

Groupe de travail Réseau
Request for Comments : 4348
 Catégorie : Sur la voie de la normalisation

S. Ahmadi
 janvier 2006
 Traduction Claude Brière de L'Isle

Format de charge utile RTP pour codec vocal multimode large bande à débit variable (VMR-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 Internet Society (2006).

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 le codec vocal multimode large bande à débit variable (VMR-WB, *Variable-Rate Multimode Wideband*). Le format de charge utile est destiné à être capable d'interopérer avec les formats existants de transport VMR-WB sur des réseaux non IP. Un enregistrement de type de support est inclus pour le format de charge utile VMR-WB RTP.

VMR-WB est un codec large bande multimode à débit variable qui a un certain nombre de modes de fonctionnement, dont un est interopérable avec le codec audio AMR-WB (c'est-à-dire, de la RFC 3267) à certains débits. Donc, des dispositions ont été prises dans le présent document pour faciliter et simplifier l'échange de paquets de données entre VMR-WB et AMR-WB dans le mode interopérable sans impliquer de fonction de transcodage.

Table des matières

1. Introduction.....	2
2. Conventions et acronymes.....	2
3. Codec vocal multimode large bande à débit variable (VMR-WB).....	3
3.1 Traitement de la parole en bande étroite.....	3
3.2 Transmission continue contre discontinue.....	4
3.3 Prise en charge de session multi canaux.....	4
4. Robustesse contre la perte de paquet.....	4
4.1 Correction d'erreur directe (FEC).....	4
4.2 Entrelacement de trame et encapsulation multi trames.....	5
5. Scénarios de voix sur IP VMR-WB.....	5
5.1 De terminal IP à terminal IP.....	5
5.2 De passerelle à terminal IP.....	6
5.3 De passerelle à passerelle (entre terminaux à capacité VMR-WB et à capacité AMR-WB).....	6
5.4 De passerelle à passerelle (entre deux terminaux à capacité VMR-WB).....	7
6. Formats de charge utile RTP VMR-WB.....	7
6.1 Usage de l'en-tête RTP.....	7
6.2 Format de charge utile sans en-tête.....	8
6.3 Format de charge utile alignée sur l'octet.....	9
6.4 Considérations de mise en œuvre.....	13
7. Contrôle d'encombrement.....	13
8. Considérations sur la sécurité.....	14
8.1 Confidentialité.....	14
8.2 Authentification et intégrité.....	14
9. Paramètres de format de charge utile.....	14
9.1 Enregistrement MIME de charge utile RTP VMR-WB.....	15
9.2. Transposition des paramètres MIME en SDP.....	16
9.3 Considérations sur le modèle offre-réponse.....	16
10. Considérations relatives à l'IANA.....	17
11. Remerciements.....	17

12. Références.....	17
12.1 Références normatives.....	17
12.2 Références pour information.....	18
Adresse de l'auteur.....	18
Déclaration complète de droits de reproduction.....	18

1. Introduction

Le présent document spécifie le format de charge utile pour la mise en paquets des signaux de parole codés en VMR-WB dans le protocole de transport en temps réel (RTP, *Real-time Transport Protocol*) [RFC3550]. Le format de charge utile VMR-WB prend en charge la transmission de canaux seuls et multiples, d'entrelaçage de trame, de plusieurs trames par charge utile, de charge utile sans en-tête, l'utilisation de la commutation de mode, et l'interopération avec les formats existants de transport VMR-WB sur des réseaux non IP, comme décrit à la Section 3.

Le format de charge utile est décrit à la Section 6. Le format de fichier VMR-WB (c'est-à-dire, pour le transport de données de parole VMR-WB dans des applications en mode mémorisation comme la messagerie électronique) est spécifié dans [C.S0050-A]. À la Section 9, un enregistrement de type de support est fourni pour le format de charge utile VMR-WB RTP.

Comme VMR-WB est interopérable avec AMR-WB à certain débits, le présent document a tenté de maximiser les similitudes avec la RFC 3267 tout en optimisant le format de charge utile pour les modes non interopérables du codec VMR-WB.

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] et indiquent les niveaux d'exigence pour les mises en œuvre RTP conformes..

Les acronymes suivants sont utilisés dans le document :

3GPP (*Third Generation Partnership Project*) : projet en partenariat de troisième génération
 3GPP2 (*Third Generation Partnership Project 2*) : projet n° 2 de partenariat de troisième génération
 CDMA (*Code Division Multiple Access*) : accès multiple par répartition en code
 WCDMA (*Wideband Code Division Multiple Access*) : accès multiple par répartition en code large bande
 GSM (*Global System for Mobile Communications*) : système mondial de communications mobiles
 AMR-WB (*Adaptive Multi-Rate Wideband Codec*) : codec large bande multi débits adaptatif
 VMR-WB (*Variable-Rate Multimode Wideband Codec*) : codec large bande multi modes à débit variable
 CMR (*Codec Mode Request*) : demande de mode codec
 GW (*Gateway*) : passerelle
 DTX (*Discontinuous Transmission*) : transmission discontinue
 FEC (*Forward Error Correction*) : correction d'erreur directe
 SID (*Silence Descriptor*) : descripteur de silence
 TrFO (*Transcoder-Free Operation*) : fonctionnement sans transcodeur
 UDP (*User Datagram Protocol*) : protocole de datagramme d'utilisateur
 RTP (*Real-Time Transport Protocol*) : protocole de transport en temps réel
 RTCP (*RTP Control Protocol*) : protocole de contrôle du protocole de transport en temps réel
 MIME (*Multipurpose Internet Mail Extension*) : extensions de messagerie Internet multi objets
 SDP (*Session Description Protocol*) : protocole de description de session
 VoIP (*Voice-over-IP*) : voix sur IP

Le terme de "mode interopérable" dans ce document se réfère au mode 3 VMR-WB, qui est interopérable avec les codecs AMR-WB modes 0, 1, et 2.

Le terme "modes non interopérables" dans ce document se réfère aux modes VMR-WB 0, 1, et 2.

Le terme "bloc de trames" est utilisé dans ce document pour décrire l'ensemble synchronisé de trames de parole dans une

session VMR-WB multi canaux. 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 représentent exactement la même durée.

3. Codec vocal multimode large bande à débit variable (VMR-WB)

VMR-WB est la norme de codage de la parole en large bande développée par le projet n° 2 en partenariat de troisième génération (3GPP2) pour le codage/décodage de contenus de parole en large bande/bande étroite dans les services multimédia des systèmes cellulaires en CDMA de troisième génération [C.S0052-0]. VMR-WB est un codec vocal large bande multi mode à débit variable à source contrôlée. Il a un certain nombre de modes de fonctionnement, où chaque mode est un compromis entre la qualité vocale et le débit de données moyen. Le mode de fonctionnement dans VMR-WB (comme le montre le Tableau 2) est choisi sur la base des conditions de trafic du réseau et de la qualité de service désirée. Le débit de données moyen désiré dans chaque mode est obtenu en codant les trames de parole aux débits permis (comme montré dans les Tableaux 1 et 3) conformes au système CDMA2000, selon les caractéristiques instantanées de l'entrée de parole et des contraintes de débit maximum et minimum imposées par l'opérateur du réseau.

Alors que VMR-WB est un codec CDMA natif conforme à toutes les exigences du système CDMA, il est de plus interopérable avec AMR-WB [RFC3267], [TS26.193] à 12,65, 8,85, et 6,60 kbit/s. Ceci est dû au fait que VMR-WB et AMR-WB partagent la même technologie sous-jacente. Cette caractéristique permet des interconnexions sans transcodeur (TrFO, *Transcoder-Free*) entre VMR-WB et AMR-WB à travers des systèmes filaires/sans fil différents (par exemple, GSM/WCDMA et CDMA2000) sans utiliser de conversion de format de support inutilement complexe.

