

Groupe de travail Réseau

A. Sollaud, France Telecom

Request for Comments : 5391

Catégorie : Sur la voie de la normalisation

Traduction Claude Brière de L'Isle

novembre 2008

Format de charge utile RTP pour la Recommandation UIT-T G.711.1

Statut du présent mémoire

Le présent document spécifie un protocole de l'Internet sur la voie de la 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 droits de reproduction

Copyright (c) 2008 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.

Résumé

Le présent document spécifie un format de charge utile du protocole de transport en temps réel (RTP, *Real-time Transport Protocol*) à utiliser pour le codec audio de la Recommandation UIT-T G.711.1. Deux enregistrements de types de supports sont aussi inclus.

Table des matières

1. Introduction.....	1
2. Fondements.....	2
3. Usage de l'en-tête RTP.....	2
4. Format de charge utile.....	3
4.1 En-tête de charge utile.....	3
4.2 Données audio.....	3
5. Paramètres de format de charge utile.....	4
5.1 Enregistrement de type de support MICA-LB.....	4
5.2 Enregistrement de type de support PCMU-WB.....	5
5.3 Transposition en paramètres RTP.....	6
6. Interopérabilité de G.711.....	7
7. Contrôle d'encombrement.....	7
8. Considérations sur la sécurité.....	8
9. Considérations relatives à l'IANA.....	8
10. Références.....	8
10.1 Références normatives.....	8
10.2 Références pour information.....	9
Adresse de l'auteur.....	9

1. Introduction

La Recommandation UIT-T G.711.1 [G.711.1] est une extension de large bande incorporée du codec audio de la Recommandation G.711 [G.711]. Le présent document spécifie un format de charge utile pour la mise en paquets des signaux audio codés selon G.711.1 dans le protocole de transport en temps réel (RTP).

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

G.711.1 est un algorithme de codage audio et parole large bande incorporée G.711 fonctionnant à 64, 80, et 96 kbit/s. À 64 kbit/s, G.711.1 est pleinement interopérable avec G.711. Donc, un déploiement efficace dans les infrastructures de voix sur IP (VoIP, *Voice over IP*) existantes fondées sur G.711 est prévu.

Le codec opère sur des trames de 5 ms, et le taux d'échantillonnage par défaut est de 16 kHz. L'entrée et la sortie à 8 kHz est aussi prise en charge pour les modes en bande étroite.

Le codeur produit un flux binaire incorporé structuré en trois couches correspondant aux trois débits binaires disponibles : 64, 80, et 96 kbit/s. Le flux binaire peut être tronqué du côté du décodeur ou par tout composant du système de communication pour ajuster, "au vol", le débit binaire à la valeur désirée.

Le tableau suivant donne plus de détails sur ces couches.

Couche	Description	Débit binaire
L0	compatible G.711	64 kbit/s
L1	amélioration de bande étroite	16 kbit/s
L2	amélioration de large bande	16 kbit/s

Tableau 1 : Description des couches

Les combinaisons de ces trois couches résultent en la définition de quatre modes, dans le tableau suivant :

Mode	couche 0	couche 1	couche2	Bande audio	Débit binaire
R1	x			bande étroite	64 kbit/s
R2a	x	x		bande étroite	80 kbit/s
R2b	x		x	large bande	80 kbit/s
R3	x	x	x	large bande	96 kbit/s

Table 2 : Description des modes

3. Usage de l'en-tête RTP

Le format de l'en-tête RTP est spécifié dans la [RFC3550]. Le format de charge utile défini dans le présent document utilise les champs de l'en-tête de façon cohérente avec cette spécification.

Marqueur (M) : G.711.1 ne définit rien de spécifique concernant la transmission discontinue (DTX, *Discontinuous Transmission*) autrement dit la suppression de silence. Des mécanismes indépendants du codec peuvent être utilisés, comme le format générique de charge utile de bruit de confort défini dans la [RFC3389]. Pour les applications qui n'envoient pas de paquets ou des paquets occasionnels de bruit de confort durant les silences, le premier paquet d'une salve de parole -- c'est-à-dire, le premier paquet après une période de silence durant laquelle des paquets n'ont pas été transmis de façon contiguë -- DEVRAIT être distingué en réglant à un le bit marqueur dans l'en-tête de données RTP. Le bit marqueur dans tous les autres paquets est à zéro. Le début d'une salve de parole PEUT être utilisé pour ajuster le délai d'exécution à refléter le changement des délais du réseau. Les applications sans suppression du silence DOIVENT régler le bit marqueur à zéro.

