

Groupe de travail Réseau
Request for Comments : 4352
 Catégorie : Sur la voie de la normalisation

J. Sjoberg, Ericsson
 M. Westerlund, Ericsson
 A. Lakaniemi, Nokia
 S. Wenger, Nokia
 janvier 2006

Traduction Claude Brière de L'Isle

Format de charge utile RTP pour le codec audio large bande multi débits adaptatif étendu (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 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*) pour les signaux audio en large bande multi débits adaptatif étendu (AMR-WB+, *Extended Adaptive Multi-Rate Wideband*). Le codec AMR-WB+ est une extension audio du codec vocal AMR-WB. Il englobe les types de trame AMR-WB et un certain nombre de nouveaux types de trame conçus pour prendre en charge la musique et la parole de haute qualité. Un enregistrement de type de support pour AMR-WB+ est inclus dans la présente spécification.

Table des matières

1. Introduction.....	2
2. Définitions.....	2
2.1 Glossaire.....	2
2.2 Terminologie.....	3
3. Fondements de AMR-WB+ et principes de conception.....	3
3.1 Codec audio AMR-WB+.....	3
3.2 Codage multi débits et adaptation de débit.....	5
3.3 Détection d'activité vocale et transmission discontinue.....	5
3.4 Prise en charge de session multi canaux.....	5
3.5 Détection et protection d'erreur binaire inégale.....	5
3.6 Robustesse contre la perte de paquet.....	5
3.7 Audio AMR-WB+ sur scénarios IP.....	7
3.8 Signalisation hors bande.....	7
4. Format de charge utile RTP pour AMR-WB+.....	7
4.1 Usage de l'en-tête RTP.....	8
4.2 Structure de charge utile.....	8
4.3 Définitions de charge utile.....	9
4.4 Considérations d'entrelacement.....	15
4.5 Considérations de mise en œuvre.....	15
5. Contrôle d'encombrement.....	17
6. Considérations sur la sécurité.....	17
6.1 Confidentialité.....	17
6.2 Authentification et intégrité.....	18
7. Paramètres de format de charge utile.....	18
7.1 Enregistrement de type de support.....	18
7.2 Transposition des paramètres de type de support en SDP.....	19
8. Considérations relatives à l'IANA.....	20
9. Contributeurs.....	20
10. Remerciements.....	20

11. Références.....	21
11.1 Références normatives.....	21
11.2 Références pour information.....	21
Adresse des auteurs.....	22
Déclaration complète de droits de reproduction.....	22

1. Introduction

Le présent document spécifie le format de charge utile pour la mise en paquets des signaux audio codés en large bande multi débits adaptatif étendu [AMR-WB+] dans le protocole de transport en temps réel (RTP) [RFC3550]. Le format de charge utile prend en charge la transmission d'audio mono ou stéréo, agrégeant plusieurs trames par charge utile, et des mécanismes qui améliorent la robustesse du flux de paquets contre les pertes de paquet.

Le codec AMR-WB+ est une extension du codec vocal large bande multi débits adaptatif (AMR-WB, *Adaptive Multi-Rate Wideband*). Les nouvelles caractéristiques incluent une bande passante audio étendue pour permettre une haute qualité pour les signaux non vocaux (par exemple, la musique) la prise en charge native de l'audio stéréophonique, et l'option de fonctionner, et commuter, plusieurs fréquences internes d'échantillonnage (ISF, *Internal Sampling Frequency*). Le principal scénario d'usage pour AMR-WB+ est le transport sur IP. Donc, l'interfonctionnement avec les autres réseaux de transport, comme discuté pour AMR-WB dans la [RFC3267], n'est pas un souci majeur et n'est donc pas traité dans ce mémoire.

L'application clé prévue pour AMR-WB+ est l'écoulement de flux en direct. Pour rendre le processus de mise en paquets sur un serveur de flux en direct aussi efficace que possible, un format de charge utile aligné sur l'octet est désirable. Donc, un mode efficace en bande passante (comme défini pour AMR-WB dans la [RFC3267]) n'est pas spécifié ici ; les économies de bande passante du mode efficace en bande passante seraient de toutes façons très faibles, car tous les types de trames d'extension sont alignés sur l'octet.

La capacité de codage stéréo de AMR-WB+ rend obsolète la prise en charge du transport multi canaux au niveau du format de charge utile RTP, comme spécifié pour AMR-WB [RFC3267]. Donc, cette caractéristique n'est pas incluse dans le présent mémoire.

Cette spécification n'inclut pas de définition d'un format de fichier pour AMR-WB+. À la place, elle se réfère au format de fichier 3GP fondé sur l'ISO [TS26.244], qui prend en charge AMR-WB+ et fournit toutes les fonctionnalités requises. Le format 3GP prend aussi en charge la mémorisation de AMR, AMR-WB, et nombreux autres formats multimédia, permettant par là un playback synchronisé.

Le reste du document est organisé comme suit : on trouvera dans la Section 3 les informations de base sur le codec AMR-WB+, et les principes de conception. Le format de charge utile est lui-même spécifié à la Section 4. Les Sections 5 et 6 discutent respectivement du contrôle d'encombrement et des considérations sur la sécurité. Dans la Section 7, on fournit un enregistrement de type de support.

2. Définitions

2.1 Glossaire

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 large bande multi débits adaptatif
 AMR-WB+ (*Extended Adaptive Multi-Rate Wideband (Codec)*) codec large bande multi débits adaptatif étendu
 CN (*Comfort Noise*) bruit de fond
 DTX (*Discontinuous Transmission*) transmission discontinue
 FEC (*Forward Error Correction*) correction d'erreur directe (CED)
 FT (*type de trame*) type de trame
 ISF (*Internal Sampling Frequency*) fréquence d'échantillonnage interne
 SCR (*Source-Controlled Rate Operation*) fonctionnement à débit contrôlé par la source
 SID (*Silence Indicator*) indicateur de silence (trames contenant seulement des paramètres de CN)
 TFI (*Transport Frame Index*) indice de trame de transport
 TS (*Timestamp*) horodatage
 VAD (*Voice Activity Detection*) détection d'activité vocale

UED (*Unequal Error Detection*) détection d'erreur inégale

UEP (*Unequal Error Protection*) protection d'erreur inégale

2.2 Terminologie

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].

3. Fondements de AMR-WB+ et principes de conception

Le codec audio large bande multi débits adaptatif étendu [AMR-WB+] est conçu pour compresser les signaux vocaux et audio à bas débit et de bonne qualité. Le codec est spécifié par le projet en partenariat de troisième génération (3GPP, *Third Generation Partnership Project*). Les principales applications ciblées sont 1) le service de paquets commutés en flux directs (PSS, *packet-switched streaming*) [TS26.233], 2) le service de messagerie multimédia (MMS, *multimedia messaging service*) [TS26.140], et 3) le service de diffusion et de diffusion groupée multimédia (MBMS, *multimedia broadcast and multicast service*) [MBMS]. Cependant, du fait de sa souplesse et sa robustesse, AMR-WB+ convient bien aussi pour les services en flux directs dans d'autres environnements de transport très variés, par exemple, l'Internet.

3.1 Codec audio AMR-WB+

À l'origine, le 3GPP a développé le codec audio AMR-WB+ pour les services de flux directs et de messagerie dans le système mondial pour les communications mobiles (GSM, *Global System for Mobile communications*) et les systèmes cellulaires de troisième génération (3G). Le codec est conçu comme une extension audio du codec vocal AMR-WB. L'extension ajoute de nouvelles fonctionnalités au codec afin de fournir une qualité audio élevée pour une large gamme de signaux incluant la musique. Le fonctionnement stéréophonique a aussi été ajouté. Un nouvel algorithme de codage stéréo hybride de grande efficacité permet un fonctionnement en stéréo à des débits binaires aussi faibles que 6,2 kbit/s.

Le codec AMR-WB+ inclut les neuf types de trames spécifiés pour AMR-WB, étendus par de nouveaux débits binaires allant de 5,2 à 48 kbit/s. Les types de trame AMR-WB peuvent employer seulement une fréquence d'échantillonnage de 16 000 Hz et fonctionnent seulement sur des signaux monophoniques. Les nouvelles extensions de type de trame introduites peuvent cependant opérer à diverses fréquences d'échantillonnage internes (ISF, *Internal Sampling Frequency*) aussi bien en mono qu'en stéréo. Voir les détails au Tableau 24 de [AMR-WB+]. La fréquence d'échantillonnage de sortie du décodeur est limitée à 8, 16, 24, 32, ou 48 kHz.

Les opérations générales de codage de AMR-WB+ sont les suivantes. Le codeur reçoit l'audio échantillonné à, par exemple, 48 kHz. Le processus de codage commence avec le pré traitement et le rééchantillonnage à l'ISF choisie par l'utilisateur. Le codage est effectué sur des super-trames de taille égale. Chaque super-trame correspond à 2048 échantillons par canal, à l'ISF. Le codec prend un certain nombre de décisions de codage pour chaque super trame, choisissant par là entre les différents algorithmes de codage et longueurs de blocs, afin de réaliser un codage à fidélité optimale adapté aux caractéristiques de signal de la source. Le codage stéréo (si utilisé) s'exécute séparément du codage du cœur monophonique, permettant donc la sélection de différentes combinaisons de taux de codage de cœur et stéréo. L'audio codé résultant est produit dans quatre trames de transport de longueur égale. Chaque trame de transport correspond à 512 échantillons à l'ISF et est individuellement utilisable par le décodeur, pourvu que sa position dans la structure de la super trame soit connue.

