

Équipe d'ingénierie de l'Internet (IETF)
Request for Comments: 5686
Catégorie : Sur la voie de la normalisation
ISSN: 2070-1721

Y. Hiwasaki, NTT Corporation
H. Ohmuro, NTT Corporation
octobre 2009
Traduction Claude Brière de L'Isle

Format de charge utile RTP pour les codecs incorporés de loi μ pour les codecs vocaux de communication IP à faible délai (UEMCLIP)

Résumé
Le présent document décrit le format de charge utile RTP pour un codeur incorporé de loi μ pour la communication IP à faible délai (UEMCLIP, *mu-law Embedded Coder for Low-delay IP communication*) de codec vocal amélioré de la Recommandation UIT-T G.711. Ce flux binaire a une structure adaptable avec un flux binaire de loi incorporé, aussi appelé PCMU, fournissant donc une opération de transcodage élégante en parole à bande étroite et à large bande.

Statut de ce mémoire
Ceci est un document de l'Internet sur la voie de la normalisation.

Le présent document a été produit par l'équipe d'ingénierie de l'Internet (IETF). Il représente le consensus de la communauté de l'IETF. Il a subi une révision publique et sa publication a été approuvée par le groupe de pilotage de l'ingénierie de l'Internet (IESG). Tous les documents approuvés par l'IESG ne sont pas candidats à devenir une norme de l'Internet ; voir la Section 2 de la RFC5741.

Les informations sur le statut actuel du présent document, tout errata, et comment fournir des réactions sur lui peuvent être obtenues à <http://www.rfc-editor.org/info/rfc5686>

Notice de droits de reproduction
Copyright (c) 2009 IETF Trust et les personnes identifiées comme auteurs du document. Tous droits réservés.

Le présent document est soumis au BCP 78 et aux dispositions légales de l'IETF Trust qui se rapportent aux documents de l'IETF (<http://trustee.ietf.org/license-info>) en vigueur à la date de publication de ce document. Prière de revoir ces documents avec attention, car ils décrivent vos droits et obligations par rapport à ce document. Les composants de code extraits du présent document doivent inclure le texte de licence simplifié de BSD comme décrit au paragraphe 4.e des dispositions légales du Trust et sont fournis sans garantie comme décrit dans la licence de BSD simplifiée.

Le présent document peut contenir des matériaux provenant de documents de l'IETF ou de contributions à l'IETF publiées ou rendues disponibles au public avant le 10 novembre 2008. La ou les personnes qui ont le contrôle des droits de reproduction sur tout ou partie de ces matériaux peuvent n'avoir pas accordé à l'IETF Trust le droit de permettre des modifications de ces matériaux en dehors du processus de normalisation de l'IETF. Sans l'obtention d'une licence adéquate de la part de la ou des personnes qui ont le contrôle des droits de reproduction de ces matériaux, le présent document ne peut pas être modifié en dehors du processus de normalisation de l'IETF, et des travaux dérivés ne peuvent pas être créés en dehors du processus de normalisation de l'IETF, excepté pour le formater en vue de sa publication comme RFC ou pour le traduire dans une autre langue que l'anglais.

Table des matières

1. Introduction.....	2
1.1 Terminologie.....	2
2. Fondements du format de support.....	2
3. Format de charge utile.....	3
3.1 Usage de l'en-tête RTP.....	4
3.2 Trames multiples dans un paquet RTP.....	4
3.3 Données de charge utile.....	4
4. Transcodage entre UEMCLIP et G.711.....	7
5. Considérations de contrôle d'encombrement.....	7
6. Paramètres de format de charge utile.....	8
6.1 Enregistrement de type de support.....	8
6.2 Transposition en paramètres SDP.....	8

6.3 Considérations sur le modèle d'offre-réponse.....	9
7. Considérations sur la sécurité.....	11
8. Considérations relatives à l'IANA.....	11
9. Références.....	11
9.1 Références normatives.....	11
9.2 Références pour information.....	12
Adresse des auteurs.....	12

1. Introduction

Le présent document spécifie le format de charge utile pour l'envoi de parole codée dans un codeur de loi μ incorporé pour une communication IP à faible délai (UEMCLIP, *mU-law EMbedded Coder for Low-delay IP communication*) en utilisant le protocole de transport en temps réel (RTP, *Real-time Transport Protocol*) [RFC3550]. UEMCLIP est un codec propriétaire qui améliore la loi μ de la Recommandation [UIT-T G.711] et qui est conçu pour aider le marché à une transition en douceur vers l'environnement de communications à haut débit qui s'annonce tout en réalisant une très faible charge de transcodage des supports avec les terminaux existants, dans lesquels la mise en œuvre de G.711 est obligatoire.

On devrait noter que, d'une manière générale, les codecs sont négociés et changés en utilisant un échange SDP. Aussi, la [RFC3550] définit des modèles généraux de mixeur et traducteur RTP, où le transcodage de supports ne peut pas avoir lieu au nœud. Dans ces cas, le concept de structure incorporée n'est pas utile. Cependant, il y a d'autres cas où un transcodage coûteux est inévitable dans les types couramment déployés d'unités de contrôle multi-points (MCU, *Multi-point Control Unit*) qui terminent les supports et les paquets RTCP [RFC5117], et quand des terminaux à bande étroite et à large bande coexistent. Cette structure de flux binaires incorporée peut réduire le transcodage de supports à une simple troncature du flux binaire.

