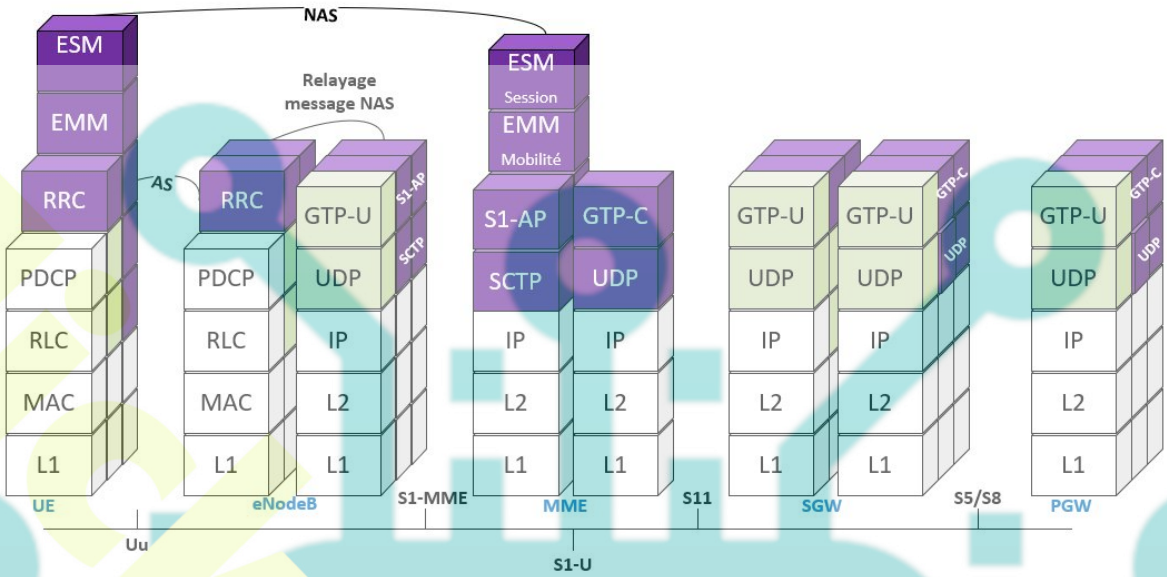
Le protocole SCTP (stream Control Transport Protocol)

Ce protocole est utilisé en remplacement de UDP (non fiable) et TCP (orienté flux et pas message)

Il permet le transport fiable mais évitant les retransmissions inutiles.

Les piles de protocoles

Les couches vertes correspondent aux messages utilisateur (User Plane) et les couches violette aux messages de contrôle (Control Plane).