Noter que le concept de mode dans VMR-WB est différent de celui de AMR-WB où chaque mode de codec AMR-WB à débit fixe est adapté aux conditions de canal prévalentes par un compromis entre le nombre total de bits de codage de source et de bits de codage de canal.

VMR-WB est capable de transitions entre divers modes sans dégradation de la qualité vocale attribuable à la commutation de mode elle-même. Le mode de fonctionnement du codeur VMR-WB peut être commuté sans à-coups sans connaissance préalable du décodeur. Tout mode non interopérable (c'est-à-dire, les modes VMR-WB 0, 1, ou 2) peut être choisi selon les conditions de trafic (par exemple, encombrement du réseau) et la qualité de service désirée.

En mode interopérable (c'est-à-dire, en mode VMR-WB 3) le changement de mode entre les modes VMR-WB n'est pas permis parce que il y a seulement un mode AMR-WB interopérable en VMR-WB. Comme le codec AMR-WB peut demander un changement de mode, selon les conditions du canal, les données dans la bande incluses dans les structures de trames de parole VMR-WB (voir les détails à la Section 8 de [C.S0052-0]) sont utilisées durant une interconnexion interopérable pour commuter entre les trames de parole VMR-WB des types 0, 1, et 2 en VMR-WB mode 3 (correspondant aux modes de codec AMR-WB 0, 1, ou 2).

Comme mentionné précédemment, VMR-WB est conforme au système CDMA2000 avec les taux de codage permis montrés dans le Tableau 1.

Type de trame	Bits par paquet (taille de trame)	Taux de codage (kbit/s)
Plein débit	266	13,3
Demi débit	124	6,2
Quart de débit	54	2,7
Huitième de débit	20	1,0
Blanc	0	0
Écrasement	0	0

Tableau 1 : Types de trames de parole permises en système CDMA2000 et taux de codage associé

VMR-WB est robuste à de forts pourcentages de perte de trames de parole et aux trames de paroles avec des informations de débit corrompues. La réception d'un type de trames de parole de Écrasement (SPEECH_LOST) au décodeur invoque le mécanisme incorporé de dissimulation d'erreur de trames de parole. Le mécanisme incorporé de dissimulation d'erreur de trames de parole dans VMR-WB cache les effets des trames de paroles perdues en exploitant les données dans la bande et les informations disponibles dans les trames de paroles précédentes.

3.1 Traitement de la parole en bande étroite

VMR-WB a la capacité de fonctionner avec des signaux de parole en entrée/sortie échantillonnés à 16000 Hz ou 8000 Hz dans tous les modes de fonctionnement [C.S0052-0]. Le décodeur VMR-WB n'exige pas une connaissance a priori du taux d'échantillonnage du support original (c'est-à-dire, signaux parole/audio échantillonnés à 8 ou 16 kHz) à l'entrée du codeur. Le décodeur VMR-WB génère, par défaut, une sortie à large bande de 16000 Hz sans considération de la fréquence d'échantillonnage du codeur d'entrée. Selon l'application, le décodeur peut aussi être configuré à générer une sortie à 8000 Hz.

Donc, bien que la présente spécification définit un débit d'horloge RTP de 16000 Hz pour le codec VMR-WB, l'injection et le traitement de supports de bande étroite à 8000 Hz durant une session est aussi permis ; cependant, un débit d'horloge RTP de 16000 Hz DOIT toujours être utilisé.

Le choix d'une fréquence d'échantillonnage de sortie de VMR-WB dépend de la mise en œuvre et des capacités audio acoustiques du côté receveur.

3.2 Transmission continue contre discontinue

Le fonctionnement en commutation de circuits de VMR-WB au sein d'un réseau CDMA exige la transmission continue des données de parole durant une conversation. La caractéristique intrinsèque de débit variable à source contrôlée des codecs de parole CDMA est requise pour un fonctionnement optimal du système CDMA et du contrôle d'interférences. Cependant, VMR-WB a la capacité de fonctionner en mode de transmission discontinue pour certaines applications de commutation de paquets sur les réseaux IP (par exemple, VoIP) où le nombre de bits et paquets transmis durant les périodes de silence est réduit au minimum. Le fonctionnement des DTX VMR-WB est similaire à celui de AMR-WB [RFC3267], [TS26.193].

3.3 Prise en charge de session multi canaux

Le format de charge utile RTP aligné sur l'octet défini dans le présent document prend en charge le contenu audio multi canaux (par exemple, une session de parole stéréophonique). Bien que le codec VMR-WB lui-même ne prenne pas en charge le codage du contenu audio multi canaux dans un seul flux binaire, il peut être utilisé pour coder et décoder chacun des canaux individuels séparément.

Pour transporter le contenu multi canaux codé séparément, les trames de parole pour tous les canaux qui sont des trames de parole et codés pour les mêmes périodes de 20 ms sont logiquement collectés dans un bloc de trames de parole.

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 des différents canaux dans chaque bloc de trames de parole. Quand on utilise SDP pour la signalisation (voir les détails au paragraphe 9.2) le nombre de canaux est spécifié dans l'attribut rtpmap, et l'ordre des canaux porté dans chaque bloc de trames de parole est impliqué par le nombre de canaux comme spécifié au paragraphe 4.1 de la [RFC3551].

4. Robustesse contre la perte de paquet

Le format de charge utile aligné sur l'octet décrit dans le présent document (voir les détails à la Section 6) prend en charge plusieurs caractéristiques, incluant la correction d'erreur directe (FEC) et l'entrelacement de trames de parole, afin d'augmenter la robustesse contre la perte de paquets.

4.1 Correction d'erreur directe (FEC)

Le simple schéma de répétition des données envoyées précédemment est une façon de réaliser la FEC. Un autre schéma possible, qui est plus efficace en bande passante, est d'utiliser la FEC de charge utile externe ; par exemple, de la [RFC2733], qui génère des paquets supplémentaires contenant les données de réparation.

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

Dans cet exemple, chaque bloc de trames de parole est retransmis une fois dans le paquet de charge utile RTP suivant. Ici,

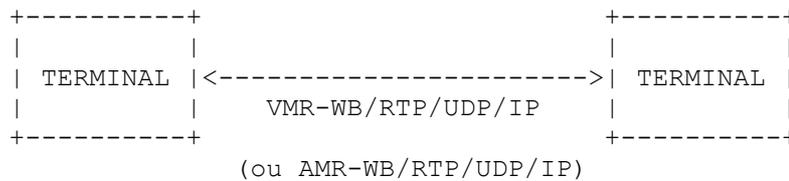


Figure 2 : terminal IP à terminal IP

Un service de conversation fait peser des exigences sur le format de charge utile. Un faible délai est un facteur très important, c'est-à-dire, moins de blocs de trames de parole par paquet de charge utile. De faibles frais généraux sont aussi nécessaires quand le format de charge utile traverse des liaisons à faible bande passante, en particulier si la fréquence des paquets va être élevée.

Le service de flux directs a des exigences de temps réel moins strictes et donc peut utiliser un plus grand nombre de blocs de trames de parole 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 de parole par paquet rend la transmission plus vulnérable à la perte de paquet, de sorte que l'entrelacement peut être utilisé pour réduire l'effet de la perte de paquet sur la qualité de 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.

Pour les terminaux IP à capacité VMR-WB aux deux extrémités, selon la mise en œuvre, tous les modes de codec VMR-WB peuvent être utilisés dans ce scénario. Aussi, les deux formats de charge utile sans en-tête et aligné sur l'octet (voir les détails à la Section 6) peuvent être utilisés. Pour l'interconnexion interopérable entre VMR-WB et AMR-WB, seul le mode 3 VMR-WB est utilisé, et toutes les restrictions décrites au paragraphe 9.3 s'appliquent.