Les fondements et l'idée de base du format de support sont décrits à la Section 2. Les détails du format de charge utile sont donnés à la Section 3. Les questions de transcodage avec G.711 sont discutées à la Section 4, et les considérations de contrôle d'encombrement sont dans la Section 5. À la Section 6, sont fournis les paramètres de format de charge utile pour l'enregistrement d'un type de support pour les transpositions du format de charge utile UEMCLIP de RTP et du protocole de description de session (SDP, *Session Description Protocol*). Les considérations sur la sécurité et les considérations relatives à l'IANA sont traitées respectivement dans les Section 7 et 8,.

1.1 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].

2. Fondements du format de support

UEMCLIP est une version améliorée de la loi μ de la Recommandation UIT-T G.711, autrement appelée PCMU (*MIC loi μ*) [RFC4856]. Il vise les applications du protocole de voix sur Internet (VoIP, *Voice on Internet Protocol*) et son but principal est de fournir une plate-forme de communication large bande très interopérable avec les terminaux existants équipés avec G.711 et de stimuler le glissement graduel du marché vers l'utilisation de communications large bande. Dans les systèmes de conférence multi-points largement déployés, les paquets passent généralement à travers des MCU à terminaison sur le protocole de contrôle de RTP (RTCP, *RTP Control Protocol*), "Topo-RTCP-terminating-MCU" (*point à multipoints utilisant des MCU à terminaison RTCP*) comme défini dans la [RFC5117]. Parce que le flux binaire G.711 est incorporé dans le flux binaire, un transcodage coûteux des supports peut être évité dans ce cas.

Le présent document ne discute pas les détails de mise en œuvre du codeur et du décodeur, mais décrit seulement le format du flux binaire.

À cause de sa nature adaptable, il y a de nombreux sous flux binaires (sous couches) dans un flux binaire UEMCLIP. En choisissant les sous couches appropriées, le codec peut s'adapter aux exigences suivantes :

- o fréquence d'échantillonnage,
- o nombre de canaux,
- o qualité de la parole, et

- o débit binaire.

Le codec UEMCLIP opère avec une trame de 20 ms, et inclut trois sous codeurs comme montré au Tableau 1. La couche centrale est une loi μ G.711 à 64 kbit/s, et les deux autres sont des couches d'amélioration de la qualité et de la bande passante avec un débit binaire de 16 kbit/s chacune.

Couche	Description	Débit binaire	Algorithme de codage
a	cœur G.711	64	MIC loi μ
b	amélioration de la bande inférieure	16	quantification de bloc de domaine de temps
c	amélioration de la bande supérieure	16	quantification de bloc MDCT

Table 1 : Description des sous couches

Sur la base de ces sous-couches,le codec UEMCLIP opère dans quatre modes comme le montre le Tableau 2. Ici, "Ch" est le nombre de canaux et "Fs" est la fréquence d'échantillonnage en kHz. On devrait noter que la version courante prend seulement en charge le fonctionnement sur un seul canal et qu'il pourrait y avoir des extensions futures avec des capacités multi-canaux. Les modes 2 et 5 absents sont réservés pour une possible extension future de modes d'échantillonnage à 32 kHz. Comme la définition de mode est supposée croître, des autres modes non définis dans ce tableau NE DOIVENT PAS être utilisés, pour des raisons de compatibilité et d'interopérabilité.

Mode	Ch	Fs	Couche a	Couche b	Couche c	Débit en-tête (kbit/s)	Débit total (kbit/s)
0	1	8	x	-	-	64	67,2
1	1	16	x	-	x	80	84,0
2	-	-	-	-	-	-	-
3	1	8	x	x	-	80	84,0
4	1	16	x	x	x	96	100,8
5	-	-	-	-	-	-	-

Table 2 : Description du mode

Le flux binaire UEMCLIP contient des en-têtes internes et d'autres informations auxiliaires en plus des données de couche. Il en résulte un débit total supérieur à la somme des couches montrées dans le tableau ci-dessus. Les détails des en-têtes internes et des informations auxiliaires sont décrits au paragraphe 3.3.1.

Définir la fréquence d'échantillonnage et le nombre de canaux ne résulte pas en un mode singulier, c'est-à-dire, il peut y avoir plusieurs modes pour la même fréquence d'échantillonnage ou le même nombre de canaux. Les modes pris en charge vont différer entre les mises en œuvre ; donc, l'envoyeur et le receveur doivent négocier quel mode utiliser pour la transmission.

3. Format de charge utile

Comme charge utile RTP, le flux binaire UEMCLIP peut contenir une ou plusieurs trames, comme montré à la Figure 1.

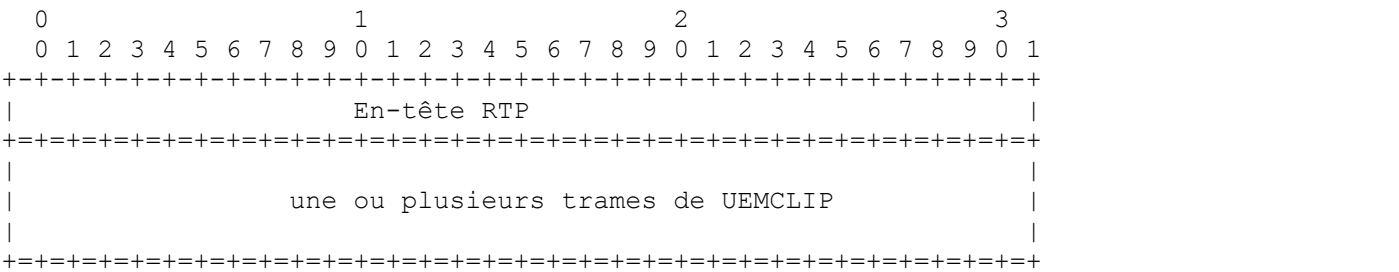


Figure 1 : Format de charge utile RTP

Le flux binaire UEMCLIP a une structure adaptable ; donc, il est possible de reconstruire le signal en en décodant une partie. Une trame UEMCLIP est composée d'un en-tête principal (MH, *main header*) suivi par une ou plusieurs (jusqu'à trois) sous couches (SL, *sub-layer*) comme le montre la Figure 2.

```

+---+-----+//--+
|MH| SL n°1 |...|
+---+-----+//--+

```

Figure 2 : Trame UEMCLIP (format de débit binaire)

Comme sous couche, la couche centrale, c'est-à-dire, la "couche a", DOIT toujours être incluse. On devrait noter que la localisation de la couche centrale peut ou non suivre immédiatement le champ MH. Le décodeur DOIT toujours se référer aux indices des couches pour un décodage approprié parce que l'ordre des sous couches est arbitraire.