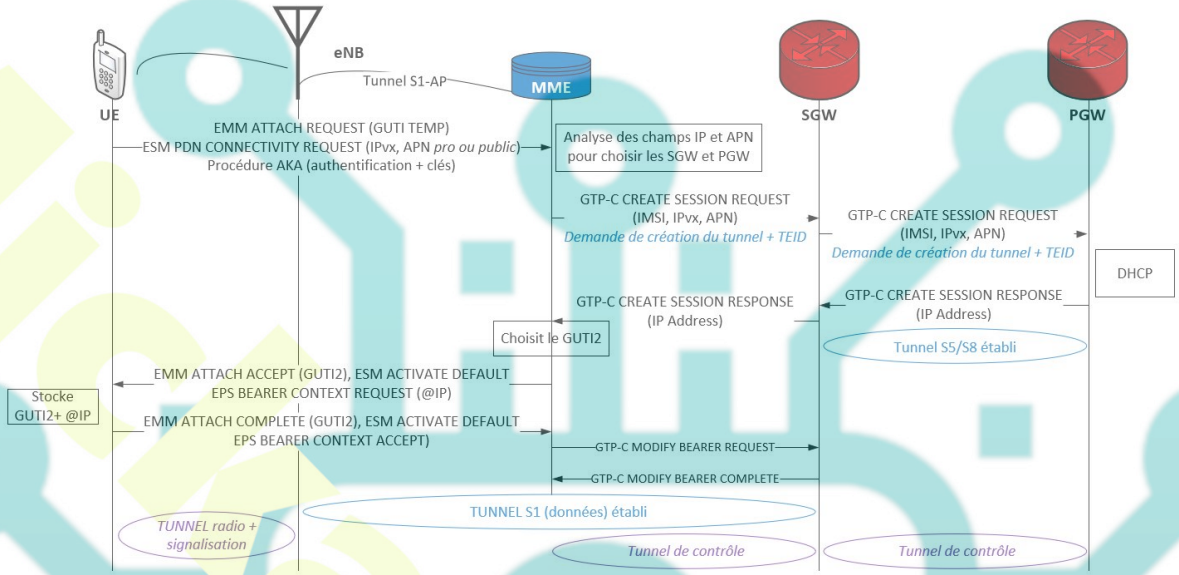


Mise sous tension d'un terminal

Si le mobile n'est pas en mode avion, il établit un tunnel par défaut (Bearer). Il demande un établissement via le tunnel par défaut. Pour établir la connectivité, il envoie un **EPS Connectivity Request** au travers d'un message EMM Attachement.

NB. Il est important pour le réseau de connaitre l'état du mobile, une notion d'état est mémorisée pour chaque terminal.

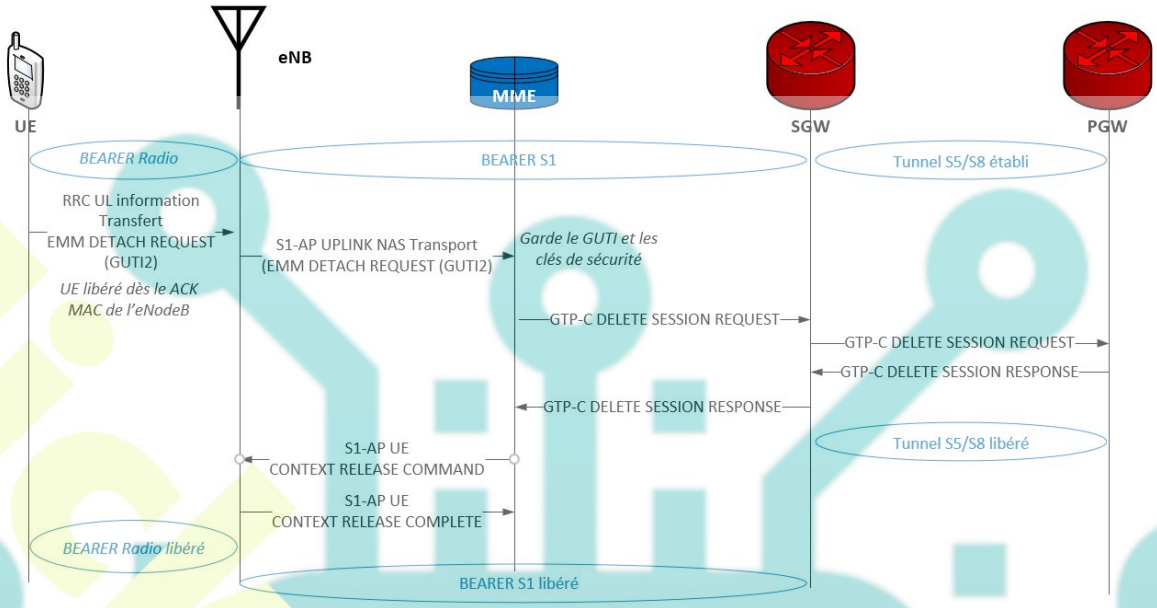
Demande d'attachement



- Une fois la connexion effectuée, L'UE est dans l'état EMM Register et possède l'IMSI, le GUTI, l'association de sécurité et une adresse IP.
- Le MME possède lui l'état de l'UE, l'association de sécurité, le GUTI, l'adresse IP, la localisation… du terminal.
- Les SGW et PGW possèdent eux le TEID.

Demande de détachement

Si on active le mode avion ou une déconnexion. L'UE fait une demande de détachement.



