

Groupe de travail Réseau
Request for Comments : 4867
 RFC rendue obsolète : 3267
 Catégorie : Sur la voie de la normalisation
 Traduction Claude Brière de L'Isle

J. Sjoberg, Ericsson
 M. Westerlund, Ericsson
 A. Lakaniemi, Nokia
 Q. Xie, Motorola
 avril 2007

Format de charge utile RTP et format de mémorisation de fichier pour les codecs audio multi débit adaptif (AMR) et multi débit adaptif large bande (AMR-WB)

Statut du présent mémoire

Le présent document spécifie un protocole de l'Internet en cours de normalisation pour la communauté de l'Internet, et appelle à des discussions et suggestions pour son amélioration. Prière de se référer à l'édition en cours des "Protocoles officiels de l'Internet" (STD 1) pour voir l'état de normalisation et le statut de ce protocole. La distribution du présent mémoire n'est soumise à aucune restriction.

Notice de Copyright

Copyright (C) The IETF Trust (2007).

Résumé

Le présent document spécifie un format de charge utile du protocole de transport en temps réel (RTP, *Real-time Transport Protocol*) à utiliser pour les signaux de parole codés en multi débits adaptatifs (AMR, *Adaptive Multi-Rate*) et multi débits adaptatifs large bande (AMR-WB, *Adaptive Multi-Rate Wideband*). Le format de charge utile est conçu pour être capable d'interopérer avec les formats de transport AMR et AMR-WB existants sur des réseaux non IP. De plus, un format de fichier est spécifié pour le transport de données de parole AMR et AMR-WB dans des applications en mode de mémorisation des données comme la messagerie électronique. Deux enregistrements de type de support distincts sont inclus, un pour AMR et un pour AMR-WB, spécifiant l'utilisation des deux formats de charge utile RTP et de mémorisation. Le présent document rend obsolète la RFC 3267.

Table des matières

1. Introduction.....	2
2. Conventions et acronymes.....	2
3. Généralités et principes de conception de AMR/AMR-WB.....	3
3.1 Le codec vocal multi débits adaptatifs (AMR).....	3
3.2 Le codec vocal multi débits adaptatifs large bande (AMR-WB).....	3
3.3. Codage multi débits et adaptation de mode.....	3
3.4 Détection d'activité vocale et transmission discontinue.....	4
3.5 Prise en charge de session multi canaux.....	4
3.6 Détection et protection inégales des erreurs binaires.....	4
3.7 Robustesse à la perte de paquets.....	6
3.8 Mode efficace en bande passante ou alignement sur l'octet.....	7
3.9 Scénarios AMR ou AMR-WB de voix sur IP.....	7
4. Formats de charge utile RTP AMR et AMR-WB.....	9
4.1 Usage de l'en-tête RTP.....	9
4.2 Structure de charge utile.....	10
4.3 Mode efficace en bande passante.....	10
4.4 Mode aligné sur l'octet.....	15
4.5. Considérations de mise en œuvre.....	20
5. Format de mémorisation AMR et AMR-WB.....	21
5.1 En-tête d'un seul canal.....	21
5.2 En-tête multi canaux.....	21
5.3 Trames de parole.....	22
6. Contrôle d'encombrement.....	23
7. Considérations sur la sécurité.....	23
7.1 Confidentialité.....	23
7.2 Authentification et intégrité.....	24
8. Paramètres de format de charge utile.....	24

8.1 Enregistrement de type de support AMR.....	24
8.2. Enregistrement de type de support AMR-WB.....	26
8.3 Transposition de paramètres de type de support en SDP.....	28
9. Considérations relatives à l'IANA.....	31
10. Changements par rapport à la RFC 3267.....	31
11. Remerciements.....	32
12. Références.....	33
12.1 Références normatives.....	33
12.2 Références pour information.....	33
Adresse des auteurs.....	34
Déclaration complète de droits de reproduction.....	35

1. Introduction

Le présent document rend obsolète la RFC 3267 et étend cette spécification avec les règles d'offre/réponse. Voir à la Section 10 les changements apportés à ce format en relation avec la RFC 3267.

Le présent document spécifie le format de charge utile pour la mise en paquets des signaux de parole codés AMR et AMR-WB dans le protocole de transport en temps réel (RTP) [RFC3550]. Le format de charge utile prend en charge la transmission de plusieurs canaux, de plusieurs trames par charge utile, l'utilisation de l'adaptation de mode codec rapide, la robustesse contre la perte de paquet et les erreurs binaires, et l'interopération avec les formats existants de transport AMR et AMR-WB sur des réseaux non IP, comme décrit à la Section 3.

Le format de charge utile lui-même est spécifié à la Section 4. Un format de fichier en rapport est spécifié à la Section 5 pour le transport de données de parole AMR et AMR-WB dans les applications en mode mémorisation comme la messagerie électronique. Dans la Section 8, deux enregistrements de type de support séparés sont fournis, un pour AMR et un pour AMR-WB.

Bien que cette définition de format de charge utile RTP prenne en charge le transport de la parole en AMR et en AMR-WB, il est important de se souvenir que AMR et AMR-WB sont deux codecs différents et qu'ils sont toujours traités comme des types de charge utile différents dans RTP.

2. Conventions et acronymes

Les mots clés "DOIT", "NE DOIT PAS", "EXIGE", "DEVRA", "NE DEVRA PAS", "DEVRAIT", "NE DEVRAIT PAS", "RECOMMANDE", "PEUT", et "FACULTATIF" en majuscules dans ce document sont à interpréter comme décrit dans le BCP 14, [RFC2119].

Les acronymes suivants sont utilisés dans le présent document :

3GPP (*Third Generation Partnership Project*) projet en partenariat de troisième génération
 AMR (*Adaptive Multi-Rate Codec*) codec multi débits adaptatif
 AMR-WB (*Adaptive Multi-Rate Wideband Codec*) codec multi débits adaptatif large bande
 CMR (*Codec Mode Request*) demande de mode codec
 CN (*Comfort Noise*) bruit de confort
 DTX (*Discontinuous Transmission*) transmission discontinue
 ETSI (*European Telecommunications Standards Institute*) Institut européen des normes de télécommunication
 FEC (*Forward Error Correction*) correction d'erreur directe
 SCR (*Source Controlled Rate Operation*) (fonctionnement) à un taux contrôlé par la source
 SID (*Silence Indicator*) indicateur de silence (trames contenant seulement des paramètres CN)
 UED (*Unequal Error Detection*) détection d'erreur inégale
 UEP (*Unequal Error Protection*) protection d'erreur inégale
 VAD (*Voice Activity Detection*) détection d'activité vocale

Le terme "bloc de trames" est utilisé dans le présent document pour décrire l'ensemble de trames de parole synchronisées dans une session multi canaux AMR ou AMR-WB. En particulier, dans une session à N canaux, un bloc de trames va contenir N trames de parole, une provenant de chacun des canaux, et toutes les N trames de parole représentant exactement la même période.

L'ordre des octets utilisé dans le présent document est l'ordre des octets du réseau, c'est-à-dire l'octet de poids fort en premier. L'ordre des bits est aussi le bit de poids fort en premier. Ceci est présenté dans toutes les figures comme ayant le bit de poids fort le plus à gauche sur la ligne et avec le numéro le plus faible. Certains champs binaires peuvent se dérouler sur plusieurs lignes et dans ce cas les bits de la première ligne sont de poids plus fort que les bits de la ligne suivante.

3. Généralités et principes de conception de AMR/AMR-WB

AMR et AMR-WB ont été à l'origine conçus pour les systèmes radioélectriques mobiles à commutation de circuit. Du fait de leur souplesse et leur robustesse, ils conviennent aussi pour d'autres services de communication de parole en temps réel sur des réseaux à commutation de paquets comme l'Internet.

À cause de la souplesse de ces codecs, le comportement dans une application particulière est contrôlé par plusieurs paramètres qui choisissent des options ou spécifient les valeurs acceptables pour une variable. Ces options et variables sont décrites en termes généraux aux endroits appropriés dans le texte de la présente spécification comme des paramètres à établir par des moyens hors bande. Dans la Section 8, tous les paramètres sont spécifiés sous la forme d'enregistrements de sous types de supports pour les codages AMR et AMR-WB. La méthode utilisée pour signaler ces paramètres à l'établissement de la session ou pour les établir par un accord préalable entre les participants sort du domaine d'application de ce document ; cependant, le paragraphe 8.3 donne une transposition des paramètres dans le protocole de description de session (SDP, *Session Description Protocol*) [RFC4566] pour les applications qui utilisent SDP.

3.1 Le codec vocal multi débits adaptifs (AMR)

Le codec AMR a été à l'origine développé et normalisé par l'Institut européen des normes de télécommunications (ETSI) pour les systèmes cellulaires GSM. Il est maintenant choisi par le projet en partenariat de troisième génération (3GPP) comme codec obligatoire pour les systèmes cellulaires de troisième génération (3G) [AMR-Code].

Le codec AMR est un codec multi modes qui prend en charge huit modes de codage de parole en bande étroite avec des débits binaires entre 4,75 et 12,2 kbit/s. La fréquence d'échantillonnage utilisée dans l'AMR est 8000 Hz et le codage de la parole est effectué sur des trames de parole de 20 ms. Donc, chaque trame de parole codée en AMR représente 160 échantillons de la parole originale.

Parmi les huit modes de codage AMR, trois sont déjà séparément adoptés comme normes autonomes. En particulier, le mode 6,7 kbit/s est adopté comme PDC-EFR [ARIB], le mode à 7,4 kbit/s comme codec IS-641 en AMRT [ACELP], et le mode 12,2 kbit/s comme GSM-EFR [GSM06.60].

3.2 Le codec vocal multi débits adaptifs large bande (AMR-WB)

Le codec de parole multi débits adaptifs large bande [AMR-WB] a été à l'origine développé par le 3GPP pour être utilisé dans les systèmes cellulaires GSM et 3G.

Similaire à l'AMR, le codec AMR-WB est aussi un codec de parole multi modes. AMR-WB prend en charge neuf modes de codage de la parole en large bande avec des débits binaires respectifs allant de 6,6 à 23,85 kbit/s. La fréquence d'échantillonnage utilisée en AMR-WB est de 16 000 Hz et le traitement de la parole est effectué sur des trames de 20 ms. Cela signifie que chaque trame codée en AMR-WB représente 320 échantillons de parole.

3.3. Codage multi débits et adaptation de mode

La capacité de codage multi débits (c'est-à-dire, multi modes) de AMR et AMR-WB est destinée à préserver une haute qualité de la parole sous une large gamme de conditions de transmission.

Avec AMR ou AMR-WB, les systèmes radio mobiles sont capables d'utiliser la bande passante disponible aussi efficacement que possible. Par exemple, dans le GSM il est possible d'ajuster dynamiquement le taux de codage de la parole durant une session afin de s'adapter continuellement aux conditions variables de la transmission en divisant la bande passante globale fixée entre les données de parole et le codage protecteur contre les erreurs. Cela permet le meilleur compromis possible entre le taux de compression de la parole et la tolérance à l'erreur. Pour effectuer l'adaptation de mode, le décodeur (receveur de la parole) a besoin de signaler au codeur (envoyeur de la parole) le nouveau mode préféré. Ce

signal de changement de mode est appelé une demande de mode codec (CMR, *Codec Mode Request*).

Comme dans la plupart des sessions la parole est envoyée dans les deux directions entre les deux extrémités, les demandes de mode provenant du décodeur à une extrémité au codeur à l'autre extrémité sont portées sur les trames de parole dans la direction inverse. En d'autres termes, il n'y a pas de signalisation hors bande nécessaire pour envoyer les CMR.

Il est exigé de chaque mise en œuvre de codec AMR ou AMR-WB qu'elle prenne en charge tous les modes respectifs de codage de la parole définis par le codec et elle doit être capable de traiter la commutation de mode à tous les modes à tout moment. Cependant, certains systèmes de transport peuvent imposer des limitations au nombre de modes pris en charge et au nombre de fois que le mode peut changer à cause de limitations de bande passante ou d'autres contraintes. Pour cette raison, il est permis au décodeur d'indiquer son accord pour un mode particulier ou un sous ensemble des modes définis pour la session en utilisant des moyens hors bande.

Par exemple, la liaison radio GSM peut seulement utiliser un sous ensemble d'au plus quatre modes différents dans une certaine session. Ce sous ensemble peut être toute combinaison des huit modes AMR pour une session AMR ou toute combinaison des neuf modes AMR-WB pour une session AMR-WB.

De plus, pour une meilleure interopérabilité avec le GSM à travers une passerelle, il est permis au décodeur d'utiliser des moyens hors bande pour régler le nombre minimum de trames entre deux changements de mode et de limiter les changements de mode seulement parmi les modes voisins.

La Section 8 spécifie un ensemble de paramètres de type de supports qui peut être utilisé pour signaler ces contrôles d'adaptation de mode à l'établissement de la session.

3.4 Détection d'activité vocale et transmission discontinue

Les deux codecs prennent en charge la détection d'activité vocale (VAD, *Voice Activity Detection*) et la génération des paramètres de bruit de confort (CN, *Comfort Noise*) durant les périodes de silence. Donc, les codecs ont l'option de réduire à un minimum le nombre de bits et de paquets transmis durant les périodes de silence. L'opération d'envoi des paramètres de CN à des intervalles réguliers durant les périodes de silence est généralement appelé l'opération de transmission discontinue (DTX, *Discontinuous Transmission*) ou de taux de source contrôlée (SCR, *Source Controlled Rate*). Les trames AMR ou AMR-WB contenant des paramètres de CN sont appelées des trames d'indicateur de silence (SID, *Silence Indicator*). Voir plus de détails sur les fonctions de VAD et DTX dans [TS26.092] et [TS26.192].

3.5 Prise en charge de session multi canaux

Les deux formats de charge utile RTP et de mémorisation définis dans le présent document prennent en charge le contenu audio multi canaux (par exemple, une session de parole stéréophonique).

Bien que les codecs AMR et AMR-WB eux-mêmes ne prennent pas en charge le codage de contenu audio multi canaux dans un seul flux binaire, ils peuvent être utilisés pour coder et décoder séparément chaque canal individuel.

Pour transporter (ou mémoriser) les contenus multi canaux codés séparément, les trames de parole pour chaque canal tramées et codées pour les mêmes périodes de 20 ms sont logiquement collectées dans un bloc de trames.

À l'établissement de la session, une signalisation hors bande doit être utilisée pour indiquer le nombre de canaux dans la session, et l'ordre des trames de parole provenant de canaux différents dans chaque bloc de trames. Quand SDP est utilisé pour la signalisation, le nombre de canaux est spécifié dans l'attribut `rtpmap` et l'ordre des canaux portés dans chaque bloc de trames est impliqué par le nombre de canaux comme spécifié au paragraphe 4.1 de la [RFC3551].

3.6 Détection et protection inégales des erreurs binaires

Les bits de parole codés dans chaque trame AMR ou AMR-WB ont une sensibilité de perception différente aux erreurs binaires. Cette propriété a été exploitée dans les systèmes cellulaires pour réaliser une meilleure qualité vocale en utilisant les mécanismes de protection et de détection d'erreur inégale (UEP et UED).

Les mécanismes d'UEP/UED se concentrent sur la protection et la détection de bits corrompus sur les bits perçus comme les plus sensibles dans une trame AMR ou AMR-WB. En particulier, les bits de parole dans une trame AMR ou AMR-WB sont divisés en classes A, B, et C, où les bits de classe A sont les plus sensibles et les bits de classe C les moins sensibles

(voir le Tableau 1 ci-dessous pour AMR et [AMR-WB-Struc] pour AMR-WB). Une trame AMR ou AMR-WB est seulement déclarée endommagée si des erreurs de bit sont trouvées dans les bits les plus sensibles, c'est-à-dire, de classe A. Par ailleurs, il est acceptable d'avoir des erreurs de bit sur les autres bits, c'est-à-dire, de classes B et C.

Indice	Mode	Bits de classe A	Bits de parole totaux
0	AMR à 4,75	42	95
1	AMR à 5,15	49	103
2	AMR à 5,9	55	118
3	AMR à 6,7	58	134
4	AMR à 7,4	61	148
5	AMR à 7,95	75	159
6	AMR à 10,2	65	204
7	AMR à 12,2	81	244
8	AMR SID	39	39

Tableau 1 : nombre de bits de classe A pour le codec AMR

De plus, une trame endommagée est encore utile pour la dissimulation d'erreur au décodeur car certains des bits les moins sensibles peuvent encore être utilisés. Cette approche peut améliorer la qualité de la parole par rapport à l'élimination de la trame endommagée.