Le flux binaire UEMCLIP n'inclut pas explicitement les informations suivantes : mode et fréquence d'échantillonnage (Fs). Comme décrit précédemment, ces informations DOIVENT être échangées lors de l'établissement d'une connexion, par exemple, au moyen de SDP.

3.1 Usage de l'en-tête RTP

Chaque paquet RTP commence par un en-tête RTP fixe, comme expliqué dans la [RFC3550]. Les champs suivants de l'en-tête RTP fixe utilisés spécifiquement pour les flux UEMCLIP sont développés :

Type de charge utile : l'allocation d'un type de charge utile RTP pour ce format de paquet sort du domaine d'application de ce document ; cependant, on suppose qu'un type de charge utile dans la gamme dynamique devra être alloué.

Horodatage : il code l'instant d'échantillonnage du premier échantillon de signal de parole dans le paquet de données RTP.

Pour les flux UEMCLIP, l'horodatage RTP DOIT être fondé sur une horloge à 8000 ou 16000 Hz. Dans les cas où le taux d'échantillonnage audio peut changer durant une session, le taux d'horodatage RTP DOIT être égal au taux maximum (en Hz) donné dans la gamme de mode (voir le paragraphe 6.2.1). Cela implique que le taux d'horodatage RTP pour le type de charge utile UEMCLIP NE DOIT PAS changer durant une session. Par exemple, pour un flux UEMCLIP avec un échantillonnage audio à 8 kHz, où une transition à un mode d'échantillonnage audio à 16 kHz est permis, l'horodatage RTP doit toujours avancer en utilisant le rythme d'horloge à 16 kHz. Pour un mode d'échantillonnage audio fixe, le taux d'horodatage RTP devrait être 8 ou 16 kHz, selon le taux d'échantillonnage.

Bit marqueur : si le codec est utilisé pour des applications avec transmission discontinue (DTX, ou compression de silence) le premier paquet après une période de silence durant laquelle des paquets n'ont pas été transmis de façon contiguë DEVRAIT avoir le bit marqueur réglé à un dans l'en-tête de données RTP. Le bit marqueur dans tous les autres paquets DOIT être à zéro. Les applications sans DTX DOIVENT régler le bit marqueur à zéro.

3.2 Trames multiples dans un paquet RTP

Plus d'une trame UEMCLIP peut être incluse dans un seul paquet RTP par un envoyeur. Cependant, les envoyeurs ont les restrictions supplémentaires suivantes :

- o un seul paquet RTP NE DEVRAIT PAS inclure plus de trames UEMCLIP qu'il n'en tient dans la MTU de chemin,
- o toutes les trames contenues dans un seul paquet RTP DOIVENT être du même mode,
- o les trames NE DOIVENT PAS être partagées entre des paquets RTP.

Il est RECOMMANDÉ que le nombre de trames contenues dans un paquet RTP soit cohérent avec l'application. Comme UEMCLIP est conçu pour des applications de téléphonie où le délai a un grand impact sur la qualité, moins de trames par paquet est préférable pour un moindre délai.

3.3 Données de charge utile

Dans un flux binaire UEMCLIP, tous les numéros sont codés dans l'ordre des octets du réseau (*gros boutien*).

3.3.1 En-tête principal

L'en-tête principal (MH) est placé au sommet d'une trame et a une taille de 6 octets. Le contenu de l'en-tête principal est montré dans la Figure 3.

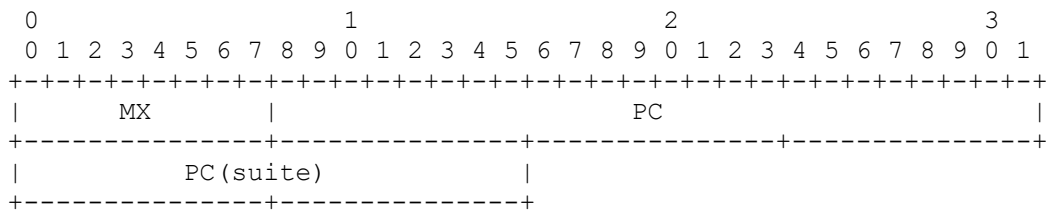


Figure 3 : Format d'en-tête principal UEMCLIP

Informations de mixage (MX, *Mixing information*) : 8 bits. Ce champ est seulement pertinent quand des Topo-RTCP-terminating-MCU sont utilisées pour interpréter ces champs. Voir au paragraphe 3.3.1.1 le détail des champs.

Informations de dissimulation de perte de paquet (PC, *Packet-loss Concealment*) : 40 bits. Voir au paragraphe 3.3.1.2.