Le terminal en état d'inactivité

Un eNodeB utilise un temporisateur (RRC inactivity timer) qu'il lance à la fin d'un échange et si le temporisateur tombe à zéro, on libère la connexion RRC et l'UE perd son RNTI.

- On libère au bout de 10 à 15 secondes environ

En cas d'inactivité réseau, on coupe le tunnel entre l'UE et l'eNodeB et on libère la connexion S1-AP et le S1 Bearer. Cependant, on ne sait plus où se trouve l'UE (état ECM IDLE) mais l'UE reste attaché au réseau.

Si le mobile redemande des données on repasse en mode ECM Connected.

Les différents états

EMM-deregistered (nécessairement ECM_idle)

- Terminal non connecté réseau, pas d'adresse IP, pas de RNTI

EMM-registered et ECM-connected

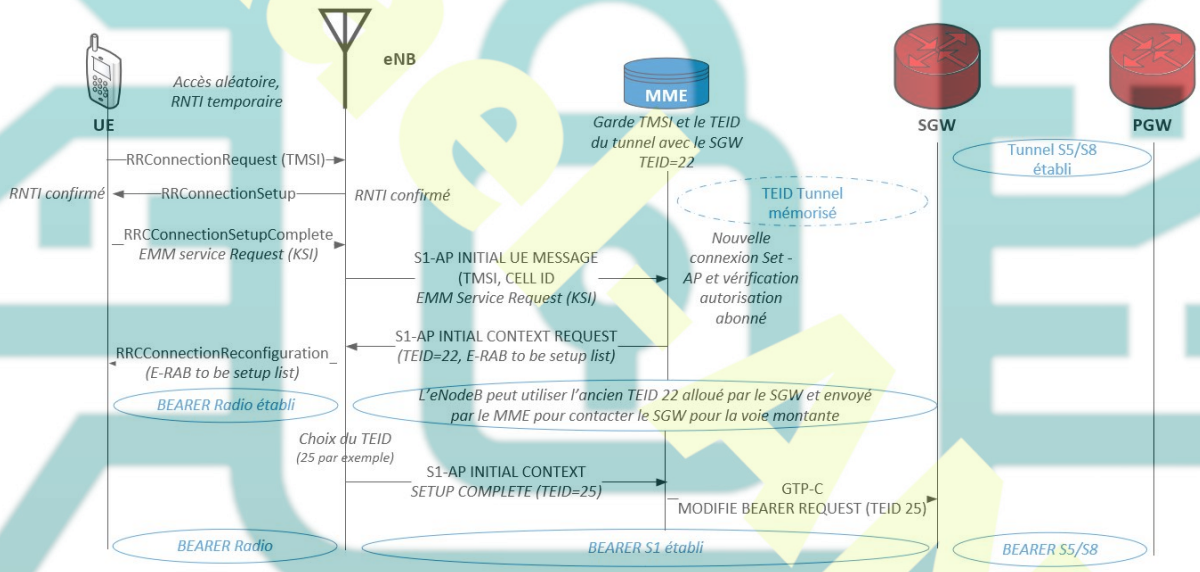
- UE connecté au réseau, avec une adresse IP et un RNTI
- Localisation de la cellule de l'UE connue par le MME
- Tous les tunnels et les connexions sont établies

EMM-registered et ECM-idle

- Terminal connecté au réseau, avec une adresse IP mais plus de RNTI
- Pas de tunnels et de connexions avec l'eNodeB
- Tunnels maintenus entre SGW et PGW et entre SGW et MME
- Localisation approximative de l'UE (ensemble de cellules)

