

Groupe de travail Réseau  
**Request for Comments : 2658**  
 Catégorie : En cours de normalisation

K. McKay, QUALCOMM Incorporated  
 août 1999  
 Traduction Claude Brière de L'Isle

## Format de charge utile RTP pour l'audio vocal pur

### Statut du présent mémoire

Le présent document spécifie un protocole en cours de normalisation de l'Internet 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 "Normes officielles des protocoles de l'Internet" (STD 1) pour connaître l'état de la 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 (1999). Tous droits réservés.

### Résumé

Le présent document décrit le format de charge utile RTP pour l'audio en voix pure. Le format de paquet accepte un entrelacement variable pour réduire les effets des pertes de paquet sur la qualité audio.

## 1. Introduction

Le présent document décrit comment l'audio de voix pure compressée tel que produit par le codec Qualcomm de voix pure [TIA/EIA/IS-733] peut être formaté pour une utilisation comme type de charge utile RTP. Une méthode est fournie pour entrelacer le résultat de la compression pour réduire la dégradation de qualité due à la perte de paquets. De plus, l'envoyeur peut choisir divers réglages d'entrelacement sur la base de l'importance de faibles délais de bout en bout contre une plus grande tolérance à la perte de paquets.

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

## 2. Fondements

La norme IS-733 [TIA/EIA/IS-733] de l'association des industries électroniques (EIA, *Electronic Industries Association*) et de l'association de l'industrie des télécommunications (TIA, *Telecommunications Industry Association*) définit un algorithme de compression audio à utiliser dans les applications CDMA. En plus d'être le codec standard pour tous les terminaux CDMA sans fil, le codec Qualcomm PureVoice (aussi appelé. Qcelp) est utilisé dans plusieurs applications Internet dont les plus notables sont JFax(tm), Apple(r) QuickTime(tm), et Eudora(r).

Le codec Qcelp [TIA/EIA/IS-733] compresse chaque 20 millisecondes de parole entrée à 8000 Hz, par échantillons de 16 bits en une des quatre tailles de trame de sortie différentes : taux 1 (266 bits), taux 1/2 (124 bits), taux 1/4 (54 bits) ou taux 1/8 (20 bits). Le codec choisit le taux de la trame de sortie sur la base de l'analyse de l'entrée de parole et le mode de fonctionnement en cours (taux normal ou réduit). Pour les schémas de parole normaux, il en résulte une sortie moyenne de 6,8 kbit/s pour le mode normal et 4,7 kbit/s pour le mode à taux réduit.