Le codec prend en charge 13 ISF différentes, allant de 12,8 à 38,4 kHz, comme décrit dans le Tableau 24 de [AMR-WB+]. Le grand nombre d'ISF permet un compromis entre la bande passante audio et le débit binaire de la cible. Comme le codage est effectué sur 2048 échantillons à l'ISF, la durée d'une super trame et le débit binaire effectif du type de trame utilisé varie.

L'ISF de 25600 Hz a une durée de super trame de 80 ms. C'est la valeur "nominale" utilisée pour décrire dorénavant les débits binaires de codage. En supposant cette normalisation, le choix de l'ISF résulte en des variations de débit binaire de 1/2 à 3/2 du débit binaire nominal.

Le codage pour les modes d'extension est effectué comme un codage de cœur monophonique et un codage stéréo. Le codage de cœur est exécuté en partageant le signal monophonique entre une bande de fréquence inférieure et une bande de fréquence supérieure. La bande inférieure est codée en employant soit une prédiction linéaire avec excitation par séquences

codées à structure algébrique (ACELP, *algebraic code excited linear prediction*) soit une excitation codée à transformation (TCX, *transform coded excitation*). Ce choix peut être fait une fois par trame de transport, mais doit obéir à certaines limitations de combinaisons légales au sein de la super trame. La bande supérieure est codée en utilisant une approche d'extension de bande passante paramétrique à bas débit.

Le signal stéréo est codé en employant une décomposition similaire de bande de fréquence ; cependant, le signal est divisé en trois bandes qui sont paramétrées individuellement.

Le débit binaire total produit par l'extension est le résultat de la combinaison du débit de cœur du codeur, du débit stéréo, et de l'ISF. L'extension prend en charge 8 débits de codage de cœur différents, produisant des débits binaires entre 10,4 et 24,0 kbit/s ; voir le Tableau 22 de [AMR-WB+]. Il y a 16 débits de codage stéréo qui génèrent des débits binaires entre 2,0 et 8,0 kbit/s ; voir le Tableau 23 de [AMR-WB+]. Le type de trame identifie de façon univoque les modes AMR-WB, 4 débits d'extension fixes (voir ci-dessous), 24 combinaisons de débits de cœur et stéréo pour les signaux stéréo, et les 8 débits de cœur pour les signaux mono, comme indiqué dans la liste du Tableau 25 de [AMR-WB+]. Cela implique que AMR-WB+ prend en charge des débits de codage entre 10,4 et 32 kbit/s, en supposant une ISF de 25 600 Hz.

Des ISF différentes permettent une liberté supplémentaire des débits binaires produits et de la qualité audio. Le choix d'une ISF change la bande passante audio disponible du signal reconstruit, et aussi le débit binaire total. Le débit binaire pour une combinaison donnée de type de trame et d'ISF est déterminé en multipliant le débit binaire du type de trame par le facteur de débit binaire de l'ISF utilisée ; voir le Tableau 24 de [AMR-WB+].

L'extension a aussi quatre types de trame qui ont des ISF fixées. Voir les types de trame 10 à 13 dans le Tableau 21 de [AMR-WB+]. Ces quatre types de trames prédéfinis ont une fréquence d'échantillonnage d'entrée fixée au codeur, qui peut être réglée à 16 ou 24 kHz. Comme les types de trame AMR-WB, les trames de transport codées en utilisant ces types de trame représentent exactement 20 ms du signal audio. Cependant, elles font aussi partie de super trames de 80 ms. Les types de trame 0 à 13 (AMR-WB et taux d'extension fixes) figurant dans le Tableau 21 de [AMR-WB+], n'exigent pas une indication d'ISF explicite. Les autres types de trame, 14 à 47, exigent que l'ISF employée soit indiquée.

Les 32 différents types de trame de l'extension, en combinaison avec les 13 ISF, permettent une grande souplesse de débit binaire et de choix de la qualité audio désirée. Il existe un certain nombre de combinaisons qui produisent le même débit binaire du codec. Par exemple, un flux audio de 32 kbit/s peut être produit en utilisant le type de trame 41 (c'est-à-dire, 25,6 kbit/s) et l'ISF de 32 kHz ($5/4 * (19,2 + 6,4) = 32$ kbit/s) ou le type de trame 47 et l'ISF de 25,6 kHz ($1 * (24 + 8) = 32$ kbit/s). La combinaison qui est la plus bénéfique pour la qualité audio perçue dépend du contenu. Dans l'exemple ci-dessus, le premier cas fournit une bande passante audio plus élevée, tandis que le second dépense le même nombre de bits sur une bande passante audio un peu plus étroite mais fournit une fidélité plus élevée. Les codeurs sont libres de choisir la combinaison qu'ils estiment la meilleure.

Comme une trame de transport correspond toujours à 512 échantillons à l'ISF utilisée, sa durée est limitée à la gamme 13,33 à 40 ms ; voir le Tableau 1. Un débit d'horloge d'horodatage RTP de 72 000 Hz, comme exigé par la présente spécification, résulte en des longueurs de trame de transport AMR-WB+ de 960 à 2880 tics d'horodatage, qui dépend seulement de l'ISF choisie.

Indice	ISF	Durée (ms)	Durée (tics TS à 72 kHz)
0	N/A	20	1440
1	12800	40	2880
2	14400	35,55	2560
3	16000	32	2304
4	17067	30	2160
5	19200	26,67	1920
6	21333	24	1728
7	24000	21,33	1536
8	25600	20	1440
9	28800	17,78	1280
10	32000	16	1152
11	34133	15	1080
12	36000	14,22	1024
13	38400	13,33	960

Tableau 1 : Nombre normatif de tics d'horodatage RTP pour chaque trame de transport selon l'ISF (ISF et durée sont arrondies en ms)

Le codeur est libre de changer l'ISF et le type de codage de trame (mono et stéréo) durant une session. Pour les types de trame d'extension avec les indices 10-13 et 16-47, les changements d'ISF et de type de trame sont obligés d'intervenir à des limites de super trame. Cela implique que, pour les types de trame mentionnés, l'ISF est constante sur toute une super trame. Cette limitation ne s'applique pas pour les types de trame aux indices 0 à 9, 14, et 15; c'est-à-dire, les types originaux de trame AMR-WB.

Un certain nombre de caractéristiques du codec AMR-WB+ exigent une attention particulière du point de vue du transport, et de solutions qui pourraient peut-être être vues comme non orthodoxes. D'abord, il y a des contraintes sur l'horodatage RTP, à cause de la relation entre la durée de trame et les ISF. Ensuite, chaque trame d'audio codé doit conserver des informations sur son type de trame, l'ISF, et la position dans la super trame.

3.2 Codage multi débits et adaptation de débit

La capacité de codage multi débits de AMR-WB+ est conçue pour préserver une haute qualité audio dans une large gamme d'exigences de bande passante et de conditions de transmission.

AMR-WB+ permet une commutation continue entre des types de trames qui utilisent le même nombre de canaux audio et la même ISF. Il est exigé de chaque mise en œuvre de codec AMR-WB+ qu'elle prenne en charge tous les types de trames définis par le codec et elle doit être capable de traiter la commutation entre toutes paires de type de trames. La commutation entre des types de trames employant un nombre différent de canaux audio ou une ISF différente doit aussi être prise en charge, mais elle peut n'être pas complètement continue. Donc, il est recommandé d'effectuer rarement de telles commutations et, si possible, durant des périodes de silence.

3.3 Détection d'activité vocale et transmission discontinue

AMR-WB+ prend en charge les mêmes algorithmes que AMR-WB pour les paramètres de détection d'activité vocale (VAD, *voice activity detection*) et de génération de bruit de fond (CN, *comfort noise*) durant les périodes de silence. Cependant, ces fonctionnalités ne peuvent seulement être utilisées qu'en conjonction avec le type de trame AMR-WB (FT=0-8). Cette option permet de réduire à un minimum le nombre de bits et 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 appelée la transmission discontinue (DTX, *discontinuous transmission*) ou le fonctionnement à débit de source contrôlé (SCR, *source controlled rate*). Les trames AMR-WB+ qui contiennent des paramètres de CN sont appelées des trames d'indicateur de silence (SID, *silence indicator*). Des détails sur les fonctions de VAD et DTX sont fournies dans [TS26.192] et [TS26.193].

3.4 Prise en charge de session multi canaux

Certains types de trame AMR-WB+ prennent en charge le codage d'audio stéréophonique. À cause de cette prise en charge native d'un signal à deux canaux stéréophoniques, il ne semble pas nécessaire de prendre en charge le transport multi canaux avec des instances de codec séparées, comme spécifié dans la charge utile RTP AMR-WB [RFC3267]. Le codec a la capacité de dégradation de la stéréo en mono au titre du processus de décodage. Donc, un receveur qui est seulement capable de jouer de l'audio monophonique doit quand même être capable de décoder et exécuter des signaux originellement codés et transmis en stéréo. Cependant, pour éviter de gaspiller des bits à du codage stéréo qui ne va pas être utilisé, un mécanisme est défini dans la présente spécification pour signaler l'audio en mono seul.