3.3.1.1 Champ Informations de mixage

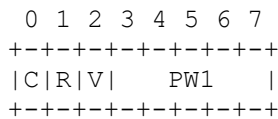


Figure 4 : Champ Informations de mixage (MX)

C : bit de vérification : 1 bit. Fanion de validité de V et PW1. Ce bit à "1" indique que les deux paramètres sont valides, et "0" indique que les paramètres devraient être ignorés. Si un de ces paramètres est invalide, ce bit devrait être réglé à "0". Ce fanion est principalement destiné à un point à multipoints utilisant des MCU à terminaison RTCP à capacité UEMCLIP. Ce fanion devrait être réglé à "0" en cas de transcodage à partir de G.711 (voir la Section 4).

Bit réservé (R) : 1 bit. Ce bit devrait être ignoré. Il est à 0 par défaut.

V : 1 bit. Fanion de détection d'activité vocale de la trame en cours, conçu pour être utilisé pour les opérations de MCU. Ce fanion à "1" indique que la trame est un segment actif (vocal) et "0" indique que c'est un segment inactif (non vocal) ou silencieux. Ce fanion est spécifiquement conçu pour les informations de mixage. Le jugement de DTX fondé sur ce fanion n'est pas recommandé.

PW1 : 5 bits. Code de puissance du signal de la trame en cours. Le code est obtenu en calculant la moyenne quadratique (RMS, *root mean square*) de "couche a" et en codant cette RMS en utilisant la loi μ G.711 [G.711]. En notant la RMS codée R, PW1 est alors obtenu par $PW1 = ((\sim R) \gg 2) \& 0x1F$, où " $\sim$ ", " $\gg$ ", "&" sont, respectivement l'arithmétique de complément à un, le décalage à droite, et l'opérateur ET au bit près.

3.3.1.2 Champ Informations de PLC (PC)

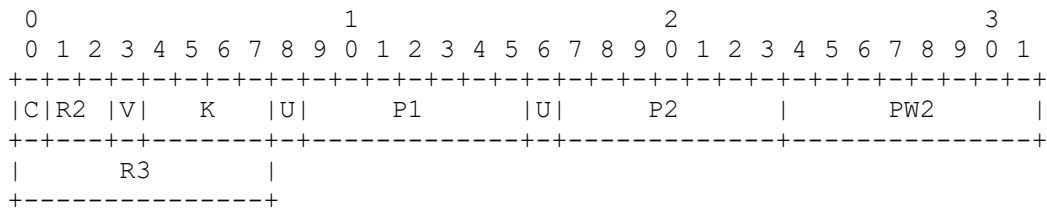


Figure 5 : Champ Informations de PLC (PC)

Bit C : 1 bit. Fanion de validité de V, K, U, P1, P2, et PW2. Si le fanion est "1", cela signifie que tous ces paramètres sont valides, et "0" signifie que les paramètres devraient être ignorés. Si un de ces paramètres est invalide, ce bit devrait être réglé à "0". Comme pour C en 3.3.1.1, ce fanion devrait être réglé à "0" en cas de transcodage à partir de G.711 (voir la Section 4).

Réservé (R2) : 2 bits. Ces bits devraient être ignorés. Ils sont à 0 par défaut.

Fanion V : 1 bit. Fanion de détection d'activité vocale de la trame en cours, conçu pour être utilisé pour la dissimulation de perte de paquet. Il pourrait ne pas être le même que dans les informations de mixage, et ne pas être synchrone au bit marqueur dans l'en-tête RTP. Le jugement de DTX fondé sur ce fanion n'est pas recommandé.

Indicateur de trame (K) : 4 bits. Cette valeur indique le décalage de trame de U, P, et PW2. Comme c'est une meilleure idée de porter les paramètres de caractéristiques de parole comme des informations de PLC dans une trame différente pour maintenir la qualité de la parole, cette valeur de décalage de trame dit à quelle trame les paramètres sont à associer. Les valeurs vont de "0" à "15". Si le numéro de trame courant est N, par exemple, la valeur K indique que U, P, et PW2 sont associés à la trame N-K. L'indicateur de trame est égal à la différence de numéro de séquence RTP quand une trame UEMCLIP est contenue dans un seul paquet RTP.

Fanion U : 1 bit. Indicateur de signal voix/non voix de la trame courante. Ce fanion à "0" indique que la trame est un segment de signal vocal, et "1" indique qu'il est un segment de signal non vocal.

P : 7 bits. Code tonal de la trame courante. Le délai tonal réel est calculé comme $P1 + 20$ échantillons au taux d'échantillonnage à 8 kHz. Le délai tonal doit être $20 \leq \text{longueur tonale} \leq 120$. Les codes allant de "0x65" à "0x7F" ne sont pas utilisés. Pour obtenir le délai tonal, toute méthode d'estimation tonale peut être utilisée, comme celle utilisée dans l'Appendice I de G.711 [G.711 A1].

Fanion U : 1 bit. Indicateur de signal voix/non voix de la trame de décalage. Ce fanion à "0" indique que la trame est un segment de signal vocal, et "1" indique que c'est un segment de signal non vocal. La valeur de décalage est définie comme K.

P2 : 7 bits. Code tonal de la trame décalée. La valeur de décalage est définie comme K. La méthode de calcul est identique à celle de "P" au paragraphe 3.3.1.1, sauf qu'il est fondé sur le signal de la trame de décalage.

PW2 : 8 bits. Code de puissance du signal de la trame de décalage. La valeur de décalage est définie comme K.

Bits réservés : 8 bits. Ces bits devraient être ignorés. Tous les bits sont "0" par défaut.

3.3.2 Sous couche

Sous-couche (SL) est un sous en-tête suivi par des flux binaires de couche, comme montré à la Figure 6. Le sous en-tête indique la localisation de la couche et le nombre d'octets.