5.2 De passerelle à terminal IP

Un autre scénario se présente quand la parole codée en VMR-WB va être transmise d'un système non IP (par exemple, un réseau 3GPP2/CDMA2000) à un terminal IP, et/ou vice versa, comme montré à la Figure 3.

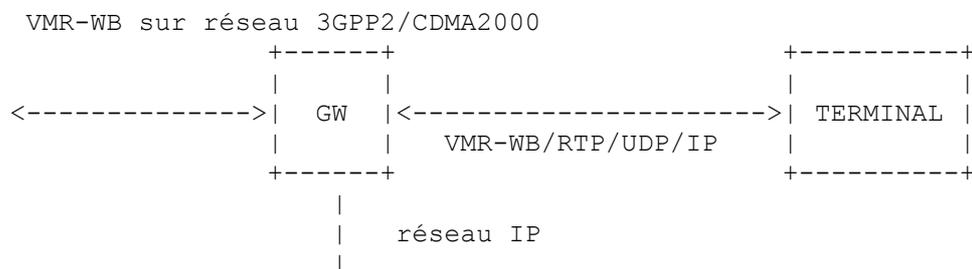


Figure 3 : scénario de terminal GW à VoIP

La capacité de VMR-WB de commuter en douceur entre modes de fonctionnement est exploitée dans les réseaux CDMA (non IP) pour optimiser la qualité de la parole pour une certaine condition de trafic. Pour préserver cette fonctionnalité dans les scénarios qui incluent une passerelle à un réseau IP en utilisant le format de charge utile aligné sur l'octet, un champ demande de mode codec (CMR, *codec mode request*) est proposé. 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 parole qui va au décodeur non IP. L'algorithme de contrôle de mode dans la passerelle DEVRAIT s'accommoder du délai imposé par le réseau IP sur la réponse à la CMR par le terminal IP.

Le terminal IP NE DEVRAIT pas établir la CMR (voir le paragraphe 6.3.2) mais la passerelle peut régler la valeur de la CMR sur les trames de paroles qui vont vers le codeur dans la partie non IP pour optimiser la qualité de parole de ce codeur à la passerelle et pour effectuer le contrôle d'encombrement sur le réseau IP.

5.3 De passerelle à passerelle (entre terminaux à capacité VMR-WB et à capacité AMR-WB)

Un troisième scénario probable est que RTP/UDP/IP est utilisé comme transport entre deux systèmes non IP, c'est-à-dire, IP est à l'origine et se termine sur des passerelles des deux côtés du transport IP, comme illustré à la Figure 4. C'est le scénario le plus probable pour une interconnexion interopérable entre des stations mobiles 3GPP/(GSM-WCDMA)/AMR-

WB et à capacité 3GPP2/CDMA2000/VMR-WB. Dans ce scénario, le terminal à capacité VMR-WB se déclare aussi capable de AMR-WB avec le mode restreint réglé comme décrit au paragraphe 9.3. La valeur de CMR peut être établie 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 : (1) la valeur de CMR qu'elle reçoit sur le côté IP ; (2) une valeur de CMR qu'elle peut choisir pour le contrôle d'encombrement de la transmission sur le côté IP ; et (3) la valeur de CMR fondée sur son estimation de la qualité de réception sur le côté non IP. Les détails de l'algorithme de contrôle du trafic sont laissés à la mise en œuvre.

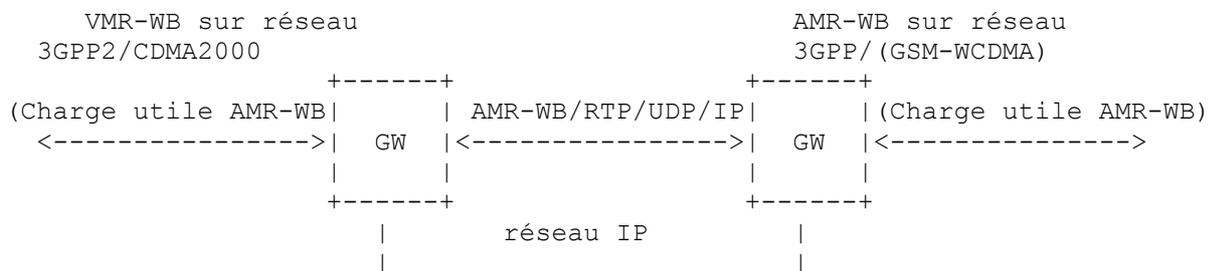


Figure 4 : scénario GW à GW (AMR-WB <-> interconnexion interopérable VMR-WB)

Durant et à l'initiation d'une interconnexion interopérable entre VMR-WB et AMR-WB, seul le mode VMR-WB 3 peut être utilisé. Il y a trois types de trames (c'est-à-dire, FT=0, 1, ou 2 ; voir le Tableau 3) au sein de ce mode qui sont compatibles avec le codec AMR-WB modes 0, 1, et 2, respectivement. Si le codec AMR-WB est engagé dans une interconnexion interopérable avec VMR-WB, l'ensemble de modes de codec actifs AMR-WB doit être limité à 0, 1, et 2.

5.4 De passerelle à passerelle (entre deux terminaux à capacité VMR-WB)

Le quatrième exemple de scénario VoIP est composé d'un transport RTP/UDP/IP entre deux systèmes non IP ; c'est-à-dire, IP est à l'origine et terminé sur des passerelles des deux côtés du transport IP, comme illustré à la Figure 5. C'est le scénario le plus probable pour une interconnexion de station mobile à station mobile (MS-à-MS) sans transcodeur (TrFO, *Transcoder-Free*) entre deux terminaux 3GPP2/CDMA2000 qui utilisent tous deux le codec VMR-WB.

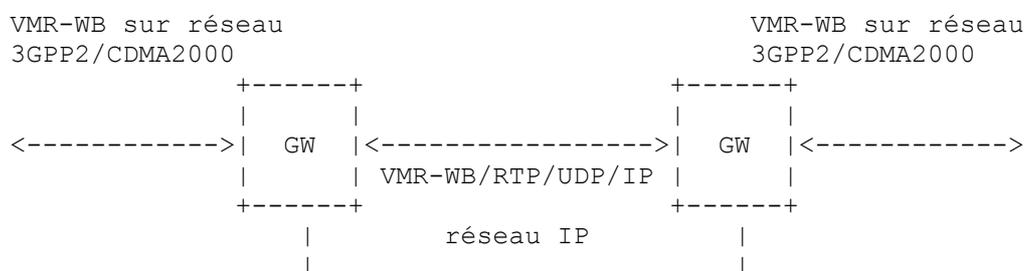


Figure 5 : scénario GW à GW (scénario VoIP CDMA2000 de MS-à-MS)

6. Formats de charge utile RTP VMR-WB

Pour une certaine session, le format de charge utile peut être soit sans en-tête, soit aligné sur l'octet, selon le mode de fonctionnement qui est établi pour la session via des moyens hors bande, et l'application.

Le format de charge utile sans en-tête est conçu pour une efficacité maximum de la bande passante, la simplicité, et une faible latence. Une seule trame de parole de données du codec peut être envoyée dans chaque paquet de format de charge utile sans en-tête. Aucun champ d'en-tête de la charge utile ni entrée du tableau de contenus (ToC) n'est présent (la même considération est aussi faite dans la [RFC3558]).

Dans le format de charge utile aligné sur l'octet, tous les champs d'une charge utile, incluant l'en-tête de charge utile, les entrées de tableau des contenus, et les trames de parole elles-mêmes, sont individuellement alignés sur les limites d'octet pour rendre les mises en œuvre efficaces.

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

Entre les deux formats de charge utile, seul le format aligné sur l'octet a la capacité d'utiliser l'entrelacement pour rendre le transport de parole robuste à la perte de paquet.

Le format de charge utile VMR-WB aligné sur l'octet dans le mode interopérable est identique à celui de AMR-WB (c'est-à-dire, de la RFC 3267).

