

Groupe de travail Réseau  
**Request for Comments : 4396**  
 Catégorie : Sur la voie de la normalisation  
 Traduction Claude Brière de L'Isle

J. Rey, Panasonic  
 Y. Matsui, Panasonic  
 février 2006

## Format de charge utile RTP pour texte synchronisé du projet en partenariat de 3e génération (3GPP)

### 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 RTP pour la transmission de textes synchronisés 3GPP (projet en partenariat de 3ème génération). Le texte synchronisé 3GPP est un format de texte décoré qui a des contraintes de temps et une mémorisation définie dans un fichier 3GP. Le texte synchronisé peut être synchronisé avec des contenus audio/vidéo et utilisé dans des applications comme l'ajout d'en-têtes, de titres, et de présentations multimédia. Dans les sections suivantes, on traite les problèmes du texte en flux synchronisés, et du format des charges utiles pour les flux de texte synchronisé 3GPP sur RTP.

### Table des matières

1. Introduction.....	2
2. Motivation, exigences, et raisons du concept.....	2
2.1 Motivation.....	2
2.2 Composants de base du format de support de texte synchronisé 3GPP.....	2
2.3 Exigences.....	3
2.4 Limitations.....	4
2.5 Raisons du concept.....	4
3. Terminologie.....	5
4. Format de charge utile RTP pour texte synchronisé 3GPP.....	6
4.1 Définitions d'en-tête de charge utile.....	7
4.2 Mise en mémoire tampon des descriptions d'échantillon.....	14
4.3 Trouver les valeurs d'en-tête de charge utile dans les fichiers 3GP.....	15
4.4 Fragmentation des échantillons de texte synchronisé.....	17
4.5 Réassemblage des échantillons de texte chez le receveur.....	18
4.6. Charges utiles agrégées.....	19
4.7 Exemples de charges utiles.....	21
4.8 Relations avec la RFC 3640.....	24
4.9 Relations avec la RFC 2793.....	25
5. Transport résilient.....	25
6. Contrôle d'encombrement.....	26
7. Description d'un scénario.....	26
7.1 Position et composition d'un rendu de texte.....	26
7.2 Usage de SMIL.....	27
7.3 Trouver les valeurs de présentation dans un fichier 3GP.....	28
8. Type de support texte synchronisé 3GPP.....	28
9. Usage de SDP.....	30
9.1 Transposition en SDP.....	30
9.2 Usage de paramètre dans le modèle SDP d'offre/réponse.....	30
9.3 Exemples d'offre/réponse.....	32

9.4 Usage de paramètres en dehors de offre/réponse.....	34
10. Considérations relatives à l'IANA.....	34
11. Considérations sur la sécurité.....	34
12. Références normatives.....	34
12.1 Références normatives.....	34
12.2 Références pour information.....	35
13. Bases de la structure de fichier 3GP.....	36
14. Remerciements.....	36
Adresse des auteurs.....	37
Déclaration complète de droits de reproduction.....	37

## 1. Introduction

Le texte synchronisé 3GPP est un format de support pour du texte décoré qui a des contraintes de temps spécifié dans la spécification technique du 3GPP TS 26.245, "Service transparent de flux de paquets commutés de bout en bout (PSS, *packet switched streaming service*) ; format de texte synchronisé (TTF, *Timed Text Format*) (Livraison 6)" [PSS-TTF]. À côté du texte en clair, le format de texte synchronisé 3GPP permet la création de texte décoré comme pour les applications de karaoke, de texte déroulant pour les nouvelles, ou du texte avec des hyperliens. Ces contenus peuvent être ou non synchronisés avec d'autres supports, comme de l'audio ou de la vidéo.

L'objet du présent document est de fournir le moyen d'écouler des contenus de texte synchronisé 3GPP en utilisant RTP [RFC3550]. Cela inclut l'écoulement de texte synchronisé lu à partir d'un fichier 3GP, ainsi que l'écoulement de texte synchronisé généré en temps réel, autrement dit de l'écoulement en direct.

La Section 2 contient la motivation du document, une vue d'ensemble du format de support, les exigences, et les raisons de conception. La Section 3 définit la terminologie utilisée. La Section 4 spécifie les en-têtes de charges utiles, les règles de fragmentation et de réassemblage pour les échantillons de texte, les règles d'agrégation des charges utiles, et les relations de ce document avec la [RFC3640] et la [RFC2793]. La Section 5 spécifie des schémas simples pour un transport résilient et donne des pointeurs sur d'autres mécanismes possibles. La Section 6 traite du contrôle d'encombrement. La Section 7 spécifie une description de scénario. La Section 8 définit le type de support. La Section 9 spécifie SDP pour les sessions en envoi individuel et en diffusion groupée, incluant l'usage du modèle d'offre/réponse [RFC3264]. Les Sections 10 et 11 traitent des considérations relatives à l'IANA et à la sécurité. La Section 12 fait la liste des références. Les bases de la structure de fichier 3GP forment la Section 13.

## 2. Motivation, exigences, et raisons du concept

### 2.1 Motivation

Le format de texte synchronisé 3GPP a été développé pour l'usage des services spécifiés dans la spécification des services de flux de paquets commutés transparents de bout en bout du 3GPP (3GPP PSS, *3GPP Transparent End-to-end Packet-switched Streaming Services*) [PSS6].

À ce jour, PSS permet le téléchargement de contenus de texte synchronisé 3GPP mémorisés dans des fichiers 3GP. Cependant, du fait de l'absence d'un format de charge utile RTP, il n'est pas possible d'écouler des contenus de texte synchronisé 3GPP sur RTP.

Le présent document spécifie un tel format de charge utile.

### 2.2 Composants de base du format de support de texte synchronisé 3GPP

Avant d'entrer dans les détails de la conception, il est nécessaire de savoir comment est construit le format de support. On peut identifier quatre composants fonctionnels différents : les informations de présentation, le formatage par défaut, les chaînes de texte, et la décoration. Dans ce qui suit, on explique brièvement cela et on les confronte à leurs désignations dans un fichier 3GP :

- o Informations initiales de représentation spatiale relatives aux chaînes de texte : ce sont la hauteur et la largeur de la région de texte où le texte est affiché, la position de la région de texte dans l'affichage, et la couche ou proximité du texte pour l'utilisateur. Dans les fichiers 3GP, ces informations sont contenues dans la Boîte d'en-tête de piste (*Track*

*Header Box*) (les désignations de fichier 3GP sont en majuscules pour être clair).