3.6.1 Application de UEP et UED dans un réseau IP

Pour tirer pleinement parti de la robustesse aux erreurs de bits des codecs AMR et AMR-WB, le format de charge utile RTP est conçu pour faciliter l'UEP/UED dans un réseau IP. On devrait noter cependant que l'utilisation de UEP et UED discutée ci-dessous est FACULTATIVE.

L'UEP/UED dans un réseau IP peut être réalisée par la détection des erreurs de bit dans les bits de classe A et en tolérant les erreurs de bit dans les classes B/C de la trame AMR ou AMR-WB dans chaque charge utile RTP.

Il existe des protocoles de couche de liaison qui n'éliminent pas les paquets qui contiennent des erreurs de bit, par exemple, SLIP et certaines liaisons sans fil. Avec le schéma de trafic Internet glissant vers une approche plus centrée sur le multimédia, plus de couches de liaison de cette nature pourront émerger à l'avenir. Avec la prise en charge par la couche de transport de sommes de contrôle partielles (par exemple, celles prises en charge par UDP léger [RFC3828]) le trafic AMR et AMR-WB tolérant aux erreurs de bits pourrait réaliser de meilleures performances sur ces types de liaisons. La relation entre la somme de contrôle partielle de UDP léger à la couche de transport et la couverture de somme de contrôle fournie par la trame de couche de liaison est décrite dans la spécification de UDP léger [RFC3828].

Il y a au moins deux approches de base pour porter le trafic AMR et AMR-WB sur des réseaux IP tolérants aux erreurs de bit :

- Utiliser une somme de contrôle partielle pour couvrir les en-têtes IP, de protocole de transport (par exemple, UDP léger) RTP et de charge utile, et les bits de parole les plus importants de la charge utile. Les en-têtes IP, UDP et RTP ont besoin d'être protégés, et il est recommandé qu'au moins tous les bits de classe A soient couverts par la somme de contrôle.
- Utiliser une somme de contrôle partielle pour couvrir seulement les en-têtes IP, de protocole de transport, RTP et de charge utile, mais un CRC de trame AMR ou AMR-WB pour couvrir les bits de classe A de chaque trame de parole dans la charge utile RTP.

Dans l'une et l'autre approche, au moins une partie des bits de classes B/C sont laissés sans vérification d'erreur et donc la tolérance aux erreurs de bits est réalisée.

Note : il est quand même important que les concepteurs de réseaux fassent attention au taux résiduel d'erreurs de bits de classes B et C. Bien que moins sensibles aux erreurs que les bits de classe A, les bits des classes B et C ne sont pas insignifiants, et des erreurs non détectées dans ces bits causent une dégradation de la qualité de la parole. Un exemple de taux d'erreurs résiduelles considérés comme acceptables pour AMR dans le système universel de télécommunications mobiles (UMTS, *Universal Mobile Telecommunications System*) se trouve dans [TS26.102] et pour AMR-WB dans [TS26.202].

L'interface d'application au protocole de transport UEP/UED (par exemple, UDP léger) peut ne pas fournir de contrôle sur le taux d'erreur de la liaison, en particulier dans un scénario de passerelle. Donc, il appartient au concepteur d'un nœud avec une interface de liaison de ce type de choisir un taux résiduel d'erreur de bit qui soit assez bas pour prendre en charge des applications telles que le codage AMR quand elles transmettent des paquets d'un protocole de transport UEP/UED.

L'approche 1 est efficace du point de vue des bits, souple et simple, mais présente deux inconvénients, a) les erreurs de bits dans les bits de parole protégés vont causer l'élimination de la charge utile, et b) quand plusieurs trames AMR ou AMR-WB sont transportées dans une charge utile RTP, il y a la possibilité qu'une seule erreur de bit dans les bits protégés cause l'élimination de toutes les trames.

Ces inconvénients peuvent être évités, si nécessaire, avec quelques frais généraux sous la forme d'un CRC par trame (approche 2). Dans le problème a), le CRC rend possible la détection des erreurs de bits dans la classe A et il utilise la trame pour la dissimulation d'erreur, ce qui donne une petite amélioration de la qualité de la parole. Pour b), lors du transport de plusieurs trames dans une charge utile, les CRC suppriment la possibilité qu'une seule erreur de bit dans la classe A causent l'élimination de toutes les trames. Éviter cela améliore la qualité de la parole quand plusieurs trames AMR ou AMR-WB sont transportées sur des liaisons sujettes à des erreurs de bits.

Le choix entre les deux approches ci-dessus doit être fait sur la base de la bande passante disponible, et de la tolérance désirée aux erreurs de bits. Aucune des solutions n'est appropriée pour tous les cas. La Section 8 définit les paramètres qui peuvent être utilisés à l'établissement de session pour choisir entre ces approches.

3.7 Robustesse à la perte de paquets

Le format de charge utile prend en charge plusieurs moyens, incluant la correction d'erreur directe (FEC, *Forward Error Correction*) et l'entrelacement de trames, pour augmenter la robustesse à la perte de paquets.

3.7.1 Utilisation de la correction d'erreur directe (FEC)

Le schéma simple de répétition des données envoyées précédemment est une façon de réaliser la FEC. Un autre schéma possible qui est plus efficace pour la bande passante est d'utiliser la FEC externe à la charge utile, par exemple, de la [RFC2733], qui génère des paquets supplémentaires contenant les données de réparation. La charge utile entière peut aussi être triée dans un ordre approprié pour prendre en charge les schémas de FEC externes qui utilisent l'UEP. Il y a aussi un travail en cours sur une version générique d'un tel schéma [RFC5109] qui peut être appliqué au transport de charge utile AMR ou AMR-WB.

Avec AMR ou AMR-WB, il est possible d'utiliser la capacité multi débits du codec pour envoyer des copies redondantes d'une trame en utilisant le même mode ou un autre, par exemple, un avec une bande passante inférieure. On décrit ci-après un tel schéma.

Ceci implique la simple retransmission des blocs de trames précédemment transmis avec le ou les blocs de trames actuels. Ceci est fait en utilisant une fenêtre glissante pour grouper les blocs de trames de parole pour les envoyer dans chaque charge utile. La Figure 1 en montre un exemple.

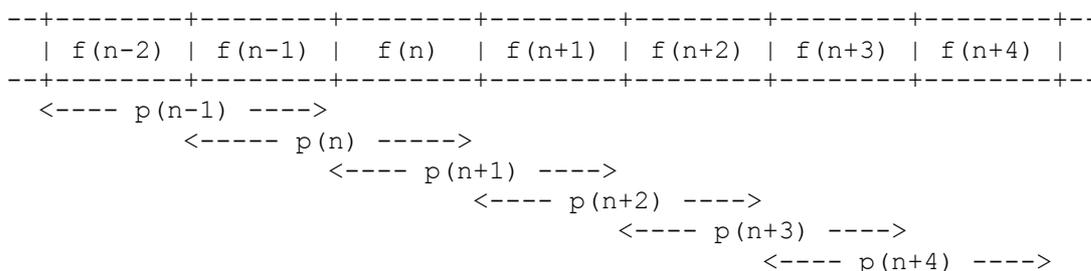


Figure 1 : exemple de transmission redondante

Dans cet exemple, chaque bloc de trames est retransmis une fois dans le paquet de charge utile RTP suivant. Ici, $f(n-2)..f(n+4)$ note une séquence de blocs de trames de parole, et $p(n-1)..p(n+4)$ une séquence de paquets de charge utile.

L'utilisation de cette approche n'exige pas de signalisation à l'établissement de la session. Cependant, un paramètre pour fournir un délai maximum de transmission de toute trame redondante est défini à la Section 8. En d'autres termes,

L'envoyeur de parole peut choisir d'utiliser ce schéma sans consulter le receveur. C'est parce qu'un paquet contenant des trames redondantes ne va pas paraître différent d'un paquet avec seulement de nouvelles trames. Le receveur peut recevoir plusieurs copies ou versions (codées avec des modes différents) d'une trame pour un certain horodatage si aucun paquet n'est perdu. Si plusieurs versions de la même trame de parole sont reçues, il est recommandé que le mode avec le plus haut débit soit utilisé par le décodeur de parole.

Ce schéma de redondance fournit la même fonctionnalité que celui décrit dans la [RFC2198], "Charge utile RTP pour données audio redondantes". Dans la plupart des cas, le mécanisme dans ce format de charge utile est plus efficace et plus simple que d'exiger des deux points d'extrémité qu'ils prennent en plus en charge la RFC 2198. Il y a deux situations dans lesquelles l'utilisation de la RFC 2198 est indiquée : si l'étalement dans le temps requis entre le codage primaire et le codage redondant est supérieur à la durée de 5 trames, les frais généraux de bande passante de la RFC 2198 vont être inférieurs ; ou, si un codec non AMR est désiré pour le codage redondant, le format de charge utile AMR ne va pas être capable de le porter.

L'envoyeur est responsable du choix d'une quantité appropriée de redondance sur la base des retours sur le canal, par exemple, dans les rapports de receveur RTCP. Un envoyeur ne devrait pas fonder le choix de la FEC sur la CMR, car ce paramètre a très probablement été établi sur la base d'informations non IP, par exemple, des mesures de performances de liaison radio. L'envoyeur est aussi responsable de l'évitement d'encombrement, qui peut être exacerbé par la redondance (voir les détails à la Section 6).

3.7.2 Utilisation de l'entrelacement de trames

Pour diminuer les frais généraux du protocole, la conception de la charge utile permet que plusieurs blocs de trames de parole soient encapsulés dans un seul paquet RTP. Un des inconvénients d'une telle approche est que des pertes de paquets peuvent causer la perte de plusieurs blocs consécutifs de trames de parole, ce qui cause généralement une distorsion clairement audible dans la parole reconstruite. L'entrelacement de blocs de trames peut améliorer la qualité de la parole dans de tels cas en répartissant les pertes consécutives dans une série de pertes d'un seul bloc de trames. Cependant, l'entrelacement et la mise en bouquets de plusieurs blocs de trames par charge utile vont aussi augmenter le délai de bout en bout et ce n'est donc pas approprié pour tous les types d'applications. Les applications de flux directs vont très probablement être capables d'exploiter l'entrelacement pour améliorer la qualité de la parole dans des conditions de transmission à pertes.

Cette conception de charge utile prend en charge l'utilisation de l'entrelacement de trames comme une option. Pour que le codeur (envoyeur de parole) utilise l'entrelacement de trames dans ses paquets RTP sortants pour une certaine session, le décodeur (receveur de la parole) doit indiquer qu'il le prend en charge via des moyens hors bande (Section 8).

3.8 Mode efficace en bande passante ou alignement sur l'octet

Pour une session donnée, le format de charge utile peut être efficace en bande passante ou aligné sur l'octet, selon le mode de fonctionnement qui est établi pour la session via des moyens hors bande.

Dans le format aligné sur l'octet, tous les champs d'une charge utile, incluant l'en-tête de charge utile, les entrées de table des matières, et les trames de parole elles-mêmes, sont individuellement alignés sur les limites d'octet pour rendre les mises en œuvre efficaces. Dans le format efficace en bande passante, c'est seulement la charge utile entière qui est alignée sur l'octet, donc moins de bits de bourrage sont ajoutés.

Note : l'alignement sur l'octet d'un champ ou d'une charge utile signifie que le dernier octet est bourré avec des zéros dans les bits de moindre poids pour compléter l'octet. On notera aussi que ce bourrage est distinct du bourrage indiqué par le bit P dans l'en-tête RTP.

Entre les deux modes de fonctionnement, seul le mode aligné sur l'octet a la capacité d'utiliser le tri robuste, l'entrelacement, et le CRC de trame pour rendre le transport de parole plus robuste à la perte de paquet et aux erreurs de bits.

3.9 Scénarios AMR ou AMR-WB de voix sur IP

Le principal scénario pour ce format de charge utile est IP de bout en bout entre deux terminaux, comme le montre la Figure 2. Ce format de charge utile est supposé être utile pour les deux services de conversation et de flux directs.

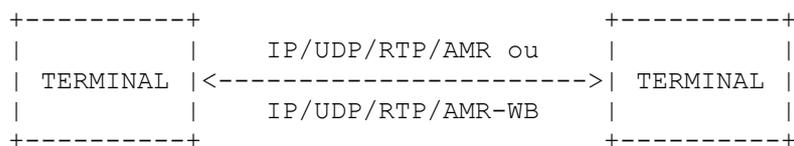


Figure 2 : scénario de terminal IP à terminal IP

Un service de conversation met des exigences sur le format de charge utile. Un faible délai est un facteur très important, c'est-à-dire, peu de blocs de trames de parole par paquet de charge utile. De faibles frais généraux sont aussi requis quand le format de charge utile traverse des liaisons à faible bande passante, en particulier quand la fréquence des paquets va être élevée. Pour les liaisons à faible bande passante, il est aussi avantageux de prendre en charge l'UED, qui permet à un fournisseur de liaison de réduire les délais et la perte de paquets, ou de réduire l'utilisation des ressources de liaisons.

Un service de flux directs a moins d'exigences strictes de temps réel et donc peut utiliser un plus grand nombre de blocs de trames par paquet qu'un service de conversation. Cela réduit les frais généraux des en-têtes IP, UDP, et RTP. Cependant, inclure plusieurs blocs de trames par paquet rend la transmission plus vulnérable à la perte de paquets, de sorte que l'entrelacement peut être utilisé pour réduire l'effet que la perte de paquet va avoir sur la qualité de la parole. Un serveur de flux directs qui traite un grand nombre de clients a aussi besoin d'un format de charge utile qui exige aussi peu de ressources que possible quand il fait la mise en paquets. Les modes aligné sur l'octet et d'entrelacement exigent le moins de ressources, tandis que les modes CRC, de tri robuste, et d'efficacité de la bande passante ont des exigences supérieures.

Un autre scénario est quand la parole codée en AMR ou AMR-WB est transmise d'un système non IP (par exemple, un réseau GSM ou 3GPP UMTS) à un terminal VoIP IP/UDP/RTP, et/ou vice versa, comme décrit par la Figure 3.

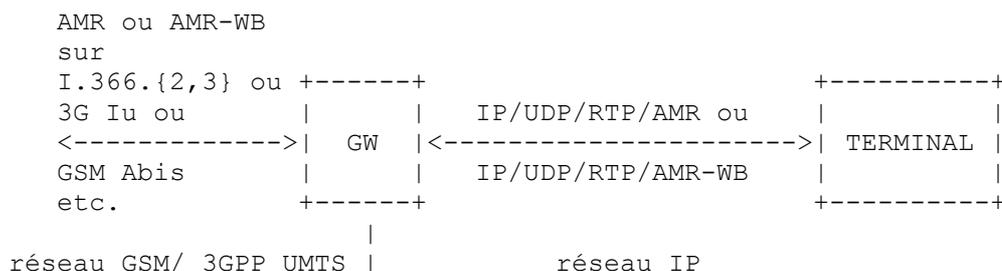


Figure 3 : scénario de passerelle à terminal VoIP

Dans ce cas, il est probable que la trame AMR ou AMR-WB est mise en paquets d'une façon différente dans le réseau non IP et va devoir être remise en paquets dans RTP à la passerelle (*GW*). Aussi, les trames de parole provenant du réseau non IP peuvent venir avec des informations d'UEP/UED (par exemple, un indicateur de qualité de trame) qui vont devoir être préservées et transmises au décodeur avec les bits de parole. C'est ce qui est précisé au paragraphe 4.3.2.

La capacité de AMR de faire une commutation de mode rapide est exploitée dans certains réseaux non IP pour optimiser la qualité de la parole. Pour préserver cette fonctionnalité dans des scénarios incluant une passerelle vers un réseau IP, un champ Demande de mode codec (CMR) est nécessaire. La passerelle va être chargée de transmettre la CMR entre les parties non IP et IP dans les deux directions. Le terminal IP devrait suivre la CMR transmise par la passerelle pour optimiser la qualité de la parole allant au décodeur non IP. L'algorithme de contrôle de mode dans la passerelle doit s'accommoder du délai imposé par le réseau IP à la réponse du terminal IP à la CMR.

Le terminal IP ne devrait pas établir la CMR (voir au paragraphe 4.3.1) mais la passerelle peut régler la valeur de la CMR sur les trames qui vont vers le codeur dans la partie non IP pour optimiser la qualité de la parole du codeur à la passerelle. La passerelle peut autrement établir une valeur de CMR inférieure, si désiré, comme moyen de contrôler l'encombrement sur le réseau IP.

Un troisième scénario probable est que IP/UDP/RTP est utilisé comme transport entre deux systèmes non IP, c'est-à-dire, IP est à l'origine et est terminé dans des passerelles des deux côtés du transport IP, comme illustré dans la Figure 4.