3.5 Détection et protection d'erreur binaire inégale

Les bits audio codés dans chaque trame AMR-WB sont triés en accord avec leurs différentes sensibilités perceptibles aux erreurs binaires. Dans les systèmes cellulaires, par exemple, cette propriété peut être exploitée pour obtenir une meilleure qualité vocale, en utilisant les mécanismes de détection et de protection d'erreur inégale (UEP et UED). Cependant, les bits des types de trame d'extension du codec AMR-WB+ n'ont pas une propriété cohérente de signification perceptive et ne sont pas triés dans cet ordre. Donc, l'UEP ou la UED n'ont pas de sens avec les types de trame d'extension. Si il est besoin d'utiliser UEP ou UED pour les types de trame AMR-WB, il est recommandé que la [RFC3267] soit utilisée.

3.6 Robustesse contre la perte de paquet

Le format de charge utile prend en charge deux mécanismes pour améliorer la robustesse contre la perte de paquet : la simple correction d'erreur directe (FEC, *forward error correction*) et l'entrelacement de trames.

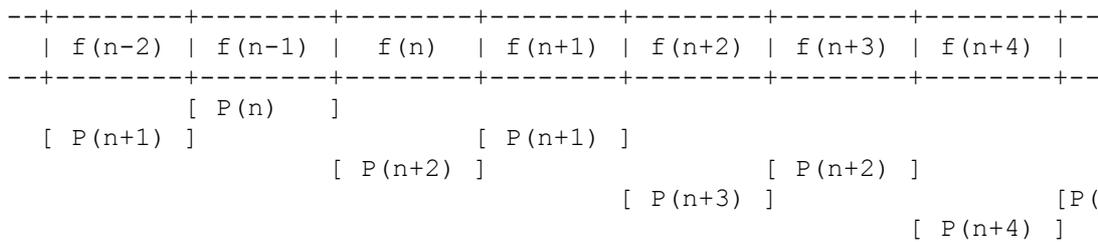


Figure 2 : Exemple de schéma d'entrelacement avec délai constant

Dans la Figure 2, les trames consécutives, notées $f(n-2)$ à $f(n+4)$ sont agrégées en paquets $P(n)$ à $P(n+4)$ chaque paquet portant deux trames. Cette approche fournit un schéma d'entrelacement qui permet un délai constant dans les deux processus d'entrelacement et de désentrelacement. La mémoire tampon de désentrelacement a besoin de place pour au moins trois trames, incluant celle qui est prête à être consommée. L'espace de mémorisation pour trois trames est nécessaire, par exemple, quand $f(n)$ est la prochaine trame à être décodée : comme la trame $f(n)$ a été reçue dans le paquet $P(n+2)$, qui porte aussi la trame $f(n+3)$, ces deux trames sont mémorisées dans la mémoire tampon. De plus, la trame $f(n+1)$ reçue dans le paquet précédent, $P(n+1)$, est aussi dans la mémoire tampon de désentrelacement. Noter aussi que dans cet exemple l'occupation de la mémoire tampon varie : quand la trame $f(n+1)$ est la prochaine à décoder, il y a seulement deux trames, $f(n+1)$ et $f(n+3)$, dans la mémoire tampon.

3.7 Audio AMR-WB+ sur scénarios IP

Comme la principale application cible pour le codec AMR-WB+ est de flux de direct sur des réseaux de paquets, le scénario d'usage le plus pertinent pour ce format de charge utile est IP de bout en bout entre un serveur et un terminal, comme le montre la Figure 3.

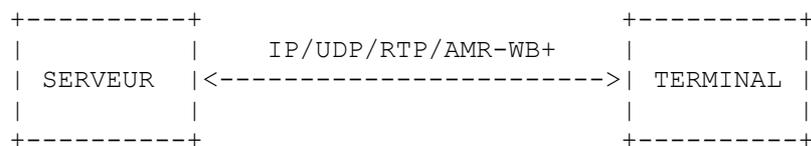


Figure 3 : Scénario de serveur à terminal IP

3.8 Signalisation hors bande

Certaines des options de ce format de charge utile restent constantes à travers une session. Donc, elles peuvent être contrôlées/négociées à l'établissement de session. Tout au long de la présente spécification, ces options et variables sont notées comme des "paramètres à établir par des moyens hors bande". Dans la Section 7, tous les paramètres sont formellement spécifiés sous la forme d'un enregistrement de type de support pour le codage AMR-WB+. La méthode utilisée pour signaler ces paramètres à l'établissement de session ou pour arranger un accord préalable des participants sort du domaine d'application de ce document ; cependant, le paragraphe 7.2 fournit une transposition des paramètres dans le protocole de description de session (SDP, *Session Description Protocol*) [RFC2327] pour les applications qui utilisent SDP.

4. Format de charge utile RTP pour AMR-WB+

La conception de la charge utile pour AMR-WB+ met principalement l'accent sur la minimisation des frais généraux dans les cas d'utilisation normaux, tout en fournissant une complète souplesse avec des frais généraux légèrement supérieurs. Afin de garder la spécification raisonnablement simple, on s'est abstenu de définir des paramètres spécifiques de trames pour chaque type de trame. Quelques paramètres communs ont plutôt été spécifié pour couvrir tous les types de trames.

Le format de charge utile a deux modes : mode de base et mode entrelacé. Les principales différences structurelles entre les deux modes sont l'extension des entrées du tableau des contenus avec des champs de déplacement de trame quand on fonctionne en mode entrelacé. Le mode de base prend en charge l'agrégation de plusieurs trames consécutives dans une charge utile. Le mode entrelacé prend en charge l'agrégation de plusieurs trames qui sont non consécutives dans le temps. Dans les deux modes, il est possible d'avoir des trames codées avec différents types de trame dans la même charge utile. L'ISF doit rester constante dans la charge utile d'un seul paquet.

Le format de charge utile est conçu autour de la propriété des trames AMR-WB+ qu'elles sont consécutives dans le temps et partagent la même durée de trame (en l'absence de changement d'ISF). Cela permet au receveur de déduire l'horodatage pour une trame individuelle dans une charge utile. En mode de base, le processus de déduction se fonde sur l'ordre des trames. En mode entrelacé, il se fonde sur les champs de déplacement compacts. Les horodatages de trame sont utilisés pour régénérer l'ordre correct des trames après réception, identifier les dupliqués, et détecter les trames perdues qui exigent la dissimulation.

Le schéma d'entrelacement de ce format de charge utile est significativement plus souple que celui spécifié dans la RFC 3267. Le format de charge utile AMR et AMR-WB est seulement capable d'utiliser des schémas périodiques avec des trames prises dans un groupe entrelacé à des intervalles fixes. Le schéma d'entrelacement de la présente spécification permet quant à lui tout schéma d'entrelacement, pour autant que la distance dans l'ordre de décodage entre deux trames adjacentes ne soit pas de plus de 256 trames. Noter que même à la plus haute ISF cela permet une profondeur d'entrelacement de jusqu'à 3,41 secondes.

Pour permettre la résilience aux erreurs par la redondance de transmission, les périodes couvertes par plusieurs paquets PEUVENT se chevaucher. Un receveur DOIT être prêt à recevoir toute trame audio plusieurs fois. Toutes les trames envoyées de façon redondante DOIVENT utiliser le même type de trame et d'ISF, et DOIVENT avoir le même horodatage RTP, ou DOIVENT être une trame NO_DATA (FT=15).

La charge utile consiste en éléments alignés sur l'octet (en-tête, ToC, et trames audio). Seules les trames audio de type AMR-WB (0 à 9) requièrent un bourrage pour l'alignement sur l'octet. Si un bourrage supplémentaire est désiré, alors le bit P dans l'en-tête RTP PEUT être établi, et un bourrage PEUT être ajouté comme spécifiée dans la [RFC3550].

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 cette spécification.

L'horodatage RTP correspond à l'instant d'échantillonnage du premier échantillon codé pour la première trame dans le paquet. La fréquence de l'horloge d'horodatage DEVRA être 72 000 Hz. Cette fréquence permet que la durée de trame soit un nombre entier de tics d'horodatage RTP pour les ISF spécifiés au Tableau 1. Elle fournit aussi des facteurs de conversion raisonnables aux fréquences d'échantillonnage audio d'entrée/sortie prises en charge par le codec. Voir au paragraphe 4.3.2.3 des indications sur la façon de déduire l'horodatage RTP pour toute trame audio après la première.

Le bit marqueur d'en-tête RTP (M) DEVRA être réglé à 1 chaque fois que la première trame portée dans le paquet est la première trame dans un jet de parole (voir la définition d'un jet de parole au paragraphe 4.1 de la [RFC3551]). 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 le format défini dans le présent mémoire sort du domaine d'application de ce document. Le profil RTP utilisé alloue un type de charge utile statique ou rend obligatoire de lier de façon dynamique le type de charge utile.