- o Réglages par défaut pour le formatage et le positionnement du texte : style (fonte, taille, couleur,...) couleur du fond, justification horizontale et verticale, largeur de ligne, enroulement, etc. Pour les fichiers 3GP, cela correspond aux descriptions d'échantillon.
- o Chaînes de texte réelles : caractères codés avec UTF-8 [RFC3629] ou UTF-16 [RFC2781].
- o Décoration : si des caractères ont un style, délai, brillance, différents, etc., cela doit être indiqué. La décoration n'est présente dans les échantillons de texte que si elle est réellement nécessaire. Autrement, les réglages par défaut ci-dessus s'appliquent. Dans les fichiers 3GP, au sein de chaque échantillon de texte, la décoration (c'est-à-dire, les boîtes de modificateur) est ajoutée aux chaînes de texte, si nécessaire. Au moment de la rédaction de ce format de charge utile, les modificateurs suivants sont spécifiés dans la spécification de format de support de texte synchronisé 3GPP [PSS-TTF] :
  - surbrillance du texte
  - surbrillance de couleur
  - texte clignotant
  - caractéristiques de karaoke
  - hyperlien
  - texte retardé
  - style de texte
  - positionnement de boîte de texte
  - indication d'enveloppement de texte

### 2.3 Exigences

Une fois que les composants de base sont connus, il est nécessaire de définir à quelles exigences doit satisfaire le format de charge utile :

1. Il doit permettre à la fois les flux en direct et les flux à partir d'un fichier 3GP.

Note : pour les besoins de ce document, le terme "flux en direct" se réfère aux scénarios où le flux de texte synchronisé est envoyé à partir d'un codeur direct. À réception, le contenu peut ou non être mémorisé dans un fichier 3GP. Normalement, dans les applications de flux directs, l'expéditeur encapsule le contenu de texte synchronisé en paquets RTP suivant les lignes directrices données dans le présent document. Du côté receveur, une mémoire tampon est utilisée pour annuler les retards du réseau et la gigue de retard. Si le receveur et l'expéditeur prennent en charge les mécanismes de résilience à la perte de paquet (voir la paragraphe 5) il peut aussi être possible de récupérer des pertes de paquets. Noter que la façon dont l'expéditeur et le receveur gèrent et dimensionnent en fait les mémoires tampon est un choix de conception des mises en œuvre.

2. De plus, il devra être possible à un receveur RTP qui utilise ce format de charge utile, et capable de mémoriser en format 3GP, d'obtenir toutes les informations nécessaires des paquets RTP pour mémoriser les contenus de texte reçus conformément au format du fichier 3GP. Ce fichier peut ou non être le même que le fichier d'origine.

Note : le format de fichier 3GP lui-même se fonde sur la recommandation ISO sur le format de base de support de fichier [ISO14496-12]. Le paragraphe 13.1 donne des indications sur la structure du fichier 3GP. De plus, les paragraphes 4.3 et 7.3 spécifient où se trouvent les informations nécessaires pour remplir les en-têtes de charge utile dans un fichier 3GP. Pour les flux directs, les valeurs appropriées conformes au format et unités décrits dans [PSS-TTF] devront être utilisées. Lorsque nécessaire, des précisions sur les valeurs appropriées sont données dans ce document.

3. Il devra permettre un transport efficace et résilient des contenus de texte synchronisé sur RTP. En particulier :
  - a. Permettre la transmission des descriptions d'échantillons par des moyens aussi bien dans la bande que hors bande. Les descriptions d'échantillon sont des informations importantes, qui s'appliquent potentiellement à plusieurs échantillons de texte. Ces réglages de formatage par défaut sont normalement transmis hors bande (de façon fiable) une fois à la phase d'initialisation. Si des descriptions d'échantillons supplémentaires sont nécessaires dans le cours d'une session, elles peuvent aussi être envoyées hors bande ou dans la bande. La transmission dans la bande bien que non fiable, peut être plus appropriée pour l'envoi des descriptions d'échantillons si elles doivent être fréquentes, par opposition à l'établissement d'un canal de communication supplémentaire pour SDP, par exemple. Elle est aussi utile dans les cas où un canal hors bande ne peut pas être disponible et pour les flux de direct, où les contenus ne sont pas connus a priori. Donc, le format de charge utile devra permettre la transmission hors bande et dans la bande des descriptions d'échantillons. Le paragraphe 4.1.6 spécifie un en-tête de charge utile pour transmettre les

descriptions d'échantillons dans la bande. La Section 9 spécifie comment les descriptions d'échantillons sont transposées en SDP.

- b. Permettre la fragmentation d'un échantillon de texte en plusieurs paquets RTP afin de couvrir une large gamme d'applications et environnements de réseau. En général, la fragmentation devrait être un événement rare, étant donné les faibles débits binaires et les relativement petites tailles d'échantillon de texte. Cependant, le format de support de texte synchronisé 3GPP permet de plus grands échantillons de texte. Donc, le format de charge utile devra prendre en compte cela et donner le moyen de s'accommoder de la fragmentation et du réassemblage. Le paragraphe 4.4 traite de la fragmentation.
- c. Permettre l'agrégation des unités dans un paquet RTP pour rendre le transport plus efficace. Dans un environnement de communication mobile, une taille d'échantillon de texte typique est d'environ 100 à 200 octets. Si le débit binaire disponible et la taille de paquet le permettent, les unités devraient être agrégées en un paquet RTP. Le paragraphe 4.6 traite de l'agrégation.
- d. Permettre l'utilisation de mécanismes de transport résilients, comme la répétition, la retransmission [RFC4588], et la FEC [RFC2733] (voir la paragraphe 5). Pour une discussion plus générale, se référer à la [RFC2354], qui discute des mécanismes disponibles pour la réparation des flux.

## 2.4 Limitations

La taille des en-têtes de charge utile a été optimisée pour RTP. Au lieu d'utiliser des champs d'en-tête (S)LEN, SDUR, et SIDX de 32 bits, qui porteraient de nombreux bits inutilisés la plupart du temps, il a été choisi dès la conception de réduire la taille de ces champs. Par conséquent, ce format de charge utile a des valeurs maximum réduites par rapport aux tailles et durées des échantillons (de texte) et aux descriptions d'échantillon. Ces valeurs maximum diffèrent de celles allouées dans les fichiers 3GP, où elles sont exprimées en utilisant des entiers de 32 bits (non signés). Dans certains cas, des mécanismes d'extension sont fournis pour traiter de plus grandes valeurs. Cependant, on note que les valeurs utilisés ici devraient être suffisantes pour les applications de flux directs ciblées.