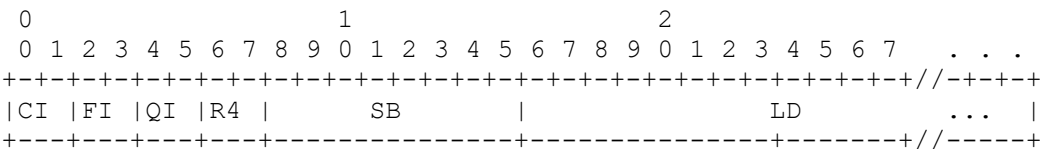


Figure 6 : Format de sous couche (SL)

Indice de canal (CI, *Channel index*) : 2 bits. Indique le nombre de canaux. Pour tous les modes donnés au Tableau 2, ce devrait être "0". Les détails sont donnés au Tableau 3.

Indice de fréquence (FI, *Frequency index*) : 2 bits. Indique le nombre de fréquences. "0" signifie que la couche est dans la bande de fréquence de base, un numéro supérieur signifie que la couche est dans la bande de fréquence correspondante. Les détails sont donnés au Tableau 3.

Indice de qualité (QI, *Quality index*) : 2 bits. Indique le numéro de la couche de qualité. "0" signifie que la couche est dans la couche de base, et un numéro supérieur signifie que la couche est dans la couche de qualité correspondante. Les détails sont donnés au Tableau 3.

Réservé n° 4 (R4) : 2 bits. Non utilisé (réservé). La valeur par défaut est "0".

SB (taille de sous couche) : 8 bits. Indique la taille en octets des données de la sous couche suivante.

Données de couche (LD, *Layer Data*) : SB*8 bits. Données réelles de la sous couche.

Pour toutes les couches montrées dans le Tableau 1, les indices des couches sont montrés dans le Tableau 3.

Couche	CI	FI	QI
a	0	0	0
b	0	0	1
c	0	1	0

Table 3 : Indices de couche

4. Transcodage entre UEMCLIP et G.711

Comme indiqué à la Section 2, le flux binaire G.711 codé en loi μ (couche a) est la couche cœur d'un flux binaire UEMCLIP, et est toujours incorporé. Cela signifie que le transcodage de supports du flux binaire UEMCLIP en G.711 n'a pas à subir de procédure de décodage et recodage, mais qu'une simple extraction va suffire. Cependant, cela ne s'applique pas à la procédure inverse, c'est-à-dire, au transcodage de G.711 en UEMCLIP, parce que les informations auxiliaires dans l'en-tête principal (MH) doivent être allouées séparément. On devrait noter que ce transcodage de supports est utile pour un traducteur de supports (*Topo-Media-Translator*) ou un point à multipoints utilisant des MCU à terminaison RTCP (*Topo-RTCP-terminating-MCU*) en la [RFC5117], et toutes les exigences s'appliquent. Cela signifie qu'un appareil de transcodage de cette sorte DOIT réécrire les paquets RTCP, ainsi que les paquets de supports RTP.

Le transcodage de UEMCLIP en G.711 loi μ peut être fait aisément en trouvant une sous couche appropriée. Dans une trame, le transcodeur devrait chercher une sous-couche avec un indice de couche de "0x00", et où le LD suivant a une taille de SB*8 bits (UEMCLIP a une trame de 20 ms donc, SB = 160) sont les données réelles du flux binaire G.711. On devrait noter que le transcodeur ne devrait pas toujours s'attendre à ce que la couche centrale soit située juste après l'en-tête principal.

Par ailleurs, le transcodage de G.711 en UEMCLIP n'est pas entièrement direct. Comme il n'y a pas de moyen de générer des sous couches d'amélioration, un flux binaire G.711 peut seulement être converti en un flux binaire UEMCLIP de mode 0. Si le flux binaire original G.711 est codé en loi A, il devrait d'abord être converti en loi μ pour devenir la couche centrale. Parce que la taille d'une trame UEMCLIP est de 20 ms, un flux binaire G.711 codé en loi μ DOIT être un tronçon de 160 échantillons pour devenir une couche centrale. Pour le contenu de l'en-tête principal, quand le codeur UEMCLIP n'est pas disponible, il devrait suivre ces lignes directrices :

- o Les bits de vérification pour le mixage et PLC (C1 et C2) sont réglés à 0.
- o Les bits réservés (R1 à R3) dans MH sont réglés aux valeurs par défaut respectives.

Pour la couche centrale (c'est-à-dire, le flux binaire G.711 de loi μ) elle devrait avoir l'en-tête de sous couche suivant :

- o Tous les champs CI, FI, QI, et R4 DOIVENT être à 0.
- o La taille de sous couche (SB) DOIT être 160 pour une trame de 20 ms.

5. Considérations de contrôle d'encombrement

Les considérations générales de contrôle d'encombrement pour transporter les données RTP s'appliquent aussi à UEMCLIP sur RTP [RFC3550] ainsi qu'à tout profil RTP applicable comme le profil audiovisuel (AVP, *Audio-Visual Profile*) [RFC3551].

La bande passante d'un flux binaire UEMCLIP peut être réduite en passant à des modes à débits binaires inférieurs. La structure en couches incorporée de UEMCLIP peut aider à contrôler l'encombrement, quand un changement dynamique de mode (voir le paragraphe 6.2.1) est disponible, et la gamme des modes est obtenue par une négociation offre-réponse comme indiqué au paragraphe 6.3. On devrait noter que cela implique un traitement RTCP approprié quand le débit binaire est modifié dans un traducteur ou un mixeur RTP [RFC3550].

Mettre en paquets plus de trames dans chaque charge utile RTP peut réduire le nombre de paquets envoyés, et donc les frais généraux des en-têtes IP/UDP/RTP, au prix d'un délai accru et d'une robustesse réduite à l'erreur contre les pertes de paquet. Cela devrait être traité avec soin parce que des délais accrus signifient une qualité réduite.