## 3. Format de paquet RTP/Qcelp

L'horodatage RTP est en unités de 1/8000 de seconde. Les données de charge utile RTP du codec Qcelp ont ce format :

```

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
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     En-tête RTP [RFC1889]                                     |
+=====+=====+=====+=====+=====+=====+=====+=====+
|RR | LLL | NNN |                                                                |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     une ou plusieurs trames de données de codec                                     |
|                                     ....                                         |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

L'en-tête RTP a leurs valeurs prévues telles que décrites dans la [RFC1889]. Le bit d'extension n'est pas établi et ce type de

charge utile n'établit jamais le bit marqueur. Les trames de données de codec sont alignées sur les limites d'octet. Lorsque l'entrelacement est utilisé et/ou lorsque plusieurs trames de données de codec sont présentes dans un seul paquet RTP, l'horodatage est, comme toujours, celui des plus anciennes données représentées dans le paquet RTP. Les autres champs ont la signification suivante :

Réservé (RR) : 2 bits DOIT être réglé à zéro par l'expéditeur, DEVRAIT être ignoré par le récepteur.

Entrelacement (LLL) : 3 bits

DOIT avoir une valeur comprise entre 0 et 5 inclus. Les deux valeurs restantes (6 et 7) NE DOIVENT PAS être utilisées par les expéditeurs. Si ce champ n'est pas zéro, l'entrelacement est activé. Tous les récepteurs DOIVENT prendre en charge l'entrelacement. Les expéditeurs PEUVENT accepter l'entrelacement. Les expéditeurs qui ne prennent pas en charge l'entrelacement DOIVENT régler les champs LLL et NNN à zéro.

Indice d'entrelacement (NNN) : 3 bits

DOIT avoir une valeur inférieure ou égale à la valeur de LLL. Les valeurs de NNN supérieures à la valeur de LLL sont invalides.

### 3.1 Valeurs invalides reçues

À réception d'un paquet RTP qui a une valeur invalide du champ LLL ou NNN, le paquet RTP DOIT être traité comme perdu par le récepteur pour les besoins de la génération de trames d'écrasement, comme décrit à la section 4.

### 3.2 Format de trame de données de codec

Le résultat du codec Qcelp doit être converti en trames de données de codec pour inclusion dans la charge utile RTP comme suit :

- a. L'octet 0 de la trame de données de codec indique le taux et la taille totale de la trame de données de codec comme indiqué dans le tableau suivant :

Octet 0	Taux	Taille totale de trame de données de codec (en octets)
0	blanc	1
1	1/8	4
2	1/4	8
3	1/2	17
4	1	35
5	réservé	8 (DEVRAIT être traité comme une valeur réservée)
14	écrasement	1 (NE DEVRAIT PAS être transmise par l'expéditeur)
autre	n/a	réservé

La réception d'une trame de données de codec avec une valeur réservée dans l'octet 0 DOIT être considérée comme donnée invalide comme décrit en 3.1.

- b. Les bits numérotés selon le standard [TIA/EIA/IS-733] de celui de poids fort à celui de moindre poids sont mis en paquet en octets. Le bit de plus fort numéro (265 pour le taux 1, 123 pour le taux 1/2, 53 pour le taux 1/4 et 19 pour le taux 1/8) est placé dans le bit de poids fort (bit Internet 0) de l'octet 1 de la trame d'octets de codec. Le bit du second plus fort numéro (264 pour le taux 1, etc.) est placé dans le second bit de plus fort poids (bit Internet 1) de l'octet 1 de la trame de données. Cela continue ainsi jusqu'à ce que le bit 258 de la trame standard au taux 1 soit placé dans le bit de moindre poids de l'octet 1. Le bit 257 du standard est placé dans le bit de poids fort de l'octet 2 et ainsi de suite jusqu'à ce que le bit 0 de la trame standard au taux 1 soit placé dans le bit Internet 1 de l'octet 34 de la trame de données de codec. Les bits inutilisés restants du dernier octet de la trame de données de codec DOIVENT être réglés à zéro.