Les limitations suivantes s'appliquent :

1. La taille maximum d'échantillons de texte portés dans les paquets RTP est restreinte à des entiers de 16 bits non signés (cela inclut les chaînes de texte et modificateurs). Cela signifie une taille maximum pour l'unité de l'ordre de 64 k octets. Aucun mécanisme d'extension n'est fourni.
2. Les valeurs d'indice de description d'échantillon sont restreintes à des entiers de 8 bits non signés. Un mécanisme d'extension est donné au paragraphe 4.3.
3. La durée de l'échantillon de texte est restreinte à des entiers de 24 bits non signés. Cela donne une durée maximum au débit d'horloge d'horodatage de 1000 Hz d'environ 4,6 heures. Néanmoins, un mécanisme d'extension est donné au paragraphe 4.3.
4. Les descriptions d'échantillon sont aussi réduites en taille : si la taille ne peut pas être exprimée par un entier non signé de 16 bits, la description d'échantillon ne devra pas être transportée. Comme dans le cas de la taille d'échantillon, aucun mécanisme d'extension n'est fourni.
5. Une limitation supplémentaire concerne la prise en charge des codages UTF-16 : seul le transport de chaînes de texte suivant l'ordre gros boutien est accepté. Voir les détails au paragraphe 4.1.1.

## 2.5 Raisons du concept

Les choix de conception suivants ont été faits.

1. Approche de "l'unité" : les formats de charge utile spécifiés dans le présent document suivent un schéma simple : un en-tête commun de 3 octets suivi par un en-tête spécifique pour chaque type d'échantillon de texte (fragment). À la suite de ces en-têtes est placé le contenu de l'échantillon de texte (paragraphe 4.1.1 et suivants). Cette structure est appelée une "unité".

Les unités suivantes ont été conçues pour se conformer aux exigences mentionnées au paragraphe 2.3 :

- a. l'unité de type 1 qui contient un échantillon de texte complet,
- b. l'unité de type 2 qui contient une chaîne de texte complète ou un fragment de celle-ci,
- c. l'unité de type 3 qui contient les modificateurs complets ou seulement leur premier fragment,

- d. l'unité de type 4 qui contient un fragment de modificateur autre que le premier,
- e. l'unité de type 5 qui contient une description d'échantillon.

Cette approche par "unité" est motivée par les raisons suivantes :

- Elle permet une classification simple des échantillons de texte et des fragments d'échantillon de texte qui peuvent être portés par le format de charge utile.
- Elle permet une interopérabilité facile avec la [RFC3640]. Durant le développement de ce format de charge utile, de l'intérêt a été montré par les participants à la normalisation de MPEG-4 pour développer une structure commune de charge utile pour le transport du texte synchronisé 3GPP. Bien que l'interopérabilité ne soit pas strictement nécessaire pour que ce format de charge utile fonctionne, elle a été recherchée dans ce format de charge utile. Le paragraphe 4.8 explique comment cela est fait.

2. Le compte de caractères n'est pas mis en œuvre. Ce format de charge utile ne détecte pas les pertes de fragments d'échantillons de texte, mais il ne permet pas à un receveur RTP de trouver le nombre exact de caractères de texte perdus. En fait, la taille de fragment incluse dans les en-têtes de charge utile n'aide pas à trouver le nombre de caractères perdus parce que les codages UTF-8/UTF-16 [RFC3629], [RFC2781] utilisés donnent un nombre variable d'octets par caractère.

Pour trouver le nombre exact de caractères perdus, un champ supplémentaire reflétant le compte de caractères (et éventuellement le décalage du caractère) lors d'une fragmentation serait nécessaire. Cela exigerait de plus que l'entité qui effectue la fragmentation compte les caractères inclus dans chaque fragment de texte.

Un avantage d'avoir un compte de caractères serait que l'application d'affichage serait capable de remplacer les caractères manquants par d'autres caractères représentant les caractères perdus. Par exemple, si on prend le texte "Du texte a été perdu maintenant" et qu'on suppose la perte d'un paquet contenant le texte du milieu, cela pourrait être affiché (avec le compte de caractères) : "Du #####maintenant" au lieu de : "Du #maintenant", qui est ce que ce format de charge utile permet ("#" indique un caractère ou paquet manquant).

Cependant, le consensus du groupe travail est que pour des applications comme le sous titrage ou les présentations multimédia qui utilisent ce format de charge utile, une correction d'erreur partielle ne vaut pas le coût d'inclure deux champs supplémentaires, à savoir le compte de caractères et le décalage de caractères. À la place, il est recommandé que des frais généraux soient investis pour fournir une correction d'erreur complète en protégeant le moins de fragments d'échantillon de texte en utilisant les mesures définies à la paragraphe 5.

3. Le réassemblage de fragment : afin de ré-assembler les échantillons de texte, les informations de décalage sont nécessaires. Au lieu d'un décalage de caractère ou d'octet, un seul octet, Total/CeluiCi, est utilisé. Ces deux valeurs indiquent le nombre total et l'indice courant des fragments d'un échantillon de texte. C'est plus simple que d'avoir un champ de décalage de caractères dans chaque fragment. Les détails sont au paragraphe 4.1.3.
4. Un champ de longueur, Longueur, est présent dans les champs d'en-tête communs. Bien que la longueur dans le format de charge utile RTP ne soit pas nécessaire pour la plupart des applications RTP (normalement, les couches inférieures, comme UDP, fournissent ces informations) cela facilite l'interopérabilité avec la RFC 3640. C'est parce que les unités d'accès (AU, *Access Unit*) utilisées pour le portage des données dans la RFC 3640 doivent inclure une indication de longueur. Les détails sont au paragraphe 4.8.
5. Les champs d'en-tête dans les en-têtes de charge utile spécifiques (en-têtes Type dans les paragraphes 4.1.2 à 4.1.6) ont été arrangés pour un traitement facile sur les machines à 32 bits. Pour cette raison, les champs SIDX et SDUR sont passés à l'unité de Type 1, par rapport aux autres unités.

### 3. 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].

De plus, les termes suivants ont une signification spécifique dans le contexte du présent document :

échantillon de texte ou échantillon de texte complet : dans le format de support de texte synchronisé 3GPP [PSS-TTF], ces termes se réfèrent à une unité de données de texte synchronisé comme contenue dans le fichier source (3GP). Cela

inclut le compte d'octets de la chaîne de texte, éventuellement une marque d'ordre d'octet (BOM, *Byte Order Mark*) la chaîne de texte et tous les modificateurs qui peuvent suivre. Son équivalent en audio/vidéo serait une trame. Dans le présent document, cependant, un échantillon de texte contient seulement des chaînes de texte suivies de zéro, un ou plusieurs modificateurs. Cette définition d'échantillon de texte exclut le compte de 16 bits des octets de la chaîne de texte et la BOM de 16 bits de la marque d'ordre d'octet (BOM) présents dans les échantillons de texte de fichier 3GP (voir le paragraphe 4.3 et la Figure 9). La BOM de 16 bits n'est pas transportée dans RTP, comme expliqué au paragraphe 4.1.1.

chaînes de texte : caractères réels du texte codés en UTF-8 ou UTF-16. Quand on utilise ce format de charge utile, la chaîne de texte ne contient aucune marque d'ordre d'octet (BOM). Voir les détails à la Figure 9.