6. Paramètres de format de charge utile

6.1 Enregistrement de type de support

Cet enregistrement est fait en utilisant le gabarit défini dans la [RFC4288] et selon la [RFC4855].

Nom de type de support : audio

Nom de sous type de support : UEMCLIP

Paramètres exigés :

Rate : définit le taux d'échantillonnage, et DOIT être 8000 ou 16000. Voir les détails au paragraphe 6.2.1 "Spécification du mode" de la RFC 5686 (cette RFC).

Paramètres facultatifs :

ptime : voir la [RFC4566].

maxptime : voir la [RFC4566].

mode : indique la gamme des modes qui peuvent changer dynamiquement durant une session. Les valeurs possibles sont une liste séparée de virgules des modes dans l'ensemble des modes pris en charge. Si seulement un mode est spécifié, cela signifie que le mode ne doit pas être changé durant la session. Quand il n'est pas spécifié, le mode de transmission devient par défaut un mode singulier comme spécifié au Tableau 4. Voir les détails au paragraphe 6.2.1 "Spécification du mode" de la RFC 5686 (cette RFC).

Considérations de codage : ce type de support est tramé et contient des données binaires. Voir le paragraphe 4.8 de la RFC 4288.

Considérations de sécurité : voir la Section 7 "Considérations sur la sécurité" de la RFC 5686 (cette RFC).

Considérations d'interopérabilité : ce support peut être directement transcodé en G.711 codé en loi μ . Voir la Section 4 "Transcodage entre UEMCLIP et G.711" de la RFC 5686 (cette RFC).

Spécification publiée : RFC 5686 (cette RFC)

Applications qui utilisent ce type de support : flux audio et vidéo et outils de conférence.

Informations supplémentaires : aucune

Usage prévu : COMMUN

Restrictions d'usage : ce type de support dépend du tramage RTP, et est donc seulement défini pour le transfert via RTP.

Personne & adresse de messagerie à contacter pour plus d'informations : Yusuke Hiwasaki
<hiwasaki.yusuke@lab.ntt.co.jp>

Auteur : Yusuke Hiwasaki

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

6.2 Transposition en paramètres SDP

Les types de supports audio/UEMCLIP sont transposés en champs dans le protocole de description de session (SDP) [RFC4566] comme suit :

Nom de support : la ligne "m=" du SDP DOIT être audio.

Nom de codage : le nom du sous type de support enregistré devrait être utilisé pour la ligne "a=rtpmap".

Fréquence d'échantillonnage : selon le mode, le débit d'horloge (fréquence d'échantillonnage) spécifié dans "a=rtpmap" DOIT être choisi parmi ceux définis dans le Tableau 2. Voir les détails au paragraphe 6.2.1.

Paramètres de codage : comme c'est un flux audio, les paramètres de codage indiquent le nombre de canaux audio, et cela DEVRAIT être "1" par défaut, comme choisi parmi ceux définis dans le Tableau 2. Ceci est FACULTATIF.

Heure de paquet : la longueur de trame de tout UEMCLIP est 20 ms, donc l'argument d'une "a=ptime" DEVRAIT être un multiple de "20". Quand elle n'est pas mentionnée dans SDP, elle devrait aussi prendre par défaut la taille minimum de "20".

Spécificités de UEMCLIP : toute description spécifique de UEMCLIP est définie dans les paramètres de spécification de format ("a=fmtp"). Chaque paramètre DOIT être séparé par ";", et si il existe un attribut (valeur) il DOIT être défini avec "=". Pour des raisons de compatibilité, toute application/terminal DOIT ignorer les paramètres qui ne sont pas compris. C'est pour assurer la compatibilité aval avec les paramètres ajoutés dans de futures améliorations. La spécification de mode devrait être faite ici (voir le paragraphe 6.2.1).

6.2.1 Spécification du mode

Comme le codec UEMCLIP peut fonctionner dans de nombreux modes (débits binaires) il est souhaitable de spécifier la gamme de modes auxquels un codeur ou un décodeur peut opérer. Quand il échange des messages SDP, un offreuseur devrait spécifier toutes les combinaisons possibles de numéros de mode comme arguments à "mode=" dans une ligne "a=fmtp", délimités par des virgules ",". Dans le cas où il spécifie plusieurs modes, ceux-ci DEVRAIENT apparaître en ordre de

priorité décroissante.

Bien que les décodeurs UEMCLIP DEVRAIENT accepter des flux binaires dans n'importe quels modes, une mise en œuvre peut échouer à s'adapter aux changements de mode dynamiques durant une session. Pour cette raison, une application peut choisir de fonctionner avec un mode fixe ou avec plusieurs modes qui peuvent être changés dynamiquement. Si le mode doit être fixe et que les changements ne sont pas permis, cela peut être indiqué en spécifiant un seul mode par type de charge utile.

Les numéros de mode qui peuvent être spécifiés dans un type de charge utile comme arguments à "mode" sont restreints par une combinaison de débit d'horloge et de nombre de canaux audio. C'est parce que SDP lie un type de charge utile à une combinaison d'une fréquence d'échantillonnage et d'un nombre de canaux audio. Le Tableau 4 donne les numéros de mode choissables qui sont attribués avec des débits d'horloge. Quand les spécifications de mode ne sont pas données, un type de charge utile DOIT prendre par défaut un seul mode en utilisant la valeur par défaut spécifiée dans ce tableau.