Le paramètre de type de support "canaux" est utilisé pour indiquer le nombre maximum de canaux permis pour un certain type de charge utile. Un type de charge utile où "canaux=1" (mono) DEVRA seulement porter du contenu mono. Un type de charge utile pour lequel "canaux=2" a été déclaré PEUT porter du contenu à la fois mono et stéréo. Noter que cette définition est différente de celle de la [RFC3551]. Comme mentionné précédemment, le codec AMR-WB+ traite la prise en charge de contenu stéréo et le mixage (éventuel) de stéréo en mono en interne. Cela rend inutile de négocier le nombre de canaux pour des raisons autres que l'efficacité du débit binaire.

4.2 Structure de charge utile

La charge utile consiste en un en-tête de charge utile, un tableau de contenu, et les données audio qui représentent une ou plusieurs trames audio. Le diagramme qui suit montre la présentation générale du format de charge utile :

```
+-----+-----+-----+
|En-tête charge utile| Tableau de contenus | Données audio...
+-----+-----+-----+
```

Les charges utiles qui contiennent plus d'une trame audio sont appelées des charges utiles composées.

Les paragraphes qui suivent décrivent les variantes du format de charge utile selon le mode utilisé : mode de base ou mode entrelacé.

4.3 Définitions de charge utile

4.3.1 En-tête de charge utile

L'en-tête de charge utile porte des données communes pour toutes les trames de la charge utile. La structure de l'en-tête de charge utile est décrite ci-dessous.

```

 0 1 2 3 4 5 6 7
+-----+-----+-----+-----+
|   ISF   | TFI | L |
+-----+-----+-----+

```

ISF (5 bits) : Indique la fréquence d'échantillonnage interne employée pour toutes les trames dans cette charge utile. La valeur d'indice correspond à la fréquence d'échantillonnage interne spécifiée au Tableau 24 de [AMR-WB+]. Ce champ DEVRA être réglé à 0 pour les charges utiles qui contiennent des trames avec des valeurs de type de trame de 0 à 13.

TFI (*Transport Frame Index*) (2 bits) : indice de trame de transport, de 0 (premier) à 3 (dernier), indiquant la position de la première trame de transport de cette charge utile dans la structure de super-trame AMR-WB+. Pour les charges utiles avec des trames de seulement des valeurs de type de trame de 0 à 9, ce champ DEVRA être réglé à 0 par l'expéditeur. La valeur du TFI pour une trame de type 0 à 9 DEVRA être ignorée par le receveur. Noter que le type de trame est codé dans le tableau de contenus (comme discuté plus loin) ; donc, la dépendance mentionnée du type de trame peut être appliquée facilement en interprétant seulement les valeurs portées dans l'en-tête de charge utile. Il n'est pas nécessaire d'interpréter le flux binaire audio lui-même.

L (1 bit) : fanion du champ Long déplacement pour les charges utiles en mode entrelacé. Si il est réglé à 0, des champs de déplacement de quatre bits sont utilisés pour indiquer le décalage d'entrelacement ; si il est à 1, des champs de déplacement de huit bits sont utilisés (voir le paragraphe 4.3.2.2). Pour les charges utiles en mode de base, ce bit DEVRA être à 0 et DEVRA être ignoré par le receveur.

Noter que les trames qui emploient des valeurs d'ISF différentes exigent une encapsulation dans des paquets séparés. Donc, des considérations particulières s'appliquent quand on génère des paquets entrelacés et qu'un changement d'ISF est exécuté. En particulier, les trames qui, en accord avec le schéma d'entrelacement précédemment utilisé, devraient être agrégées dans un seul paquet doivent être séparées dans différents paquets, afin que la condition susmentionnée (toutes les trames dans un paquet partagent l'ISF) reste vraie. Une mise en œuvre naïve qui sépare les trames d'ISF différentes dans des paquets différents peut avoir pour résultat jusqu'à deux fois le nombre de paquets RTP, par rapport à une solution d'entrelacement optimale. L'altération de l'entrelacement avant et après le changement d'ISF peut réduire le besoin de paquets RTP supplémentaires.

4.3.2 Charge utile Tableau de contenus

Le tableau de contenus (ToC) consiste en une liste d'entrées, correspondant chacune à un groupe de trames audio portées dans la charge utile, comme décrit ci-dessous.

```

+-----+-----+-----+-----+
| Entrée de ToC n°1 | Entrée de ToC n°2 | ... | Entrée de ToC n°N |
+-----+-----+-----+-----+

```

Lorsque plusieurs groupes de trames sont présents dans une charge utile, les entrées de ToC DEVRONT être placées dans le paquet dans l'ordre croissant des valeurs d'horodatage RTP (modulo 2^{32}) de la première trame de transport que l'entrée de TOC représente.

4.3.2.1 Entrée de ToC en mode de base

Une entrée de ToC de charge utile en mode de base a le format suivant :

```

0                               1
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|F|Type de trame| Nb de trames |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

F (1 bit) : Réglé à 1, il indique que cette entrée de ToC est suivie par une autre entrée de ToC ; réglé à 0, il indique que cette entrée de ToC est la dernière du ToC.

Type de trame (FT) (7 bits) : Indique le type de trame de codec audio utilisé pour le groupe de trames référencé par cette entrée de ToC. Le FT désigne la combinaison de débits de cœur et stéréo AMR-WB+, un des types spéciaux de trames AMR-WB+, le débit AMR-WB, ou le bruit de fond, comme spécifié au Tableau 25 de [AMR-WB+].

Nombre de trames (8 bits) : Indique le nombre de trames dans le groupe référencé par cette entrée de ToC. Les entrées de ToC dont le champ est égal à 0 (qui indiquerait zéro trame) NE DEVRONT PAS être utilisées, et les paquets reçus avec une telle entrée de TOC DEVRONT être éliminés.

4.3.2.2 Entrée de ToC en mode entrelacement

Deux différents formats d'entrée de ToC sont définis en mode entrelacé. Ils diffèrent par la longueur du champ Déplacement, 4 bits ou 8 bits. Le bit L dans l'en-tête de charge les différencie.

Si L=0, une entrée de ToC a le format suivant :

```

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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|F| Type trame | Nb de trames | DIS1 | ... | DISi | ... |
+-----+-----+-----+-----+-----+-----+-----+-----+
| ... | ... | DISn |Bourrag|
+-----+-----+-----+-----+-----+-----+

```

F (1 bit) : voir la définition en 4.3.2.1.

Type de trame (FT) (7 bits) : voir la définition en 4.3.2.1.

Nb de trames (8 bits) : voir la définition en 4.3.2.1.

DIS1...DISn (4 bits) : liste de n (n = nombre de trames) champs de déplacement indiquant le déplacement de la ième (i = 1..n) trame audio par rapport à la trame audio précédente dans la charge utile, en unités de trames. Les valeurs de quatre bits d'entier non signé de déplacement peuvent être entre 0 et 15, indiquant le nombre de trames audio dans l'ordre du décodage entre la (i-1)ème et la ième trame dans la charge utile. Noter que pour la première entrée de ToC de la charge utile, la valeur de DIS1 n'a pas de signification. Elle DEVRA être réglée à zéro par un envoyeur et DEVRA être ignorée par un receveur. Cette localisation de trame dans l'ordre du décodage est définie de façon univoque par l'horodatage RTP et le TFI dans l'en-tête de charge utile. Noter aussi que pour les entrées de ToC suivantes, DIS1 indique le nombre de trames entre la dernière trame du groupe précédent et la première trame de ce groupe.

Bourrage (4 bits) : pour assurer l'alignement sur l'octet, quatre bits de bourrage DEVRONT être inclus à la fin de l'entrée de ToC au cas où il y aurait un nombre impair de trames dans le groupe référencé par cette entrée. Ces bits DEVRONT être réglés à zéro et DEVRONT être ignorés par le receveur. Si un groupe contenant un nombre pair de trames est référencé par cette entrée de ToC, ces bits de bourrage NE DEVRONT PAS être inclus dans la charge utile.

Si L=1, une entrée de ToC a le format suivant :