Reprise de la connexion de l'UE après inactivité

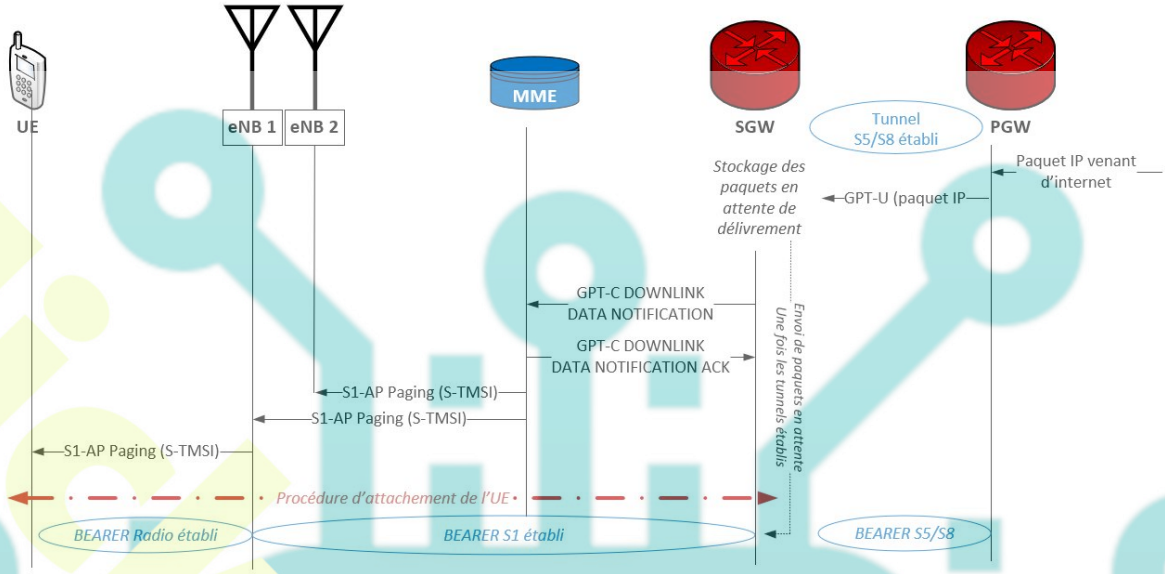
Pour éviter que la reconnexion soit trop longue, l'astuce consiste de conserver les informations du tunnel S1 (TEID) dans le MME. Ce principe évite de relancer des requêtes de tunnel entre le MME et le SGW.



Envoi des données d'un serveur vers un UE en mode veille

Le paquet envoyé d'un serveur vers l'UE va passer vers le PGW, puis le PGW envoie le paquet au SGW via le tunnel S1/S8. Ensuite le SWG demande au MME d'établir la connexion (le SWG utilise le canal de contrôle pour cela).

Cependant, le MME ne connait qu'approximativement la localisation de l'UE. Alors, le MME envoie une demande de S1-AP Pagging (contenant le S-TMSI de l'UE) vers les eNodeB qui pourraient contacter l'UE. L'UE voyant son S-TMSI passer récupère le message et déclenche sa procédure de connexion (RRCConnexion Request).



QOS sur la 4G

Les tunnels créés par défaut conviennent à la plupart des applications sur IP en utilisant le *Best Effort*. Cependant, pour la voix sur IP, ces tunnels ne sont pas suffisants.

Pour pallier le fait que la QOS n'est pas suffisante, on créé des bearers dédiés avec de nouveaux TEID. Les bearers dédiés sont privilégiés par rapport aux tunnels par défaut.

Gestion de la localisation

Un UE pouvant se trouver dans n'importe quelle cellule, comment faire pour trouver un terminal ?

Solution 1

Un UE envoie un message dès qu'il change de cellule (en s'apercevant qu'il change de voie balise donc d'eNodeB)

Avantages – le fait que l'UE se signale, le système n'a pas à s'en occuper
Inconvénients – quand on se déplace, on change souvent de cellule et du coup l'UE doit travailler régulièrement et consommer beaucoup de batterie.

Par exemple, si on considère que dans une zone à forte densité où un eNodeB (station de base) couvre une zone de 600 mètres de diamètre. Un piéton qui se déplace à 3.6 km/h soit 1 m/s, implique que l'UE va envoyer des données de localisation toutes les 600 secondes. Si on se déplace en voiture à 36 Km/h soit 10 mètres par secondes, l'UE va envoyer ses données toutes les minutes.

La mise à jour de localisation échange des messages et la batterie va être rapidement à plat.

Solution 2

L'astuce consiste à regrouper les eNodeB pour agrandir la zone. Ce système s'appelle une zone de suivi (Tracking Area). Cette zone possède un identifiant contenant le MMC (code pays), MNC (code opérateur), TAC (Tracking Area Code).