Débit d'horloge	Canaux	Modes choissables	Mode par défaut
16000	1	0, 1, 3, 4	1

Table 4 : Modes par défaut

On devrait noter qu'un mode attribué avec une plus grande fréquence d'échantillonnage (Fs) n'est pas utilisé en conjonction avec de plus faibles débits d'horloge spécifiés dans "a=rtpmap". Cela signifie que les modes 0 et 3 peuvent être spécifiés dans un type de charge utile ayant un débit d'horloge de 8000 et 16000 dans "a=rtpmap", mais les modes 1 et 4 ne peuvent pas être spécifiés avec l'un ayant un débit d'horloge de 8000.

6.3 Considérations sur le modèle d'offre-réponse

6.3.1 Lignes directrices sur l'offre-réponse

Les procédures relatives à l'échange de messages SDP DOIVENT suivre la [RFC3264]. On donne ci-après une liste détaillée de la sémantique d'utilisation du format de charge utile UEMCLIP dans un échange d'offre-réponse.

- o Un offreur DEVRAIT offrir toutes les combinaisons possibles de type de charge utile UEMCLIP qu'il peut traiter, c'est-à-dire, fréquence d'échantillonnage, nombre de canaux, et paramètres de fntp, dans l'ordre de préférence. Quand la bande passante de transmission est restreinte, elle DOIT être offerte en accord avec la restriction.
- o Quand plusieurs types de charge utile UEMCLIP sont offerts, il est RECOMMANDÉ que le répondeur choisisse un seul type de charge utile UEMCLIP dans sa réponse.
- o Dans un type de charge utile UEMCLIP, un répondeur DOIT renvoyer un ou des numéros de mode convenables comme un sous ensemble de ce qui a été offert. Cela signifie qu'il y a une hypothèse symétrique sur les flux envoyés et reçus, et l'offreur NE DOIT PAS envoyer dans des modes qu'il n'a pas offerts.
- o Dans un SDP d'offre/réponse, tous les paramètres fntp qui ne sont pas connus DOIVENT être ignorés. Si des paramètres inconnus/indéfinis devaient être offerts, un répondeur DOIT supprimer l'entrée du message de réponse.
- o Un receveur d'un message SDP DOIT seulement utiliser des types et modes spécifiés de charge utile. Quand une spécification de mode manque, c'est-à-dire, si un mode n'est pas spécifié du tout, la session DOIT par défaut prendre un seul mode sans changement de mode durant une session. Dans ce cas, les valeurs de mode par défaut, comme montré au Tableau 4, DOIVENT être utilisées sur la base de la fréquence d'échantillonnage et du nombre de canaux. On ne doit chercher dans ce tableau que quand il n'y a pas de spécification de mode ; donc, l'offreur/répondeur NE DOIT PAS supposer que le mode par défaut est toujours disponible quand il n'est pas dans la liste des modes spécifiés.
- o Quand une condition offerte ne tient pas dans les capacités d'un répondeur, il NE DOIT naturellement répondre à aucune des conditions, et la session PEUT procéder à un ré-INVITE, si possible. Si une condition (mode) est décidée, un offreur et un répondeur DOIVENT transmettre à cette condition.

6.3.2 Exemples

Quand un offreur indique qu'il souhaite changer dynamiquement entre modes (0, 1, 3, et 4) durant une session, un exemple d'offre SDP pourrait être :

```
v=0
o=john 51050101 51050101 IN IP4 offhost.exemple.com
s=-
c=IN IP4 offhost.exemple.com
t=0 0
m=audio 5004 RTP/AVP 96
a=rtpmap:96 UEMCLIP/16000/1
a=fmtp:96 mode=4,1,3,0
```

On devrait noter que la liste des modes apparaît dans l'ordre de préférences de l'offreur.

Quand un répondeur peut seulement opérer dans les modes 1 et 0 mais peut commuter dynamiquement entre ces modes durant une session, un répondeur DOIT supprimer les entrées de mode 3 et 4, et répondre par :

```
v=0
o=lena 549947322 549947322 IN IP4 anshost.exemple.org
s=-
c=IN IP4 anshost.exemple.org
t=0 0
m=audio 5004 RTP/AVP 96
a=rtpmap:96 UEMCLIP/16000/1
a=fmtp:96 mode=1,0
```

Par suite, tous deux vont commencer à communiquer en mode 1 ou 0, et peuvent commuter dynamiquement entre ces modes durant la session.

Par ailleurs, quand le répondeur est capable de communiquer dans les modes 1 ou 0, et ne peut pas commuter entre modes durant une session, un exemple d'une telle réponse est comme suit :

```
v=0
o=lena 549947322 549947322 IN IP4 anshost.exemple.org
s=-
c=IN IP4 anshost.exemple.org
t=0 0
m=audio 5004 RTP/AVP 96
a=rtpmap:96 UEMCLIP/16000/1
a=fmtp:96 mode=1
```

Par suite, tous deux vont commencer à communiquer en mode 1. On devrait noter que le changement de mode durant cette session n'est pas permis parce que le répondeur a répondu avec un seul mode, et le répondeur a choisi le mode 1 plutôt que le mode 0 conformément à l'ordre offert.

Si un offreur ne veut pas de changement de mode durant une session mais est capable de recevoir les flux binaires en mode 4 ou 1, le SDP devrait ressembler à :

```
v=0
o=john 51050101 51050101 IN IP4 offhost.exemple.com
s=-
c=IN IP4 offhost.exemple.com
t=0 0
m=audio 5004 RTP/AVP 96 97
a=rtpmap:96 UEMCLIP/16000/1
a=fmtp:96 mode=4
a=rtpmap:97 UEMCLIP/16000/1
a=fmtp:97 mode=1
```

et si le répondeur préfère communiquer en mode 1, la réponse serait :

```
v=0
o=lena 549947322 549947322 IN IP4 anshost.exemple.org
s=-
c=IN IP4 anshost.exemple.org
t=0 0
m=audio 5004 RTP/AVP 97
a=rtpmap:97 UEMCLIP/16000/1
a=fmtp:97 mode=1
```

On notera qu'il est RECOMMANDÉ de choisir un seul type de charge utile UEMCLIP pour les réponses.