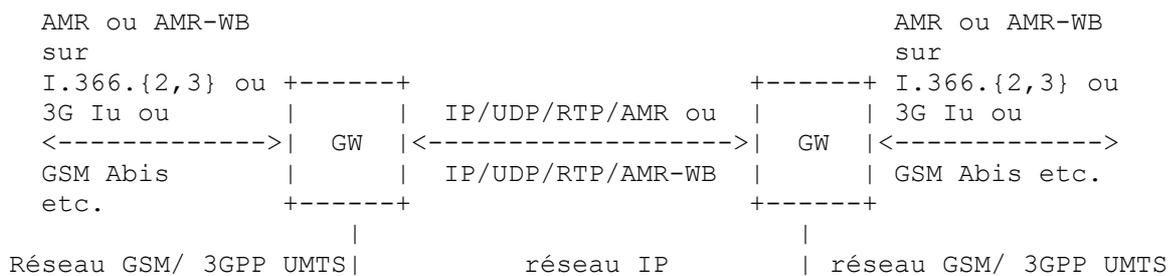


Figure 4 : scénario de passerelle à passerelle

Ce scénario exige les mêmes mécanismes pour préserver les informations UED/UEP et CMR que dans le scénario d'une seule passerelle. De plus, la valeur de CMR peut être réglée dans les paquets reçus par les passerelles sur le côté réseau IP. La passerelle devrait transmettre au côté non IP une valeur de CMR qui est le minimum de trois valeurs :

- la valeur de CMR qu'elle reçoit sur le côté IP ;
- la valeur de CMR qu'elle calcule sur la base de sa qualité de réception sur le côté non IP ; et
- une valeur de CMR qu'elle peut choisir pour le contrôle d'encombrement de transmission sur le côté IP.

Les détails de l'algorithme de contrôle relèvent de la mise en œuvre.

4. Formats de charge utile RTP AMR et AMR-WB

Les formats de charge utile AMR et AMR-WB ont une structure identique, et ils sont spécifiés ensemble. Les seules différences sont dans les types de trames de codec contenues dans la charge utile. Le format de charge utile consiste en l'en-tête RTP, l'en-tête de charge utile, et les données de charge utile.

4.1 Usage de l'en-tête RTP

Le format de l'en-tête RTP est spécifié dans la [RFC3550]. Ce format de charge utile utilise les champs de l'en-tête d'une manière cohérente avec la présente spécification.

L'horodatage RTP correspond à l'instant d'échantillonnage du premier échantillon codé pour le premier bloc de trames dans le paquet. La fréquence d'horloge de l'horodatage est la même que la fréquence d'échantillonnage, de sorte que l'unité d'horodatage est en échantillons.

La durée d'un bloc de trames de parole est de 20 ms pour AMR et AMR-WB. Pour AMR, la fréquence d'échantillonnage est 8 kHz, correspondant à 160 échantillons de parole codés par trame provenant de chaque canal. Pour AMR-WB, la fréquence d'échantillonnage est 16 kHz, correspondant à 320 échantillons par trame provenant de chaque canal. Donc, l'horodatage est augmenté de 160 pour AMR et de 320 pour AMR-WB pour chaque bloc de trames consécutif.

Un paquet peut contenir plusieurs blocs de trames de parole codées ou de paramètres de bruit de confort. Si l'entrelacement est employé, les blocs de trames encapsulés dans une charge utile sont pris en accord avec les règles d'entrelacement comme défini au paragraphe 4.4.1. Autrement, chaque paquet couvre une période d'un ou plusieurs intervalles de bloc de trame contigus de 20 ms. Dans le cas où les données provenant de tous les canaux pour un bloc de trames particulier dans la période manquent (par exemple, à une passerelle provenant de quelque autre format de transport) il est possible d'indiquer qu'aucune donnée n'est présente pour ce bloc de trames plutôt que de couper en deux un paquet de bloc multi trames, comme expliqué au paragraphe 4.3.2.

Pour permettre la résilience à l'erreur par une transmission redondante, les périodes couvertes par plusieurs paquets PEUVENT se chevaucher. Un receveur DOIT être prêt à recevoir toute trame de parole plusieurs fois, en dupliqués exacts, dans différents modes de débit AMR, ou avec des données présentes dans un paquet et pas dans l'autre. Si plusieurs versions de la même trame de parole sont reçues, il est RECOMMANDÉ que le mode avec le débit le plus élevé soit utilisé par le décodeur de la parole. Une certaine trame NE DOIT PAS être codée comme parole dans un paquet et comme paramètres de bruit de confort dans une autre.

La longueur de charge utile est toujours transformée en un nombre entier d'octets en bourrant avec des bits de zéro si nécessaire. Si du bourrage additionnel est requis pour amener la longueur de la charge utile à un plus grand multiple

d'octets ou pour une autre raison, alors le bit P dans l'en-tête RTP peut être établi et du bourrage ajouté comme spécifié dans la [RFC3550].

Le bit marqueur (M) de l'en-tête RTP DEVRA être réglé à 1 si le premier bloc de trames porté dans le paquet contient une trame de parole qui est la première d'une salve de paroles. Pour tous les autres paquets, le bit marqueur DEVRA être réglé à zéro (M=0).

L'allocation d'un type de charge utile RTP pour ce nouveau format de paquet sort du domaine d'application du présent document, et ne sera pas spécifiée ici. Il est prévu que le profil RTP sous lequel ce format de charge utile est utilisé allouera un type de charge utile pour ce codage ou spécifiera que le type de charge utile est à fixer de façon dynamique.

4.2 Structure de charge utile

La charge utile complète consiste en un en-tête de charge utile, une table des matières de charge utile, et en données de parole représentant un ou plusieurs blocs de trames de parole. Le diagramme qui suit montre la disposition du format général de charge utile :

```
+-----+-----+-----+
| en-tête ch. ut. | table des matières | données de parole ...
+-----+-----+-----+
```

Les charge utiles qui contiennent plus d'un bloc de trames de parole sont appelées des charges utiles composées.

Les paragraphes qui suivent décrivent les variations du format de charge utile selon que la session AMR est établie pour utiliser le mode efficace en bande passante ou le mode aligné sur l'octet et toutes fonctions FACULTATIVES pour un tri robuste, pour l'entrelacement, et les CRC de trame. Les mises en œuvre DEVRAIENT prendre en charge le fonctionnement des deux modes efficace en bande passante et aligné sur l'octet pour accroître l'interopérabilité.

4.3 Mode efficace en bande passante

4.3.1 En-tête de charge utile

En mode efficace en bande passante, l'en-tête de charge utile consiste simplement en 4 bits de demande de mode codec :

```
 0 1 2 3
+++++++
| CMR |
+++++++
```

CMR (4 bits) : indique une demande de mode codec envoyée au codeur de parole au site du receveur de cette charge utile.

La valeur du champ CMR est réglée à l'indice de type de trame du mode de parole correspondant demandé. L'indice de type de trame peut être de 0 à 7 pour AMR, comme défini au Tableau 1a de [AMR-struct], ou de 0 à 8 pour AMR-WB, comme défini au Tableau 1a de [AMR-WB-Struct]. La valeur de CMR de 15 indique qu'aucune demande de mode n'est présente, et les autres valeurs sont pour des utilisations futures.

La demande de mode codec reçue dans le champ CMR est valide jusqu'à la réception de la prochaine demande de mode codec, c'est-à-dire, d'une nouvelle valeur de CMR reçue correspondant à un mode de parole, ou qu'un NO_DATA outrepassa la valeur de CMR reçue précédemment correspondant à un mode de parole ou à NO_DATA. Donc, si un terminal souhaite recevoir continuellement des trames dans le même mode X, il doit régler CMR=X pour toutes ses charges utiles sortantes, et si un terminal n'a pas de préférence sur le mode à recevoir, il DEVRAIT régler CMR=15 dans toutes ses charges utiles sortantes.

Si une charge utile est reçue avec une valeur de CMR qui n'est pas un mode de parole ou NO_DATA, la CMR DOIT être ignorée par le receveur.

Dans une session multi canaux, la demande de mode codec DEVRAIT être interprétée par le receveur de la charge utile comme étant le mode de codage désiré pour tous les canaux dans la session.

Un point d'extrémité IP NE DEVRAIT PAS régler la demande de mode codec sur la base des pertes de paquet ou autre indication d'encombrement, pour plusieurs raisons :

- l'autre extrémité du chemin IP peut être une passerelle avec un réseau non IP (comme une liaison radio) qui a besoin d'établir le champ CMR pour optimiser les performances sur ce réseau ;
- l'encombrement sur le réseau IP est géré par l'expéditeur IP, dans ce cas, à l'autre extrémité du chemin IP. Les retours sur l'encombrement DEVRAIENT être fournis à cet expéditeur IP par RTCP ou autres moyens, et ensuite l'expéditeur peut choisir d'éviter l'encombrement en utilisant le mécanisme le plus approprié. Cela peut inclure d'ajuster le mode de codec, mais aussi d'ajuster le niveau de redondance ou le nombre de trames par paquet.

Le codeur DEVRAIT suivre une demande de mode codec reçue, mais PEUT changer pour un mode de numéro inférieur si il le choisit, par exemple, pour contrôler l'encombrement.

Le champ CMR DOIT être réglé à 15 pour les paquets envoyés à un groupe de diffusion groupée. Le codeur de l'expéditeur de parole DEVRAIT ignorer les demandes de mode de codec lors de l'envoi de parole à une session de diffusion groupée mais PEUT utiliser les informations de retour de RTCP comme indication qu'un changement de mode de codec est nécessaire.

Le choix du mode de codec PEUT être restreint par un paramètre de session à un sous ensemble des modes disponibles. Si il en est ainsi, le mode demandé DOIT être dans le sous ensemble signalé (voir à la Section 8). Si la valeur de CMR reçue est en dehors du sous ensemble signalé de modes, elle DOIT être ignorée.

4.3.2 Charge utile Table des matières

La table des matières (ToC) consiste en une liste des entrées de ToC, chacune représentant une trame de parole.

Dans le mode efficace en bande passante, une entrée de ToC a le format suivant :

```

0 1 2 3 4 5
+-----+
| F |  FT  | Q |
+-----+
```

F (1 bit) : réglé à 1, il indique que cette trame est suivie par une autre trame de parole dans cette charge utile ; réglé à 0, il indique que cette trame est la dernière dans cette charge utile.

FT (4 bits) : indice de type de trame, indiquant le mode de codage de parole AMR ou AMR-WB ou le bruit de confort (SID) de la trame correspondante portée dans cette charge utile.

La valeur de FT est définie au Tableau 1a de [AMR-struct] pour AMR et au Tableau 1a de [AMR-WB-Struct] pour AMR-WB. FT=14 (SPEECH_LOST, seulement disponible pour AMR-WB) et FT=15 (NO_DATA) sont utilisées pour indiquer, respectivement, des trames qui sont perdues ou non transmises dans cette charge utile.

Une trame NO_DATA (FT=15) pourrait signifier qu'aucune donnée pour cette trame n'a été produite par le codeur de parole ou qu'aucune donnée pour cette trame n'est transmise dans la charge utile en cours (c'est-à-dire, des données valides pour cette trame pourraient avoir été envoyées dans un paquet antérieur ou pourraient l'être dans un paquet ultérieur).

Si une entrée de ToC avec une valeur de FT dans la gamme de 9 à 14 est reçue pour AMR ou de 10 à 13 pour AMR-WB, le paquet entier DEVRAIT être éliminé. C'est pour éviter la perte de synchronisation des données dans le processus de dépaquetisation, qui peut résulter en une énorme dégradation de la qualité de la parole.

Noter que les paquets contenant seulement des trames NO_DATA NE DEVRAIENT PAS être transmis dans une configuration de format de charge utile, sauf dans le cas d'entrelacement. Aussi, les blocs de trames contenant seulement des trames NO_DATA à la fin d'un paquet NE DEVRAIENT PAS être transmises dans une configuration de format de charge utile, sauf dans le cas d'entrelacement. Le SCR/DTX AMR est décrit dans [AMR-WB-SCR] et le SCR/DTX AMR-WB dans le [TS26.193].

Les types de trame de bruit de confort supplémentaires spécifiés dans le Tableau 1a de [AMR-struct] (c'est-à-dire, GSM-EFR CN, IS-641 CN, et PDC-EFR CN) NE DOIVENT PAS être utilisés dans ce format de charge utile parce que le codec AMR normalisé est seulement obligé de mettre en œuvre le type de trame SID général AMR et non ceux qui sont

natifs des codages incorporés.

Q (1 bit) : indicateur de qualité de trame. Réglé à 0, il indique que la trame correspondante est sévèrement endommagée, et que le receveur devrait régler le RX_TYPE (voir [AMR-WB-SCR]) à SPEECH_BAD ou à SID_BAD selon le type de trame (FT).

L'indicateur de qualité de trame est inclus pour l'interopérabilité avec le format de charge utile ATM décrit dans la Recommandation UIT-T I.366.2, avec l'interface UMTS Iu [TS25.415], ainsi qu'avec d'autres formats de transport. L'indicateur de qualité de trame permet que des trames endommagées soient transmises au décodeur de parole pour la dissimulation d'erreurs. Cela peut améliorer la qualité de la parole mieux que d'éliminer les trames endommagées. Voir plus de détails au paragraphe 4.4.2.1.

Pour les sessions multi canaux, les entrées de ToC de toutes les trames d'un bloc de trames sont placées dans la ToC en ordre consécutif comme défini au paragraphe 4.1 de la [RFC3551]. Quand plusieurs blocs de trames sont présents dans un paquet en mode d'efficacité de bande passante, ils vont être placés dans le paquet dans l'ordre de leur instant de création.

Donc, avec N canaux et K blocs de trames de parole dans un paquet, il DOIT y avoir N*K entrées dans la ToC, et les N premières entrées vont être du premier bloc de trames, les N secondes entrées vont être du second bloc de trames, et ainsi de suite.

La figure qui suit montre un exemple de ToC de trois entrées dans une session d'un seul canal en utilisant le mode d'efficacité de bande passante.

```

0                               1
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7
+-----+-----+-----+-----+-----+
|1| FT | Q|1| FT | Q|0| FT | Q|
+-----+-----+-----+-----+

```

Voici un exemple de la façon dont les entrées de ToC vont apparaître dans la ToC d'un paquet portant trois blocs de trames consécutifs dans une session avec deux canaux (L et R).

```

+-----+-----+-----+-----+-----+
| 1L | 1R | 2L | 2R | 3L | 3R |
+-----+-----+-----+-----+
|<----->|<----->|<----->|
  Bloc de   Bloc de   Bloc de
  trames 1   trames 2   trames 3

```

4.3.3 Données de parole

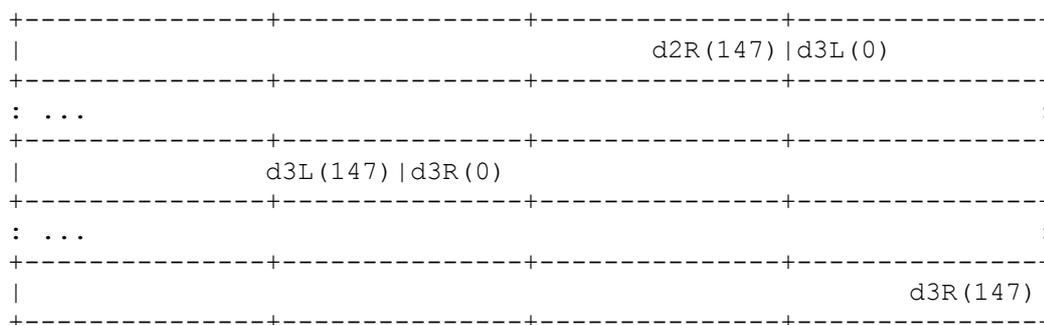
Les données de parole d'une charge utile contiennent zéro, une ou plusieurs trames de parole ou trames de bruit de confort, comme décrit dans la ToC de la charge utile.

Note : pour les entrées de ToC avec FT=14 ou 15, il ne va pas y avoir de trame de parole correspondante présente dans les données de parole.

Chaque trame de parole représente 20 ms de parole codée avec le mode indiqué dans le champ FT de l'entrée de ToC correspondante. La longueur de la trame de parole est implicitement définie par le mode indiqué dans le champ FT. La notation de l'ordre et de la numérotation des bits est comme spécifié pour le format d'interface 1 (IF1, *Interface Format 1*) dans [AMR-struct] pour AMR et [AMR-WB-Struct] pour AMR-WB. Comme spécifié ici, les bits des trames de parole ont été réarrangés en ordre de sensibilité décroissante, tandis que les bits de trames de bruit de confort sont dans l'ordre produit par le codeur. La séquence de bits résultante pour une trame de longueur K bits est notée d(0), d(1), ..., d(K-1).