6.1 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 cette spécification.

L'horodatage RTP correspond à l'instant du premier échantillonnage codé pour les premiers blocs de trames de parole dans le paquet. La fréquence de l'horloge d'horodatage est la même que la fréquence d'échantillonnage par défaut (c'est-à-dire, 16 kHz) de sorte que l'unité d'horodatage est en échantillons.

La durée d'un bloc de trames de parole est de 20 ms pour VMR-WB. Pour le fonctionnement normal à large bande de VMR-WB, la fréquence d'échantillonnage du support en entrée/sortie est de 16 kHz, correspondant à 320 échantillons par trame de parole provenant de chaque canal. Donc, l'horodatage est augmenté de 320 pour VMR-WB pour chaque bloc consécutif de trames de parole.

Le codec VMR-WB est capable de traiter les signaux parole/audio échantillonnés à 8 kHz. Par défaut, la fréquence d'échantillonnage de sortie du décodeur VMR-WB est de 16 kHz. Selon l'application, le décodeur peut être configuré à générer aussi une fréquence d'échantillonnage de sortie de 8 kHz. Comme les formats de charge utile RTP de VMR-WB pour les supports échantillonnés à 8 et 16 kHz sont identiques et que le décodeur VMR-WB n'a pas besoin de savoir à priori la fréquence d'échantillonnage d'entrée du codeur, un débit d'horloge RTP fixe de 16000 Hz est défini pour le codec VMR-WB. Cela permet l'injection ou le traitement de supports parole/audio échantillonnés à 8 kHz sans avoir à changer le débit d'horloge RTP durant une session. Noter que l'horodatage est aussi incrémenté de 320 par bloc de trames de parole pour les supports échantillonnés à 8 kHz.

Un paquet peut contenir plusieurs blocs de trames de parole de parole codée ou de paramètres de bruit de confort. Si l'entrelacement est employé, les blocs de trames de parole encapsulés dans une charge utile sont pris conformément aux règles d'entrelacement définies au paragraphe 6.3.2. Autrement, chaque paquet couvre une période d'un ou plusieurs intervalles contigus de blocs de trames de parole de 20 ms. Dans le cas où des données provenant de tous les canaux pour un bloc particulier de trames de parole sont manquantes dans la période (par exemple, à une passerelle avec un autre format de transport) il est possible d'indiquer qu'aucune donnée n'est présente pour ce bloc de trames de parole, au lieu de casser en deux un paquet de bloc multi-trames de parole, comme expliqué au paragraphe 6.3.2.

Sans considération du format de charge utile utilisé, la charge utile RTP est toujours d'un nombre entier d'octets avec un bourrage de bits à zéro si nécessaire. Si un bourrage supplémentaire est requis pour amener la longueur de la charge utile à un multiple d'octets plus grand ou pour quelque autre raison, alors le bit P dans l'en-tête RTP PEUT être établi, et un bourrage ajouté, comme spécifié dans la [RFC3550].

Le bit marqueur (M) de l'en-tête RTP DEVRA toujours être réglé à 0 si le codec VMR-WB opère en transmission continue. Quand il fonctionne en transmission discontinue (DTX) le bit marqueur de l'en-tête RTP DEVRA être réglé à 1 si le premier bloc de trames de parole porté dans le paquet contient une trame de parole, qui est la première dans une émission de parole. Pour tous les autres paquets, le bit marqueur DEVRA être réglé à zéro (M=0).

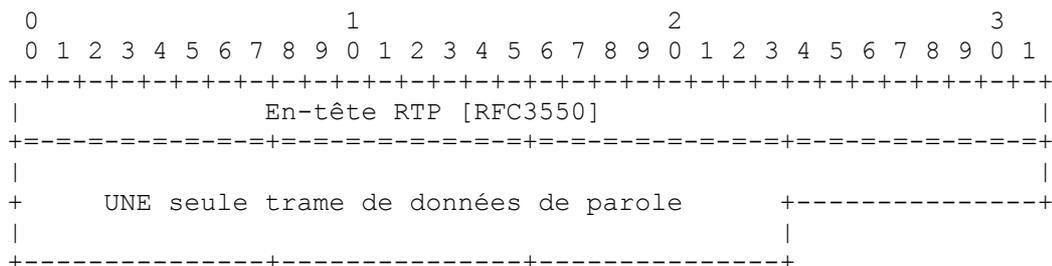
L'allocation d'un type de charge utile RTP pour ce format de charge utile sort du domaine d'application du présent document et ne sera pas spécifiée ici. On suppose que le profil RTP sous lequel ce format de charge utile va être utilisé allouera un type de charge utile pour ce codage ou spécifiera que le type de charge utile est alloué dynamiquement (Section 9).

6.2 Format de charge utile sans en-tête

Le format de charge utile sans en-tête est conçu pour une efficacité maximale de la bande passante, la simplicité, et le

minimum de délai. Seulement une trame de parole de données de parole est présente dans chaque paquet de format de charge utile sans en-tête. Aucun champ de l'en-tête de charge utile ni d'entrée de ToC n'est présent. Le taux de codage pour les trames de parole peut être déterminé à partir de la longueur des trames de parole de données de parole, car il y a seulement une trame de parole de données de parole dans chaque format de charge utile sans en-tête.

L'utilisation des champs d'en-tête RTP pour le format de charge utile sans en-tête est le même que son correspondant pour le format de charge utile aligné sur l'octet. La transposition binaire détaillée des paquets de données de parole permise pour ce format de charge utile est décrit à la Section 8 de [C.S0052-0]. Comme le format de charge utile sans en-tête n'est pas compatible avec la charge utile RTP AMR-WB, seuls des modes non interopérables de VMR-WB DEVRONT être utilisés avec ce format de charge utile. C'est-à-dire que FT=0, 1, 2, et 9 NE DEVRA PAS être utilisé avec le format de charge utile sans en-tête.

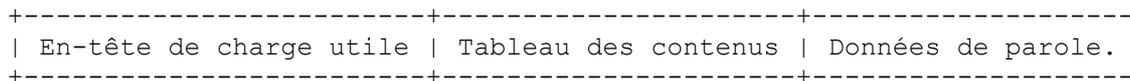


Noter que le mode de fonctionnement, en utilisant ce format de charge utile, est décidé par le site d'émission (codeur). Le mode de fonctionnement pas défaut pour le codeur VMR-WB est le mode 0 [C.S0052-0]. La demande de changement de mode PEUT aussi être envoyée par des moyens non RTP, qui sortent du domaine d'application de cette spécification.

6.3 Format de charge utile alignée sur l'octet

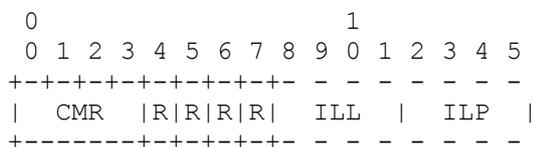
6.3.1 Structure de charge utile

La charge utile complète consiste en un en-tête de charge utile, un tableau des contenus de la charge utile, et en données de parole représentant un ou plusieurs blocs de trames de parole. Le diagramme suivant montre la disposition générale du format de charge utile :



6.3.2 En-tête de charge utile

Dans le format de charge utile aligné sur l'octet, l'en-tête de 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) : cela indique une demande de mode codec envoyée au codeur de parole au site du receveur de cette charge utile. La valeur de CMR de 15 indique qu'aucune demande de mode n'est présente, et les autres valeurs non utilisées sont réservée pour une future utilisation.