Voici le détail de la façon dont une trame de taux 1/8 est convertie en une trame de données de codec :



LLL = L, NNN = L

Trame L, trame L + L + 1, trame L + 2 (L + 1), trame L + 3 (L + 1), ... pour un total de B trames

Les envoyeurs DOIVENT transmettre dans l'ordre des horodatages croissants. De plus, au sein de chaque groupe d'entrelacement, le paquet RTP qui constitue le groupe d'entrelacement DOIT être transmis dans l'ordre des valeurs croissantes du champ NNN. Bien que cela ne garantisse pas un délai de bout en bout réduit à l'extrémité de réception, lorsque des paquets sont livrés dans l'ordre par le transport sous-jacent, le délai sera réduit au minimum possible.

De plus, les envoyeurs sont soumis aux restrictions suivantes :

- o Une fois commencée une transmission avec un certain SSRC et une certaine valeur d'entrelacement, ils NE DOIVENT PAS augmenter la valeur d'entrelacement. Si la valeur d'entrelacement a besoin d'être augmentée, un nouveau numéro de SSRC DOIT être utilisé.
- o Ils PEUVENT diminuer la valeur d'entrelacement seulement entre les groupes d'entrelacement. Si la valeur d'entrelacement est diminuée, elle NE DOIT PAS être augmentée (même à la valeur d'origine) bien qu'elle puisse encore être diminuée ultérieurement.

### 3.5 Trouver les limites de groupe entrelacé

Étant donné un paquet RTP avec le numéro de séquence S, la valeur d'entrelacement (champ LLL) L, et la valeur d'indice d'entrelacement (champ NNN) N, le groupe d'entrelacement consiste en paquets RTP avec des numéros de séquence de S-N à S-N+L inclus. En d'autres termes, le groupe d'entrelacement consiste toujours en L+1 paquets RTP avec des numéros de séquence qui se suivent. La valeur du faisceau pour tous les paquets RTP dans un groupe d'entrelacement DOIT être la même.

Le receveur détermine la valeur de faisceau attendue pour tous les paquets RTP dans un groupe d'entrelacement par le nombre de trames de données de codec en faisceau dans le premier paquet RTP du groupe d'entrelacement reçu. Noter que ce peut n'être pas le premier paquet RTP du groupe d'entrelacement si les paquets sont livrés déclassés par le transport sous-jacent.

À réception d'un paquet RTP dans un groupe d'entrelacement avec d'autres valeurs de faisceau que celles attendues, le receveur PEUT éliminer les trames de données de codec de la fin du paquet RTP ou ajouter des trames de données de codec d'écrasement à la fin du paquet afin de fabriquer un paquet de substitution avec la valeur de faisceau attendue. Le receveur PEUT aussi choisir de plutôt éliminer le groupe d'entrelacement tout entier et faire un silence.

### 3.6 Reconstruction de l'audio entrelacé

Soit un ensemble de paquets RTP ordonné par numéro de séquence RTP dans un groupe d'entrelacement numéroté de 0 à L, où L est la valeur d'entrelacement et B est la valeur de faisceau, et des trames de données de codec au sein de chaque paquet RTP qui sont numérotées dans l'ordre de la première à la dernière avec les numéros de 1 à B, la séquence originale, ordonnée selon le temps, des trames de sortie du codec, peut être reconstruite comme suit :

L + 1 premières trames :

Trame 0 provenant du paquet 0 du groupe d'entrelacement

Trame 0 provenant du paquet 1 du groupe d'entrelacement

et ainsi de suite jusqu'à ...

Trame 0 provenant du paquet L du groupe d'entrelacement

Secondes L + 1 trames :

Trame 1 du groupe d'entrelacement 0 du groupe d'entrelacement

Trame 1 du groupe d'entrelacement 1 du groupe d'entrelacement

et ainsi de suite jusqu'à ...

Trame 1 du groupe d'entrelacement L du groupe d'entrelacement

et ainsi de suite jusqu'à ...

B<sup>èmes</sup> L + 1 trames:

Trame B du groupe d'entrelacement 0 du groupe d'entrelacement

Trame B du groupe d'entrelacement 1 du groupe d'entrelacement

et ainsi de suite jusqu'à ...

Trame B du groupe d'entrelacement L du groupe d'entrelacement