fragment ou fragment d'échantillon de texte : fraction d'un échantillon de texte. Un fragment peut contenir des chaînes de texte ou un contenu de modificateur (décoration) mais pas les deux en même temps.

contenu d'échantillon : terme général pour identifier les données de texte synchronisé transportées en utilisant ce format de charge utile. Le contenu d'échantillon peut être un ou plusieurs échantillons de texte, descriptions d'échantillons, et fragments d'échantillon (noter que, selon le paragraphe 4.6, il n'y a qu'un seul cas où plus d'un fragment peut être inclus dans une charge utile).

décoration ou modificateurs : ces termes sont utilisés de façon interchangeable dans le document pour noter le contenu de l'échantillon de texte qui modifie le formatage de texte par défaut. Les modificateurs peuvent, par exemple, spécifier une taille de fonte différente pour une séquence de caractères particulière ou définir le rythme de karaoke pour l'échantillon.

description d'échantillon : informations qui sont potentiellement partagées par plus d'un échantillon de texte. Dans un fichier 3GP, une description d'échantillon est mémorisée dans un endroit où elle peut être partagée. Elle contient des informations d'établissement et par défaut comme la direction de déroulement, la position de la boîte de texte, la valeur du retard, la fonte par défaut, la couleur de fond, etc.

unités ou unités de transport : les en-têtes de charge utile spécifiés dans ce document encapsulent des échantillons de texte, leurs fragments, et les descriptions d'échantillons en plaçant devant eux un en-tête commun et un en-tête spécifique de charge utile (paragraphe 4.1.1 à 4.1.6) construisant ainsi ce qu'on appelle ici une unité (de transport).

agrégation ou paquet agrégé : la charge utile d'un paquet agrégé (RTP) consiste en plusieurs unités (de transport).

piste ou flux : les fichiers 3GP contiennent des pistes audio/vidéo et du texte. Le présent document permet l'écoulement de pistes de texte avec RTP. Donc, ces termes sont utilisés de façon interchangeable dans ce document dans le contexte des fichiers 3GP.

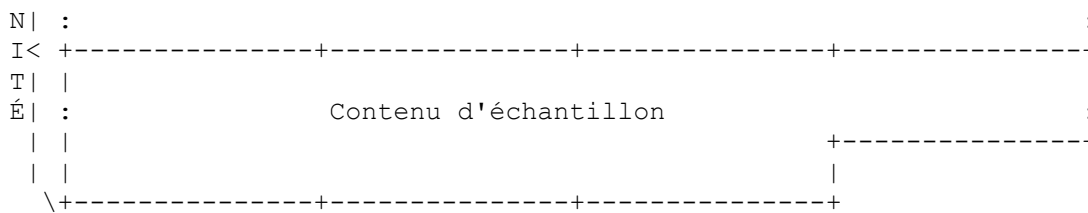
Boîte d'en-tête de support / boîte d'en-tête de piste / ... : le format de fichier 3GP utilise ces structures définies dans le format de fichier de base ISO [ISO14496-12]. Dans ce document les références à ces termes commencent par une majuscule par souci de clarté.

#### 4. Format de charge utile RTP pour texte synchronisé 3GPP

Le format d'un paquet RTP contenant du texte synchronisé 3GPP est comme ci-dessous :

```

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
+-----+-----+-----+-----+-----+-----+-----+-----+
|V=2|P|X| CC   |M|   PT       |       Numéro de séquence       |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Horodatage                               |
+-----+-----+-----+-----+-----+-----+-----+-----+
|   Identifiant de source de synchronisation source (SSRC)   |
/-----+-----+-----+-----+-----+-----+-----+
| |U|   R   | Type|           Longueur           |           :
| +-----+-----+-----+-----+-----+-----+           :
U| :           (champs d'en-tête variables selon le Type)           :
```



**Figure 1. Format de paquet RTP de texte synchronisé 3GPP**

Bit marqueur (M) : il DEVRA être établi à 1 si le paquet RTP inclut un ou plusieurs échantillons de texte complets ou le dernier fragment d'un échantillon de texte ; autrement, il est réglé à zéro (0).

Horodatage : il DOIT indiquer l'instant d'échantillonnage de la première (ou seule) unité contenue dans le paquet RTP. La valeur initiale DEVRAIT être déterminée au hasard, comme spécifié dans RTP [RFC3550]. La valeur de l'horodatage devrait fournir assez de résolution temporelle pour exprimer la durée des échantillons de texte, pour synchroniser le texte avec les autres supports, et pour effectuer les mesures du protocole de contrôle RTP (RTCP) comme la gigue de délai inter arrivées ou le bloc de rapport de réception de paquet RTCP (paragraphe 4.3 de la [RFC3611]). Ceci est conforme à RTP, paragraphe 5.1 : "La résolution de l'horloge DOIT être suffisante pour la précision de synchronisation désirée et pour mesurer la gigue d'arrivée des paquets (un tic par trame vidéo n'est normalement pas suffisant)". Cette observation s'applique aux pistes de texte synchronisé incluses dans un fichier 3GP et aux sessions de direct. Dans le cas d'une piste de texte synchronisé 3GP, la granularité de l'horodatage est la valeur du paramètre "échelle de temps" dans la boîte d'en-tête de support pour cette piste de texte. Chaque piste dans un fichier 3GP PEUT avoir son propre débit d'horloge comme spécifié dans la boîte d'en-tête de support. De même, les applications de flux de direct DEVRONT utiliser un débit d'horloge d'horodatage approprié. Une valeur par défaut de 1000 Hz est RECOMMANDÉE. D'autres débits d'horloge d'horodatage PEUVENT être utilisés. Dans ce cas, le comportement normal est ici de confronter le débit d'horloge du texte synchronisé 3GPP avec celui utilisé par un flux audio ou vidéo associé. Dans une charge utile agrégée, les unités DOIVENT être placées dans l'ordre d'exécution, c'est-à-dire, les plus tôt les premières dans la charge utile. Si des unités de type 1 sont agrégées, l'horodatage des unités suivantes DOIT être obtenu en ajoutant la durée de l'échantillon de texte synchronisé des échantillons précédents à la valeur de l'horodatage RTP. Il y a deux exceptions à cette règle : les unités de type 5 et une charge utile agrégée contenant deux fragments du même échantillon de texte. Les détails du calcul de l'horodatage sont donnés au paragraphe 4.6. Finalement, les débits d'horloge d'horodatage DOIT être signés par des moyens hors bande à l'établissement de la session, par exemple, en utilisant le paramètre de type de support "rate" dans SDP. Voir les détails à la Section 9.