La valeur cu champ CMR est réglé en accord avec le tableau suivant :

CMR	Modes de fonctionnement VMR-WB
0	mode VMR-WB 3 (mode AMR-WB interopérable à 6,60 kbit/s)
1	mode VMR-WB 3 (mode AMR-WB interopérable à 8,85 kbit/s)
2	mode VMR-WB 3 (mode AMR-WB interopérable à 12,65 kbit/s)

3	mode VMR-WB 2
4	mode VMR-WB 1
5	mode VMR-WB 0
6	mode VMR-WB 2 avec codage demi débit maximum
7-14	(réservé)
15	pas de préférence (aucune demande de mode n'est présente)

Tableau 2 : Liste des valeurs de CMR valides et leur mode de fonctionnement VMR-WB associé

R : c'est un bit réservé qui DOIT être réglé à zéro. Le receveur DOIT ignorer tous bits R.

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 d'entrelacement est L+1, en nombre de blocs de trames de parole.

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 de parole dans cette charge utile dans le groupe d'entrelacement. Si la valeur de ILP se trouve supérieure à ILL, la charge utile DEVRAIT être éliminée.

Les champs ILL et ILP DOIVENT être présents dans chaque paquet d'une session si l'entrelacement est signalé pour la session.

La demande de mode reçue dans le champ CMR est valide jusqu'à la réception de la prochaine CMR, c'est-à-dire, jusqu'à ce que la nouvelle valeur de CMR reçue écrase la précédente. Donc, si un terminal souhaite recevoir continuellement des trames de paroles 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 valide, la CMR DOIT être ignorée par le receveur.

Dans une session multi canaux, la CMR DEVRAIT être interprétée par le receveur de la charge utile comme le mode de codage désiré pour tous les canaux de la session, si le réseau le permet.

Il y a deux facteurs qui affectent le choix de mode VMR-WB : (i) les performances de toute liaison CDMA connectée via une passerelle (par exemple, dans un scénario de GW à terminal IP), et (ii) l'état d'encombrement d'un réseau IP. Les performances de la liaison CDMA sont signalées via le champ CMR, qui n'est pas utilisé par les points d'extrémité IP seulement. L'état du réseau IP est surveillé en utilisant, par exemple, RTCP. Un envoyeur doit choisir le mode de fonctionnement de façon à satisfaire ces deux contraintes (voir la Section 7).

Le codeur DEVRAIT suivre une demande de mode reçue, mais PEUT changer en un mode différent si le réseau l'exige, 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 chez l'envoyeur de parole DEVRAIT ignorer les demandes de mode quand il envoie de la parole à une session de diffusion groupée mais PEUT utiliser les informations de rétroaction RTCP comme indication qu'un changement de mode est nécessaire.

Si l'option d'entrelacement est utilisée, l'entrelacement DOIT être effectué sur la base du bloc de trames de parole, par opposition à sur la base de la trame de parole, 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 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 que le nombre de blocs de trames de parole portés dans chaque charge utile est N. On a alors :

Charge utile s (le premier paquet de ce groupe d'entrelacement) : ILL=L, ILP=0,
 Blocs de trames de parole portés : n, n+(L+1), n+2*(L+1),..., n+(N-1)*(L+1)
 Charge utile s+1 (le second paquet de ce groupe d'entrelacement) : ILL=L, ILP=1,
 Blocs de trames de parole portés : n+1, n+1+(L+1), n+1+2*(L+1),..., n+1+(N-1)*(L+1)
 ...
 Charge utile s+L (dernier paquet de ce groupe d'entrelacement) : ILL=L, ILP=L,

Blocs de trames de parole portés : $n+L$, $n+L+(L+1)$, $n+L+2*(L+1)$, ..., $n+L+(N-1)*(L+1)$

Le prochain groupe d'entrelacement va commencer au bloc de trames de parole $n+N*(L+1)$. Il n'y aura aucun effet d'entrelacement tant que le nombre de blocs de trames de parole par paquet (N) n'est pas au moins 2. De plus, le nombre de blocs de trames de parole par charge utile (N) et la valeur de ILL NE DOIT PAS être changée dans 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'expéditeur de la charge utile DOIT seulement appliquer l'entrelacement 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 MIME "entrelacement=I" pour régler le nombre maximum de blocs de trames de parole permis dans un groupe d'entrelacement à I.

Quand il effectue l'entrelacement, l'expéditeur DOIT utiliser un nombre approprié de blocs de trames de parole 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, $N*(L+1) \leq I$.

L'exemple suivant montre le ToC de trois paquets consécutifs, chacun portant 3 blocs de trames de parole, dans une session entrelacée à deux canaux.

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). Cela fait un groupe d'entrelacement de 9 blocs de trames de parole.

Paquet n° 1

ILL=2, ILP=0 :

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

Paquet n° 2

ILL=2, ILP=1 :

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

Paquet n° 3

ILL=2, ILP=2 :

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

6.3.3 Charge utile Tableau des contenus

Le tableau des contenus (ToC) dans le format de charge utile aligné sur l'octet consiste en entrées d'une liste de ToC où chaque entrée correspond à une trame de parole portée dans la charge utile, c'est-à-dire que quand l'entrelacement est utilisé, les blocs de trames de parole dans le ToC ne vont presque jamais être placés dans des instants consécutifs. À la place, la présence et l'ordre des blocs de trames de parole dans un paquet vont suivre le schéma décrit en 6.3.2.

```
+-----+
| Liste des entrées de ToC |
+-----+
```

Pour le format de charge utile alignée sur l'octet, une entrée de ToC est comme suit :

```

 0 1 2 3 4 5 6 7
+--+--+--+--+--+--+--+
| F |  FT  | Q | P | P |
+--+-----+--+--+--+

```

Le tableau des contenus (ToC) consiste en entrées de liste de ToC, chacune représentant une trame de parole.

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

FT (4 bits) : indice de type de trame dont la valeur est choisie en accord avec le Tableau 3.

Dans le mode interopérable, FT=14 (SPEECH_LOST) et FT=15 (NO_DATA) sont utilisés pour indiquer des trames de paroles qui ne sont respectivement ni perdues ni en cours de transmission dans cette charge utile. FT=14 ou 15 PEUT être utilisé dans les modes non interopérables pour indiquer respectivement une suppression de trame de parole ou une trame de parole blanche (voir au paragraphe 2.1 de [C.S0052-0]).

Si une charge utile avec une valeur de FT invalide est reçue, la charge utile DOIT être éliminée. Noter que pour les entrées de ToC avec FT=14 ou 15, il n'y aura pas de trame de parole correspondante dans la charge utile.

Selon l'application et le mode de fonctionnement de VMR-WB, toute combinaison des types permis de trame de parole (FT) montrés au Tableau 3 PEUT être utilisée.

Q (1 bit) : indicateur de qualité de trame. Réglé à 0, il indique que la trame de parole correspondante est corrompue. En mode interopérable, le côté receveur (avec codec AMR-WB) devrait régler le RX_TYPE à SPEECH_BAD ou SID_BAD selon le type de trame de parole (FT), si Q=0. Le codeur VMR-WB règle toujours le bit Q à 1. Le décodeur VMR-WB peut ignorer le bit Q.

Bits P : les bits de bourrage DOIVENT être réglés à zéro et DOIVENT être ignorés par le receveur.

FT	Taux de codage	Taille de trame (en bits)
0	Plein débit interopérable (AMR-WB 6,60 kbit/s)	132
1	Plein débit interopérable (AMR-WB 8,85 kbit/s)	177
2	Plein débit interopérable (AMR-WB 12,65 kbit/s)	253
3	Plein débit 13,3 kbit/s	266
4	Demi débit 6,2 kbit/s	124
5	Quart de débit 2,7 kbit/s	54
6	Huitième de débit 1,0 kbit/s	20
7	(réservé)	-
8	(réservé)	-
9	CNG (AMR-WB SID)	40
10	(réservé)	-
11	(réservé)	-
12	(réservé)	-
13	(réservé)	-
14	Ecrasement (AMR-WB SPEECH_LOST)	0
15	Blanc (AMR-WB NO_DATA)	0

Tableau 3 : types de charge utile de trames de parole VMR-WB pour transport en temps réel

Pour les sessions multi canaux les entrées de ToC de toutes les trames de paroles d'un bloc de trames de parole sont placées dans le ToC en ordre consécutif. Donc, avec N canaux et K blocs de trames de parole dans un paquet, il DOIT y avoir N*K entrées dans le ToC, et les N premières entrées vont être du premier bloc de trames de parole, les N secondes entrées vont venir du second bloc de trames de parole, et ainsi de suite.