Type de charge utile (PT, *Payload Type*) : l'allocation d'un type de charge utile RTP pour ce format de paquet sort du domaine d'application du présent document, et ne sera pas spécifiée ici. Il est prévu que le profil RTP selon lequel ce format de charge utile va être utilisé allouera un type de charge utile pour ce codec ou spécifiera que le type de charge utile devra être lié dynamiquement (voir le paragraphe 5.3).

Horodatage : la fréquence d'horloge d'horodatage RTP est la même que celle de la fréquence d'échantillonnage par défaut

de 16 kHz.

G.711.1 a aussi la capacité de fonctionner avec des signaux d'entrée/sortie échantillonnés à 8 kHz. Cela n'affecte pas le flux binaire, et le décodeur n'exige pas une connaissance à priori du taux d'échantillonnage du signal original à l'entrée du codeur. Donc, selon la mise en œuvre et les capacités acoustiques audio des appareils, l'entrée du codeur et/ou la sortie du décodeur peuvent être configurées à 8 kHz ; cependant, un débit d'horloge RTP de 16 kHz DOIT toujours être utilisé.

La durée d'une trame est 5 ms, correspondant à 80 échantillons à 16 kHz. Donc, l'horodatage est augmenté de 80 à chaque trame consécutive.

4. Format de charge utile

La charge utile complète consiste en un en-tête de charge utile de 1 octet, suivi par une ou plusieurs trames audio G.711.1 consécutives de même mode.

Le mode peut changer entre les paquets, mais pas au sein d'un paquet.

4.1 En-tête de charge utile

L'en-tête de charge utile est illustré ci-dessous :

```

 0 1 2 3 4 5 6 7
+-----+
| 0 0 0 0 0 | MI |
+-----+
```

Les cinq bits de poids fort sont réservés pour une future extension et DOIVENT être réglés à zéro et DOIVENT être ignorés à réception.

Le champ indice de mode (MI, *Mode Index*) (3 bits) donne le mode de la ou des trames suivantes conformément au tableau suivant :

Indice de mode	Mode G.711.1	Taille de trame
1	R1	40 octets
2	R2a	50 octets
3	R2b	50 octets
4	R3	60 octets

Tableau 3 : Modes dans l'en-tête de charge utile

Toutes les autres valeurs de MI sont réservées pour une future utilisation et NE DOIVENT PAS être utilisées.

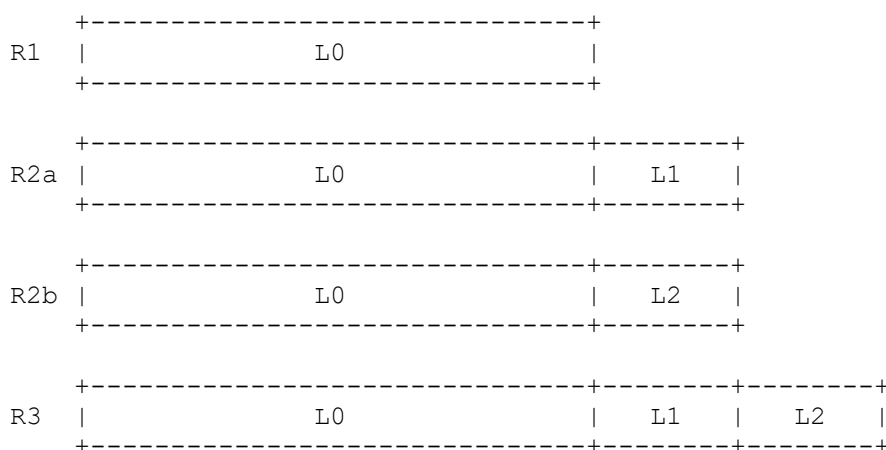
Les charges utiles reçues avec une valeur de MI non définie DOIVENT être éliminées.