```

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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|F| Type trame | Nb de trames | DIS1 | ... |
+-----+-----+-----+-----+-----+-----+-----+-----+
| ... | DISn |
+-----+-----+-----+-----+-----+-----+

```

F (1 bit) : voir la définition en 4.3.2.1.

Type de trame (FT) (7 bits) : voir la définition en 4.3.2.1.

Nb de trames (8 bits) : voir la définition en 4.3.2.1.

DIS1...DISn (8 bits) : liste de n (n= nombre de trames) champs de déplacement indiquant le déplacement de la ième (i=1..n) trame audio par rapport à la trame audio précédente dans la charge utile, en unités de trames. Les valeurs de huit bits d'entier non signé de déplacement peuvent être entre 0 et 255, indiquant le nombre de trames audio dans l'ordre du décodage entre la (i-1)ème et la ième trame dans la charge utile. Noter que pour la première entrée de ToC de la charge utile, la valeur de DIS1 n'a pas de signification. Elle DEVRA être réglée à zéro par un envoyeur et DEVRA être ignorée par un receveur. Cette localisation de trame dans l'ordre du décodage est définie de façon univoque par l'horodatage RTP et le TFI dans l'en-tête de charge utile. Noter aussi que pour les entrées de ToC suivantes, DIS1 indique le déplacement entre la dernière trame du groupe précédent et la première trame de ce groupe.

4.3.2.3 Déduction de l'horodatage RTP

La valeur de l'horodatage RTP pour une trame DEVRA être la valeur d'horodatage du premier échantillon audio codé dans la trame. La valeur d'horodatage pour une trame est déduite différemment selon le mode de la charge utile, de base ou entrelacé. Dans les deux cas, la première trame dans un paquet composé a un horodatage RTP égal à celui reçu dans l'en-tête RTP. Dans le mode de base, l'heure RTP pour toute trame suivante est déduite en deux étapes. D'abord, on calcule la somme des durées de trame (voir le Tableau 1) de toutes les trames précédentes dans la charge utile. Ensuite, cette somme est ajoutée à la valeur de l'horodatage de l'en-tête RTP. Par exemple, supposons que la valeur de l'horodatage de l'en-tête RTP soit 12 345, la charge utile porte quatre trames, et la durée de trame est 16 ms (ISF = 32 kHz) correspondant à 1152 tics d'horodatage. Alors l'horodatage RTP de la quatrième trame de la charge utile est $12\ 345 + 3 * 1152 = 15\ 801$.

En mode entrelacé, l'horodatage RTP pour chaque trame dans la charge utile est déduit de l'horodatage de l'en-tête RTP et de la somme des décalages de temps de toutes les trames qui précèdent dans cette charge utile. Les horodatages de trame sont calculés sur la base des champs de déplacement et de la durée de trame déduite de la valeur de l'ISF. Noter que le déplacement en temps entre la trame i-1 et la trame i est $(DIS_i + 1) * \text{durée de trame}$, parce que la durée de la (i-1)ème doit aussi être pris en compte. L'horodatage de la première trame du premier groupe de trames (TS(1)) (c'est-à-dire, la première trame de la charge utile) est l'horodatage de l'en-tête RTP. Pour les trames suivantes dans le groupe, l'horodatage est calculé par : $TS(i) = TS(i-1) + (DIS_i + 1) * \text{durée de trame}$, $2 < i < n$

Pour les groupes de trames suivants, l'horodatage de la première trame est calculé par :

$$TS(1) = TS_{\text{préc}} + (DIS_1 + 1) * \text{durée de trame},$$

où TS_{préc} note l'horodatage de la dernière trame du groupe précédent. Les horodatages des trames suivantes dans le groupe sont calculées de la même façon que pour le premier groupe.

L'exemple suivant déduit les horodatages RTP pour les trames dans une charge utile en mode entrelacé ayant l'en-tête et les information de ToC suivants :

Horodatage d'en-tête RTP : 12345

ISF = 32 kHz

Champ de déplacement de la trame 1 : DIS1 = 0

Champ de déplacement de la trame 2 : DIS2 = 6

Champ de déplacement de la trame 3 : DIS3 = 4

Champ de déplacement de la trame 4 : DIS4 = 7

En supposant une ISF de 32 kHz, qui implique une durée de trame de 16 ms, une trame dure 1152 tics. L'horodatage de la première trame dans la charge utile est l'horodatage RTP, c'est-à-dire, $TS(1) = \text{RTP TS}$. Noter que la valeur du champ de déplacement pour cette trame doit être ignorée. Pour la seconde trame dans la charge utile, l'horodatage peut être calculé comme $TS(2) = TS(1) + (DIS_2 + 1) * 1152 = 20\ 409$. Pour la troisième trame, l'horodatage est $TS(3) = TS(2) + (DIS_3 + 1) * 1152 = 26\ 169$. Finalement, pour la quatrième trame de la charge utile, on a $TS(4) = TS(3) + (DIS_4 + 1) * 1152 = 35\ 385$.

4.3.2.4 Considérations de type de trame

La valeur du type de trame (FT) est définie au Tableau 25 de [AMR-WB+]. FT=14 (AUDIO_LOST) est utilisé pour noter les trames perdues. Une trame NO_DATA (FT=15) pourrait résulter de deux situations : d'abord qu'aucune donnée n'a été produite par le codeur audio; et ensuite, qu'aucune donnée n'est transmise dans la charge utile actuelle. Un exemple de ce dernier cas serait que la trame en question a été ou va être envoyée dans un paquet antérieur ou ultérieur. La durée pour ces trames non incluses dépend de la fréquence d'échantillonnage interne indiquée par le champ ISF.

Pour les types de trame avec les indices de 0 à 13, le champ ISF DEVRA être réglé à 0. La durée de trame pour ces types de trames est fixée à 20 ms, c'est-à-dire, 1440 tics en 72 kHz. Pour les charges utiles contenant seulement des trames de type 0 à 9, le champ TFI DEVRA être réglé à 0 et DEVRA être ignoré par le receveur. Dans une charge utile qui combine des trames de type 0 à 9 et 10 à 13, les valeurs de TFI doivent être réglées de façon à correspondre aux trames de transport de type 10 à 13. Donc, les trames de type 0 à 9 vont aussi avoir un TFI déduit, qui est ignoré.

4.3.2.5 Autres considérations de ToC

Si une entrée de ToC est reçue avec une valeur de FT indéfinie, tout le paquet DEVRA être éliminé. C'est pour éviter la perte de la synchronisation des données dans le processus de dépaquetisation, qui peut résulter en une sévère dégradation de la qualité audio.

Les paquets contenant seulement des trames NO_DATA NE DEVRAIENT PAS être transmis. Aussi, les trames NO_DATA à la fin d'une séquence de trames à porter dans une charge utile NE DEVRAIENT PAS être incluses dans le paquet transmis. Le SCR/DTX AMR-WB+ est identique au SCR/DTX AMR-WB décrit dans [TS26.193] et peut seulement être utilisé en combinaison avec les types de trames AMR-WB (0 à 8).

Quand plusieurs groupes de trames sont présents, leurs entrées de ToC DEVRONT être placées dans le ToC dans l'ordre croissant des valeurs d'horodatage RTP (modulo 2^{32}) de la première trame de transport que représente l'entrée de TOC, indépendamment du mode de charge utile. En mode de base, les trames DEVRONT être consécutives en temps, tandis qu'en mode entrelacé, les trames PEUVENT non seulement être non consécutives en temps mais PEUVENT même avoir des distances variables inter trames.

4.3.2.6 Exemples de ToC

L'exemple suivant illustre une ToC pour trois trames audio en mode de base. Noter que dans ce cas, toutes les trames audio sont codées en utilisant le même type de trame, c'est-à-dire, il y a seulement une entrée de ToC.

```

0                               1
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|0| Type trame1 | Nb trames = 3 |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

L'exemple suivant décrit une ToC de trois entrées en mode de base. Noter que dans ce cas, la charge utile porte aussi trois trames, mais trois entrées de ToC sont nécessaires parce que les trames de la charge utile sont codées en utilisant des types de trame différents.

```

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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|1| Type trame1 | Nb trames = 1 | |1| Type trame2 | Nb trames = 1 | |
+-----+-----+-----+-----+-----+-----+-----+-----+
|0| Type trame3 | Nb trames = 1 |
+-----+-----+-----+-----+-----+-----+

```

L'exemple suivant illustre une ToC avec deux entrées en mode entrelacé utilisant des champs de déplacement de quatre bits. La charge utile inclut deux groupes de trames, le premier inclut une seule trame, et l'autre comporte deux trames.

```

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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|1| Type trame1 | Nb trames = 1 | DIS1 | Bourg. |0| Type trame2 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Nb trames = 2 | DIS1 | DIS2 |
+-----+-----+-----+-----+-----+-----+

```

4.3.3 Données audio

Les données audio d'une charge utile consistent en zéro, une ou plusieurs trames audio, comme décrit dans le ToC de la charge utile.

Les entrées de ToC avec FT = 14 ou 15 représentent des types de trame de longueur 0. Donc, aucune donnée NE DEVRA être placée dans la section Données audio pour représenter des trames de ce type.

Comme mentionné précédemment, chaque trame audio d'un type de trame d'extension représente une trame de transport AMR-WB+ correspondant au codage de 512 échantillons d'audio, échantillonnés avec la fréquence d'échantillonnage interne spécifiée par l'indicateur d'ISF. À une exception près : les types de trame d'indice 10 à 13 sont seulement capables d'utiliser une seule fréquence d'échantillonnage interne (25 600 Hz). Les taux de codage (combinaison du débit binaire de cœur et du débit binaire stéréo) sont indiqués dans le champ Type de trame de l'entrée de ToC correspondante. La longueur en octets de la trame audio est implicitement définie par le champ Type de trame et est donnée aux Tableaux 21 et 25 de [AMR-WB+]. La notation de l'ordre et de la numérotation des bits est spécifiée dans [AMR-WB+]. Pour les types de trame d'extension AMR-WB+ et les trames de bruit de fond, les bits sont dans l'ordre produit par le codeur. Le dernier octet de chaque trame audio DOIT être bourré avec des zéros à la fin si tous les bits de l'octet ne sont pas utilisés. En d'autres termes, chaque trame audio DOIT être alignée sur une limite d'octet.

4.3.4 Méthodes de formation de la charge utile

La charge utile commence avec l'en-tête de charge utile, suivi par le tableau des contenus, qui consiste en une liste d'entrées de ToC.