6.3.4 Données de parole

Les données de parole d'une charge utile contiennent une ou plusieurs trames de parole comme décrit dans le ToC de la charge utile.

Chaque trame de parole représente 20 ms de parole codée dans un des taux de codage disponibles selon le mode de fonctionnement. La longueur des trames de parole est définie par le type de trame de parole dans le champ FT, avec les considérations suivantes :

- Le dernier octet de chaque trame de parole DOIT être bourré de zéros à la fin si tous les bits de l'octet ne sont pas utilisés. 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, les trames de parole DOIVENT être arrangées une pleine trame de parole après l'autre.

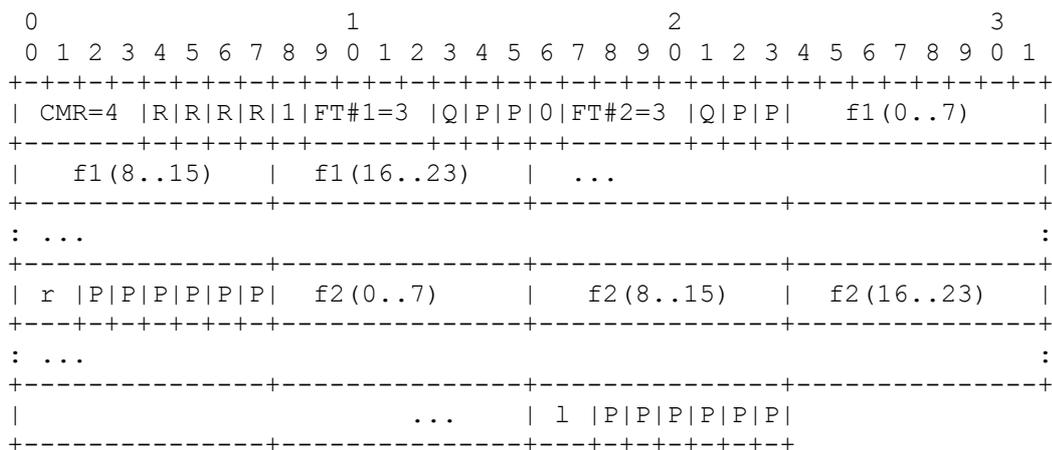
La notation de l'ordre et de la numérotation des bits des données de parole sont comme spécifié dans la norme VMR-WB [C.S0052-0].

La charge utile commence par l'en-tête de charge utile d'un octet, ou deux si l'entrelacement des trames de parole est choisi. L'en-tête de charge utile est suivi par le tableau des contenus consistant en une liste d'entrées de ToC de un octet.

Les données de parole suivent le tableau des contenus. Pour les besoins de la mise en paquets, tous les octets composant une trame de parole sont ajoutés à la charge utile comme un tout. Les trames de parole sont mises en paquet dans le même ordre que leur entrée de ToC correspondante dans la liste du ToC, avec l'exception que si une certaine trame de parole a une entrée de ToC avec FT=14 ou 15, il n'y aura pas d'octet de données présent pour cette trame de parole.

6.3.5 Exemple de charge utile : charge utile de base d'un seul canal portant plusieurs trames

Le diagramme suivant montre un format de charge utile aligné sur l'octet provenant d'une session à un seul canal qui porte deux trames de paroles VMR-WB plein débit (FT=3). Dans la charge utile, une demande de mode codec est envoyée (par exemple, CMR=4) demandant que le codeur du côté receveur utilise le mode VMR-WB 1. Aucun entrelacement n'est utilisé. Noter que dans cet exemple le dernier octet dans les deux trames de parole est bourré avec des zéros pour les aligner sur l'octet.



$r = f1(264,265)$

$l = f2(264,265)$

6.4 Considérations de mise en œuvre

Une application qui met en œuvre ce format de charge utile DOIT comprendre tous les paramètres de charge utile. Toute transposition des paramètres dans un protocole de signalisation DOIT prendre en charge tous les paramètres. Donc, une mise en œuvre de ce format de charge utile dans une application utilisant SDP est obligée de comprendre tous les paramètres de charge utile dans leur forme transposée en SDP. Cette exigence assure qu'une mise en œuvre peut toujours décider si elle est capable de communiquer.

Pour permettre une interconnexion interopérable efficace avec AMR-WB et assurer qu'un terminal VMR-WB se déclare de façon appropriée comme terminal à capacité AMR-WB (voir au paragraphe 9.3) il est aussi RECOMMANDÉ qu'une mise en œuvre de charge utile RTP VMR-WB comprenne la signalisation AMR-WB pertinente.

Pour assurer l'interopérabilité entre diverses mises en œuvre de VMR-WB, les mises en œuvre DEVRONT prendre en charge les deux formats de charge utile sans en-tête et aligné sur l'octet. La prise en charge de l'entrelacement est facultative.

6.4.1 Validation du décodage et dispositions pour les paquets perdus ou retardés

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 de la session et des valeurs trouvées dans les champs d'en-tête de la charge utile, ne correspond pas à la taille du paquet reçu, le receveur DEVRAIT éliminer le paquet pour éviter une potentielle dégradation de la qualité de parole, et invoquer le mécanisme de dissimulation d'erreur de trame de parole incorporé dans VMR-WB. Donc, les paquets invalides DEVRONT être traités comme des paquets perdus.

Les paquets en retard (c'est-à-dire, l'indisponibilité d'un paquet quand il est nécessaire au décodage chez le receveur) devraient être traités comme des paquets perdus. De plus, si le paquet en retard fait partie d'un groupe d'entrelacement, selon la disponibilité des autres paquets dans ce groupe d'entrelacement, le décodage doit être repris à partir de la prochaine trame de parole disponible (en ordre séquentiel). En d'autres termes, l'indisponibilité d'un paquet dans un groupe d'entrelacement à un certain moment ne devrait pas invalider les autres paquets dans ce groupe d'entrelacement, qui peuvent arriver plus tard.

7. Contrôle d'encombrement

Les considérations générales de contrôle d'encombrement pour transporter les données de RTP s'appliquent aussi à la parole VMR-WB sur RTP. Cependant, la capacité multimode du codec de parole VMR-WB peut fournir un avantage sur les autres formats de charge utile pour le contrôle d'encombrement dans la mesure où la demande de bande passante peut être ajustée en choisissant un mode de fonctionnement différent.

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

Si la correction d'erreur directe (FEC) est utilisée pour alléger la perte de paquet, la quantité de frais généraux ajoutée par la FEC va devoir être réglée afin que l'utilisation de la FEC elle-même ne cause pas un problème d'encombrement.

Le contrôle d'encombrement pour RTP DEVRA être utilisé conformément à la [RFC3550] et tout profil RTP applicable, par exemple, de la [RFC3551]. Cela signifie que le contrôle d'encombrement est exigé pour toute transmission sur des réseaux au mieux non gérés.

L'encombrement sur le réseau IP est géré par l'expéditeur IP. Des retours sur l'encombrement DEVRAIENT être fournis à cet expéditeur IP par RTCP ou d'autres moyens, et ensuite l'expéditeur peut choisir d'éviter l'encombrement en utilisant le mécanisme le plus approprié. Cela peut inclure de choisir un mode de fonctionnement approprié, mais aussi d'ajuster le niveau de redondance ou le nombre de trames de paroles par paquet.