L'identité du TAI est régulièrement diffusée par les stations de base.

Avantages

Lorsqu'un UE se trouve dans une zone, il n'échange plus rien, il n'échangera des données que lorsqu'il changera de zone.

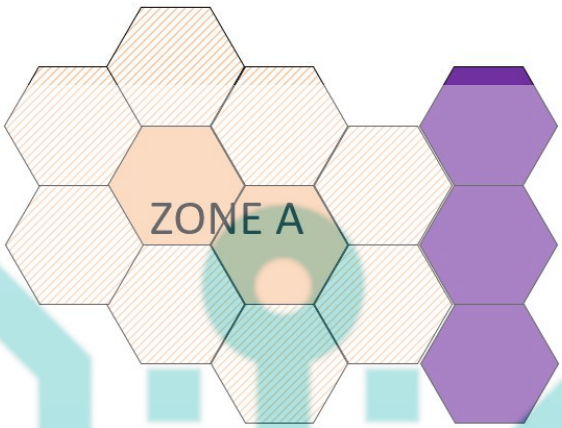
Par exemple, si on considère que dans une zone à forte densité ou un eNodeB (station de base) couvre une zone de 1800 mètres de diamètre. Un piéton se déplace à 3.6 km/h soit 1 m/s, implique que l'UE va envoyer des données de localisation toutes les 30 minutes. Si on se déplace en voiture à 36 Km/h soit 10 mètres par secondes, l'UE va envoyer ses données toutes les 3 minutes.

Inconvénients

On ne sait plus exactement où se trouve le terminal, et lorsque l'on cherche à le joindre, le MME doit diffuser un message de paging à tous les eNodeB de la zone.

Gestion des zones

Le problème des zones est que les cellules frontières peuvent être plus nombreuses que les cellules centrales. De ce fait, la charge de signalisation des eNodeB (zone de bordure) est plus forte que dans les cellules centrales. En effet, on est obligé d'échanger beaucoup de blocs de transport (signalisation) et on perd en capacité pour les données utiles elles-mêmes.



La solution consiste à envoyer une liste de zone de suivi (TA List), permettant de limiter les mises à jour de localisation. L'astuce consiste ensuite à envoyer une liste différente aux terminaux d'une même zone.

Par exemple un mobile 1 peut recevoir la liste TA1, TA2, TA5, TA6 et TA7 et le mobile 2 recevoir la liste TA1, TA2, TA3, TA4 et Ta6. Un opérateur peut ainsi équilibrer la charge.

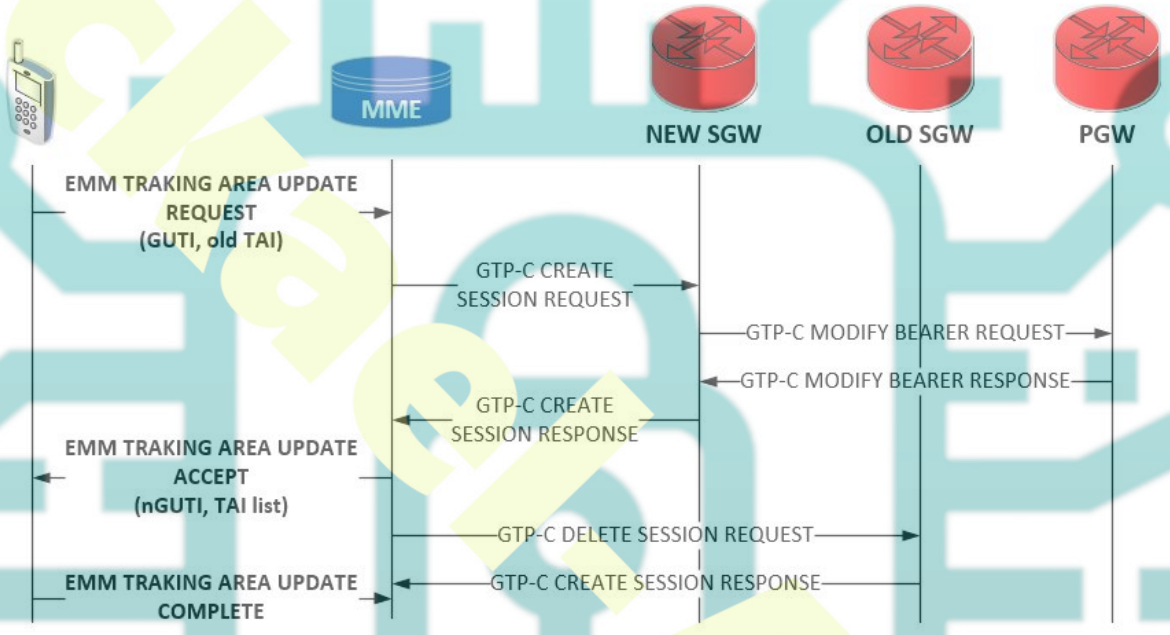
Dans la procédure de changement de cellules, on utilise le même schéma que précédemment, la différence est que le MME possède la liste des TA de l'UE.