4.3.4 Algorithme de formation de la charge utile

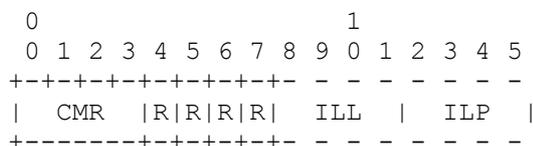
La charge utile RTP complète en mode d'efficacité de la bande passante est formée en mettant en paquet les bits provenant de l'en-tête de la charge utile, de la table des matières, et des trames de parole dans l'ordre (comme défini par leurs entrées de ToC correspondantes dans la liste de la ToC) et en alignant la charge utile sur l'octet, avec de 0 à 7 bits de bourrage. Les bits de bourrage DOIVENT être réglés à zéro et DOIVENT être ignorés à réception. Ils sont empaquetés de façon contiguë



4.4 Mode aligné sur l'octet

4.4.1 En-tête de charge utile

Dans le mode aligné sur l'octet, l'en-tête de la charge utile consiste en une CMR de 4 bits, 4 bits réservés, et facultativement, un en-tête d'entrelacement de huit bits, comme montré ci-dessous :



CMR (4 bits) : comme défini au paragraphe 4.3.1.

R : est un bit réservé qui DOIT être réglé à zéro. Tous les bits R DOIVENT être ignorés par le receveur.

ILL (4 bits, entier non signé) : c'est un champ FACULTATIF qui n'est présent que si l'entrelacement est signalé hors bande pour la session. ILL=L indique au receveur que la longueur de l'entrelacement est L+1, en nombre de blocs de trames.

ILP (4 bits, entier non signé) : c'est un champ FACULTATIF qui n'est présent que si l'entrelacement est signalé. ILP DOIT prendre une valeur entre 0 et ILL, inclus, indiquant l'indice d'entrelacement pour les blocs de trames dans cette charge utile dans le groupe d'entrelacement. Si la valeur de ILP est supérieure à ILL, la charge utile DEVRAIT être éliminée.

Les champs ILL et ILP DOIT être présents dans chaque paquet d'une session si l'entrelacement est signalé pour la session. L'entrelacement DOIT être effectué sur la base d'un bloc de trames (c'est-à-dire, PAS d'une trame) dans une session multi canaux.

L'exemple suivant illustre l'arrangement des blocs de trames de parole dans un groupe d'entrelacement durant une session d'entrelacement. On suppose ici que ILL=L pour le groupe d'entrelacement qui commence au bloc de trames de parole n. On suppose aussi que le premier paquet de charge utile du groupe d'entrelacement est s, et le nombre de blocs de trame de parole portés dans chaque charge utile est N. On va alors avoir :

La charge utile s (le premier paquet de ce groupe d'entrelacement) : ILL=L, ILP=0,
porte les blocs de trames : n, n+(L+1), n+2*(L+1), ..., n+(N-1)*(L+1)

La charge utile s+1 (le second paquet de ce groupe d'entrelacement) : ILL=L, ILP=1,
blocs de trames : n+1, n+1+(L+1), n+1+2*(L+1), ..., n+1+(N-1)*(L+1)

...

La charge utile s+L (le dernier paquet de ce groupe d'entrelacement) : ILL=L, ILP=L,
blocs de trames : n+L, n+L+(L+1), n+L+2*(L+1), ..., n+L+(N-1)*(L+1)

Le prochain groupe d'entrelacement va commencer au bloc de trames n+N*(L+1).

Il ne va pas y avoir d'effet d'entrelacement si le nombre de blocs de trames par paquet (N) n'est pas d'au moins 2. De plus, le nombre de blocs de trames par charge utile (N) et la valeur de ILL NE DOIVENT PAS être changés à l'intérieur d'un groupe d'entrelacement. En d'autres termes, toutes les charges utiles dans un groupe d'entrelacement DOIVENT avoir le

même ILL et DOIVENT contenir le même nombre de blocs de trames de parole.

L'envoyeur de la charge utile DOIT appliquer l'entrelacement seulement si le receveur a signalé son utilisation par des moyens hors bande. Comme l'entrelacement va augmenter les exigences de mise en mémoire tampon chez le receveur, le receveur utilise le paramètre de type de support "entrelacement=I" pour régler le nombre maximum de blocs de trames permis dans un groupe d'entrelacement à I.

Quand il effectue l'entrelacement, l'envoyeur DOIT utiliser un nombre de blocs de trames approprié par charge utile (N) et ILL afin que la taille résultante d'un groupe d'entrelacement soit inférieure ou égale à I, c'est-à-dire, $N*(L+1) \leq I$.

4.4.2 Charge utile Table des matières et CRC de trames

La table des matières (ToC) en mode aligné sur l'octet consiste en une liste d'entrées de ToC où chaque entrée correspond à une trame de parole portée dans la charge utile et, facultativement, une liste de CRC de trame de parole. C'est-à-dire que la ToC est comme suit :

```
+-----+
| liste des entrées de ToC |
+-----+
| liste des CRC de trame   | (facultatif)
- - - - -
```

Note : pour les entrées de ToC avec FT=14 ou 15, il ne va pas y avoir de trame de parole ou trame de CRC correspondante présente dans la charge utile.

La liste des entrées de ToC est organisée de la même façon que décrit pour le mode efficace en bande passante du paragraphe 4.3.2, avec l'exception suivante : quand l'entrelacement est utilisé, les blocs de trames dans la ToC ne vont presque jamais être placés dans un ordre consécutif dans le temps. La présence et l'ordre des blocs de trames dans un paquet va plutôt suivre le schéma décrit au paragraphe 4.4.1.

Voici un exemple qui montre la ToC de trois paquets consécutifs, chacun portant trois blocs de trames, dans une session à deux canaux entrelacés. Ici, les deux canaux sont gauche (L) et droit (R) avec L qui vient avant R, et la longueur d'entrelacement est 3 (c'est-à-dire, ILL=2). Il en résulte une taille de groupe d'entrelacement de 9 blocs de trames.

Paquet n° 1

ILL=2, ILP=0 :

```
+-----+-----+-----+-----+-----+
| 1L | 1R | 4L | 4R | 7L | 7R |
+-----+-----+-----+-----+
|<----->|<----->|<----->|
  Bloc de   Bloc de   Bloc de
  trames 1   trames 4   trames 7
```

Paquet n° 2

ILL=2, ILP=1 :

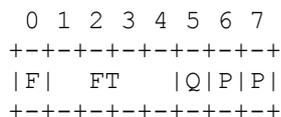
```
+-----+-----+-----+-----+-----+
| 2L | 2R | 5L | 5R | 8L | 8R |
+-----+-----+-----+-----+
|<----->|<----->|<----->|
  Bloc de   Bloc de   Bloc de
  trames 2   trames 5   trames 8
```

Paquet n° 3

ILL=2, ILP=2 :

```
+-----+-----+-----+-----+-----+
| 3L | 3R | 6L | 6R | 9L | 9R |
+-----+-----+-----+-----+
|<----->|<----->|<----->|
  Bloc de   Bloc de   Bloc de
  trames 3   trames 6   trames 9
```

Une entrée de ToC prend le format suivant en mode aligné sur l'octet :



F (1 bit) : voir la définition au paragraphe 4.3.2.

FT (4 bits, entier non signé) : voir la définition au paragraphe 4.3.2.

Q (1 bit) : voir la définition au paragraphe 4.3.2.

P : bits de bourrage, DOIVENT être réglés à zéro, et ignorés à réception.

La liste des CRC est FACULTATIVE. Elle n'existe que si l'utilisation de CRC est signalée hors bande pour la session. Quand elle est présente, chaque CRC de la liste est long de 8 bits et correspond à une trame de parole (PAS à un bloc de trames) porté dans la charge utile. Le calcul et l'utilisation du CRC sont spécifiés au paragraphe qui suit.

4.4.2.1 Utilisation de CRC de trame pour UED sur IP

Le concept général de UED/UEP sur IP est discuté au paragraphe 3.6. Ce paragraphe donne plus de détails sur la façon d'utiliser le CRC de trame dans l'en-tête de charge utile aligné sur l'octet avec une somme de contrôle partielle de couche de transport pour réaliser l'UED.

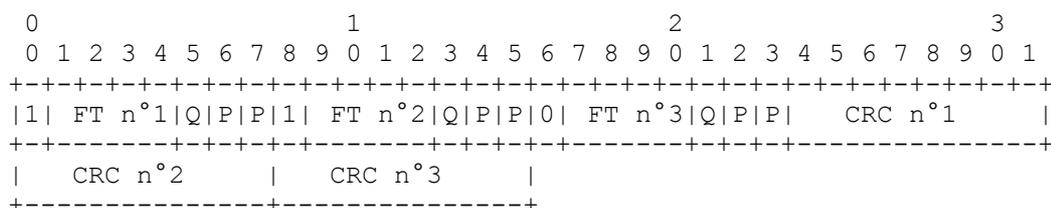
Pour réaliser l'UED, on DEVRAIT utiliser une somme de contrôle de couche transport (par exemple, celle définie dans UDP léger [RFC3828]) pour protéger les en-têtes IP, de protocole de transport (par exemple, UDP léger) et RTP, ainsi que l'en-tête de charge utile et la table des matières dans la charge utile. Le CRC de trame, quand il est utilisé, DOIT être calculé seulement sur tous les bits de classe A dans la trame AMR ou AMR-WB. Les bits des classes B et C dans la trame AMR ou AMR-WB NE DOIVENT PAS être inclus dans le calcul de CRC et NE DEVRAIENT PAS être couverts par la somme de contrôle de transport.

Note : le nombre de bits de classe A pour les divers modes de codage dans le codec AMR est spécifié pour information dans [AMR-struc] et est donc copié dans le Tableau 1 du paragraphe 3.6 pour le rendre normatif pour ce format de charge utile. Le nombre de bits de classe A pour les divers modes de codage dans le codec AMR-WB est spécifié comme normatif dans le Tableau 2 de [AMR-WB-Struc], et la trame de SID (FT=9) a 40 bits de classe A. Ces définitions de bits de classe A DOIVENT être utilisées pour ce format de charge utile.

Si la somme de contrôle de couche de transport ou de couche de liaison détecte des erreurs dans la partie protégée (sensible) on suppose que le paquet entier va être éliminé comme défini par UDP léger [RFC3828].

Le receveur de la charge utile DEVRAIT examiner l'intégrité des données des bits de classe A reçus en recalculant le CRC sur les bits de classe A reçus et en comparant le résultat à la valeur trouvée dans l'en-tête de charge utile reçu. Si les deux valeurs ne correspondent pas, le receveur DEVRA considérer que les bits de classe A de la trame reçue comme endommagés et DOIT mettre à zéro le fanion Q de la trame. Cela va ensuite causer le marquage de la trame comme SPEECH_BAD, si le FT de la trame est 0 à 7 pour AMR ou 0 à 8 pour AMR-WB, ou SID_BAD si le FT de la trame est 8 pour AMR ou 9 pour AMR-WB, avant qu'elle soit passée au décodeur de parole. Voir les détails dans [AMR-WB-SCR] et [TS26.193].

L'exemple suivant montre une ToC alignés sur l'octet avec une liste de CRC pour une charge utile contenant 3 trames de parole provenant d'une session à un seul canal (en supposant qu'aucun des FT n'est égal à 14 ou 15) :



Chacun des CRC prend 8 bits

```

  0   1   2   3   4   5   6   7
+---+---+---+---+---+---+---+---+
| c0| c1| c2| c3| c4| c5| c6| c7|
+---+---+---+---+---+---+---+---+
(MSB)                               (LSB)

```

et est calculé par le polynome générateur cyclique : $C(x) = 1 + x^2 + x^3 + x^4 + x^8$

où $^$ est l'opérateur d'exponentiation.

En forme binaire, le polynome apparaît comme suit : 101110001 (MSB..LSB).

Le calcul réel du CRC est fait comme suit : d'abord, un registre de CRC de 8 bits est remis à zéro : 00000000. Pour chaque bit sur lequel le CRC devra être calculé, une opération OUX est faite entre le bit le plus à droite (de moindre poids) du registre de CRC et le bit. Le registre de CRC est alors décalé à droite d'un cran (la signification de chaque bit est réduite de un) imputant un "0" au bit le plus à gauche (de plus fort poids, MSB). Si le résultat de l'opération OUX mentionnée ci-dessus est un "1", alors "10111000" est OUXé au bit près dans le registre de CRC. Cette opération est répétée pour chaque bit que le CRC devrait couvrir. Dans ce cas, le premier bit va être d(0) pour la trame de parole que le CRC devrait couvrir. Quand le dernier bit (par exemple, d(54) pour AMR 5,9 conformément au Tableau 1 du paragraphe 3.6) a été utilisé dans ce calcul de CRC, le contenu du registre de CRC devrait simplement être copié dans le champ correspondant dans la liste des CRC.

Le calcul rapide du CRC sur une CPU généraliste est possible en utilisant un algorithme de tableau.

4.4.3 Données de parole

En mode aligné sur l'octet, les données de parole sont portées de manière similaire à celle du mode efficace en bande passante discuté au paragraphe 4.3.3, avec les exceptions suivantes :

- Le dernier octet de chaque trame de parole DOIT être bourré avec des bits à zéro à la fin si tous les bits de l'octet ne sont pas utilisés. Les bits de bourrage DOIVENT être ignorés à réception. En d'autres termes, chaque trame de parole DOIT être alignée sur l'octet.
- Quand plusieurs trames de parole sont présentes dans les données de parole (c'est-à-dire, une charge utile composée) les trames de parole sont arrangées en une pleine trame après l'autre comme d'habitude, ou avec les octets de toutes les trames entrelacés ensemble au niveau de l'octet, selon le type de paramètres de support négociés pour le type de charge utile. Comme les bits au sein de chaque trame sont ordonnés avec les bits les plus sensibles à l'erreur en premiers, l'entrelacement des octets collecte les bits sensibles de toutes les trames pour qu'ils soient le plus près du début du paquet. C'est ce qui est appelé "ordre de tri robuste" qui permet l'application des mécanismes d'UED (comme UDP léger [RFC3828]) ou d'UEP (comme ULP [RFC5109]) aux données de la charge utile. Les détails de l'assemblage de la charge utile sont donnés au paragraphe suivant.

L'utilisation de l'ordre de tri robuste pour un type de charge utile DOIT faire l'objet d'un accord par des moyens hors bande. La Section 8 spécifie un paramètre de type de support à cette fin.

Note : l'ordre de tri robuste DOIT seulement être effectué au niveau trame et est donc indépendant de l'entrelacement, qui est au niveau du bloc de trames, comme décrit au paragraphe 4.4.1. En d'autres termes, le tri robuste peut être appliqué au type de charge utile non entrelacé ou entrelacé.

4.4.4 Méthodes de formation de la charge utile

Il existe deux méthodes différentes de mise en paquet, à savoir l'ordre normal et l'ordre de tri robuste, pour former une charge utile en mode aligné sur l'octet. Dans les deux cas, l'en-tête de la charge utile et la table des matières sont empaquetés dans la charge utile de la même façon ; la différence est dans l'empaquetage des trames de parole.

La charge utile commence avec l'en-tête de un octet de la charge utile, ou de deux octets si l'entrelacement de trame est choisi. L'en-tête de charge utile est suivi par la table des matières qui consiste en une liste d'entrées de ToC de un octet. Si

les CRC de trame sont à inclure, ils suivent la table des matières avec un CRC de 8 bits remplissant chaque octet. Noter que si une certaine trame a une entrée de ToC avec FT=14 ou 15, il n'y aura pas de CRC présent.

Les données de parole suivent la table des matières, ou les CRC si il y en a. Pour la mise en paquets dans l'ordre normal, tous les octets comprenant une trame de parole sont ajoutés à la charge utile comme un tout. Les trames de parole sont empaquetées dans le même ordre que leur entrées de ToC correspondants sont rangées dans la liste de la ToC, à l'exception que si une certaine trame a une entrée de ToC avec FT=14 ou 15, il n'y aura pas d'octets de données présents pour cette trame.