Si un ensemble de mode restreint a été établi par la signalisation (voir la Section 5) les charges utiles reçues avec une valeur de MI qui n'est pas dans cet ensemble DOIVENT être éliminées.

4.2 Données audio

Après cet en-tête de charge utile, les trames audio consécutives sont mises en paquet dans l'ordre chronologique, c'est-à-dire, la plus ancienne d'abord. Toutes les trames DOIVENT être du même mode, indiqué par le champ MI de l'en-tête de charge utile.

Au sein d'une trame, les couches sont toujours mises en paquet dans le même ordre : couche 0 puis couche 1 pour le mode R2a, couche 0 puis couche 2 pour le mode R2b, couche 0 puis couche 1 puis couche 2 pour le mode R3. Ceci est illustré ci-dessous .



La taille d'une trame est donnée par le mode, conformément au tableau 3, et le nombre réel de trames est facile à déduire de la taille de la partie données audio :

$$\text{nb_trames} = (\text{taille_des_données_audio}) / (\text{taille_d'une_trame}).$$

Seules les trames complètes doivent être considérées. Donc si il y a un reste dans la division ci-dessus, les octets correspondants restants dans la charge utile reçue DOIVENT être ignorés.

5. Paramètres de format de charge utile

Cette section définit les paramètres qui peuvent être utilisés pour configurer les caractéristiques facultatives dans la transmission RTP G.711.1.

G.711 en loi A et en loi mu sont toutes deux prises en charge dans la couche centrale L0, mais il n'y a pas d'interopérabilité entre loi A et loi mu. De sorte que deux types de supports avec les mêmes paramètres vont être définis : audio/MICA-LB pour le cœur loi A, et audio/MICU-LB pour le cœur loi mu. Ceci est cohérent avec la séparation des types de support audio/MICA et audio/MICU pour l'audio G.711.

Les paramètres sont définis ici au titre des enregistrements de sous type de supports pour le codec G.711.1. Une transposition des paramètres dans le protocole de description de session (SDP, *Session Description Protocol*) [RFC4566] est aussi fournie pour les applications qui utilisent SDP. Dans les protocoles de contrôle qui n'utilisent pas MIME ou SDP, les paramètres de type de supports doivent être transposés en le format approprié utilisé avec ce protocole de contrôle.

5.1 Enregistrement de type de support MICA-LB

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

Nom de type : audio

Nom de sous type : PCMA-WB (*MICA-LB en français*)

Paramètres exigés : aucun

Paramètres facultatifs :

mode-set : restreint l'ensemble de modes de codec actifs à un sous ensemble de tous les modes. Les valeurs possibles sont une liste séparée de virgules dans l'ensemble : 1, 2, 3, 4 (voir les indices de mode dans le Tableau 3 de la RFC 5391).

Les modes sont donnés par ordre de préférence ; le premier est le préféré. Si mode-set est spécifié, les trames codées avec des modes en dehors du sous ensemble NE DOIVENT PAS être envoyées dans une charge utile RTP. Si il n'est pas présent, tous les modes de codec sont permis.

ptime : durée recommandée (en millisecondes) représentée par le support dans un paquet. Ce devrait être un multiple entier de 5 ms (la taille de trame). Voir la Section 6 de la RFC 4566.

maxptime : durée maximum (en millisecondes) qui peut être encapsulée dans un paquet. Ce devrait être un multiple entier de 5 ms (taille de trame). Voir la Section 6 de la RFC 4566.

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

Considérations de sécurité : voir la Section 8 de la RFC 5391.