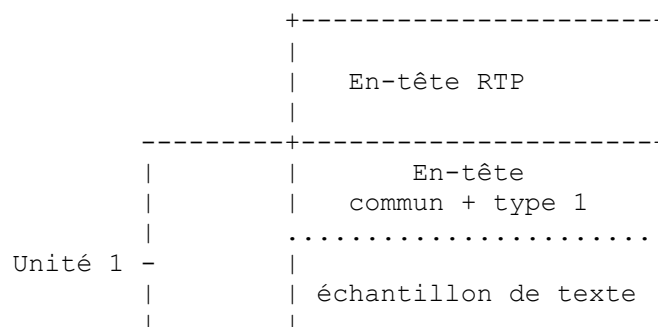
Type de charge utile (PT) : le type de charge utile est réglé dynamiquement et envoyé par des moyens hors bande.

L'usage du reste des champs d'en-tête RYP (à savoir, V, P, X, CC, SN et SSRC) suit les règles de RTP et du profil utilisé.

#### 4.1 Définitions d'en-tête de charge utile

Les unités (de transport) spécifiées dans le présent document consistent en un ensemble de champs communs (U, R, Type, Longueur) suivi par des champs d'en-tête spécifiques (Types 1 à 5) et un contenu d'échantillon de texte. Voir les Figure 1 et Figure 2.

Dans la Figure 2 sont décrits deux exemples de paquets RTP. Le premier contient une charge utile RTP agrégée avec deux échantillons de texte complets, et le second contient un fragment d'échantillon de texte. Après l'explication de chaque en-tête d'unité, les exemples détaillés de charge utile suivent au paragraphe 4.7.



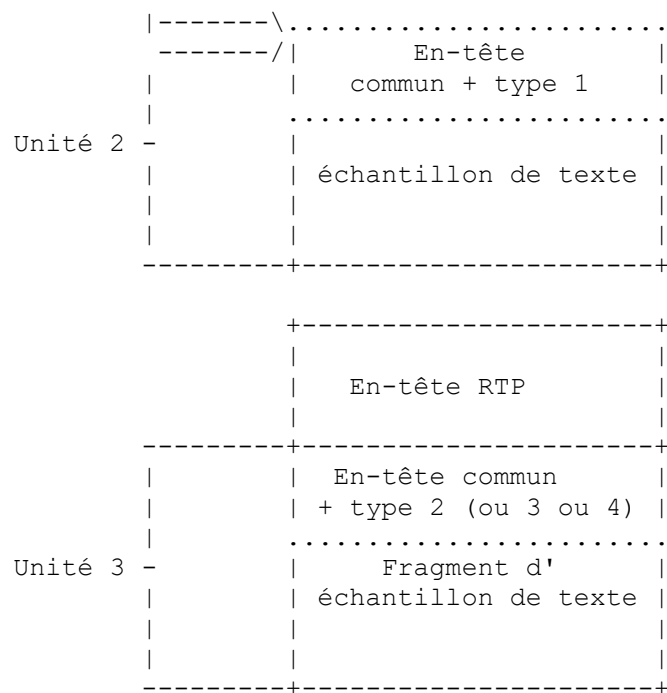


Figure 2. Exemple de paquets RTP

#### 4.1.1 Champs communs d'en-tête de charge utile

Les champs communs à tous les en-têtes de charge utile ont le format suivant :

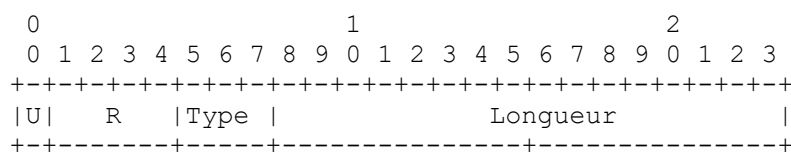


Figure 3. Champs d'en-tête commun de charge utile

Où :

- o U (1 bit) "Fanion de transformation UTF" : il est utilisé pour informer les receveurs RTP de l'utilisation de UTF-8 (U=0) ou UTF-16 (U=1) pour coder la chaîne de texte. Les chaînes de texte UTF-16 transportées par ce format de charge utile DOIVENT être mises en série en ordre gros boutien, autrement dit dans l'ordre des octets du réseau.

Note : les clients de texte synchronisé conformes au format de texte synchronisé 3GPP [PSS-TTF] sont seulement obligés de comprendre la mise en série gros boutienne. Donc, afin de faciliter l'interopérabilité, la mise en série inverse (petit boutienne) n'est pas prise en charge par ce format de charge utile.

Pour les formats de charge utile définis dans le présent document, le bit U n'est utilisé que dans les en-têtes de type 1 et de type 2. Les envoyeurs DOIVENT régler le bit U à zéro dans les en-têtes de type 3, 4, et 5. Par conséquent, les receveurs DOIVENT ignorer le bit U dans les en-têtes des types 3, 4, et 5.

- o R (4 bits) "Bits réservés" : pour de futures extensions. Ce champ DOIT être réglé à zéro (0x0) et DOIT être ignoré à réception.
- o Type (3 bits) "Champ de type" : ce champ spécifie quels champs d'en-tête spécifiques suivent. Les valeurs de type suivantes sont définies :
  - type 1, pour un échantillon de texte complet.
  - type 2, pour un fragment de chaîne de texte (sans modificateur).
  - type 3, pour une boîte de modificateur complète sur le premier fragment d'une boîte de modificateur.
  - type 4, pour un fragment de modificateur autre que le premier.



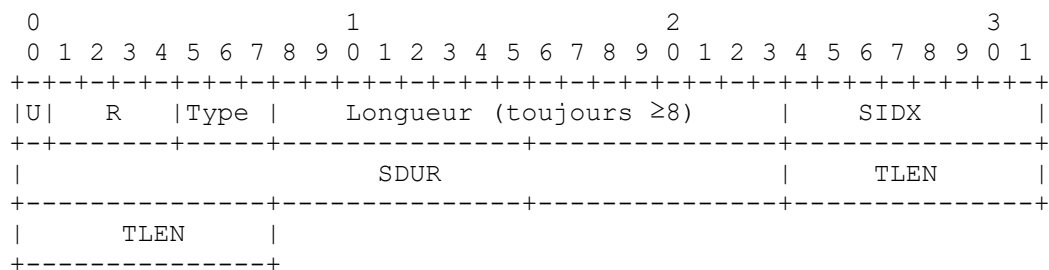
- type 5, pour une description d'échantillon. Exactement un en-tête par description d'échantillon.
  - Les types 0, 6, et 7 sont réservés pour de futures extensions. Noter que de futures extensions sont possibles, par exemple, une unité qui signale explicitement le nombre de caractères présents dans un fragment (voir le paragraphe 2.5). Afin de garantir la rétro-compatibilité, il DEVRA être possible que les anciens clients ignorent les unités (plus récentes) qu'ils ne comprennent pas, sans invalider les mécanismes de calcul d'horodatage ou sans les empêcher de décoder les autres unités.
- o Longueur (16 bits) : indique la taille (en octets) de ce champ d'en-tête et de tous les champs qui suivent, c'est-à-dire, le champ Longueur suivi par l'unité de charge utile : chaînes de texte et modificateurs (si il en est). Cette définition exclut seulement l'octet initial U/R/TYPE de l'en-tête commun. Le champ Longueur suit l'ordre des octets du réseau.