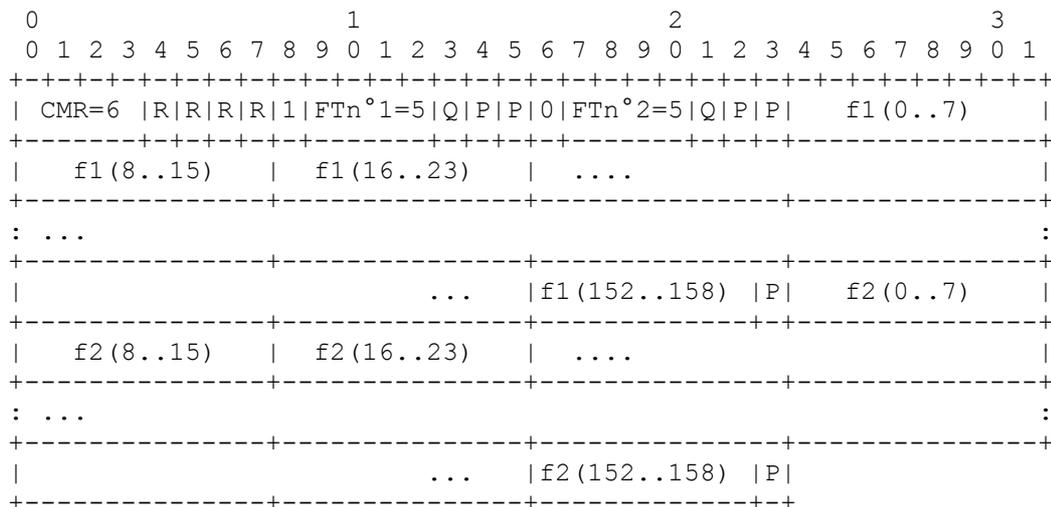
Pour la mise en paquet dans l'ordre de tri robuste, les octets de toutes les trames de parole sont entrelacés au niveau de l'octet. C'est-à-dire que la portion de données de la charge utile commence avec le premier octet de la première trame, suivi par le premier octet de la seconde trame, puis le premier octet de la troisième trame, et ainsi de suite. Après l'ajout du premier octet de la dernière trame, le cycle se répète avec le second octet de chaque trame. Le processus continue pour autant d'octets qui sont présents dans la plus longue trame. Si les trames n'ont pas toutes la même longueur d'octets, une trame plus courte est sautée une fois que tous ses octets ont été ajoutés. L'ordre des trames dans le cycle va être séquentiel si l'entrelacement de trame n'est pas utilisé, ou en accord avec le schéma d'entrelacement spécifié dans l'en-tête de la charge utile si l'entrelacement de trame est utilisé. Noter que si une certaine trame a une entrée de ToC avec FT=14 ou 15, aucun octet de données ne sera présent pour cette trame, de sorte qu'elle est sautée dans le cycle de tri robuste.

L'UED et/ou UEP est RECOMMANDÉE pour couvrir au moins l'en-tête RTP, l'en-tête de charge utile, la table des matières, et les bits de classe A d'une charge utile triée. Combien exactement d'octets ont besoin d'être couverts dépend du réseau et de l'application. Si les CRC sont utilisés avec le tri robuste, seuls l'en-tête RTP, l'en-tête de charge utile, et la ToC DEVRAIENT être couverts par UED/UEP. Le moyen de communiquer le nombre d'octets à couvrir aux autres couches qui effectuent UED/UEP sort du domaine d'application de la présente spécification.

4.4.5 Exemples de charge utile

4.4.5.1 Charge utile de base d'un seul canal portant plusieurs trames

Le diagramme qui suit montre une charge utile alignée sur l'octet provenant d'un type de charge utile à un seul canal qui porte deux trames AMR de mode de codage de 7,95 kbit/s (FT=5). Dans la charge utile, une demande de mode codec est envoyée (CMR=6) demandant au codeur côté receveur d'utiliser le mode de codage AMR à 10,2 kbit/s. Aucun CRC de trame, entrelacement, ou tri robuste n'est utilisé.



Note : dans cet exemple, le dernier octet des deux trames de parole est bourré avec un bit à zéro pour l'alignement sur l'octet.

4.4.5.2 Charge utile de deux canaux avec CRC, entrelacement, et tri robuste

Cet exemple montre une charge utile alignée sur l'octet sur un type de charge utile à deux canaux. Deux blocs de trames, chacun contenant deux trames de parole en mode de codage à 7,95 kbit/s (FT=5) sont portés dans cette charge utile.

Les deux canaux sont gauche (L) et droit (R) avec L venant avant R. Dans la charge utile, une demande de mode codec est aussi envoyée (CMR=6) demandant que le codeur côté receveur utilise le mode de codage AMR à 10,2 kbit/s.

De plus, le CRC de trame, le tri robuste, et l'entrelacement de bloc de trames sont tous activés pour le type de charge utile. La longueur d'entrelacement est 2 (ILL=1) et cette charge utile est la première dans un groupe d'entrelacement (ILP=0).

Les deux premières trames dans la charge utile sont les trames de parole des canaux L et R du bloc de trames n° 1, consistant en les bits f1L(0..158) et f1R(0..158), respectivement. Les deux trames suivantes sont les trames de parole des canaux L et R du bloc de trames n° 3, consistant en les bits f3L(0..158) et f3R(0..158), respectivement, du fait de l'entrelacement. Pour chacune des quatre trames de parole, un CRC est calculé comme CRC1L(0..7), CRC1R(0..7), CRC3L(0..7), et CRC3R(0..7), respectivement. Finalement, un tri robuste est fait sur la charge utile.

```

0                               1                               2                               3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
| CMR=6 |R|R|R|R| ILL=1 | ILP=0 |1|FT#1L=5|Q|P|P|1|FT#1R=5|Q|P|P|
+-----+-----+-----+-----+-----+-----+-----+-----+
|1|FT#3L=5|Q|P|P|0|FT#3R=5|Q|P|P|          CRC1L          |          CRC1R          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|          CRC3L          |          CRC3R          | f1L(0..7) | f1R(0..7) |
+-----+-----+-----+-----+-----+-----+-----+-----+
| f3L(0..7) | f3R(0..7) | f1L(8..15) | f1R(8..15) |
+-----+-----+-----+-----+-----+-----+-----+-----+
| f3L(8..15) | f3R(8..15) | f1L(16..23) | f1R(16..23) |
+-----+-----+-----+-----+-----+-----+-----+-----+
: ... :
+-----+-----+-----+-----+-----+-----+-----+-----+
| f3L(144..151) | f3R(144..151) |f1L(152..158)|P|f1R(152..158)|P|
+-----+-----+-----+-----+-----+-----+-----+-----+
|f3L(152..158)|P|f3R(152..158)|P|
+-----+-----+-----+-----+-----+-----+-----+

```

Note : dans cet exemple, le dernier octet des quatre trames de parole est bourré avec un bit à zéro pour l'alignement sur l'octet.

4.5. Considérations de mise en œuvre

Une application qui met en œuvre ce format de charge utile DOIT comprendre tous les paramètres de charge utile dans la signalisation hors bande utilisée. Par exemple, si une application utilise SDP, tous les paramètres SDP et de type de support du présent document DOIVENT être compris. Cette exigence assure qu'une mise en œuvre peut toujours décider si elle est capable ou non de communiquer.

Aucun mode de fonctionnement du format de charge utile n'est de mise en œuvre obligatoire. Les exigences de l'application qui utilise le format de charge utile devraient être utilisées pour déterminer ce qui est mis en œuvre. Pour réaliser l'interopérabilité de base, une mise en œuvre DEVRAIT au moins appliquer les deux modes d'efficacité de la bande passante et d'alignement sur l'octet pour un seul canal audio. Les autres modes de fonctionnement : entrelacement, tri robuste, et CRC de trame (en canal simple et multiple) sont de mise en œuvre FACULTATIVE.

Les paramètres mode-change-period, mode-change-capability, et mode-change-neighbor sont destinés à la signalisation avec les points d'extrémité GSM. Quand l'interopérabilité avec le GSM est désirée, les codeurs DEVRAIENT seulement effectuer des changements de mode de codec vers des modes voisins et en multiples entiers de 40 ms (deux blocs de trames) mais les décodeurs DEVRAIENT accepter les changements de mode de codec à tout moment, c'est-à-dire, pour tout bloc de trames. Le codeur peut choisir arbitrairement la phase initiale (bloc de trames pair ou impair) où les changements de mode de codec sont effectués, mais DEVRAIT alors s'en tenir à cette phase autant que possible. Cependant, dans des cas rares, les transferts inter-cellulaires ou autres événements (par exemple, renvoi d'appel) peuvent changer cette phase et peuvent aussi causer des changements de mode pour des modes non voisins. Le décodeur DEVRA donc être prêt à accepter des changements aussi dans l'autre phase et vers d'autres modes. La Section 8 spécifie l'usage des paramètres mode-change-period et mode-change-capability pour indiquer le comportement désiré des applications.

Voir dans [TS26.103] les configurations préférées de AMR et AMR-WB pour le fonctionnement dans les réseaux GSM et

3GPP UMTS. Dans les scénarios de passerelle, il peut être demandé aux codeurs par le paramètre "mode-set" d'utiliser un ensemble de modes limité qui est pris en charge par la liaison au delà de la passerelle. De plus, pour éviter l'encombrement sur cette liaison, le codeur DEVRAIT limiter le mode de codec initial pour une session à un mode inférieur, jusqu'à ce qu'au moins un bloc de trames soit reçu avec des informations de contrôle de débit.

4.5.1 Validation du décodage

Lors du traitement d'un paquet de charge utile reçu, si le receveur trouve que la longueur de charge utile calculée, sur la base des informations sur le type de la charge utile et les valeurs trouvées dans les champs d'en-tête de la charge utile, ne correspondent pas à la taille du paquet reçu, le receveur DEVRAIT éliminer le paquet. C'est parce que décoder un paquet qui a des erreurs dans son champ Longueur pourrait sévèrement dégrader la qualité de la parole.

5. Format de mémorisation AMR et AMR-WB

Le format de mémorisation est utilisé pour mémoriser les trames de parole AMR ou AMR-WB dans un fichier ou comme une pièce jointe d'un message électronique. Le contenu de plusieurs canaux est accepté.

En général, un fichier AMR ou AMR-WB a la structure suivante :

```
+-----+
| En-tête           |
+-----+
| trame de parole 1 |
+-----+
: ...              :
+-----+
| trame de parole n |
+-----+
```

Note : pour préserver l'interopérabilité avec les mises en œuvre déjà déployées, le contenu d'un seul canal utilise un format d'en-tête de fichier différent de celui d'un contenu multi canaux.

Il existe aussi un autre format de mémorisation pour AMR et AMR-WB qui convient pour les applications avec des exigences plus fortes sur le format de mémorisation, comme l'accès aléatoire ou la synchronisation avec la vidéo. Ce format est spécifié par le 3GPP comme format de fichier multimédia 3GP fondé sur la norme ISO [TS26.244]. Son type de support est spécifié par la [RFC3839].

5.1 En-tête d'un seul canal

Un en-tête de fichier AMR ou AMR-WB d'un seul canal contient seulement un numéro magique. Différents numéros magiques sont définis pour distinguer AMR de AMR-WB.

Le numéro magique pour les fichiers AMR d'un seul canal DOIT consister en la chaîne de caractères ASCII :

"#!AMR\n" (ou 0x2321414d520a en hexadécimal).

Le numéro magique pour les fichiers AMR-WB d'un seul canal DOIT consister en la chaîne de caractères ASCII :

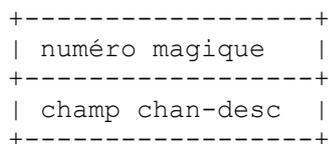
"#!AMR-WB\n" (ou 0x2321414d522d57420a en hexadécimal).

Note : le "\n" est une partie importante des numéros magiques et DOIT être inclus dans la comparaison, car autrement, les numéros magiques à un seul canal ci-dessus vont devenir indistinguables de ceux des fichiers multi canaux définis au paragraphe suivant.

5.2 En-tête multi canaux

L'en-tête multi canaux consiste en un numéro magique suivi par un champ Description de canal de 32 bits, donnant à l'en-

tête multi canaux la structure suivante :

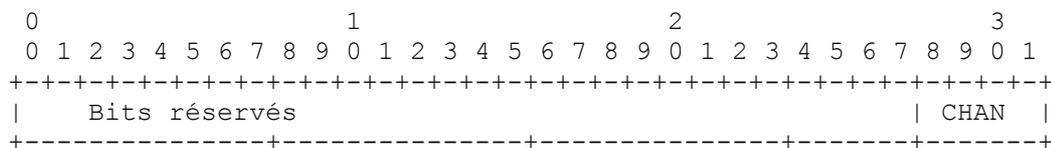


Le numéro magique pour les fichiers AMR multi canaux DOIT consister en la chaîne de caractères ASCII :
 "#!AMR_MC1.0\n" (ou 0x2321414d525F4D43312E300a en hexadécimal).

Le numéro magique pour les fichiers AMR-WB multi canaux DOIT consister en la chaîne de caractères ASCII :
 "#!AMR-WB_MC1.0\n" (ou 0x2321414d522d57425F4D43312E300a en hexadécimal).

Le numéro de version dans les numéros magiques se réfère à la version du format de fichier.

Le champ de 32 bits de description de canal est défini comme :



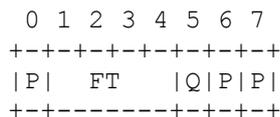
Bits réservés : DOIVENT être réglés à 0 à l'émission, et DOIVENT être ignorés à réception.

CHAN (4 bits, entier non signé) : indique le nombre de canaux audio contenus dans ce fichier de mémorisation. Les valeurs valides et l'ordre des canaux au sein d'un bloc de trames sont spécifiés au paragraphe 4.1 de la [RFC3551].

5.3 Trames de parole

Après l'en-tête de fichier, les blocs de trames de parole consécutifs dans le temps sont mémorisés dans le fichier. Chaque bloc de trames contient un nombre de trames alignées sur l'octet égal au nombre de canaux, et mémorisées en ordre croissant, en commençant par le canal 1.

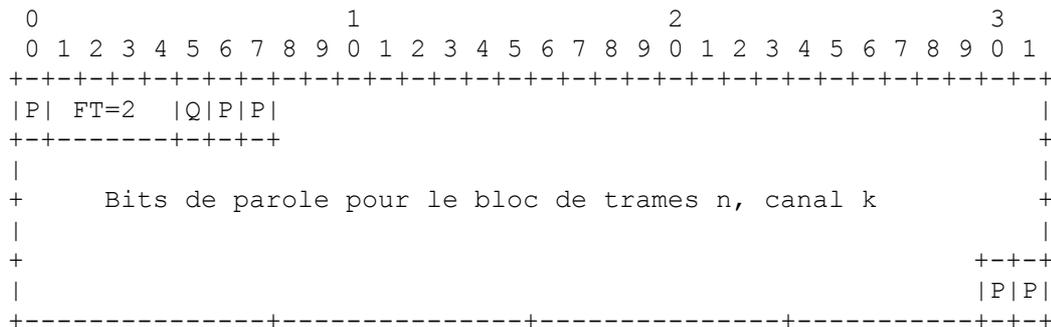
Chaque trame de parole mémorisée commence par un en-tête de trame d'un octet du format suivant :



Le champ FT et le bit Q bit sont définis de la même façon qu'au paragraphe 4.3.2. Les bits P sont du bourrage et DOIVENT être réglés à 0, et DOIVENT être ignorés à réception.

À la suite de cet en-tête d'un octet viennent les bits de parole comme définis au paragraphe 4.4.3. Le dernier octet de chaque trame est bourré avec des zéros, si nécessaire, pour réaliser l'alignement sur l'octet.

L'exemple suivant montre une trame AMR en mode de codage 5,9 kbit/s (avec 118 bits de parole) dans le format de mémorisation.



Les trames de parole non reçues ou les blocs de trames entre des mises à jour de SID durant des périodes sans parole DOIVENT être mémorisés comme des trames NO_DATA (type de trame 15, comme défini dans [AMR-struct] et [AMR-WB-Struct]). Les trames ou blocs de trames perdus dans la transmission DOIVENT être mémorisés comme des trames NO_DATA ou SPEECH_LOST (type de trame 14, seulement disponible pour AMR-WB) en blocs de trames complets pour conserver la synchronisation avec le support original.

Les trames de bruit de confort d'autres types que SID AMR (FT=8) (c'est-à-dire, les types de trame 9, 10, et 11 pour AMR) NE DEVRONT PAS être utilisées dans le format de fichier AMR.