Un message **EMM TRACKING AREA UPDATE REQUEST** contenant le GUTI et l'ancien TAI est envoyé de l'UE vers MME au travers de l'eNodeB et une réponse **EMM TRACKING AREA UPDATE ACCEPT** (nouveau GUTI, nouvelle liste TAI) est envoyée du MME vers l'UE qui stocke cette information dans sa carte USIM puis les tunnels sont libérés.

Mobilité en cas de changement de SGW

Cas 1 – les 2 SGW sont dans le même MME.

Le terminal change de cellule, il établit la connexion radio, la connexion S1-AP et on crée un nouveau tunnel de contrôle entre le MME et le SGW. Ensuite, on modifie les bearer entre le PGW et le SGW.
On demande ensuite à l'ancien SGW de vider le contexte du terminal.



Cas 2 – les 2 SWG ne sont pas dans le même MME

Le Handover

Pour gérer le handover, il faut connaitre les eNodeB de la zone, pour cela un eNodeB indique son PCI (Physical Channel Identity)

Plus le signal est faible moins le débit est important, il est donc nécessaire de transférer une connexion active d'un eNodeB vers un autre eNobeB en déterminant le bon eNodeB à utiliser.

Le terminal peut mesurer la puissance du signal émis par une station de base (les 6 plus proches). Cependant, il ne faut pas que le terminal choisisse lui-même la station de base, c'est au réseau de faire cette gestion.

Pour prévoir les changements, l'UE envoie ses résultats régulièrement au réseau (eNodeB), c'est ce qu'on appelle le **UE-assisted Network Trigerred Handover**.

Pour éviter de remonter des mesures lorsque ce n'est pas nécessaire, on attend d'avoir atteint un seuil pour le faire. En effet, lorsque l'UE est proche d'un eNodeB, le signal est

fort, il est donc inutile d'envoyer les résultats.

Gestion du handover au sein d'un même MME et SGW

Lors d'un changement d'eNodeB, il faut transférer le tunnel entre l'EnodeB et le SGW et le S1-AP entre l'eNodeB et le MME. De plus, il faut bufferiser les données à la fois sur le SGW, le PGW et surtout l'eNodeB pour ne pas perdre les données.

L'opération est complexe puisqu'il faut gérer les tunnels, les buffers et la saturation des cellules.

Pour gérer le passage d'un eNodeB source vers un eNodeB cible, on utilise une interface particulière, l'interface X2 (nouvelle en 4G).