L'attribut "ptime" est utilisé pour noter l'intervalle de mise en paquet désiré. Quand il n'est pas spécifié, il DEVRAIT par défaut être 20. Comme UEMCLIP utilise des trames de 20 ms, les valeurs de ptime de multiples de 20 impliquent plusieurs trames par paquet. Dans l'exemple ci-dessous, le ptime est réglé à 60, et cela signifie que l'offreur veut recevoir 3 trames dans chaque paquet.

```
v=0
o=kosuke 2890844730 2890844730 IN IP4 anotherhost.exemple.com
s=-
c=IN IP4 anotherhost.exemple.com
t=0 0
m=audio 5004 RTP/AVP 96
a=ptime:60
a=rtpmap:96 UEMCLIP/16000/1
```

Quand aucune spécification n'est présente, la valeur par défaut devrait être un mode fixe, et dans ce cas, le mode 1 (voir le paragraphe 6.2.1).

7. Considérations sur la sécurité

Les paquets RTP qui utilisent le format de charge utile défini dans la présente spécification sont sujets aux considérations sur la sécurité discutées dans la spécification de RTP [RFC3550] et dans tous les profils appropriés. Cela implique que la confidentialité des flux de supports est réalisée par le chiffrement sauf si le profil applicable spécifie d'autres moyens.

Une menace potentielle de déni de service existe pour le codage de données utilisant des techniques de compression qui ont une charge de calcul non uniforme à l'extrémité de réception. L'attaquant peut injecter dans le flux des datagrammes pathologiques qui sont complexes à décoder et causent une forte surcharge du résultat du receveur. Cependant, le UEMCLIP couvert dans le présent document ne présente aucune non uniformité significative.

Une autre menace potentielle est l'attaque sur la mémoire par des indices de couche ou numéros d'octets illégaux. L'utilisateur du décodeur devrait toujours avoir conscience que les numéros indiqués peuvent être corrompus et ne pas pointer sur la bonne sous couche, et ils peuvent forcer à lire au delà des limites du flux binaire. Il est conseillé qu'une mise en œuvre de décodeur rejette les couches avec de tels indices.

8. Considérations relatives à l'IANA

Un nouveau sous type de support (audio/UEMCLIP) a été enregistré par l'IANA. Voir les détails au paragraphe 6.1.

9. Références

9.1 Références normatives

[G.711] Recommandation UIT-T G.711, "Modulation par impulsions et codage (MIC) des fréquences vocales", UIT-T,

novembre 1988. Disponible à <http://www.itu.int/rec/T-REC-G.711>.

- [RFC2119] S. Bradner, "[Mots clés à utiliser](#) dans les RFC pour indiquer les niveaux d'exigence", BCP 14, mars 1997. DOI 10.17487/RFC2119, (MàJ par [RFC8174](#))
- [RFC3264] J. Rosenberg et H. Schulzrinne, "[Modèle d'offre/réponse](#) avec le protocole de description de session (SDP)", juin 2002, DOI 10.17487/RFC3264. (P.S. ; MàJ par [RFC8843](#), [9143](#))
- [RFC3550] H. Schulzrinne, S. Casner, R. Frederick et V. Jacobson, "[RTP : un protocole de transport pour les applications en temps réel](#)", STD 64, DOI 10.17487/RFC3550, juillet 2003. (MàJ par [RFC7164](#), [RFC7160](#), [RFC8083](#), [RFC8108](#), [RFC8860](#))
- [RFC3551] H. Schulzrinne et S. Casner, "[Profil RTP pour conférences audio](#) et vidéo avec contrôle minimal", STD 65, juillet 2003. (MàJ par [RFC8860](#))
- [RFC4288] N. Freed et J. Klensin, "Spécifications du [type de support et procédures d'enregistrement](#)", [BCP 13](#), décembre 2005.
- [RFC4566] M. Handley, V. Jacobson et C. Perkins, "SDP : [Protocole de description de session](#)", juillet 2006, DOI 10.17487/RFC4566. (P.S. ; remplacée par [RFC8866](#))
- [RFC4855] S. Casner, "[Enregistrement du type de support](#) des formats de charge utile RTP", février 2007. (P.S.)
- [RFC4856] S. Casner, "[Enregistrement de type de support](#) des formats de charge utile dans le profil RTP pour audio et visio conférences", février 2007. (Remplace [RFC3555](#)) (P.S.)
- [RFC5117] M. Westerlund, S. Wenger "[Topologies dans RTP](#)", janvier 2008. (Info. , Remplacée par [RFC7667](#))

9.2 Références pour information

- [G.711 A1] Recommandation UIT-T G.711, "Modulation par impulsions et codage (MIC) des fréquences vocales, Appendice I : algorithme de haute qualité et faible complexité pour la dissimulation de perte de paquet avec G.711", UIT-T, septembre 1999.

Adresse des auteurs

Yusuke Hiwasaki
NTT Corporation
3-9-11 Midori-cho,
Musashino-shi
Tokyo 180-8585
Japan
mél : hiwasaki.yusuke@lab.ntt.co.jp

Hitoshi Ohmuro
NTT Corporation
3-9-11 Midori-cho,
Musashino-shi
Tokyo 180-8585
Japan
mél : ohmuro.hitoshi@lab.ntt.co.jp