Considérations d'interopérabilité : aucune

Spécification publiée : RFC 5391

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

Informations supplémentaires : aucune

Personne & adresse de messagerie à contacter pour plus d'informations : Aurélien Sollaud, aurelien.sollaud@orange-ftgroup.com

Usage prévu : COMMUN

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

Auteur : Aurélien Sollaud

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

5.2 Enregistrement de type de support PCMU-WB

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

Nom de type : audio

Nom de sous type : PCMU-WB (*MICU-LB en français*)

Paramètres exigés : aucun

Paramètres facultatifs :

mode-set : restreint l'ensemble de modes de codec actifs à un sous ensemble de tous les modes. Les valeurs possibles sont une liste séparée de virgules dans l'ensemble : 1, 2, 3, 4 (voir les indices de mode dans le Tableau 3 de la RFC 5391). Les modes sont donnés par ordre de préférence ; le premier est le préféré. Si mode-set est spécifié, les trames codées avec des modes en dehors du sous ensemble NE DOIVENT PAS être envoyées dans une charge utile RTP. Si il n'est pas présent, tous les modes de codec sont permis.

ptime : durée recommandée (en millisecondes) représentée par le support dans un paquet. Ce devrait être un multiple entier de 5 ms (la taille de trame). Voir la Section 6 de la RFC 4566.

maxptime : durée maximum (en millisecondes) qui peut être encapsulée dans un paquet. Ce devrait être un multiple entier de 5 ms (taille de trame). Voir la Section 6 de la RFC 4566.

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

Considérations de sécurité : voir la Section 8 de la RFC 5391.

Considérations d'interopérabilité : aucune

Spécification publiée : RFC 5391

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

Informations supplémentaires : aucune

Personne & adresse de messagerie à contacter pour plus d'informations : Aurélien Sollaud, aurelien.sollaud@orange-ftgroup.com

Usage prévu : COMMUN

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

Auteur : Aurélien Sollaud

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

5.3 Transposition en paramètres RTP

Les informations portées dans la spécification de type de support ont une transposition spécifique en les champs du protocole de description de session (SDP) [RFC4566], qui est couramment utilisé pour décrire les sessions RTP. Quand SDP est utilisé pour spécifier des sessions qui emploient le codec G.711.1, la transposition est la suivante :

- o Le type de support ("audio") va dans la ligne SDP "m=" comme nom du support.
- o Le sous type de support ("PCMA-WB" ou "PCMU-WB") va dans la ligne SDP "a=rtpmap" comme nom de codage. Le débit d'horloge RTP dans "a=rtpmap" DOIT être 16000 pour G.711.1.
- o Le paramètre "mode-set" va dans l'attribut SDP "a=fmtp" en le copiant comme une chaîne "mode-set=<valeur>".
- o Les paramètres "ptime" et "maxptime" vont respectivement dans les attributs SDP "a=ptime" et "a=maxptime".

5.3.1 Considérations sur le modèle d'offre-réponse

Les considérations suivantes s'appliquent quand on utilise les procédures SDP d'offre-réponse [RFC3264] pour négocier l'utilisation de la charge utile G.711.1 dans RTP :

- o Comme G.711.1 est une extension de G.711, l'offreur DEVRAIT annoncer la prise en charge de G.711 dans sa ligne "m=audio", avec G.711.1 préféré. Cela va permettre l'interopérabilité avec les parties à capacité G.711.1 ainsi que celles seulement capables de G.711. Cela est fait en offrant le sous type de support PCMA en plus de PCMA-WB, et/ou PCMU en plus de PCMU-WB.

Voici un exemple d'une partie d'une telle offre, pour la loi A :

```
m=audio 54874 RTP/AVP 96 8
a=rtpmap:96 PCMA-WB/16000
a=rtpmap:8 PCMA/8000
```

Pour le reste, le format de charge utile pour G.711 est défini au paragraphe 4.5.14 de la [RFC3551].

- o Le paramètre "mode-set" est bidirectionnel ; c'est-à-dire, le mode-set restreint s'applique à la fois au support à recevoir et à celui envoyé par l'entité déclarante. Si un mode-set a été fourni dans l'offre, le répondeur DOIT retourner soit le même mode-set, soit un sous ensemble de ce mode-set. Le répondeur PEUT changer l'ordre de préférence. Si aucun mode-set n'a été fourni dans l'offre, le répondeur PEUT retourner un mode-set pour restreindre les modes possibles. En tous cas, le mode-set dans la réponse s'applique alors pour l'offreur et le répondeur. L'offreur NE DOIT PAS envoyer de trames d'un mode qui a été supprimé par le répondeur.