La façon dont la longueur est obtenue lors de l'écoulement d'un fichier 3GP dépend du type particulier d'unité. Ceci est expliqué pour chaque unité dans les paragraphes qui suivent.

Pour les flux de direct, la longueur d'échantillon et la valeur de Longueur pour le fragment en cours DOIVENT être calculées durant le processus d'échantillonnage ou durant la fragmentation.

En général, Longueur peut prendre les valeurs suivantes :

#### 4.1.2 En-tête de type 1



**Figure 4. Format d'en-tête de type 1**

Ce type d'en-tête est utilisé pour transporter des échantillons de texte complets. Cette unité devrait être la plus courante, c'est-à-dire, que l'échantillon de texte devrait généralement être assez petit pour être transporté dans une unité sans avoir à séparer les chaînes de texte des modificateurs. Dans une charge utile agrégée (paquet RTP) contenant plusieurs échantillons de texte, chaque échantillon est précédé de son propre en-tête de type 1 (voir la Figure 12).

Note : comme indiqué à la Section 3, "Terminologie", un échantillon de texte se compose de chaînes de texte suivies par les modificateurs (si il y en a). C'est aussi comme sont mémorisés les échantillons de texte dans les fichiers 3GP. La séparation d'un échantillon de texte en chaînes de texte et modificateurs est seulement nécessaire dans les grands échantillons (ou les petites tailles de MTU IP disponibles ; voir le paragraphe 4.4) et elle est réalisée avec les en-têtes de type 2 et 3, comme expliqué dans les paragraphes qui suivent

Noter aussi que les échantillons de texte vides sont considérés comme des échantillons de texte complets, bien qu'ils n'aient pas de contenu d'échantillon. Les échantillons de texte vides peuvent être utilisés pour nettoyer l'affichage ou pour mettre un terme à des échantillons de durée inconnue, par exemple. Les unités sans contenu d'échantillon DEVRONT avoir une valeur de champ Longueur de 8 (0x0008).

Les champs ci dessus ont la signification suivante :

- o U, R, et Type, comme défini au paragraphe 4.1.1.
- o Longueur, dans ce cas, représente la longueur de l'échantillon de texte (complet) plus huit (8) octets d'en-têtes. Pour trouver la longueur de l'échantillon de texte dans la boîte de taille d'échantillon des fichiers 3GP, voir au paragraphe 4.3.
- o SIDX (8 bits) "Indice d'entrée d'échantillon de texte" : c'est un indice utilisé pour identifier les descriptions d'échantillons. Le champ SIDX est utilisé pour trouver la description d'échantillon correspondant à la charge utile de l'unité. Il y a deux types de valeurs de SIDX : statique et dynamique. Les valeurs de SIDX statiques sont utilisées pour identifier les descriptions d'échantillons qui DOIVENT être envoyées hors bande et DOIVENT rester actives durant

toute la session. Une valeur de SIDX statique est liée sans équivoque à une description d'échantillon particulière durant toute la session. On DEVRAIT éviter de porter de nombreuses descriptions d'échantillons hors bande, car elles peuvent devenir grosses et finalement, le transport n'est pas le but du canal hors bande. Donc, cette caractéristique est RECOMMANDÉE pour transporter les descriptions d'échantillons qui fournissent un ensemble de réglages de format par défaut minimum. Les valeurs de SIDX statiques DOIVENT tomber dans l'intervalle (clos) [129,254]. Les valeurs de SIDX dynamiques sont utilisées pour les descriptions d'échantillons envoyées dans la bande. Les descriptions d'échantillon PEUVENT être envoyées dans la bande pour plusieurs raisons : parce qu'elles sont générées en temps réel, pour la résilience du transport, ou les deux. Une valeur de SIDX dynamique est liée sans équivoque à une description d'échantillon particulière durant la période dans laquelle elle est active dans la session, et elle NE DEVRA PAS être modifiée durant cette période. Cette période PEUT être inférieure ou égale à la durée de la session. Cette période n'est pas connue a priori. Un maximum de 64 valeurs de SIDX dynamiques simultanément actives est permis à tout moment. Les valeurs de SIDX dynamiques DOIVENT tomber dans l'intervalle clos [0,127]. Ceci devrait être assez pour les contenus enregistrés et les applications de flux en direct. Néanmoins, un mécanisme de retour à zéro est prévu au paragraphe 4.2.1 pour traiter les sessions où plus de 64 valeurs de SIDX pourraient être nécessaires. Les serveurs PEUVENT utiliser les descriptions d'échantillons dynamiques. Les clients DOIVENT être capables de recevoir et interpréter les descriptions d'échantillons dynamiques. Finalement, les valeurs de SIDX 128 et 255 sont réservées pour une future utilisation.

- o SDUR (24 bits) "Durée d'échantillon de texte" : indique la durée de l'échantillon de texte en unités d'horodatage RTP de l'échantillon de texte. Pour ce champ, une longueur de 3 octets est préférée à 2 octets. C'est parce que pour un débit d'horloge normal de 1000 Hz, 16 bits permettraient juste une durée maximum de 65 secondes, qui peut être trop courte pour certains flux. D'un autre côté, 24 bits à 1000 Hz permettent une durée maximum d'environ 4,6 heures, tandis que pour 90 kHz, cette valeur est d'environ 3 minutes. Ces valeurs devraient suffire pour les applications de flux en direct. Cependant, si une plus longue durée est nécessaire, le mécanisme d'extension spécifié au paragraphe 4.3 DEVRA être utilisé.

À part de définir la durée de l'affichage du texte, le champ Durée est aussi utilisé pour trouver l'horodatage des unités suivantes dans la charge utile d'un paquet RTP agrégé (si il en est). Ceci est expliqué au paragraphe 4.6.

Les échantillons de texte ont généralement une durée connue au moment de la transmission. Cependant, dans certains cas comme les flux de direct, le temps pendant lequel un morceau de texte devra être présenté pourrait n'être pas connu a priori. Donc, la valeur zéro de SDUR=0 (0x000000) est réservée pour signaler une durée inconnue. La quantité de temps pendant laquelle un échantillon de durée inconnue est présenté est déterminée par l'horodatage du prochain échantillon qui devra être affiché chez le receveur : les échantillons de texte de durée inconnue DEVRONT être affichés jusqu'à ce que le prochain échantillon de texte devienne actif, comme indiqué par son horodatage.