Pile de protocole X2

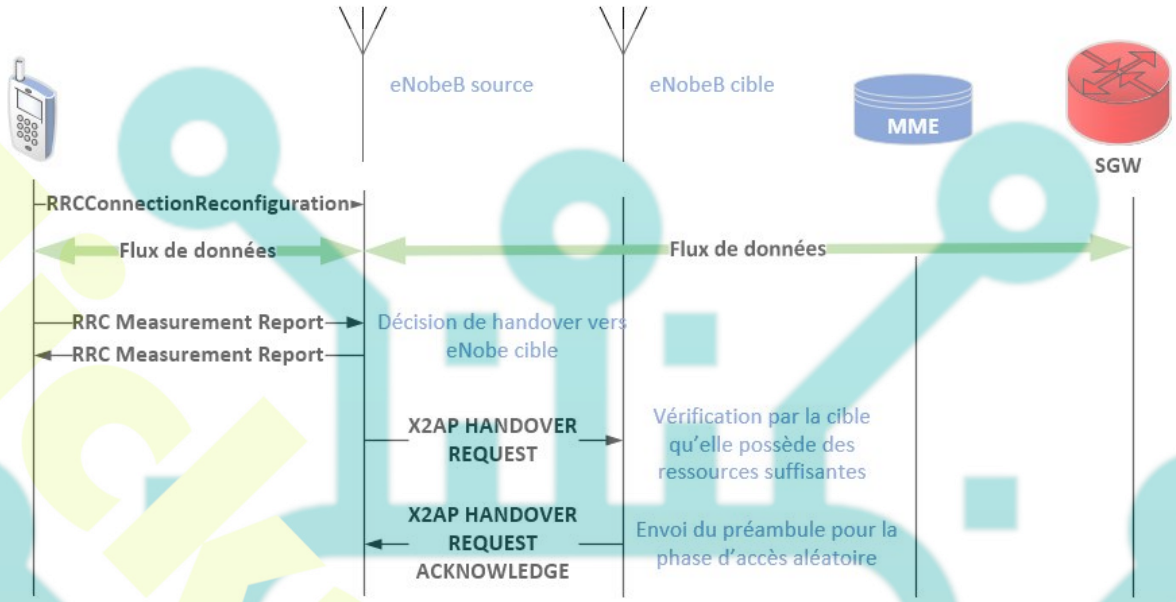
Sur le plan de contrôle, le protocole X2-AP (quasi similaire à S1-AP) fonctionne au-dessus de SCTP et sur le plan utilisateur, nous avons GTPU sur UDP.

Les phases du handover

Préalable – transmission des mesures de l'UE lorsque que le seuil est atteint, analyse par l'eNodeB des différentes mesures pour choisir le meilleur voisin.

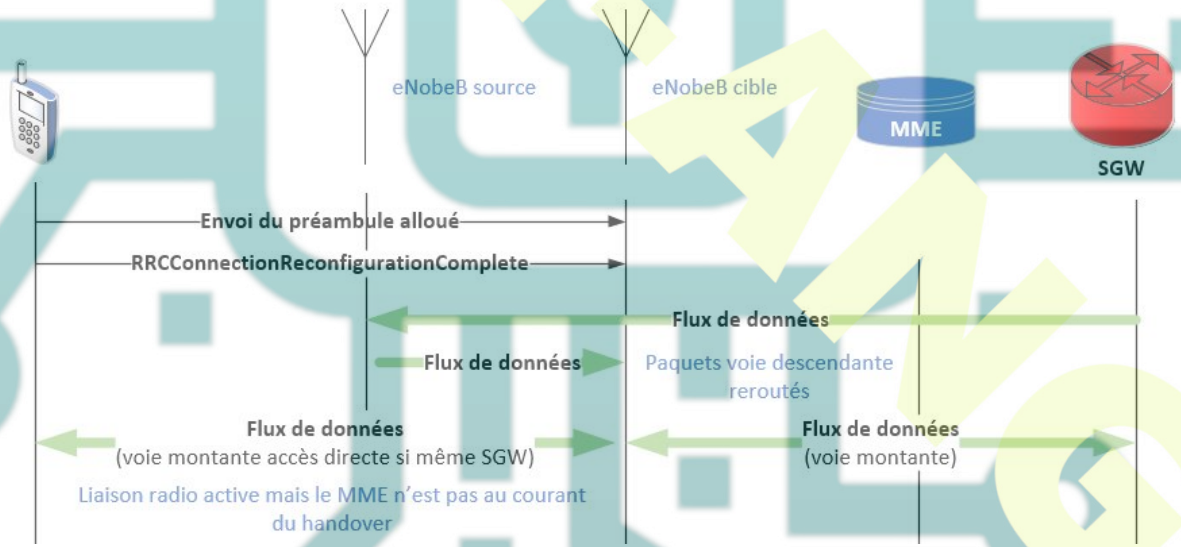
Préparation – réservation des ressources sur l'eNodeB cible