8. Considérations sur la sécurité

Les paquets RTP qui utilisent le format de charge utile défini dans la présente spécification sont soumis aux considérations de sécurité générales discutées dans RTP [RFC3550] et dans tout profil applicable comme AVP [RFC3711] ou SAVP [RFC2198].

Comme ce format transporte de l'audio codé, les principales questions de sécurité incluent la confidentialité, la protection de l'intégrité, et l'authentification de l'origine des données audio elles-mêmes. Le format de charge utile lui-même n'a pas de mécanisme de sécurité incorporé. Tout mécanisme externe convenable, comme SRTP [RFC2198], PEUT être utilisé.

Ce format de charge utile et le décodeur VMR-WB ne présentent aucune non uniformité significative dans la complexité de calcul du côté du receveur pour le traitement de paquet ; donc, ils ne posent probablement pas de problème de déni de service dû à la réception de données pathologiques.

8.1 Confidentialité

Afin d'assurer la confidentialité de l'audio codé, tous les bits de données audio DOIVENT être chiffrés. Il est moins besoin de chiffrer l'en-tête de la charge utile ou le tableau des contenus car ils portent seulement des informations sur les types de trames de parole. Ces informations pourraient aussi être utiles à un tiers, par exemple, pour la surveillance de la qualité.

L'utilisation de l'entrelacement en conjonction avec le chiffrement peut avoir un impact négatif sur la confidentialité pour une courte période. Considérons les paquets suivants (entre crochets) contenant des numéros de trames de parole: {10, 14, 18}, {13, 17, 21}, {16, 20, 24} (un schéma typique d'entrelacement diagonal continu). Le générateur souhaite refuser à certains participants la capacité d'écouter le matériel commençant à l'instant 16. Changer simplement la clé sur le paquet avec l'horodatage à ou après 16, et refuser la nouvelle clé à ces participants ne suffit pas ; les trames de paroles 17, 18, et 21 ont été fournies dans les paquets précédents sous la clé antérieure, et la dissimulation d'erreur peut rendre l'audio intelligible au moins jusqu'aux trames de parole 18 ou 19, et éventuellement plus loin.

8.2 Authentification et intégrité

Pour authentifier l'envoyeur de la parole, un mécanisme externe DOIT être utilisé. Il est RECOMMANDÉ qu'un tel mécanisme protège à la fois l'en-tête RTP complet et la charge utile (bits de parole et de données).

L'altération des données par un attaquant interposé pourrait remplacer le contenu audio et aussi résulter en une dépaquetisation/décodage erroné qui pourrait diminuer la qualité audio. Par exemple, l'altération du champ CMR peut résulter en une parole d'une qualité différente de celle désirée.

9. 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 RTP VMR-WB.

Les paramètres sont définis ici au titre de l'enregistrement de sous type MIME pour le codec vocal VMR-WB. Une transposition des paramètres dans le protocole de description de session (SDP) [RFC2327] est aussi fournie pour les applications qui utilisent SDP. Dans les protocoles de contrôle qui n'utilisent pas MIME ou SDP, les paramètres de type de support doivent être transposés dans le format approprié utilisé avec ce protocole de contrôle.

9.1 Enregistrement MIME de charge utile RTP VMR-WB

Le sous type MIME pour le codec audio large bande multimode à débit variable (VMR-WB, *Variable-Rate Multimode Wideband*) est alloué à partir de l'arborescence de l'IETF car VMR-WB est supposé être un codec vocal largement utilisé dans les flux directs multimédia et la messagerie ainsi que dans les applications VoIP. Cet enregistrement MIME ne couvre que les transferts en temps réel via RTP.

Noter que le receveur DOIT ignorer tout paramètre non spécifié et utiliser à la place les valeurs par défaut. Noter aussi que si aucune entrée de paramètre n'est définie, les valeurs par défaut vont être utilisées.

Nom de type de support : audio

Nom de sous type de support : VMR-WB

Paramètres exigés : aucun

De plus, si le paramètre entrelacement est présent, le paramètre "octet-align=1" DOIT aussi être présent.

Paramètres facultatifs :

mode-set : mode de fonctionnement demandé de VMR-WB. Restreint les modes de fonctionnement actifs à un sous ensemble de tous les modes. Les valeurs possibles sont une liste de valeurs d'entiers séparées par des virgules. Actuellement, cette liste inclut les modes 0, 1, 2, et 3 [C.S0052-0], mais elle PEUT être étendue à l'avenir. Si un tel mode-set est spécifié durant l'initiation de session, le codeur NE DOIT PAS utiliser des modes hors du sous ensemble. Si il n'est pas présent, tous les modes de fonctionnement de l'ensemble 0 à 3 sont permis pour la session.

channels : nombre de canaux audio. Les valeurs possibles et leur ordre de canal respectif est spécifié au paragraphe 4.1 de la [RFC3551]. Si il est omis, il a la valeur par défaut de 1.

octet-align : format de charge utile RTP ; les valeurs permises sont 0 et 1. Si c'est 1, le format de charge utile aligné sur l'octet DEVRA être utilisé. Si c'est 0 ou si il n'est pas présent, le format de charge utile sans en-tête est employé (par défaut).

maxptime : voir la [RFC3267]

interleaving : indique que l'entrelacement de niveau bloc de trames de parole DEVRA être utilisé pour la session. Sa valeur définit le nombre maximum de blocs de trames de parole permis dans un groupe d'entrelacement (voir le paragraphe 6.3.1). Si ce paramètre est absent, l'entrelacement NE DEVRA PAS être utilisé. La présence de ce paramètre implique aussi automatiquement que l'opération d'alignement sur l'octet DEVRA être utilisée.

ptime : voir la [RFC2327]. Il DEVRA être d'au moins la taille d'une trame de parole pour VMR-WB.

dtx : les valeurs permises sont 0 et 1. Par défaut 0 (c'est-à-dire, pas de DTX) où VMR-WB opère normalement comme un codec en continu à débit variable. Si dtx=1, le codec VMR-WB va opérer en mode de transmission discontinue où les trames de paroles de descripteur de silence (SID) sont envoyées par le codeur VMR-WB durant les intervalles de silence avec une fréquence de mise à jour ajustable. Le choix d'un taux de mise à jour de SID dépend de la mise en œuvre et d'autres considérations de réseau qui sortent du domaine d'application de la présente spécification.

Considérations de codage : ce type n'est défini que pour le transfert de données codées en VMR-WB via RTP (RFC 3550) en utilisant les formats de charge utile spécifiés à la Section 6 de la RFC 4348.

Considérations de sécurité : voir la Section 8 de la RFC 4348.

Spécification publiée : le codec vocal VMR-WB est spécifié dans les spécifications 3GPP2 C.S0052-0 version 1.0. Les méthodes de transfert sont spécifiées dans la RFC 4348.

Informations supplémentaires :

Adresse de messagerie de la personne à contacter pour plus d'informations : Sassan Ahmadi, sassan.ahmadi@ieee.org

Usage prévu : COMMUN. Il est prévu que de nombreuses applications de VoIP, messagerie multimédia et de flux directs (ainsi que des applications mobiles) vont utiliser ce type.

Auteur/contrôleur des changements : groupe de travail IETF Transport Audio/Vidéo sur délégation de l'IESG

9.2. Transposition des paramètres MIME en SDP

Les informations portées dans la spécification de type de support MIME ont une transposition spécifique en champs du protocole de description de session (SDP) [RFC2327], qui est couramment utilisé pour décrire les sessions RTP. Quand SDP est utilisé pour spécifier les sessions qui emploient le codec VMR-WB, la transposition est la suivante :

- le type de support ("audio") va dans SDP "m=" comme nom de support.
- le sous type de support (nom de format de charge utile) va dans SDP "a=rtpmap" comme nom de codage. Le débit d'horloge RTP dans "a=rtpmap" DOIT être 16000 pour VMR-WB.
- le paramètre "channels" (nombre de canaux) DOIT être soit explicitement réglé à N, soit omis, impliquant une valeur par défaut de 1. Les valeurs de N qui sont permises sont spécifiées au paragraphe 4.1 de la [RFC3551]. Le paramètre "channels", si présent, est spécifié à la suite du sous type MIME et le débit d'horloge RTP comme un paramètre de codage dans l'attribut "a=rtpmap".
- 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 type de support MIME comme une liste de paires de paramètre=valeur séparées par des points-virgules.