### 3.6.1 Responsabilités supplémentaires du receveur

Supposons que le receveur a commencé d'exécuter les trames provenant d'un groupe d'entrelacement. Le moment est venu d'exécuter la trame  $x$  provenant du paquet  $n$  du groupe d'entrelacement. Supposons de plus que le paquet  $n$  du groupe d'entrelacement n'a pas été reçu. Comme décrit à la section 4, une trame d'écrasement va être envoyée au codec Qcelp.

Maintenant, supposons que le paquet  $n$  du groupe d'entrelacement arrive avant que la trame  $x+1$  de ce paquet soit nécessaire. Les receveurs DEVRAIENT utiliser la trame  $x+1$  du paquet  $n$  nouvellement reçu plutôt que de substituer une trame d'écrasement. En d'autres termes, juste parce que le paquet  $n$  n'a pas été disponible la première fois qu'on en avait besoin pour reconstruire l'audio entrelacé, le receveur NE DEVRAIT PAS supposer qu'il ne sera disponible lorsque il sera ensuite nécessaire pour la reconstruction de l'audio entrelacé.

## 4. Traitement des paquets RTP perdus

Le codec Qcelp prend en charge la notion de trames d'écrasement. Ce sont des trames qui, pour une raison ou une autre, ne sont pas disponibles. Lors de la reconstruction de l'audio entrelacé ou de l'exécution d'audio non entrelacé, les trames d'écrasement DOIVENT être fournies au codec Qcelp pour tous les paquets manquants.

Les receveurs DOIVENT utiliser l'horloge d'horodatage pour déterminer combien de trames de données de codec sont manquantes. Chaque trame de données de codec avance l'horloge d'horodatage de EXACTEMENT 160 comptes.

Comme la valeur de faisceau peut varier (elle peut seulement diminuer) l'horloge d'horodatage est la seule façon fiable de calculer exactement combien de trames de données de codec manquent lorsque un paquet est abandonné.

Précisément, lors de la reconstruction d'audio entrelacé, un paquet RTP manquant dans le groupe d'entrelacement devrait être traité comme contenant  $B$  trames de données de codec d'écrasement où  $B$  est la valeur du faisceau pour ce groupe d'entrelacement.

## 5. Discussion

Le codec Qcelp interpole le contenu audio manquant lorsque on donne une trame d'écrasement. Cependant, la meilleure qualité est perçue par celui qui écoute lorsque les trames d'écrasement ne sont pas consécutives. Cela rend l'entrelacement souhaitable car il augmente la qualité audio lorsque des abandons de paquets sont probables.

D'un autre côté, l'entrelacement peut considérablement augmenter le délai de bout en bout. Lorsque on désire une session interactive, une valeur d'entrelacement (champ LLL) de 0 ou 1 et un facteur de mise en faisceau de 4 ou moins sont recommandées.

Lorsque le délai de bout en bout n'est pas un problème, une valeur de mise en faisceau d'au moins 4 et une valeur d'entrelacement (champ LLL) de 4 ou 5 sont recommandées sous réserve des limitations de MTU.

Les restrictions concernant l'envoyeur mentionnées aux paragraphes 3.3 et 3.4 garantissent qu'après réception du premier paquet de charge utile de la part de l'envoyeur, le receveur peut allouer une quantité bien connue d'espace de mémoire tampon qui sera suffisant pour toutes les réceptions futures provenant de la même valeur de SSRC. Moins d'espace de mémoire tampon pourrait être requis à l'avenir si l'envoyeur diminue la valeur du faisceau ou de l'entrelacement, mais jamais plus d'espace mémoire. Cela empêche la possibilité que le receveur ayant besoin d'allouer plus d'espace de mémoire tampon (avec le résultat possible qu'il n'en soit pas de disponible) si la valeur de faisceau ou d'entrelacement était augmentée par l'envoyeur. Aussi, si la valeur d'entrelacement ou de faisceau devait augmenter, le receveur pourrait être forcé de suspendre l'exécution lorsque il reçoit les paquets additionnels nécessaires pour l'exécution à une valeur plus élevée de mise en faisceau ou d'entrelacement.

## 6. Considérations pour la sécurité

Les paquets RTP qui utilisent le format de charge utile défini dans la présente spécification sont sujets aux considérations de sécurité discutées dans la spécification RTP [RFC1889], et tout profil approprié (par exemple [RFC1890]). Cela implique que la confidentialité des flux du support soit réalisée par chiffrement. Comme la compression des données utilisée avec ce format de charge utile est appliquée de bout en bout, le chiffrement peut être effectué après compression, de sorte qu'il n'y a pas de conflit entre les deux opérations.