Handover préparation



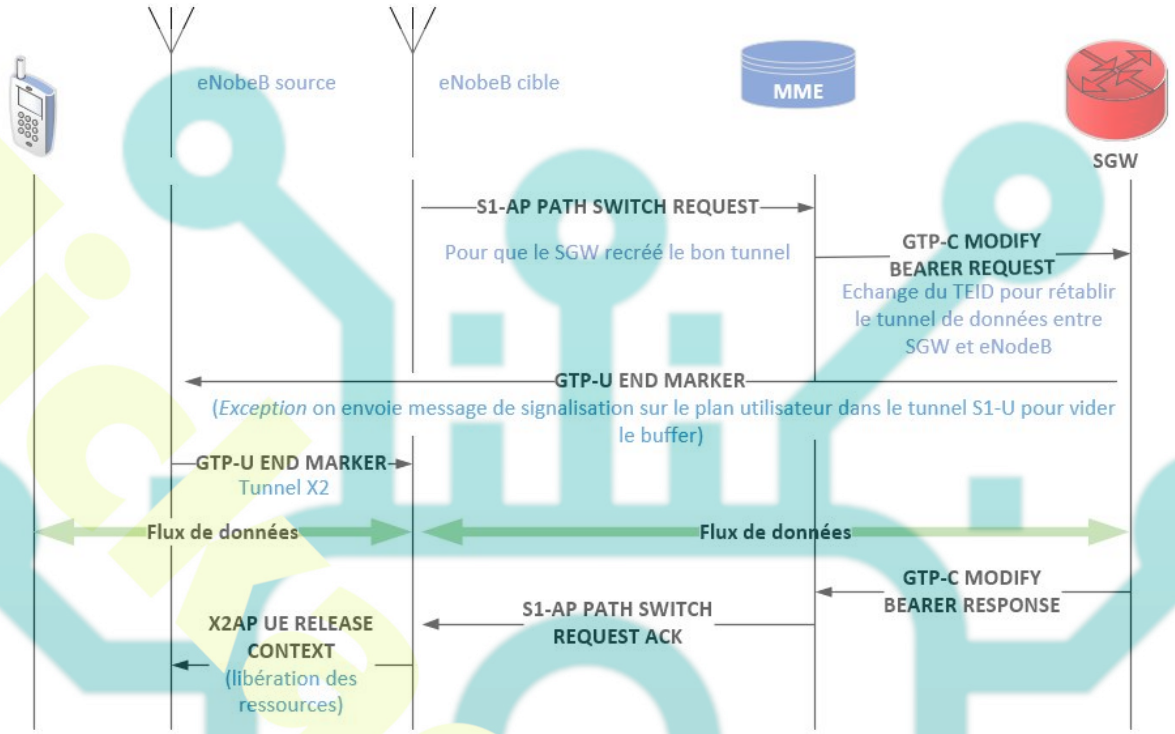
Exécution – envoi de l'ordre de changement de cellule à l'UE, reroutage des paquets, modification des tunnels et établissement de la nouvelle connexion réseau vers l'eNodeB.

Handover exécution



Terminaison – libération des ressources sur l'ancienne connexion.

Handover terminaison



Que se passe t'il si l'eNodeB cible n'a pas de ressources disponibles ?
Dans ce cas, l'eNodeB cible après avoir reçu le message X2AP Handover Request renvoie un message **X2AP HANDOVER FAILURE** pour refuser l'accès.

L'ancien eNodeB va essayer de garder la connexion de l'UE le plus longtemps possible pour trouver une solution et s'il n'en trouve pas la connexion aux données est perdue.

Sources Xavier LAGRANGE – Ecole Centrale Paris et Télécom Paristech
Christophe COUTURIER – Télécom Bretagne
Philippe MARTINS – Télécom Paristech
Alexander PELOV – Télécom Bretagne