6. Contrôle d'encombrement

Les considérations générales de contrôle d'encombrement pour le transport des données RTP s'appliquent aussi à la parole AMR ou AMR-WB sur RTP. Cependant, la capacité multi débits du codage de parole AMR et AMR-WB peut donner un avantage sur les autres formats de charge utile pour contrôler l'encombrement parce que la demande de bande passante peut être ajustée en choisissant un mode de codage différent.

Un autre paramètre qui peut impacter la demande de bande passante pour AMR et AMR-WB est le nombre de blocs de trames qui sont encapsulés dans chaque charge utile RTP. Empaqueter plus de blocs de trames dans chaque charge utile RTP peut réduire le nombre de paquets envoyés et donc les frais généraux des en-têtes IP/UDP/RTP, au prix d'un délai accru.

Si la correction d'erreur directe (FEC, *Forward Error Correction*) est utilisée pour combattre la perte de paquets, la quantité de redondance ajoutée par la FEC devra être régulée afin que l'utilisation de la FEC elle-même ne cause pas un problème d'encombrement.

Il est RECOMMANDÉ que les applications AMR ou AMR-WB qui utilisent ce format de charge utile emploient le contrôle d'encombrement. Le mécanisme réel de contrôle d'encombrement n'est pas spécifié mais devrait être convenable pour les flux en temps réel, éventuellement le "contrôle de débit convivial sur TCP" [RFC3448].

7. Considérations sur la sécurité

Les paquets RTP qui utilisent le format de charge utile défini dans la présente spécification sont soumis aux considérations générales de sécurité discutées dans la [RFC3550] et dans tout profil utilisé, comme AVP [RFC3551] ou SAVP [RFC3711].

Comme ce format transporte de la parole codée, les principaux problèmes de sécurité incluent la confidentialité, l'authentification, et l'intégrité de la parole elle-même. Le format de charge utile n'a pas par lui-même de mécanisme de sécurité incorporé. Des mécanismes externes, comme SRTP [RFC3711], doivent être utilisés pour cette fonctionnalité. Noter que le mécanisme approprié pour assurer la sécurité à RTP et aux charges utiles selon le présent mémoire peut varier. Il dépend de l'application, du transport, et du protocole de signalisation employé. Donc, un seul mécanisme n'est pas suffisant, bien que si il convient, l'usage de SRTP [RFC3711] soit RECOMMANDÉ. Les autres mécanismes connus qui peuvent être utilisés sont IPsec [RFC4301] et TLS [RFC4346] (RTP sur TCP) mais d'autres solutions peuvent aussi exister.

Ce format de charge utile ne présente aucune non uniformité significative de la complexité de calcul du côté du receveur pour le traitement des paquets, et il est peu susceptible de se prêter à une menace de déni de service due à la réception de données pathologiques.

7.1 Confidentialité

Pour réaliser la confidentialité de la parole codée AMR ou AMR-WB, tous les bits de données de parole vont devoir être chiffrés. Il est moins besoin de chiffrer l'en-tête de charge utile ou la table des matières du fait que a) ils portent seulement des informations sur le mode de parole demandé, le type de trame, et la qualité de la trame, et b) ces informations pourraient être utiles à un tiers, par exemple, pour la surveillance de la qualité.

La mise en paquets et la dépaquetisation de la charge utile AMR et AMR-WB est seulement faite aux points d'extrémité. Donc le chiffrement devrait être effectué après l'encapsulation du paquet, et le déchiffrement devrait être effectué avant la désencapsulation de paquet.

Le chiffrement peut affecter l'entrelacement. Spécifiquement, un changement de clés devrait se produire à la frontière entre les groupes d'entrelacement. Si il n'est pas fait à cette frontière sur les deux points d'extrémité, la qualité de la parole va être dégradée durant tout les groupes d'entrelacement pour tout receveur.

Le mécanisme de chiffrement peut impacter la robustesse du mécanisme de correction d'erreur. Ceci est discuté au paragraphe 9.5 de SRTP [RFC3711]. À partir de cela, l'UED/UEP fondée sur le tri robuste peut être difficile à appliquer quand les données de la charge utile sont chiffrées.

7.2 Authentification et intégrité

Pour authentifier l'envoyeur et protéger l'intégrité des paquets RTP en transit, un mécanisme externe doit être utilisé. Comme déclaré précédemment, il est RECOMMANDÉ d'utiliser SRTP [RFC3711] pour l'interopérabilité courante. Noter que l'utilisation de l'UED/UEP peut être difficile à combiner avec des mécanismes de protection de l'intégrité parce que toute erreur binaire va causer l'échec de la vérification d'intégrité.

L'altération des données par un attaquant interposé pourrait résulter en erreurs de dépaquetisation/décodage qui pourraient diminuer la qualité de la parole ou produire des communications inintelligibles. L'altération du champ CMR peut résulter en une qualité de la parole différente de celle désirée.

8. Paramètres de format de charge utile

Cette section définit les paramètres qui peuvent être utilisés pour choisir les caractéristiques facultatives des formats de charge utile AMR et AMR-WB. Les paramètres sont définis ici au titre des enregistrements de type de support pour les codecs de parole AMR et AMR-WB. Les enregistrements sont faits selon la [RFC4855] et les règles d'enregistrement de supports de la [RFC4288].

Une transposition des paramètres dans le protocole de description de session (SDP, *Session Description Protocol*) [RFC4566] est aussi fournie pour les applications qui utilisent SDP. Des paramètres équivalents pourraient être définis ailleurs pour être utilisés avec les protocoles de contrôle qui n'utilisent pas ces types de support ou SDP.

On fait deux enregistrements de type de support séparés, un pour AMR et un pour AMR-WB, parce que ce sont des codages distincts qui doivent être distingués par leur propre type de support.

Les formats de données sont spécifiés pour le transport en temps réel dans RTP et pour les applications de type mémorisation comme les pièces jointes de messagerie.

8.1 Enregistrement de type de support AMR

Le type de support pour le codec adaptatif multi débits (AMR, *Adaptive Multi-Rate*) est alloué à partir de l'arborescence de l'IETF parce que AMR est un codec de parole largement utilisé dans les applications VoIP générales et de messagerie. Cet enregistrement de type de support couvre le transfert en temps réel via RTP et les transferts non en temps réel via des fichiers mémorisés.

Noter que tout paramètre non spécifié DOIT être ignoré par le receveur.

Nom de type de support : audio

Nom de sous type de support : AMR

Paramètres exigés : aucun

Paramètres facultatifs : Ces paramètres s'appliquent seulement au transfert RTP.

octet-align : les valeurs permises sont 0 et 1. Si c'est 1, le fonctionnement aligné sur l'octet DEVRA être utilisé. Si c'est 0 ou si le paramètre n'est pas présent, le fonctionnement efficace en bande passante est employé.

mode-set : restreint l'ensemble actif de modes de codec à un sous ensemble de tous les modes, par exemple, à être capable de prendre en charge les canaux de transport comme les réseaux GSM dans les cas d'utilisation de passerelle. Des valeurs possibles sont une liste séparée de virgules des modes de l'ensemble 0 à 7 (voir le Tableau 1a de [AMR-struct]). Les types de trame SID 8 et NO_DATA (type de trame 15) ne sont jamais inclus dans l'ensemble de modes, mais

peuvent toujours être utilisés. Si mode-set est spécifié, il DOIT être respecté, et les trames codées avec des modes en dehors du sous ensemble NE DOIVENT PAS être envoyées dans une charge utile RTP ou utilisées dans des demandes de mode codec. Si il n'est pas présent, tous les modes de codec sont permis pour le type de charge utile.

mode-change-period : spécifie un nombre de blocs de trames, N (1 ou 2) qui est la période de bloc de trames à laquelle des changements de mode de codec sont permis pour l'expéditeur. La phase initiale de l'intervalle est arbitraire, mais les changements doivent être séparés par des multiples de N blocs de trames, c'est-à-dire, une valeur de 2 permet à l'expéditeur de changer de mode tous les seconds blocs de trames. La valeur de N DEVRA être 1 ou 2. Si ce paramètre n'est pas présent, les changements de mode sont permis à tout moment durant la session, c'est-à-dire que N=1.

mode-change-capability : spécifie si le client est capable de transmettre avec une période de changement de mode restreinte. Le paramètre peut prendre la valeur de 1 ou 2. Une valeur de 1 indique que le client n'est pas capable de restreindre la période de changement de mode à 2, et que le mode de codec peut être changé à tout moment. Une valeur de 2 indique que le client a la capacité de restreindre la période de changement de mode à 2, et donc que le client peut correctement interopérer avec un receveur qui demande un mode-change-period=2. Si ce paramètre n'est pas présent, la capacité de restriction de changement de mode n'est pas supportée, c'est-à-dire mode-change-capability=1. Pour être capable d'interopérer pleinement avec les passerelles avec des réseaux à commutation de circuits (par exemple, les réseaux GSM) les transmissions avec des changements de mode restreints (mode-change-capability=2) sont exigés. Donc, il est RECOMMANDÉ que les clients aient la capacité de prendre en charge la transmission selon mode-change-capability=2.

mode-change-neighbor : les valeurs permises sont 0 et 1. Si 1, l'expéditeur DEVRAIT seulement effectuer des changements de mode pour les modes voisins dans l'ensemble de modes de codec actifs. Les modes voisins sont ceux qui sont les plus proches en débit du mode courant, soit au dessus, soit en dessous. Si il est réglé à 0 ou si il n'est pas présent, le changement entre tous modes de l'ensemble actif de mode de codec est permis.

maxptime : quantité maximum du support qui peut être encapsulée dans un paquet de charge utile, exprimée comme un temps en millisecondes. Le temps est calculé comme somme du temps que le support présent dans le paquet représente. Le temps DEVRAIT être un multiple entier de la taille de trame. Si ce paramètre n'est pas présent, l'expéditeur PEUT encapsuler tout nombre de trames de parole dans un paquet RTP.

crc : les valeurs permises sont 0 et 1. Si 1, les CRC de trame DEVRONT être inclus dans la charge utile. Si 0 ou non présent, les CRC NE DEVRONT PAS être utilisés. Si crc=1, cela implique aussi automatiquement que le fonctionnement aligné sur l'octet DEVRA être utilisé pour la session.

robust-sorting : les valeurs permises sont 0 et 1. Si 1, la charge utile DEVRA employer le tri robuste de charge utile. Si 0 ou si non présent, le simple tri de charge utile DEVRA être utilisé. Si robust-sorting=1, cela implique aussi automatiquement que le fonctionnement aligné sur l'octet DEVRA être utilisé pour la session.

Interleaving : indique que l'entrelacement de niveau bloc de trames DEVRA être utilisé pour la session, et sa valeur définit le nombre maximum de blocs de trames permis dans un groupe d'entrelacement (paragraphe 4.4.1). Si ce paramètre n'est pas présent, l'entrelacement NE DEVRA PAS être utilisé. La présence de ce paramètre implique aussi automatiquement que le fonctionnement aligné sur l'octet DEVRA être utilisé.

ptime : voir la [RFC4566].

canaux : nombre de canaux audio. Les valeurs possibles (1 à 6) et leur ordre respectif de canal sont spécifiés au paragraphe 4.1 de la [RFC3551]. Si il est omis, il a la valeur par défaut de 1.

max-red : durée maximum en millisecondes qui s'écoule entre la principale (première) transmission d'une trame et toute transmission redondante que l'expéditeur va utiliser. Ce paramètre permet à un receveur d'avoir une limite de délai quand la redondance est utilisée. Les valeurs permises sont entre 0 (pas de redondance à utiliser) et 65535. Si le paramètre est omis, aucune limitation à l'utilisation de la redondance n'est présente.

Considérations de codage : Les données audio sont des données binaires, et doivent être codées pour un transport non binaire ; le codage Base64 convient pour la messagerie électronique. Quand elles sont utilisées dans le contexte de RTP, les données sont tramées comme défini dans la [RFC4288].

Considérations de sécurité : voir la Section 7 de la RFC 4867.

Spécifications publiées : RFC 4867, 3GPP TS 26.090, 26.092, 26.093, 26.101

Applications qui utilisent ce type de support : ce type de support est utilisé dans de nombreuses applications qui ont besoin de transporter ou mémoriser de la voix codée. Des exemples incluent ; la voix sur IP, les supports de flux directs, la messagerie vocale, et l'enregistrement de la voix sur les caméras numériques.

Informations supplémentaires : ce qui suit s'applique aux méthodes de transfert de fichier mémorisé :

Numéros magiques :

un seul canal : chaîne de caractères ASCII "#!AMR\n" (ou 0x2321414d520a en hexadécimal)

multi canaux : chaîne de caractères ASCII "#!AMR_MC1.0\n" (ou 0x2321414d525f4d43312e300a en hexadécimal)

Extensions de fichier : amr, AMR

Code de type de fichier Macintosh : "amr " (le quatrième caractère est espace)

Les trames de parole AMR peuvent aussi être mémorisées dans le format de fichier "3GP" défini dans [TS26.244], qui est identifié en utilisant les types de support "audio/3GPP" ou "video/3GPP" comme enregistré par la [RFC3839].

Adresse de la personne de messagerie à contacter pour plus d'informations :

Magnus Westerlund <magnus.westerlund@ericsson.com>

Ari Lakaniemi <ari.lakaniemi@nokia.com>

Utilisation prévue : COMMUNE. Ce type de support est largement utilisé dans les applications de flux directs, de VoIP, et de messagerie sur de nombreux types d'appareils.

Restrictions d'usage : quand ce type de support est utilisé dans le contexte d'un transfert sur RTP, le format de charge utile RTP spécifié à la Section 4 DEVRA être utilisé. Dans tous les autres contextes, le format de fichier défini à la Section 5 DEVRA être utilisé.

Auteurs :

Magnus Westerlund <magnus.westerlund@ericsson.com>

Ari Lakaniemi <ari.lakaniemi@nokia.com>

Contrôleur des changements : groupe de travail Transport Audio/Video de l'IETF sur délégation de l'IESG.

8.2. Enregistrement de type de support AMR-WB

Le type de support pour le codec adaptatif multi débits large bande (AMR-WB, *Adaptive Multi-Rate Wideband*) est alloué à partir de l'arborescence de l'IETF parce que AMR-WB est un codec de parole largement utilisé dans les applications VoIP générales et de messagerie. Cet enregistrement de type de support couvre le transfert en temps réel via RTP et les transferts non en temps réel via des fichiers mémorisés.

Noter que tout paramètre non spécifié DOIT être ignoré par le receveur.