Des exemples de descriptions de session SDP utilisant les codages VMR-WB suivent.

Exemple d'usage de VMR-WB dans un possible scénario VoIP (audio large bande) :

```
m=audio 49120 RTP/AVP 98
a=rtpmap:98 VMR-WB/16000
a=fmtp:98 octet-align=1
```

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

```
m=audio 49120 RTP/AVP 99
a=rtpmap:99 VMR-WB/16000/2
a=fmtp:99 octet-align=1; interleaving=30
a=maxptime:100
```

9.3 Considérations sur le modèle offre-réponse

Pour réaliser une bonne interopérabilité pour la charge utile RTP VMR-WB dans une négociation offre-réponse en usage dans SDP [RFC3264], les considérations suivantes sont faites :

- Les paramètres rate, channel, et configuration de charge utile (alignement sur l'octet et entrelacement) DEVRONT être utilisés symétriquement, c'est-à-dire que l'offre et la réponse doivent utiliser les mêmes valeurs. La taille maximum de la mémoire tampon d'entrelacement est, cependant, déclarative, et chaque agent spécifie la valeur qu'il accepte de recevoir pour les flux recvonly (*réception seule*) et sendrecv (*envoi-réception*). Pour les flux en envoi seul, la valeur indique ce que l'agent désire utiliser.
- Pour conserver l'interopérabilité parmi toutes les mises en œuvre de VMR-WB qui peuvent ou non prendre en charge tous les modes de fonctionnement de codec, les modes de fonctionnement qui sont pris en charge par une mise en œuvre PEUVENT être identifiés à l'initiation de la session. Le paramètre mode-set est déclaratif, et seuls les modes de fonctionnement qui ont été indiqués comme pris en charge par les deux extrémités DEVRONT être utilisés. Si celui qui répond ne prend pas en charge un des modes de fonctionnement fournis dans l'offre, la déclaration complète de type de charge utile DEVRAIT être rejetée en la retirant de la réponse.
- Les paramètres restants sont tous déclaratifs ; c'est-à-dire, pour un flux en envoi seul ils fournissent des paramètres que l'agent désire utiliser, tandis que pour les flux recvonly et sendrecv ils déclarent les paramètres qu'il accepte de recevoir. Le paramètre dtx est utilisé pour indiquer la prise en charge et la capacité DTX, alors qu'il est seulement RECOMMANDÉ à l'expéditeur du support d'envoyer en utilisant le DTX dans ces cas. Si DTX n'est pas pris en charge par l'expéditeur du support, il va envoyer le support sans DTX ; cela n'affectera pas l'interopérabilité mais seulement la consommation de ressource.
- Les deux configurations de format de charge utile sans en-tête et alignée sur l'octet PEUVENT être offertes par un terminal à capacité VMR-WB. Cependant, pour une interconnexion interopérable avec AMR-WB, seulement la configuration "aligné sur l'octet" sera offerte.
- Les paramètres "maxptime" et "ptime" devraient dans la plupart des cas ne pas affecter l'interopérabilité ; cependant, le réglage de ces paramètres peut affecter les performances de l'application.
- Pour conserver l'interopérabilité avec AMR-WB dans le cas où la négociation est possible en utilisant le mode interopérable VMR-WB, un terminal à capacité VMR-WB DEVRAIT aussi se déclarer capable de AMR-WB avec l'ensemble de modes limité (c'est-à-dire, seulement les modes de codec AMR-WB 0, 1, et 2 sont permis) et le mode de fonctionnement aligné sur l'octet.

Exemple :

```
m=audio 49120 RTP/AVP 98 99
a=rtpmap:98 VMR-WB/16000
a=rtpmap:99 AMR-WB/16000
a=fmtp:99 octet-align=1; mode-set=0,1,2
```

Un exemple d'échange d'offre-réponse pour le scénario VoIP décrit au paragraphe 5.3 est le suivant :

Offre de terminal CDMA2000 -> terminal WCDMA :

```
m=audio 49120 RTP/AVP 98 97
a=rtpmap:98 VMR-WB/16000
```

```
a=fmtp:98 octet-align=1
a=rtpmap:97 AMR-WB/16000
a=fmtp:97 mode-set=0,1,2; octet-align=1
```

Réponse de terminal WCDMA -> terminal CDMA2000 :

```
m=audio 49120 RTP/AVP 97
a=rtpmap:97 AMR-WB/16000
a=fmtp:97 mode-set=0,1,2; octet-align=1;
```

Pour l'usage déclaratif de SDP comme dans SAP [RFC2974] et RTSP [RFC2326], tous les paramètres sont déclaratifs et donnent les paramètres qui DEVRONT être utilisés pour recevoir et/ou envoyer le flux configuré.

10. Considérations relatives à l'IANA

L'IANA a enregistré un nouveau sous type MIME (audio/VMR-WB) ; voir la Section 9.

11. Remerciements

L'auteur remercie Redwan Salami de VoiceAge Corporation, Ari Lakaniemi de Nokia Inc., et les présidents de l'IETF/AVT Colin Perkins et Magnus Westerlund de leurs commentaires techniques qui ont amélioré ce document.

L'auteur reconnaît aussi que certaines parties des [RFC3267] et [RFC3558] ont été utilisées dans le présent document.

12. Références

12.1 Références normatives

- [C.S0052-0] Spécification technique 3GPP2 C.S0052-0 v1.0 "Source-Controlled Variable-Rate Multimode Wideband Speech Codec (VMR-WB) Service Option 62 for SpreadSpectrum Systems", juillet 2004.
- [RFC2119] S. Bradner, "[Mots clés à utiliser](#) dans les RFC pour indiquer les niveaux d'exigence", BCP 14, mars 1997. (MàJ par [RFC8174](#))
- [RFC2327] M. Handley et V. Jacobson, "SDP : [Protocole de description de session](#)", avril 1998. (*Obsolète; voir [RFC4566](#)*)
- [RFC3267] J. Sjöberg et autres, "Format de charge utile et format de mémorisation de fichier pour les codecs audio AMR et AMR-WB dans RTP", juin 2002. (*Obsolète, voir [RFC4867](#)*) (P.S.)
- [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](#))

12.2 Références pour information

- [C.S0050-A] Spécification technique 3GPP2 C.S0050-A v1.0 "3GPP2 File Formats for Multimedia Services", septembre 2005.
- [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*)
- [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](#)*)
- [RFC3558] A. Li, "Format de charge utile RTP pour les codecs à débit variable améliorés (EVRC) et les vocodeurs à mode sélectif (SMV)", juillet 2003. (*MàJ par [RFC478 8](#)*) (P.S.)
- [RFC3711] M. Baugher et autres, "Protocole de [transport sécurisé en temps réel](#) (SRTP)", mars 2004. (P.S.)
- [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).

Tous les documents du 3GPP2 peuvent être téléchargés du serveur du 3GPP2 à "<http://www.3gpp2.org/> ", voir les spécifications.

Adresse de l'auteur

Dr. Sassan Ahmadi
mél : sassan.ahmadi@ieee.org

Déclaration complète de droits de reproduction

Copyright (C) The IETF Trust (2006).

Le présent document est soumis aux droits, licences et restrictions contenus dans le BCP 78, et à www.rfc-editor.org, 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 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 pourrait 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 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 fourni par l'activité de soutien administratif (IASA) de l'IETF.