Une menace potentielle de déni de service existe pour les codages de données qui utilisent des techniques de compression qui ont une charge de calcul non uniforme à l'extrémité receveuse. L'attaquant peut injecter des datagrammes pathologiques dans le flux et qui sont complexes à décoder et causent la surcharge du receveur. Cependant, ce codage n'exhibe aucune non uniformité significative.

Comme avec tout protocole fondé sur IP, dans certaines circonstances, un receveur peut être surchargé simplement par la réception de trop de paquets, désirés ou indésirables. L'authentification de couche réseau peut être utilisée pour éliminer les paquets provenant de sources non désirées, mais le coût de traitement de l'authentification elle-même peut être trop élevé. Dans un environnement de diffusion groupée, l'élagage de sources spécifiques peut être mis en œuvre dans de futures versions de IGMP [RFC1112] et dans les protocoles d'acheminement de diffusion groupée pour permettre à un receveur de choisir quelles sources sont autorisées à le joindre.

## 7. Références

- [TIA/EIA/IS-733] TR45: High Rate Speech Service Option for Wideband Spread Spectrum Communications Systems. Disponible auprès de Global Engineering +1 800 854 7179 ou +1 303 792 2181. Peut aussi être commandé en ligne à <http://www.eia.org/eng/>.
- [RFC1112] S. Deering, "Extensions d'hôte pour [diffusion groupée sur IP](#)", STD 5, août 1989. (*MàJ par RFC2236*)
- [RFC1889] H. Schulzrinne, S. Casner, R. Frederick et V. Jacobson, "RTP : protocole de transport pour applications en temps réel", janvier 1996. (*Obsolète, voir RFC3550 STD64*)
- [RFC1890] H. Schulzrinne, "Profil RTP pour conférences audio et vidéo avec contrôle minimal", janvier 1996. (*Obsolète, voir RFC3551*) (P.S.)
- [RFC2119] S. Bradner, "[Mots clés à utiliser](#) dans les RFC pour indiquer les niveaux d'exigence", BCP 14, mars 1997.

## 8. Adresse de l'auteur

Kyle J. McKay  
QUALCOMM Incorporated  
5775 Morehouse Drive  
San Diego, CA 92121-1714  
USA  
téléphone : +1 858 587 1121  
mél : [kylem@qualcomm.com](mailto:kylem@qualcomm.com)

## 9. Déclaration complète de droits de reproduction

Copyright (C) The Internet Society (1999). Tous droits réservés.

Ce document et les traductions de celui-ci peuvent être copiés et diffusés, et les travaux dérivés qui commentent ou expliquent autrement ou aident à sa mise en œuvre peuvent être préparés, copiés, publiés et distribués, partiellement ou en totalité, sans restriction d'aucune sorte, à condition que l'avis de droits de reproduction ci-dessus et ce paragraphe soient inclus sur toutes ces copies et œuvres dérivées. Toutefois, ce document lui-même ne peut être modifié en aucune façon, par exemple en supprimant le droit d'auteur ou les références à l'Internet Society ou d'autres organisations Internet, sauf si c'est nécessaire à l'élaboration des normes Internet, auquel cas les procédures pour les droits de reproduction définis dans les procédures des normes d'Internet doivent être suivies, ou si nécessaire pour le traduire dans des langues autres que l'anglais.

Les permissions limitées accordées ci-dessus sont perpétuelles et ne seront pas révoquées par la Société Internet ou ses successeurs ou ayants droit.

Ce document et les renseignements qu'il contient sont fournis "TELS QUELS" et l'INTERNET SOCIETY et l'INTERNET ENGINEERING TASK FORCE déclinent toute garantie, expresse ou implicite, y compris mais sans s'y limiter, toute garantie que l'utilisation de l'information ici présente n'enfreindra aucun droit ou aucune garantie implicite de commercialisation ou d'adaptation à un objet particulier.

**Remerciement**

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