Les données audio suivent le tableau des contenus. Tous les octets constituant une trame audio DEVRONT être ajoutés à la charge utile comme un tout. Les trames audio sont mises en paquet dans l'ordre de l'horodatage au sein de chaque groupe de trames (par entrée de ToC). Les groupes de trames sont mis en paquet dans le même ordre que leur entrée de ToC correspondante. Noter qu'il n'y a pas d'octet de données dans un groupe qui a une entrée de ToC avec FT=14 ou FT=15.

4.3.5 Exemples de charge utile

4.3.5.1 Exemple 1 : Mode de base de charge utile portant plusieurs trames codées avec le même type de trame

La Figure 4 décrit une charge utile qui porte trois trames AMR-WB+ codées en utilisant un type de trame à 14 kbit/s (FT=26) avec une longueur de trame de 280 bits (35 octets). La fréquence d'échantillonnage interne dans cet exemple est 25,6 kHz (ISF = 8). Le TFI pour la première trame est 2, indiquant que la première trame de transport dans cette charge utile est la troisième dans la super trame. Comme cette charge utile est en mode de base, les trames suivantes de la charge utile sont des trames consécutives dans l'ordre de décodage, c'est-à-dire, la quatrième trame de transport de la super trame courante et la première trame de transport de la prochaine super trame. Noter que parce que les trames sont toutes codées en utilisant le même type de trame, seulement une entrée de ToC est requise.

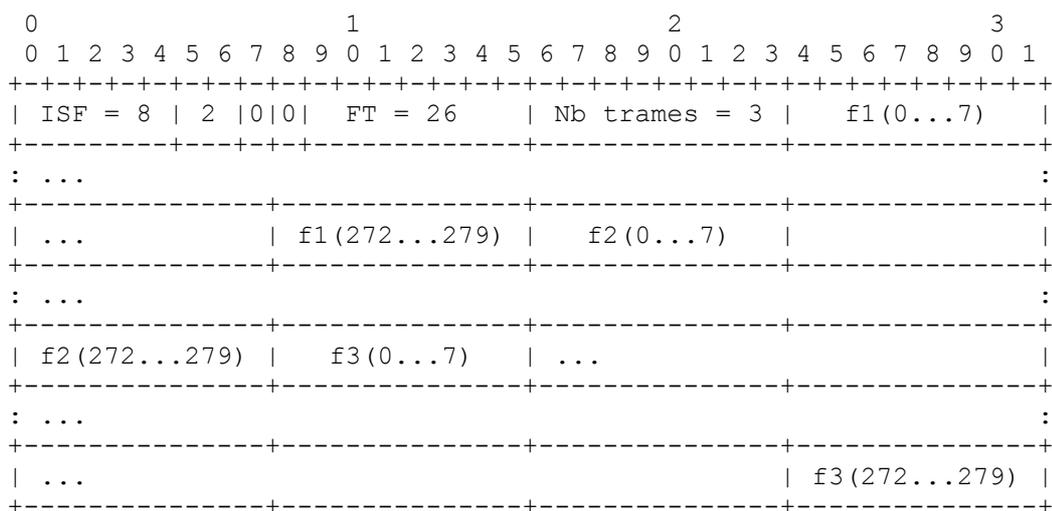


Figure 4 : Exemple de mode de base de charge utile portant trois trames de même type

4.3.5.2 Exemple 2 : Mode de base de charge utile portant plusieurs trames codées avec des types différents

La Figure 5 décrit une charge utile qui porte trois trames AMR-WB+ ; la première trame est codée en utilisant un type de trame à 18,4 kbit/s (FT=33) avec une longueur de trame de 368 bits (46 octets), et les deux trames suivantes sont codées en

utilisant un type de trame à 20 kbit/s (FT=35) d'une longueur de trame de 400 bits (50 octets). La fréquence d'échantillonnage interne dans cet exemple est 32 kHz (ISF = 10), ce qui implique un débit binaire global de 23 kbit/s pour la première trame de la charge utile, et 25 kbit/s pour les trames suivantes. Le TFI pour la première trame est 3, indiquant que la première trame de transport dans cette charge utile est la quatrième de la super trame. Comme c'est une charge utile en mode de base, les trames suivantes de la charge utile sont des trames consécutives dans l'ordre de décodage, c'est-à-dire, la première et la seconde trame de transport de la super trame en cours. Noter que comme la charge utile porte deux types différents de trame, il y a deux entrées de ToC.

```

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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| ISF=10 | 3 | 0 | 1 | FT = 33 | Nb trames = 1 | 0 | FT = 35 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Nb trames = 2 | f1(0...7) | ... |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
: ... :
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| ... | f1(360...367) | f2(0...7) |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
: ... :
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| f2(392...399) | f3(0...7) | ... |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
: ... :
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| ... | f3(392...399) |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Figure 5 : Exemple de mode de base de charge utile portant trois trames employant deux types différents

4.3.5.3 Exemple 3 : Charge utile en mode entrelacé

L'exemple de la Figure 6 décrit une charge utile en mode entrelacé, portant quatre trames codées en utilisant un type de trame de 32 kbit/s (FT=47) avec une longueur de trame de 640 bits (80 octets). La fréquence d'échantillonnage interne est 38,4 kHz (ISF = 13) impliquant un débit binaire de 48 kbit/s pour toutes les trames dans la charge utile. Le TFI pour la première trame est 0 ; donc, c'est la première trame de transport d'une super trame. Les champs de déplacement pour les trames suivantes sont DIS2=18, DIS3=15, et DIS4=10, qui indiquent que les trames suivantes ont les TFI de 3, 3, et 2, respectivement. Le fanion du champ Long déplacement dans l'en-tête de charge utile est réglé à 1, ce qui résulte en l'utilisation de huit bits pour les champs de déplacement dans l'entrée de ToC. Noter que comme toutes les trames de cette charge utile sont codées en utilisant le même type de trame, il y a besoin d'une seule entrée de ToC. De plus, le champ de déplacement pour la première trame (correspondant à la première entrée de ToC avec DIS1=0) doit être ignoré, car son horodatage et son TFI sont définis par l'horodatage RTP et le TFI trouvés dans l'en-tête de charge utile.

Les valeurs d'horodatage RTP des trames de cet exemple sont :

Trame1 : TS1 = Horodatage RTP

Trame2 : TS2 = TS1 + 19 * 960

Trame3 : TS3 = TS2 + 16 * 960

Trame4 : TS4 = TS3 + 11 * 960

```

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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| ISF=13 | 0 | 1 | 0 | FT = 47 | Nb trames = 4 | DIS1 = 0 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| DIS2 = 18 | DIS3 = 15 | DIS4 = 10 | f1(0...7) |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
: ... :
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| ... | f1(632...639) | f2(0...7) |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
: ... :

```

```

+-----+-----+-----+-----+
| ...           | f2 (632...639) | f3 (0...7) |
+-----+-----+-----+-----+
: ...
+-----+-----+-----+-----+
| ...           | f3 (632...639) | f4 (0...7) |
+-----+-----+-----+-----+
: ...
+-----+-----+-----+-----+
| ...           | f4 (632...639) |
+-----+-----+-----+-----+

```

Figure 6 : Exemple de charge utile en mode entrelacé portant quatre trames de même type

4.4 Considérations d'entrelacement

L'utilisation de l'entrelaçage exige un examen plus détaillé. Comme il est présenté dans l'exemple du paragraphe 3.6.2, un schéma d'entrelaçage exige une certaine quantité de mémoire tampon de désentrelaçage. Cet espace de mémoire tampon, exprimé en un certain nombre d'intervalles de trame de transport, est indiqué par le paramètre de type de support "entrelaçage". Le nombre de créneaux de trame nécessaires peut être converti en exigences de mémoire réelles en considérant les 80 octets par trame utilisés par les plus grandes combinaisons de débits de cœur et stéréo de AMR-WB+.

Les informations sur la taille de mémoire tampon de trames ne sont pas toujours suffisantes pour déterminer quand il est approprié de consommer des trames provenant de la mémoire tampon d'entrelaçage. Il y a deux cas dans lesquels des informations supplémentaires sont nécessaires : d'abord, quand se produit un changement de l'ISF, et ensuite, quand le schéma d'entrelaçage change. Le paramètre "int-delay" de type de support est défini pour porter ces informations. Il permet à un expéditeur d'indiquer le temps de support minimal qui doit être présent dans la mémoire tampon avant que le décodeur puisse commencer à consommer des trames provenant de la mémoire tampon. Parce que l'expéditeur a le plein contrôle sur les changements d'ISF et sur le schéma d'entrelaçage, il peut calculer cette valeur.

Dans certains cas (par exemple, si il se joint à une session en diffusion groupée avec entrelaçage à mi-session) un receveur peut initialement recevoir seulement une partie des paquets dans le schéma d'entrelaçage. Cette réception initiale partielle (dans l'ordre de suite des trames) des trames peut donner trop peu de trames pour une qualité acceptable du décodage audio. Ce problème survient aussi quand on utilise le chiffrement pour le contrôle d'accès, et que le receveur n'a pas la clé précédente.

Bien que AMR-WB+ soit robuste et donc tolérant à un fort taux d'écrasement aléatoire de trames, il aurait des difficultés à traiter des pertes de trames consécutives au démarrage. Donc, des considérations de mise en œuvre particulières sont décrites. Afin de traiter efficacement de type de démarrage, on doit noter que le décodage est seulement possible si il commence au début d'une super trame, et cela est vrai même si la première trame de transport est indiquée comme perdue. Ensuite, il est seulement RECOMMANDÉ que le décodage ne commence que si au moins deux trames de transport sont disponibles sur les quatre qui appartiennent à cette super trame.

Après la réception d'un certain nombre de paquets, dans le pire des cas, autant de paquets que couvre le schéma d'entrelaçage, les effets décrits précédemment disparaissent et le décodage normal est repris.

Des questions dimilaires se posent quand un receveur quitte une session ou a perdu l'accès au flux. Si le receveur quitte la session, cela va être un problème mineur car l'exécution est normalement arrêtée. C'est aussi un problème mineur dans le cas de l'accès perdu, car la dissimulation d'erreur de AMR-WB+ va atténuer l'audio si des pertes massives consécutives sont subies.

L'expéditeur peut éviter ce type de problème dans de nombreuses sessions en commençant et terminant correctement les schémas d'entrelaçage quand se produisent des risques de pertes. Un exemple est un changement de clé fait pour le contrôle d'accès sur un flux chiffré. Si seulement quelques clés sont fournies aux clients et qu'il y a un risque pour leur contenu reçu pour lequel ils n'ont pas la clé, il est recommandé que les schémas d'entrelaçage ne se chevauchent pas avec les changements de clés.

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. Toute