L'exemple suivant illustre comment les unités de durée inconnue DOIVENT être présentées. Si aucun échantillon de texte suivant n'est disponible, ce qui devrait être affiché relève de la mise en œuvre. Par exemple, un serveur pourrait envoyer un échantillon vide pour nettoyer la boîte de texte.

Exemple : Imaginons un aéroport où on observe le rapport des dernières nouvelles en attendant son avion. Les aéroports sont bruyants, de sorte que les nouvelles sont transcrites dans la partie inférieure de l'écran. Cette zone affiche deux lignes de texte : les titres et les mots prononcés par le journaliste. Comme il est usuel, les titres sont affichés plus longtemps que le reste. Ce temps est, en principe, inconnu du serveur de flux, qui est en direct. Un titre est juste remplacé quand le prochain titre est reçu.

Cependant, lors de la mémorisation d'un échantillon de texte avec SDUR=0 dans un fichier 3GP, la valeur de SDUR DOIT être changée en la durée effective de l'échantillon de texte, qui DOIT toujours être supérieure à zéro (noter que le format de fichier ISO [ISO14496-12] interdit explicitement une durée d'échantillon de zéro). La durée effective DOIT être calculée comme la différence d'horodatage entre l'échantillon en cours (d'une durée inconnue) et le prochain échantillon de texte affiché.

Noter que les échantillons de durée inconnue NE DEVRONT PAS utiliser de caractéristiques qui exigent la connaissance de la durée du front d'échantillon. Ces caractéristiques sont le déroulement et le karaoke dans [PSS-TTF]. Cela s'applique aussi aux futures extensions du format de texte synchronisé. De plus, seules les descriptions d'échantillons (unités de type 5) PEUVENT suivre des unités de durée inconnue dans la même charge utile agrégée. Autrement, il ne serait pas possible de calculer l'horodatage de ces autres unités.

Pour les contenus de texte mémorisés dans les fichiers 3GP, voir au paragraphe 4.3 les détails de la façon d'extraire la valeur de la durée. Pour les flux de direct, les codeurs DEVRONT allouer des valeurs et unités appropriées conformément à [PSS-TTF] et livraisons ultérieures.

- o TLEN (16 bits), "Longueur de la chaîne de texte", est un compte des octets de la chaîne de texte. Le décodeur a besoin de la longueur de la chaîne de texte pour savoir où commencent les modificateurs dans la charge utile. TLEN n'est pas présent dans les fragments de chaîne de texte (type 2) car il peut être déduit des valeurs de Longueur de chaque fragment.

La valeur TLEN est obtenue des échantillons de texte tels que contenus dans les fichiers 3GP. Voir au paragraphe 4.3. Pour les contenus de direct, TLEN DOIT être obtenu durant le processus d'échantillonnage.

- o Finalement, l'échantillon de texte réel est placé après le champ TLEN. Comme défini à la Section 3, un échantillon de texte consiste en une chaîne de caractères codés en utilisant UTF-8 ou UTF-16, suivie par zéro, un ou plusieurs modificateurs. Noter aussi qu'aucune BOM ni compte d'octet n'est inclus dans les chaînes portées dans la charge utile (par opposition aux échantillons de texte mémorisés dans les fichiers 3GP [PSS-TTF]).

#### 4.1.3 En-tête de type 2

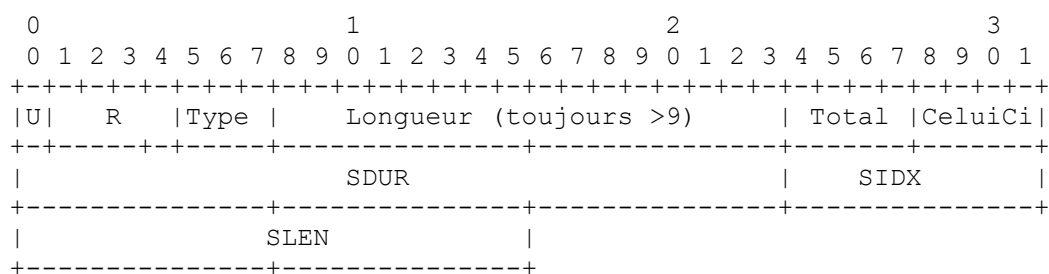


Figure 5. Format d'en-tête de type 2

Ce type d'en-tête est utilisé pour transporter une chaîne de texte complète ou un fragment. Les unités de type 2 NE DEVRONT PAS contenir de modificateur. En détails

- o U, R, et Type, comme défini au paragraphe 4.1.1.
- o SIDX et SDUR, comme défini au paragraphe 4.1.2.

Noter que les champs U, SIDX, et SDUR ont une signification parce que des chaînes de texte partielles peuvent aussi être affichées.

- o Le champ Longueur (16 bits) indique la longueur du fragment de chaîne de texte plus neuf (9) octets d'en-têtes. Sa valeur est calculée à la fragmentation. Longueur DOIT toujours être supérieur à neuf (0x0009). Autrement, l'unité DOIT être éliminée.

Selon les lignes directrices du paragraphe 4.4, les chaînes de texte DOIVENT être coupées à des limites de caractère pour permettre l'affichage des fragments de texte. Donc, un fragment de texte DOIT contenir au moins un caractère en UTF-8 ou UTF-16. En fait, ceci est de pure forme car si on observe les lignes directrices, des fragments beaucoup plus grands devraient être créés.

Noter aussi que les unités de type 2 ne contiennent pas de longueur explicite de chaîne de texte, TLEN (voir au type 1). Cela est parce que les unités de type 2 ne contiennent pas de modificateur après la chaîne de texte. Si nécessaire, la longueur de la chaîne reçue peut être obtenue en utilisant les valeurs de longueur des unités de type 2.

- o Le champ SLEN (16 bits) indique la taille (en octets) de l'échantillon de texte original (complet) auquel ce fragment appartient. Cette longueur comprend la chaîne de texte plus toutes les boîtes de modificateur présentes (et n'inclut ni la marque d'ordre des octets ni la longueur de la chaîne de texte comme mentionné à la Section 3, "Terminologie").

Concernant la longueur de l'échantillon de texte : les échantillons de texte synchronisé ne sont pas générés à des intervalles réguliers, et il n'y a pas de taille d'échantillon par défaut. Si des fichiers 3GP sont écoulés, la longueur des échantillons de texte est calculée à l'avance et incluse dans la piste elle-même, tandis que pour le codage du direct, c'est le codeur de temps réel qui DEVRA choisir une taille appropriée pour chaque échantillon de texte. Dans ce cas, la quantité de texte "capturée" dans un échantillon dépend de la source du texte et de l'application particulière (voir les exemples ci-dessous). Les échantillons peuvent, par exemple, être découpés pour correspondre à la MTU de paquet d'aussi près que possible ou pour fournir une certaine redondance pour le débit binaire disponible. L'application de codage DOIT aussi prendre en compte les contraintes de délai de la session en temps réel et évaluer si la FEC, la retransmission, ou d'autres techniques similaires sont des options raisonnables pour la réparation du flux.