Nom de type de support : audio

Nom de sous type de support : AMR-WB

Paramètres exigés : aucun

Paramètres facultatifs : Ces paramètres s'appliquent seulement au transfert RTP.

octet-align : les valeurs permises sont 0 et 1. Si c'est 1, le fonctionnement aligné sur l'octet DEVRA être utilisé. Si c'est 0 ou si le paramètre n'est pas présent, le fonctionnement efficace en bande passante est employé.

mode-set : restreint l'ensemble actif de mode de codec à un sous ensemble de tous les modes, par exemple, à être capable de prendre en charge les canaux de transport comme les réseaux GSM dans les cas d'utilisation de passerelle. Des valeurs possibles sont une liste séparée de virgules des modes de l'ensemble 0 à 8 (voir le Tableau 1a de [AMR-WB-struct]). Le type de trame SID 9, SPEECH_LOST (type de trame 14) et NO_DATA (type de trame 15) ne sont jamais inclus dans l'ensemble de modes, mais peuvent toujours être utilisés. Si mode-set est spécifié, il DOIT être respecté, et les trames codées avec des modes en dehors du sous ensemble NE DOIVENT PAS être envoyées dans une charge utile RTP ou utilisées dans des demandes de mode codec. Si il n'est pas présent, tous les modes de codec sont permis pour le type de charge utile.

mode-change-period : spécifie un nombre de blocs de trames, N (1 ou 2) qui est la période de bloc de trames à laquelle des

changements de mode de codec sont permis pour l'expéditeur. La phase initiale de l'intervalle est arbitraire, mais les changements doivent être séparés par des multiples de N blocs de trames, c'est-à-dire, une valeur de 2 permet à l'expéditeur de changer de mode tous les seconds blocs de trames. La valeur de N DEVRA être 1 ou 2. Si ce paramètre n'est pas présent, les changements de mode sont permis à tout moment durant la session, c'est-à-dire que N=1.

mode-change-capability : spécifie si le client est capable de transmettre avec une période de changement de mode restreinte. Le paramètre peut prendre la valeur de 1 ou 2. Une valeur de 1 indique que le client n'est pas capable de restreindre la période de changement de mode à 2, et que le mode de codec peut être changé à tout moment. Une valeur de 2 indique que le client a la capacité de restreindre la période de changement de mode à 2, et donc que le client peut correctement interopérer avec un receveur qui demande un mode-change-period=2. Si ce paramètre n'est pas présent, la capacité de restriction de changement de mode n'est pas supportée, c'est-à-dire mode-change-capability=1. Pour être capable d'interopérer pleinement avec les passerelles avec des réseaux à commutation de circuits (par exemple, les réseaux GSM) les transmissions avec des changements de mode restreints (mode-change-capability=2) sont exigés. Donc, il est RECOMMANDÉ que les clients aient la capacité de prendre en charge la transmission selon mode-change-capability=2.

mode-change-neighbor : les valeurs permises sont 0 et 1. Si 1, l'expéditeur DEVRAIT seulement effectuer des changements de mode pour les modes voisins dans l'ensemble de modes de codec actifs. Les modes voisins sont ceux qui sont les plus proches en débit du mode courant, soit au dessus, soit en dessous. Si il est réglé à 0 ou si il n'est pas présent, le changement entre tous modes de l'ensemble actif de mode de codec est permis.

maxptime : quantité maximum du support qui peut être encapsulée dans un paquet de charge utile, exprimée comme un temps en millisecondes. Le temps est calculé comme somme du temps que le support présent dans le paquet représente. Le temps DEVRAIT être un multiple entier de la taille de trame. Si ce paramètre n'est pas présent, l'expéditeur PEUT encapsuler tout nombre de trames de parole dans un paquet RTP.

crc : les valeurs permises sont 0 et 1. Si 1, les CRC de trame DEVRONT être inclus dans la charge utile. Si 0 ou non présent, les CRC NE DEVRONT PAS être utilisés. Si crc=1, cela implique aussi automatiquement que le fonctionnement aligné sur l'octet DEVRA être utilisé pour la session.

robust-sorting : les valeurs permises sont 0 et 1. Si 1, la charge utile DEVRA employer le tri robuste de charge utile. Si 0 ou si non présent, le simple tri de charge utile DEVRA être utilisé. Si robust-sorting=1, cela implique aussi automatiquement que le fonctionnement aligné sur l'octet DEVRA être utilisé pour la session.

entrelacement : indique que l'entrelacement de niveau bloc de trames DEVRA être utilisé pour la session, et sa valeur définit le nombre maximum de blocs de trames permis dans un groupe d'entrelacement (paragraphe 4.4.1). Si ce paramètre n'est pas présent, l'entrelacement NE DEVRA PAS être utilisé. La présence de ce paramètre implique aussi automatiquement que le fonctionnement aligné sur l'octet DEVRA être utilisé.

ptime : voir la [RFC4566].

canaux : nombre de canaux audio. Les valeurs possibles (1 à 6) et leur ordre respectif de canal sont spécifiés au paragraphe 4.1 de la [RFC3551]. Si il est omis, il a la valeur par défaut de 1.

max-red : durée maximum en millisecondes qui s'écoule entre la principale (première) transmission d'une trame et toute transmission redondante que l'expéditeur va utiliser. Ce paramètre permet à un receveur d'avoir une limite de délai quand la redondance est utilisée. Les valeurs permises sont entre 0 (pas de redondance à utiliser) et 65535. Si le paramètre est omis, aucune limitation à l'utilisation de la redondance n'est présente.

Considérations de codage : Les données audio sont des données binaires, et doivent être codées pour un transport non binaire ; le codage Base64 convient pour la messagerie électronique. Quand elles sont utilisées dans le contexte de RTP, les données sont tramées comme défini dans la [RFC4288].

Considérations de sécurité : voir la Section 7 de la RFC 4867.

Spécifications publiées : RFC 4867, 3GPP TS 26.190, 26.192, 26.193, 26.201

Applications qui utilisent ce type de support : ce type de support est utilisé dans de nombreuses applications qui ont besoin de transporter ou mémoriser de la voix codée. Des exemples incluent ; la voix sur IP, les supports de flux directs, la messagerie vocale, et l'enregistrement de la voix sur les caméras numériques.

Informations supplémentaires : ce qui suit s'applique aux méthodes de transfert de fichier mémorisé :

Numéros magiques :

un seul canal : chaîne de caractères ASCII ""#!AMR-WB\n" (ou 0x2321414d522d57420a en hexadécimal)

multi canaux : chaîne de caractères ASCII ""#!AMR-WB_MC1.0\n" (ou 0x2321414d522d57425F4D43312E300a en hexadécimal)

Extensions de fichier : awb, AWB

Code de type de fichier Macintosh : "amrw"

Identifiant d'objet ou OID : aucun

Les trames de parole AMR peuvent aussi être mémorisées dans le format de fichier "3GP" défini dans [TS26.244], qui est identifié en utilisant les types de support "audio/3GPP" ou "video/3GPP" comme enregistré par la [RFC3839].

Adresse de la personne de messagerie à contacter pour plus d'informations :

Magnus Westerlund <magnus.westerlund@ericsson.com>

Ari Lakaniemi <ari.lakaniemi@nokia.com>

Utilisation prévue : COMMUNE. Ce type de support est largement utilisé dans les applications de flux directs, de VoIP, et de messagerie sur de nombreux types d'appareils.

Restrictions d'usage : quand ce type de support est utilisé dans le contexte d'un transfert sur RTP, le format de charge utile RTP spécifié à la Section 4 DEVRA être utilisé. Dans tous les autres contextes, le format de fichier défini à la Section 5 DEVRA être utilisé.

Auteurs :

Magnus Westerlund <magnus.westerlund@ericsson.com>

Ari Lakaniemi <ari.lakaniemi@nokia.com>

Contrôleur des changements : groupe de travail Transport Audio/Video de l'IETF sur délégation de l'IESG.

8.3 Transposition de paramètres de type de support en SDP

Les informations portées dans la spécification de type de support ont une transposition spécifique en les champs du protocole de description de session (SDP) [RFC4566] qui est couramment utilisé pour décrire les sessions RTP. Quand SDP est utilisé pour spécifier des sessions qui emploient le codec AMR ou AMR-WB, la transposition est comme suit :

- Le type de support ("audio") va dans la ligne SDP "m=" comme nom de support.
- Le sous type de support (nom de format de charge utile) va dans un SDP "a=rtptime" comme nom de codage. Le débit d'horloge RTP dans un "a=rtptime" DOIT être 8000 pour AMR et 16000 pour AMR-WB, et les paramètres de codage (nombre de canaux) DOIVENT soit être explicitement réglés à N, soit être omis, impliquant une valeur par défaut de 1. Les valeurs de N permises sont spécifiées au paragraphe 4.1 de la [RFC3551].
- Les paramètres "ptime" et "maxptime" vont respectivement dans les attributs SDP "a=ptime" et "a=maxptime".
- Tous les paramètres restants vont dans l'attribut SDP "a=fmtp" en les copiant directement de la chaîne de paramètres de type de support comme une liste séparée par des caractères deux-points de paires de paramètre=valeur.

8.3.1 Considérations sur le modèle d'offre-réponse

Les considérations suivantes s'appliquent pour l'utilisation des procédures d'offre-réponse SDP pour négocier l'utilisation de charges utiles AMR ou AMR-WB dans RTP:

- Chaque combinaison de paramètres de configuration de format de transport de charge utile RTP (octet-align, crc, robust-sorting, interleaving, et channels) est unique dans son schéma binaire et non compatible avec toute autre combinaison. Quand une offre est créée dans une application qui désire utiliser les caractéristiques les plus avancées (crc, tri robuste, entrelacement, ou plus d'un canal) il est RECOMMANDÉ que l'offreur offre aussi un type de charge utile contenant seulement la configuration alignée sur l'octet ou efficacité de la bande avec un seul canal. Si plusieurs configurations intéressent l'application, elles peuvent être toutes offertes ; cependant, on devrait faire attention à ne pas

offrir trop de types de charge utile. Un répondant SDP DOIT inclure, dans la réponse SDP pour un type de charge utile, les paramètres suivants non modifiés provenant de l'offre SDP (sauf si il supprime le type de charge utile) : "octet-align"; "crc"; "robust-sorting"; "interleaving"; et "channels". L'offreur et le répondant SDP DOIVENT générer des paquets AMR ou AMR-WB comme décrit par ces paramètres.

- Le paramètre "mode-set" peut être utilisé pour restreindre l'ensemble de modes AMR/AMR-WB actifs utilisés dans une session. Cette fonction est principalement destinée aux passerelles pour accéder à des réseaux tels que GSM ou 3GPP UMTS, où le réseau d'accès peut être capable de prendre seulement en charge un sous ensemble de modes AMR/AMR-WB. Les configurations de codec préférées du 3GPP sont définies dans [TS26.202], et il est RECOMMANDÉ que les autres réseaux qui ont aussi besoin de restreindre l'ensemble de modes suivent les configurations de codec préférées définies par le 3GPP pour la plus grande interopérabilité. Le paramètre est bidirectionnel, c'est-à-dire que l'ensemble restreint s'applique aux supports à recevoir et envoyés par l'entité déclarante. Si un ensemble de modes a été fourni dans l'offre, le répondant DEVRA retourner le mode-set non modifié ou rejeter le type de charge utile. Cependant, le répondant est libre de choisir un ensemble de modes dans la réponse seulement si aucun mode-set n'a été fourni dans l'offre pour une session en envoi individuel à deux homologues. Le mode-set dans la réponse lie les deux parties, offreur et répondant. Donc, un offreur qui prend en charge tous les modes et sous ensembles NE DEVRAIT PAS inclure le paramètre mode-set. Pour tout autre offreur, il est RECOMMANDÉ d'inclure chaque mode-set qu'il peut prendre en charge comme type séparé de charge utile dans l'offre. Pour les sessions de diffusion groupée, le répondant DEVRA seulement participer à la session si il prend en charge l'ensemble de modes offert. Donc, il est RECOMMANDÉ que toute offre pour une session de diffusion groupée inclut seulement le mode-set qu'il va exiger que les répondants prennent en charge, et que le mode-set soit vraisemblablement pris en charge par tous les participants.
- Les paramètres "mode-change-period" et "mode-change-capability" sont destinés à être utilisés dans des sessions avec des passerelles, par exemple, quand il y a interopération avec des réseaux GSM. Les deux paramètres sont déclaratifs et sont combinés pour permettre au participant à une session de déterminer si le type de charge utile peut être supporté. Le paramètre mode-change-period va indiquer ce que l'offreur ou répondant exige des données qu'il reçoit, tandis que le paramètre mode-change-capability indique ses capacités de transmission. Un mode-change-period=2 dans l'offre indique l'exigence que le répondant envoie avec une période de mode-change de 2, c'est-à-dire, prenne en charge mode-change-capability=2. Si le répondant exige mode-change-period=2, il DEVRA seulement l'inclure dans la réponse si l'offreur a indiqué la prise en charge de mode-change-capability=2 ou a indiqué mode-change-period=2 ; autrement, le type de charge utile DEVRA être rejeté. Un offreur qui prend en charge mode-change-capability=2 DEVRA inclure le paramètre dans toutes les offres pour assurer la plus grande interopérabilité possible, sauf si il inclut mode-change-period=2 dans l'offre. Le paramètre mode-change-capability DEVRAIT être inclus dans les réponses. Cela indique alors la capacité du répondant de transmettre avec cette mode-change-period pour la configuration de format de charge utile fournie. L'information est utile pour de futures renégociations des formats de charge utile.
- Le paramètre "mode-change-neighbor" est une recommandation de restreindre le changement de modes de codec à son mode voisin et il DEVRAIT être suivi. Il est destiné à être utilisé dans les scénarios de passerelle (par exemple, avec un réseau GSM) où la prise en charge de ce paramètre et des opérations qu'elle implique améliore l'interopérabilité. Le paramètre "mode-change-neighbor" est déclaratif. En incluant ce paramètre, l'offreur ou le répondant indique qu'il désire recevoir des flux avec des restrictions de "mode-change-neighbor".
- Dans la plupart des cas, les paramètres "maxptime" et "ptime" ne vont pas affecter l'interopérabilité ; cependant, l'établissement de ces paramètres peut affecter les performances de l'application. Le traitement de l'offre-réponse SDP du paramètre "ptime" est décrit dans la [RFC3264]. Le paramètre "maxptime" DOIT être traité de la même façon.
- Le paramètre "max-red" est un paramètre de propriété de flux. Pour les flux de supports d'envoi individuel en envoi seul ou en envoi-réception, le paramètre déclare la limitation de redondance que l'envoyeur du flux va utiliser. Pour les flux en réception seule, il indique la valeur désirée pour le flux envoyé au receveur. Le répondant PEUT changer la valeur, mais il est RECOMMANDÉ d'utiliser la même limitation que l'offre déclare. Dans le cas de diffusion groupée, l'offreur PEUT déclarer une limitation ; il DEVRA y être répondu en utilisant la même valeur. Il est RECOMMANDÉ qu'un envoyeur de supports qui utilise ce format de charge utile inclue toujours le paramètre "max-red". Cette information va probablement simplifier le traitement du flux de supports par le receveur. Cela est particulièrement vrai si aucune redondance ne va être utilisée, et dans ce cas "max-red" est réglé à 0. Comme ce paramètre n'a pas été défini à l'origine, certains envoyeurs ne vont pas déclarer ce paramètre même si ils vont limiter ou non l'envoi de redondances.
- Tout paramètre inconnu dans une offre DEVRA être supprimé dans la réponse.

8.3.2 Usage de SDP déclaratif

En usage déclaratif, comme SDP dans RTSP [RFC2326] ou SAP [RFC2974], les paramètres DEVRONT être interprétés comme suit :

- Les paramètres de configuration de format de charge utile (alignement sur l'octet, crc, tri robuste, entrelacement, et canaux) sont tous déclaratifs, et un participant DOIT utiliser la ou les configurations fournies pour la session. Plus d'une configuration peut être fournie si nécessaire en déclarant plusieurs types de charge utile RTP ; cependant, le nombre de types devrait rester petit.
- Toute restriction du changement de mode du codeur AMR ou AMR-WB et de l'usage du mode par "mode-set", et "mode-change-period" DOIT être suivi par tous les participants à la session. La restriction indiquée par "mode-change-neighbor" DEVRAIT être suivie. Prière de noter que de telles restrictions peuvent être nécessaires si des passerelles à d'autres systèmes de transport comme le GSM participent à la session. Manquer à prendre en compte de telles restrictions peut résulter en une défaillance pour un homologue derrière une telle passerelle à recevoir correctement tout ou partie de la session. Aussi, si des restrictions différentes sont nécessaires pour des homologues différents dans la même session (sauf si un sous ensemble commn de restrictions existe) certains homologues ne vont pas être capables de participer. Noter que l'usage de mode-change-capability n'a pas de signification quant il n'y a pas de négociation, et peut donc être exclu de toute déclaration.
- Toutes les valeurs de "maxptime" et "ptime" devraient être choisies avec soin pour assurer que les participants à la session peuvent réaliser des performances raisonnables.
- L'usage de "max-red" met une limite supérieure globale à l'usage de la redondance qui doit être suivie par tous ceux qui comprennent le paramètre. Cependant, du fait de l'ajout tardif de ce paramètre, il peut être ignoré par certaines mises en œuvre.

8.3.3 Exemples

On donne ci-après des exemples de descriptions de sessions SDP qui utilisent les codages AMR et AMR-WB. Dans ces exemples, les longues lignes a=fmtp vont à la ligne pour respecter les contraintes de colonne de ce document ; le caractère barre oblique ("/") à la fin d'une ligne et le retour chariot qui le suit devraient être ignorés.

Dans un exemple d'usage d'AMR dans un possible scénario de passerelle à passerelle GSM, l'offreur est capable de prendre en charge trois ensembles de mode différents et a besoin que la mode-change-period soit 2 en combinaison avec des restrictions de mode-change-neighbor. L'autre passerelle peut seulement prendre en charge deux de ces ensembles de mode et supprime le type de charge utile 97 dans la réponse. Si la passerelle GSM offreuse prend seulement en charge un seul ensemble de modes actif en même temps, elle devrait envisager de faire les procédures de choix de 1 parmi N décrites au paragraphe 10.2 de la [RFC3264] :

Offre :

```
m=audio 49120 RTP/AVP 97 98 99
a=rtpmap:97 AMR/8000/1
a=fmtp:97 mode-set=0,2,5,7; mode-change-period=2; mode-change-capability=2; mode-change-neighbor=1
a=rtpmap:98 AMR/8000/1
a=fmtp:98 mode-set=0,2,3,6; mode-change-period=2; mode-change-capability=2; mode-change-neighbor=1
a=rtpmap:99 AMR/8000/1
a=fmtp:99 mode-set=0,2,3,4; mode-change-period=2; mode-change-capability=2; mode-change-neighbor=1
a=maxptime:20
```

Réponse :

```
m=audio 49120 RTP/AVP 98 99
a=rtpmap:98 AMR/8000/1
a=fmtp:98 mode-set=0,2,3,6; mode-change-period=2; < mode-change-capability=2; mode-change-neighbor=1
a=rtpmap:99 AMR/8000/1
a=fmtp:99 mode-set=0,2,3,4; mode-change-period=2; mode-change-capability=2; mode-change-neighbor=1
a=maxptime:20
```

L'exemple suivant montre l'usage de AMR entre un point d'extrémité non GSM et une passerelle GSM. L'offreur non GSM n'exige pas de restriction de mode-change-period ou mode-change-neighbor, mais doit signaler son mode-change-capability dans l'offre et appliquer ces restrictions dans la réponse.