```

n := n+4
fin de boucle
passer l'étape 3                                # Plus d'une ISF a changé

```

Trouvé :

Retourner les horodatages récupérés et les ISF comme :

$x(i) = t0 + i*L0$ et l'ISF associée égale à $isf0$, pour $0 < i \leq n$

$x(i) = t0 + n*L0 + (i-n)*L1$ et l'ISF associée égale à $isf1$, pour $n < i \leq n+m$

passer à Fin

Étape 3 : Plus d'un changement d'ISF s'est produit. Comme les changements d'ISF peuvent être supposés peu fréquents, une telle situation se produit seulement si de longues séquences de trames sont perdues. Dans ce cas, il n'est probablement pas utile d'essayer de récupérer les horodatages des trames perdues. Le décodeur AMR-WB+ devrait plutôt être réinitialisé, et le décodage devrait être repris en commençant par la trame avec l'horodatage $t1$.

Fin:

L'algorithme ci-dessus ne résout pas le problème quand la profondeur de la mémoire tampon du receveur est inférieure à la taille de la salve perdue. Dans ce genre de cas, lorsque la dissimulation doit être faite sans rien connaître des futures trames, la dissimulation peut résulter en la perte de l'alignement sur une limite de trame. Si cela se produit, il peut être nécessaire de réinitialiser et de redémarrer le codec pour effectuer la resynchronisation.

4.5.2 Validation du décodage

Si le receveur trouve une discordance entre la taille d'une charge utile reçue et la taille indiquée par le ToC de la charge utile, il DEVRAIT éliminer le paquet. Ceci est recommandé parce que décoder une trame analysée à partir d'une charge utile sur la base de données de ToC erronées pourrait dégrader sévèrement la qualité audio.

5. Contrôle d'encombrement

Les considérations générales de contrôle d'encombrement pour le transport des données RTP s'appliquent ; voir RTP [RFC3550] et tout profil RTP applicable comme AVP [RFC3551]. Cependant, la capacité de multi débits du codage audio AMR-WB+ fournit un mécanisme qui peut aider à contrôler l'encombrement, car la demande de bande passante peut être ajustée (dans les limites du codec) en choisissant un codage différent de type de trame ou un taux d'échantillonnage interne moindre.

Le nombre de trames encapsulées dans chaque charge utile RTP influence grandement la bande passante globale du flux RTP à cause des contraintes de frais généraux de l'en-tête. Mettre en paquet plus de trames dans chaque charge utile RTP peut réduire le nombre de paquets envoyés et donc les frais généraux d'en-tête, au prix de délais accrus et d'une robustesse à l'erreur réduite.

Si la correction d'erreur directe (FEC, *forward error correction*) est utilisée, la quantité de redondance induite par la FEC doit être réglée de telle sorte que l'utilisation de la FEC elle-même ne cause pas un problème d'encombrement.

6. 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 tout profil applicable comme AVP [RFC3551] ou SAVP [RFC3711]. 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 de l'audio lui-même. Le format de charge utile n'a par lui-même aucun mécanisme de sécurité incorporé. Tout mécanisme externe convenable, comme SRTP [RFC3711], PEUT être utilisé.

Ce format de charge utile et le décodeur AMR-WB+ ne présentent aucune non uniformité significative de complexité de calcul du côté receveur pour le traitement des paquets, et ne présentent pas de menace probable de déni de service due à la réception de données pathologiques.

6.1 Confidentialité

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

L'utilisation de l'entrelaçage en conjonction avec le chiffrement peut avoir un impact négatif sur la confidentialité, pour une brève période. Considérons les paquets suivants (entre crochets) contenant les numéros de trame indiqués : {10, 14, 18}, {13, 17, 21}, {16, 20, 24} (un schéma populaire d'entrelaçage diagonal continu). Le générateur souhaite refuser à certains participants la capacité d'écouter le matériel commençant à l'instant 16. En changeant simplement la clé sur le paquet qui a l'horodatage à ou après 16, et en refusant cette nouvelle clé à ces participants, cela ne suffit pas ; les trames 17, 18, et 21 ont été fournies dans des paquets antérieurs avec la clé précédente, et la dissimulation d'erreur peut rendre l'audio intelligible au moins jusqu'à la trame 18 ou 19, et éventuellement plus loin.

6.2 Authentification et intégrité

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

7. Paramètres de format de charge utile

Cette section définit les paramètres qui peuvent être utilisés pour choisir les caractéristiques du format de charge utile AMR-WB+. Les paramètres sont définis au titre de l'enregistrement du type de support pour le codec audio AMR-WB+. Une transposition des paramètres dans le protocole de description de session (SDP, *Session Description Protocol*) [RFC2327] est aussi fournie pour les applications qui utilisent SDP. Des paramètres équivalents pourraient être définis ailleurs pour l'usage de protocoles de contrôle qui n'utilisent pas MIME ou SDP.

Le format et les paramètres de données sont seulement spécifiés pour le transport en temps réel dans RTP.

7.1 Enregistrement de type de support

Le type de support pour le codec large bande multi débits adaptatif étendu (AMR-WB+, *Extended Adaptive Multi-Rate Wideband*) est alloué dans l'arborescence de l'IETF, car AMR-WB+ est supposé être un codec audio largement utilisé dans les applications générales de flux en direct.

Note : les paramètres non mentionnés ci-dessous DOIVENT être ignorés par le receveur.

Nom du type de support : audio

Nom du sous type de support : AMR-WB+

Paramètres exigés : aucun

Paramètres facultatifs :

Canaux : le nombre maximum de canaux audio utilisés par les trames audio. Les valeurs permises sont 1 (mono) ou 2 (stéréo). Si aucun paramètre n'est présent, le nombre maximum de canaux est 2 (stéréo). Note : quand il est réglé à 1, cela implique que les types de trames stéréo ne peuvent pas être utilisés.

Entrelaçage : indique que le mode entrelacé DEVRA être utilisé pour la charge utile. Le paramètre spécifie le nombre d'intervalles de trame de transport exigés dans une antémémoire de désentrelaçage (incluant la trame qui est prête à être consommée). Sa valeur est égale à au plus le nombre maximum de trames qui précèdent toute trame dans l'ordre de transmission et suivent la trame dans l'ordre de l'horodatage RTP. La valeur DOIT être supérieure à zéro. Si ce paramètre n'est pas présent, le mode entrelacé NE DEVRA PAS être utilisé.

Int-delay : retard minimal de support en tics d'horodatage RTP qui est nécessaire dans l'antémémoire de désentrelaçage, c'est-à-dire, la différence en tics d'horodatage RTP entre la première et la dernière trame audio présente dans

l'antémémoire de désentrelaçage.

Ptime : voir la Section 6 de la [RFC2327].

Maxptime : voir la Section 8 de la [RFC3267].

Restrictions d'usage : ce type est seulement défini pour le transfert via RTP [RFC3550].

Considérations de codage : une charge utile RTP conforme au présent format est en données binaires et peut donc devoir être codée de façon appropriée dans des environnements non binaires. Cependant, tant qu'utilisé au sein de RTP, aucun codage n'est nécessaire.

Considérations de sécurité : voir la Section 6 de la [RFC4352].

Considérations d'interopérabilité : pour conserver l'interopérabilité avec les points d'extrémité à capacité AMR-WB, dans les cas où la négociation est possible et où le point d'extrémité AMR-WB+ qui prend en charge ce format prend aussi en charge la [RFC3267] pour le transport AMR-WB, un point d'extrémité AMR-WB+ DEVRAIT se déclarer aussi comme à capacité AMR-WB (c'est-à-dire, prenant en charge aussi "audio/AMR-WB" comme spécifié dans la RFC 3267). Comme le décodeur AMR-WB+ est capable d'effectuer les conversions de stéréo en mono, tous les receveurs de AMR-WB+ devraient être capables de recevoir la stéréo et le mono, bien que le receveur soit seulement capable d'exécuter des signaux mono.

Spécification publiée : RFC 4352, 3GPP TS 26.290, voir la référence [AMR-WB+] de la RFC 4352

Informations supplémentaires : ce type MIME n'est pas applicable pour la mémorisation de fichier. La mémorisation de fichier audio codé de AMR-WB+ est spécifiée dans le format de fichier multimédia fondé sur ISO défini par le TS 26.244 du 3GPP ; voir la référence [TS26.244] de la RFC 4352. Ce format de fichier a les types MIME "audio/3GPP" ou "video/3GPP" comme défini par la [RFC3839].

Adresse de la personne à contacter pour des informations complémentaires : magnus.westerlund@ericsson.com ;
ari.lakaniemi@nokia.com

Usage prévu : COMMUN. Il est prévu que de nombreuses applications de flux en direct fondés sur IP vont utiliser ce type.

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

7.2 Transposition des paramètres de type de support en SDP

Les informations portées dans la spécification de type de support 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 une session RTP utilisant ce format de charge utile RTP, la transposition est comme suit :

- Le type de support ("audio") est utilisé dans SDP "m=" comme le nom du support.
- Le type de support (nom de format de charge utile) est utilisé dans SDP "a=rtpmap" comme nom de codage. Le débit d'horloge RTP dans un "a=rtpmap" DEVRA être 72000 pour AMR-WB+, et le paramètre de codage nombre de canaux DOIT être explicitement réglé à 1 ou 2, ou être omis, ce qui implique la valeur par défaut de 2.
- Les paramètres "ptime" et "maxptime" sont placés dans les attributs SDP "a=ptime" et "a=maxptime", respectivement.
- Tous les paramètres restants sont placés dans l'attribut SDP "a=fmtp" en les copiant directement de la chaîne MIME type de support comme une liste de paires paramètre=valeur séparées par des points-virgules.

7.2.1 Considérations sur le modèle d'offre-réponse

Pour réaliser une bonne interopérabilité dans une utilisation de négociation d'offre-réponse [RFC3264], les considérations suivantes devraient être prises en compte :

Pour un usage d'offre/réponse négociable les règles d'interprétation suivantes DEVRONT être appliquées :

- Le paramètre "entrelaçage" est symétrique, exigeant donc que le répondant l'inclut aussi pour la réponse à une offre de type de charge utile qui contient le paramètre. Cependant, la valeur d'espace de mémoire tampon est déclarative en envoi individuel. Pour l'usage en diffusion groupée, la même valeur dans la réponse est requise afin d'accepter le type de charge utile. Pour les flux déclarés comme *sendrecv* (*envoi-réception*) ou *recvonly* (*réception seule*) : le receveur va accepter la réception du flux en utilisant le mode entrelacé du format de charge utile. La valeur déclare la quantité d'espace de mémoire tampon qui est disponible chez le receveur pour que l'expéditeur l'utilise. Pour les flux en envoi seul, le paramètre indique la configuration désirée et la quantité d'espace de mémoire tampon. Il est RECOMMANDÉ de répondre en utilisant la valeur offerte, si on est capable de l'utiliser.
- Le paramètre "int-delay" est déclaratif. Pour un flux déclaré comme *sendrecv* ou *recvonly*, la valeur indique le retard initial maximum que le receveur va accepter dans la mémoire tampon de désentrelaçage. Pour un flux en envoi seul, la valeur est la quantité de temps de support que l'expéditeur désire utiliser. La valeur DEVRAIT être copiée dans toute réponse.
- Le paramètre "canaux" est déclaratif. Pour un flux "sendonly", il indique l'usage de canal désiré, stéréo et mono, ou mono seulement. Pour un flux "recvonly" et "sendrecv", le paramètre indique ce que le receveur accepte d'utiliser. Comme tout receveur va être capable de recevoir le type de trame stéréo et d'effectuer un mixage local au sein du décodeur AMR-WB+, il y a normalement seulement une raison de restreindre à mono seulement : pour éviter de dépenser du débit binaire sur des données qui ne seront pas utilisées si le frontal est seulement capable de mono.
- Le paramètre "ptime" fonctionne comme indiqué par le modèle d'offer/réponse [RFC3264] ; "maxptime" DEVRA être utilisé de la même façon.
- Pour conserver l'interopérabilité avec AMR-WB dans les cas où la négociation est possible, il est RECOMMANDÉ à un point d'extrémité à capacité AMR-WB+ qui met aussi en œuvre le format de charge utile AMR-WB [RFC3267] de se déclarer capable de AMR-WB car c'est un sous ensemble du codec AMR-WB+. En usage déclaratif, comme SDP dans RTSP [RFC2326] ou SAP [RFC2974], l'interprétation suivante des paramètres DEVRA être faite :
- Le paramètre "entrelaçage", si il est présent, configure le format de charge utile dans ce mode, et la valeur indique le nombre de trames que la mémoire tampon de désentrelaçage doit prendre en charge pour être capable de traiter correctement cette session.
- Le paramètre "int-delay" indique le délai initial de mise en mémoire tampon exigé pour recevoir ce flux correctement.
- Le paramètre "canaux" indique si le contenu transmis peut contenir à la fois des débits stéréo et mono, ou seulement mono.
- Tous les autres paramètres indiquent des valeurs qui sont utilisées par l'entité expédatrice.

7.2.2 Exemples

Voici un exemple de description de session SDP utilisant le codage mono et stéréo AMR-WB+.