Pour les sessions de diffusion groupée, si "mode-set" est fourni dans l'offre, le répondeur DEVRA seulement participer à la session si il prend en charge le mode-set offert.

- o Les paramètres "ptime" et "maxptime" vont dans la plupart des cas ne pas affecter l'interopérabilité. Le traitement de l'offre-réponse SDP dans le paramètre "ptime" est décrit dans la [RFC3264]. Le paramètre "maxptime" DOIT être traité de la même façon.
- o Tout paramètre inconnu dans une offre DOIT être ignoré par le receveur et NE DOIT PAS être inclus dans la réponse.

Voici des exemples de parties d'échanges d'offre-réponse SDP.

- o Exemple 1

Offre : tous les modes G.711.1, avec repli sur G.711 ; préfère la loi mu.

```
m=audio 54874 RTP/AVP 96 97 0 8
a=rtpmap:96 PCMU-WB/16000
a=rtpmap:97 PCMA-WB/16000
```

```
a=rtpmap:0 PCMU/8000
a=rtpmap:8 PCMA/8000
```

Réponse : tous les modes acceptés, loi mu et loi A.

```
m=audio 59452 RTP/AVP 96 97
a=rtpmap:96 PCMU-WB/16000
a=rtpmap:97 PCMA-WB/16000
```

o Exemple 2

Offre : tous les modes G.711.1, avec repli sur G.711 ; préfère la loi A.

```
m=audio 54874 RTP/AVP 96 97 8 0
a=rtpmap:96 PCMA-WB/16000
a=rtpmap:97 PCMU-WB/16000
```

Réponse : veut seulement le mode R3 avec loi A.

```
m=audio 59452 RTP/AVP 96
a=rtpmap:96 PCMA-WB/16000
a=fmtp:96 mode-set=4
```

o Exemple 3

Offre : G.711.1 loi A avec deux modes, R2b et R3, avec R3 préféré.

```
m=audio 54874 RTP/AVP 96
a=rtpmap:96 PCMA-WB/16000
a=fmtp:96 mode-set=4,3
```

Réponse : accepté

```
m=audio 59452 RTP/AVP 96
a=rtpmap:96 PCMA-WB/16000
a=fmtp:96 mode-set=4,3
```

Si le répondeur avait voulu restreindre à un mode, il aurait répondu avec seulement une valeur dans le mode-set, par exemple mode-set=3 pour le mode R2b.

5.3.2 Considérations de SDP déclaratif

Pour la déclaration d'utilisation de SDP, rien de spécifique n'est défini pour ce format de charge utile. La configuration donnée par le SDP DOIT être utilisée quand on envoie et/ou reçoit des supports dans la session.

6. Interopérabilité de G.711

La couche L0 de G.711.1 est pleinement interopérable avec G.711, et est incorporée dans tous les modes de G.711.1. Cela donne un processus facile de transcodage de G.711.1 en G.711.

Une passerelle ou tout autre appareil du réseau qui reçoit un paquet G.711.1 peut facilement extraire une charge utile compatible avec G.711, sans avoir besoin de décoder et re-coder le signal audio. Elle a simplement à prendre les données audio de la charge utile, et à supprimer les couches supérieures (L1 et/ou L2) si il en est.

Si un paquet G.711.1 contient plusieurs trames, l'enchaînement des couches L0 de chaque trame va former une charge utile compatible G.711.

7. Contrôle d'encombrement

Le contrôle d'encombrement DEVRA être utilisé pour RTP en accord avec la [RFC3550] et tout profil approprié (par exemple, de la [RFC3551]).

La nature incorporée des données audio G.711.1 peut être utile pour le contrôle d'encombrement, car un mode de codage avec un débit binaire inférieur peut être choisi quand nécessaire. Cette propriété est seulement utilisable quand plusieurs modes ont été négociés (soit pas de paramètre "mode-set" dans le SDP, soit un "mode-set" avec au moins deux modes).