Offre :

```
m=audio 49120 RTP/AVP 97
a=rtpmap:97 AMR/8000/1
a=fmtp:97 mode-change-capability=2
a=maxptime:20
```

Réponse :

```
m=audio 49120 RTP/AVP 97
a=rtpmap:97 AMR/8000/1
a=fmtp:97 mode-set=0,2,4,7; mode-change-period=2; mode-change-capability=2; mode-change-neighbor=1
a=maxptime:20
```

Exemple d'usage de AMR-WB dans un possible scénario VoIP où l'UEP peut être utilisé (99) et une déclaration de repli (98) :

```
m=audio 49120 RTP/AVP 99 98
a=rtpmap:98 AMR-WB/16000
a=fmtp:98 octet-align=1; mode-change-capability=2
a=rtpmap:99 AMR-WB/16000
a=fmtp:99 octet-align=1; crc=1; mode-change-capability=2
```

Exemple d'usage de AMR-WB dans un scénario possible de flux directs (deux canaux stéréo) :

```
m=audio 49120 RTP/AVP 99
a=rtpmap:99 AMR-WB/16000/2
a=fmtp:99 entrelacement=30
a=maxptime:100
```

Noter que les noms de format de charge utile (codage) sont couramment donnés en majuscules. Les sous types MIME sont couramment donnés en minuscules. Ces noms sont insensible à la casse dans les deux cas. De même, les noms de paramètres sont insensible à la casse à la fois dans les types MIME et dans la transposition par défaut en l'attribut SDP a=fmtp.

9. Considérations relatives à l'IANA

Deux types de supports (audio/AMR et audio/AMR-WB) ont été mis à jour ; voir la Section 8.

10. Changements par rapport à la RFC 3267

Les différences entre la RFC 3267 et le présent document sont :

- Ajout de précisions sur le comportement à l'égard de la période de changement de mode et du changement de mode voisin qui est attendu d'un client IP ; voir au paragraphe 4.5.
- Mise à jour de maxptime pour une plus grande clarté. La phrase qui disait précédemment : "Le temps DEVRAIT être un multiple de la taille de trame." dit maintenant "Le temps DEVRAIT être un multiple entier de la taille de trame". Ceci ne devrait pas avoir d'impact sur l'interopérabilité.
- Mise à jour de la définition du paramètre mode-set pour une plus grande clarté.
- Restriction des valeurs de mode-change-period à 1 ou 2, qui sont les valeurs utilisées dans les systèmes AMR à commutation de circuit.
- Ajout d'un nouveau paramètre de type de support Mode-Change-Capability qui prend 1 par défaut, qui est le comportement supposé de toute mise en œuvre non mise à jour. Cela permet à la procédure d'offre-réponse de fonctionner.

- Changement de mode-change-neighbor pour indiquer un comportement recommandé plutôt qu'exigé.
- Ajout d'un paragraphe Offre-réponse (8.3.1). Cela aura des implications sur l'interopérabilité des mises en œuvre qui ont fait des suppositions sur comment effectuer la négociation d'offre/réponse des paramètres de charge utile.
- Précision et alignement de l'usage de la détection ingale sur la spécification publiée de UDP léger dans les paragraphes 3.6.1 et 4.4.2.1. Cela inclut de remplacer une déclaration normative sur le traitement de paquet par un paragraphe pour information avec une référence à UDP léger.
- Précision de l'ordre des bits dans le calcul du CRC au paragraphe 4.4.2.1.
- Correction de la référence au paragraphe 5.3 pour les champs Q et FT.
- Changement de la définition du bit de bourrage aux paragraphes 4.4.2 et 5.3 afin qu'il soit clair qu'ils doivent être ignorés.
- Ajout d'une précision que les trames de bruit de confort avec les types de trame 9, 10, et 11 NE DEVRONT PAS être utilisées dans le format de fichier AMR.
- Précisé au paragraphe 4.3.2 que les règles sur le non envoi des trames NO_DATA s'appliquent à toutes les configurations de format de charge utile à l'exception du mode entrelacement.
- La liste de références a été mise à jour avec les RFC maintenant publiées : RFC 3448, RFC 3550, RFC 3551, RFC 3711, RFC 3828, et RFC 4566. Une référence à 3GPP TS 26.101 a aussi été ajoutée.
- Ajout de notes au paragraphe format de mémorisation et enregistrement de type de support disant que les trames AMR et AMR-WB peuvent aussi être mémorisées dans le format de fichier 3GP.
- Ajout d'un paramètre de type de support "max-red" qui permet à l'expéditeur de déclarer un usage limité de la redondance. Ce paramètre permet à un receveur d'optimiser sa fonction car il va savoir si la redondance va être utilisée ou non. Si elle est utilisée, le délai maximum supplémentaire introduit par l'expéditeur (qui est nécessaire pour être pris en considération par le receveur pour utiliser pleinement la redondance) va être connu. L'ajout de ce paramètre ne devrait pas avoir d'effet négatif sur les anciennes mises en œuvre car elles sont obligées d'ignorer les paramètres inconnus selon la RFC 3267. De plus, les anciennes mises en œuvre sont obligées de fonctionner comme si la valeur de max-red est inconnue et éventuellement infinie.
- Mise à jour de l'enregistrement de type de support pour se conformer aux nouvelles règles d'enregistrement.
- Déplacement du paragraphe sur la validation du décodage des Considérations sur la sécurité aux Considérations de mise en œuvre, ce qui est plus leur place.
- Précisions sur l'application du chiffrement, de la protection de l'intégrité, et des mécanismes d'authentification pour la charge utile.

11. Remerciements

Les auteurs tiennent à remercier Petri Koskelainen, Bernhard Wimmer, Tim Fingscheidt, Sanjay Gupta, Stephen Casner, et Colin Perkins de leurs contributions significatives pendant la rédaction et la révision de la RFC 3267 et de son présent remplacement. Les auteurs remercient également Richard Ejzak, Thomas Belling, et Gorry Fairhurst de leurs apports sur le présent document.

12. Références

12.1 Références normatives

[AMR-Code] 3GPP TS 26.090, "Adaptive Multi-Rate (AMR) speech transcoding", version 4.0.0 (2001-03), 3rd Generation Partnership Project (3GPP).

- [AMR-struct] 3GPP TS 26.101, "AMR Speech Codec Frame Structure", version 4.1.0 (2001-06), 3rd Generation Partnership Project (3GPP).
- [AMR-WB] 3GPP TS 26.190 "AMR Wideband speech codec; Transcoding functions", version 5.0.0 (2001-03), 3rd Generation Partnership Project (3GPP).
- [AMR-WB-Struc] 3GPP TS 26.201 "AMR Wideband speech codec; Frame Structure", version 5.0.0 (2001-03), 3rd Generation Partnership Project (3GPP).
- [AMR-WB-SCR] 3GPP TS 26.093, "AMR Speech Codec; Source Controlled Rate operation", version 4.0.0 (2000-12), 3rd Generation Partnership Project (3GPP).
- [RFC2119] S. Bradner, "[Mots clés à utiliser](#) dans les RFC pour indiquer les niveaux d'exigence", BCP 14, mars 1997. (MàJ par [RFC8174](#))
- [RFC3264] J. Rosenberg et H. Schulzrinne, "[Modèle d'offre/réponse](#) avec le protocole de description de session (SDP)", juin 2002. (P.S. ; MàJ par [RFC8843](#))
- [RFC3550] H. Schulzrinne, S. Casner, R. Frederick et V. Jacobson, "[RTP : un protocole de transport pour les applications en temps réel](#)", STD 64, juillet 2003. (MàJ par [RFC7164](#), [RFC7160](#), [RFC8083](#), [RFC8108](#), [RFC8860](#))
- [RFC3551] H. Schulzrinne et S. Casner, "[Profil RTP pour conférences audio](#) et vidéo avec contrôle minimal", STD 65, juillet 2003. (MàJ par [RFC8860](#))
- [RFC4288] N. Freed et J. Klensin, "Spécifications du [type de support et procédures d'enregistrement](#)", [BCP 13](#), décembre 2005.
- [RFC4566] M. Handley, V. Jacobson et C. Perkins, "SDP : [Protocole de description de session](#)", juillet 2006. (P.S. ; remplacée par [RFC8866](#))
- [RFC4855] S. Casner, "[Enregistrement du type de support](#) des formats de charge utile RTP", février 2007. (P.S.)
- [TS26.092] 3GPP TS 26.092, "AMR Speech Codec; Comfort noise aspects", version 4.0.0 (2001-03), 3rd Generation Partnership Project (3GPP).
- [TS26.192] 3GPP TS 26.192 "AMR Wideband speech codec; Comfort Noise aspects", version 5.0.0 (2001-03), 3rd Generation Partnership Project (3GPP).
- [TS26.193] 3GPP TS 26.193 "AMR Wideband Speech Codec; Source Controlled Rate operation", version 5.0.0 (2001-03), 3rd Generation Partnership Project (3GPP).

12.2 Références pour information

- [ACELP] ANSI/TIA/EIA-136-Rev.C, part 410 - "TDMA Cellular/PCS Radio Interface, Enhanced Full Rate Voice Codec (ACELP)". Anciennement IS-641. Norme TIA publiée le 1er juin 2001.
- [ARIB] ARIB, RCR STD-27H, "Personal Digital Cellular Telecommunication System RCR Standard", Association of Radio Industries et Businesses (ARIB).
- [GSM06.60] Institut Européen des normes de télécommunications, GSM 06.60, "Enhanced Full Rate (EFR) speech transcoding", version 8.0.1 (2000-11), (ETSI).
- [RFC2198] C. Perkins et autres, "[Charge utile RTP pour données audio redondantes](#)", septembre 1997. (P.S.)
- [RFC2326] H. Schulzrinne, A. Rao et R. Lanphier, "Protocole de [flux directs en temps réel](#) (RTSP)", avril 1998. (Remplacée par [RFC7826](#))
- [RFC2733] J. Rosenberg et H. Schulzrinne, "Format de charge utile RTP pour la correction d'erreur directe générique",

décembre 1999. (*Obsolète, voir [RFC5109](#)*) (*P.S.*)

- [RFC2974] M. Handley, C. Perkins, E. Whelan, "Protocole d'annonce de session (SAP)", octobre 2000. (*Expérimentale*)
- [RFC3448] M. Handley, S. Floyd, J. Padhye, J. Widmer, "[Contrôle de débit convivial](#) sur TCP (TFRC) : Spécification du protocole", janvier 2003. (*Obsolète, voir [RFC5348](#)*) (*P.S.*)
- [RFC3711] M. Baugher et autres, "Protocole de [transport sécurisé en temps réel](#) (SRTP)", mars 2004. (*P.S.*)
- [RFC3828] L-A. Larzon et autres, "[Protocole léger de datagramme d'utilisateur](#) (UDP-Lite)", juillet 2004. (*P.S, MàJ par [RFC6335](#)*)
- [RFC3839] R. Castagno, D. Singer, "Enregistrements de types MIME pour les fichiers multimédia du projet en partenariat de 3ème génération (3GPP)", juillet 2004. (*P.S.*)
- [RFC4301] S. Kent et K. Seo, "[Architecture de sécurité](#) pour le protocole Internet", décembre 2005. (*P.S.*) (*Remplace la [RFC2401](#)*)
- [RFC4346] T. Dierks et E. Rescorla, "Protocole de sécurité de la couche Transport (TLS) version 1.1", avril 2006. (*Remplace [RFC2246](#) ; Remplacée par [RFC5246](#) ; MàJ par [RFC4366](#), [4680](#), [4681](#), [5746](#), [6176](#), [7465](#), [7507](#), [7919](#)*)
- [RFC5109] A. Li, éd., "Format de charge utile RTP pour la [correction générique d'erreur directe](#)", décembre 2007. (*Remplace [RFC2733](#), [RFC3009](#)*) (*P.S.*)
- [TS25.415] 3GPP TS 25.415 "UTRAN Iu Interface User Plane Protocols", version 4.2.0 (2001-09), 3rd Generation Partnership Project (3GPP).
- [TS26.102] 3GPP TS 26.102, "AMR speech codec interface to Iu and Uu", version 4.0.0 (2001-03), 3rd Generation Partnership Project (3GPP).
- [TS26.103] 3GPP TS 26.103, "Speech codec list for GSM and UMTS", version 5.5.0 (2004-09), 3rd Generation Partnership Project (3GPP).
- [TS26.202] 3GPP TS 26.202, "AMR Wideband speech codec; Interface to Iu and Uu", version 5.0.0 (2001-03), 3rd Generation Partnership Project (3GPP).
- [TS26.244] 3GPP TS 26.244, "3GPP file format (3GP)", version 6.1.0 (2004- 09), 3rd Generation Partnership Project (3GPP).

Les documents ETSI sont disponibles à <<http://www.etsi.org/>>. Les documents 3GPP sont disponibles à <<http://www.3gpp.org/>>. Les documents TIA sont disponibles à <<http://www.tiaonline.org/>>.

Adresse des auteurs

Johan Sjoberg
Ericsson AB
SE-164 80 Stockholm, SWEDEN
téléphone : +46 8 7190000
mél : Johan.Sjoberg@ericsson.com

Magnus Westerlund
Ericsson Research
SE-164 80 Stockholm, SWEDEN
téléphone : +46 8 7190000
mél : Magnus.Westerlund@ericsson.com

Ari Lakaniemi
Nokia Research Center
P.O.Box 407
FIN-00045 Nokia Group, FINLAND
téléphone : +358-71-8008000
mél : ari.lakaniemi@nokia.com

Qiaobing Xie
Motorola, Inc.
1501 W. Shure Drive, 2-B8
Arlington Heights, IL 60004, USA
téléphone : +1-847-632-3028
mél : Qiaobing.Xie@motorola.com

Déclaration complète de droits de reproduction

Copyright (C) The IETF Trust (2007)

Le présent document est soumis aux droits, licences et restrictions contenus dans le BCP 78, et sauf pour ce qui est mentionné ci-après, les auteurs conservent tous leurs droits.

Le présent document et les informations contenues sont fournis sur une base "EN L'ÉTAT" et le contributeur, l'organisation qu'il ou elle représente ou qui le/la finance (s'il en est), la INTERNET SOCIETY, le IETF TRUST et la INTERNET ENGINEERING TASK FORCE déclinent toutes garanties, exprimées ou implicites, y compris mais non limitées à toute garantie que l'utilisation des informations encloses ne viole aucun droit ou aucune garantie implicite de commercialisation ou d'aptitude à un objet particulier.

Propriété intellectuelle

L'IETF ne prend pas position sur la validité et la portée de tout droit de propriété intellectuelle ou autres droits qui pourraient être revendiqués au titre de la mise en œuvre ou l'utilisation de la technologie décrite dans le présent document ou sur la mesure dans laquelle toute licence sur de tels droits pourrait être ou n'être pas disponible ; pas plus qu'elle ne prétend avoir accompli aucun effort pour identifier de tels droits. Les informations sur les procédures de l'ISOC au sujet des droits dans les documents de l'ISOC figurent dans les BCP 78 et BCP 79.

Des copies des dépôts d'IPR faites au secrétariat de l'IETF et toutes assurances de disponibilité de licences, ou le résultat de tentatives faites pour obtenir une licence ou permission générale d'utilisation de tels droits de propriété par ceux qui mettent en œuvre ou utilisent la présente spécification peuvent être obtenues sur le répertoire en ligne des IPR de l'IETF à <http://www.ietf.org/ipr>.

L'IETF invite toute partie intéressée à porter son attention sur tous copyrights, licences ou applications de licence, ou autres droits de propriété qui pourraient couvrir les technologies qui peuvent être nécessaires pour mettre en œuvre la présente norme. Prière d'adresser les informations à l'IETF à ietf-ipr@ietf.org.

Remerciement

Le financement de la fonction d'édition des RFC est actuellement fourni par la Internet Society.