```
m=audio 49120 RTP/AVP 99
a=rtpmap:99 AMR-WB+/72000/2
a=fmtp:99 entrelaçage=30; int-delay=86400
a=maxptime:100
```

Noter que les noms de format de charge utile (codage) sont couramment montrés en majuscules. Les sous types de supports sont couramment montrés en minuscules. Ces noms sont insensibles à la casse dans les deux cas. De même, les noms de paramètre sont insensibles à la casse dans les types MIME et dans la transposition en attribut SDP `a=fmtp`.

8. Considérations relatives à l'IANA

L'IANA a enregistré un nouveau sous type MIME (`audio/amr-wb+`) ; voir la Section 7.

9. Contributeurs

Daniel Enstrom a contribué en écrivant la section d'introduction du codec. Stefan Bruhn a contribué en écrivant

l'algorithme de récupération d'ISF.

10. Remerciements

Les auteurs tiennent à remercier Redwan Salami et Stefan Bruhn de leurs contributions significatives pendant toute la rédaction et la relecture de ce document. Dave Singer a contribué par sa relecture et ses suggestions à l'amélioration du langage. Anisse Taleb et Ingemar Johansson ont contribué par la mise en œuvre du format de charge utile et ont aidé à localiser certaines erreurs. Nous tenons aussi à remercier Qiaobing Xie, coauteur de la RFC 3267, sur laquelle se fonde le présent document.

11. Références

11.1 Références normatives

- [AMR-WB+] 3GPP TS 26.290 "Audio codec processing functions; Extended Adaptive Multi-Rate Wideband (AMR-WB+) codec; Transcoding functions", version 6.3.0 (2005-06), 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](#))
- [RFC2327] M. Handley et V. Jacobson, "SDP : [Protocole de description de session](#)", avril 1998. (*Obsolète; voir [RFC4566](#)*)
- [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](#)*)
- [RFC3267] J. Sjoberg 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](#))
- [TS26.192] 3GPP TS 26.192 "AMR Wideband speech codec; Comfort Noise aspects", version 6.0.0 (2004-12), 3rd Generation Partnership Project (3GPP).
- [TS26.193] 3GPP TS 26.193 "AMR Wideband speech codec; Source Controlled Rate operation", version 6.0.0 (2004-12), 3rd Generation Partnership Project (3GPP).

11.2 Références pour information

- [MBMS] 3GPP TS 26.140 "Multimedia Broadcast/Multicast Service (MBMS); Protocols and codecs", version 6.3.0 (2005-12), 3rd Generation Partnership Project (3GPP).
- [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*)

- [RFC3711] M. Baugher et autres, "Protocole de [transport sécurisé en temps réel](#) (SRTP)", mars 2004. (P.S.)
- [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.)
- [TS26.140] 3GPP TS 26.140 "Multimedia Messaging Service (MMS); Media formats and codes", version 6.2.0 (2005-03), 3rd Generation Partnership Project (3GPP).
- [TS26.233] 3GPP TS 26.233 "Packet Switched Streaming service", version 5.7.0 (2005-03), 3rd Generation Partnership Project (3GPP).
- [TS26.244] 3GPP TS 26.244 "Transparent end-to-end packet switched streaming service (PSS); 3GPP file format (3GP)", version 6.4.0 (2005-09), 3rd Generation Partnership Project (3GPP).

Tous les documents du 3GPP peuvent être téléchargés sur le site du 3GPP, "<http://www.3gpp.org/> ", voir les spécifications.

Adresse des auteurs

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

Magnus Westerlund
Ericsson Research
Ericsson AB
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

Stephan Wenger
Nokia Corporation
P.O. Box 100
FIN-33721 Tampere
FINLAND
téléphone : +358-50-486-0637
mél : Stephan.Wenger@nokia.com

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.