Le nombre de trames encapsulées dans chaque charge utile RTP influence la bande passante globale du flux RTP, du fait des frais généraux d'en-tête. La mise en paquet de plus de trames dans chaque charge utile RTP peut réduire le nombre de paquets envoyés et donc les frais généraux d'en-tête, aux dépens d'un délai accru et d'une robustesse à l'erreur réduite.

8. Considérations sur la sécurité

Les paquets RTP qui utilisent le format de charge utile défini dans cette spécification sont sujets aux considérations générales de sécurité discutées dans la spécification de RTP [RFC3550] et dans tout profil approprié (par exemple, de la [RFC3551]).

Comme ce format transporte de la parole/audio codé, les principales questions de sécurité incluent la confidentialité, la protection de l'intégrité, et l'authentification de la parole/audio elle-même. Le format de charge utile lui-même n'a pas de mécanisme de sécurité incorporé. Tout mécanisme externe convenable, comme le protocole sûr de transport en temps réel (SRTP, *Secure Real-time Transport Protocol*) [RFC3711], PEUT être utilisé.

Ce format de charge utile et le codage G.711.1 ne présentent aucune non uniformité significative dans la charge de calcul de l'extrémité receveuse, et donc ont peu de chances de faire peser une menace de déni de service due à la réception de datagrammes pathologiques. De plus, ils ne contiennent aucun type de contenu actif comme des scripts.

9. Considérations relatives à l'IANA

Deux nouveaux sous types de supports (audio/PCMA-WB et audio/PCMU-WB) ont été enregistrés par l'IANA. Voir les paragraphes 5.1 et 5.2.

10. Références

10.1 Références normatives

- [G.711.1] Recommandation UIT-T G.711.1, "Extension de large bande incorporée pour modulation par impulsions et codage G.711", Union Internationale des Télécommunications, mars 2008.
- [RFC2119] S. Bradner, "[Mots clés à utiliser](#) dans les RFC pour indiquer les niveaux d'exigence", BCP 14, mars 1997. (MàJ par [RFC8174](#))
- [RFC3264] J. Rosenberg et H. Schulzrinne, "[Modèle d'offre/réponse](#) avec le protocole de description de session (SDP)", juin 2002. (P.S. ; MàJ par [RFC8843](#), [9143](#))
- [RFC3550] H. Schulzrinne, S. Casner, R. Frederick et V. Jacobson, "[RTP : un protocole de transport pour les applications](#) en temps réel", STD 64, juillet 2003. (MàJ par [RFC7164](#), [RFC7160](#), [RFC8083](#), [RFC8108](#), [RFC8860](#))
- [RFC3551] H. Schulzrinne et S. Casner, "[Profil RTP pour conférences audio](#) et vidéo avec contrôle minimal", STD 65, juillet 2003. (MàJ par [RFC8860](#))
- [RFC4288] N. Freed et J. Klensin, "Spécifications du [type de support et procédures d'enregistrement](#)", [BCP 13](#), décembre 2005.
- [RFC4566] M. Handley, V. Jacobson et C. Perkins, "SDP : [Protocole de description de session](#)", juillet 2006. (P.S. ; remplacée par [RFC8866](#))
- [RFC4855] S. Casner, "[Enregistrement du type de support](#) des formats de charge utile RTP", février 2007. (P.S.)

10.2 Références pour information

- [G.711] Recommandation UIT-T G.711, "Modulation par impulsions et codage (MIC) des fréquences vocales", UIT-T, novembre 1988.
- [RFC3389] R. Zopf, "[Charge utile du protocole de transport en temps réel](#) (RTP) pour le bruit de confort", septembre 2002.
- [RFC3711] M. Baugher et autres, "Protocole de [transport sécurisé en temps réel](#) (SRTP)", mars 2004. (*P.S. ; MàJ par RFC9335*)

Adresse de l'auteur

Aurélien Sollaud
France Telecom
2 avenue Pierre Marzin
Lannion Cedex 22307
France
téléphone : +33 2 96 05 15 06
mél : aurelien.sollaud@orange-ftgroup.com