Les exemples suivants illustrent comment un codeur de temps réel peut choisir ses réglages pour s'adapter aux contraintes du scénario.

Exemple : Imaginons un scénario de diffusion de nouvelles, où les nouvelles parlées sont transcrites et synchronisées avec l'image et la voix du reporteur. On suppose que le journaliste parle à une vitesse moyenne de 5 mots par seconde avec une longueur de mot moyenne de 5 caractères plus une espace par mot, c'est-à-dire, 30 caractères par seconde. On suppose une MTU IP disponible de 576 octets et un débit binaire disponible de  $576 * 8$  bits par seconde = 4,6 kbit/s. On suppose que chaque caractère peut être codé sur deux octets en UTF-16. Dans ce scénario, plusieurs contraintes peuvent s'appliquer ; par exemple : la MTU IP disponible, la bande passante disponible, le retard admissible, et la redondance requise. Si le but est de minimiser les frais généraux de paquet, un échantillon de texte couvrant 8 secondes de texte va être le plus proche de la MTU IP :

En-tête IP/UDP/RTP/TYPE1 + (échantillon de texte de 8 secondes) =  $20 + 8 + 12 + 8 + (\sim 6 \text{ caractères/mot} * 5 \text{ mots/s} * 8 \text{ s} * 2 \text{ caract/mot}) = 528 \text{ octets} < 576 \text{ octets}$

Pour d'autres scénarios, comme des réseaux à pertes, il peut arriver que juste un paquet par échantillon ait une redondance trop faible. Dans ce cas, un choix pourrait être que le codeur "collecte" le texte chaque seconde, donnant donc des échantillons de texte (unités de type 1) de 68 octets, en-tête de type 1 inclus. On peut, par exemple, inclure trois échantillons de texte contigus dans une charge utile RTP : l'échantillon de texte en cours et les deux derniers (voir ci-dessous). Cela fait une taille totale de paquet IP de  $20 + 8 + 12 + 3 * (8 + 60) = 244$  octets. Maintenant, avec le même débit binaire disponible de 4,6 kbit/s, ces paquets de 244 octets peuvent être envoyés de façon redondante jusqu'à deux fois par seconde :

charge utile RTP (1,2,3)(1,2,3) (2,3,4)(2,3,4) (3,4,5)(3,4,5) ...  
 temps : <----1s-----> <----1s-----> <----1s-----> ...

Cela signifie que chaque échantillon de texte est envoyé au moins six fois, ce qui devrait fournir une redondance suffisante. Bien qu'elle ne soit pas aussi efficace en bande passante ( $488 * 8 < 528 * 8 < 576 * 8$  bit/s) que la mise en paquets précédente, cette option augmente la redondance du flux tout en satisfaisant encore aux contraintes de délai et de bande passante.

Un autre exemple serait celui d'un utilisateur qui envoie du texte synchronisé à partir d'une zone de frappe de l'affichage. Dans ce cas, l'échantillon de texte est créé aussitôt que l'utilisateur clique sur le bouton "envoi". Selon la longueur du paquet, la fragmentation peut être nécessaire.

Dans une application de vidéo conférence, le texte est synchronisé avec l'audio et la vidéo. Donc, les échantillons de texte doivent être affichés assez longtemps pour être lus par un humain, doivent tenir dans l'écran de vidéo, et doivent "capturer" le contenu audio rendu durant le temps de rendu de la vidéo et l'audio correspondants.

Pour un contenu mémorisé, voir au paragraphe 4.3 les détails de la façon de trouver la valeur de SLEN dans un fichier 3GP. Pour les contenus de direct, la SLEN DOIT être obtenue durant le processus d'échantillonnage.

Finalement, on note que les clients PEUVENT utiliser SLEN pour l'espace de mémoire tampon pour les fragments restants d'un échantillon de texte.

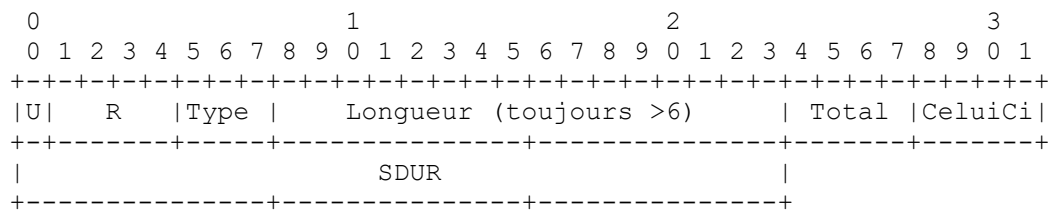
- o Les champs Total (4 bits) et CeluiCi (4 bits) indiquent, respectivement, le nombre total de fragments de l'échantillon de texte original (c'est-à-dire, la chaîne de texte et ses modificateurs) et l'ordre occupé par le fragment en cours dans cette séquence. Noter que le numéro de séquence seul ne peut pas remplacer la fonction du champ CeluiCi, car les paquets (et les fragments) peuvent être répétés, par exemple, comme dans une transmission répétée (voir la Section 5). Donc, une indication pour le "décalage de fragment" est nécessaire.

Le champ habituel "décalage d'octet" n'est pas utilisé ici pour deux raisons : a) cela prendrait un octet de plus et b) il ne fournit aucune information sur le décalage du caractère. Les chaînes de texte UTF-8/UTF-16 ont, en général, une longueur de caractère variable qui va de 1 à 6 octets. Donc, la solution Total/CeluiCi est préférée. On pourrait aussi objecter que les champs Longueur et SLEN soient utilisés à cette fin, mais bien qu'ils fournissent des informations sur la complétude de l'échantillon de texte, ils ne spécifient pas l'ordre des fragments.

Dans tous les cas (types 2, 3 et 4) si la valeur de CeluiCi est supérieure à celle de Total ou si Total est égale à zéro (0x0) le fragment DEVRA être éliminé.

- o Finalement, le contenu de l'échantillon suivant le champ SLEN consiste en un fragment de la chaîne de caractères UTF-8/UTF-16 ; aucun modificateur ne suit.

#### 4.1.4 En-tête de type 3



**Figure 6. Format d'en-tête de type 3**

Ce type d'en-tête est utilisé pour transporter le contenu entier du modificateur présent dans un échantillon de texte ou juste son premier fragment. Cela dépend de si les boîtes de modificateur tiennent dans la charge utile RTP courante.

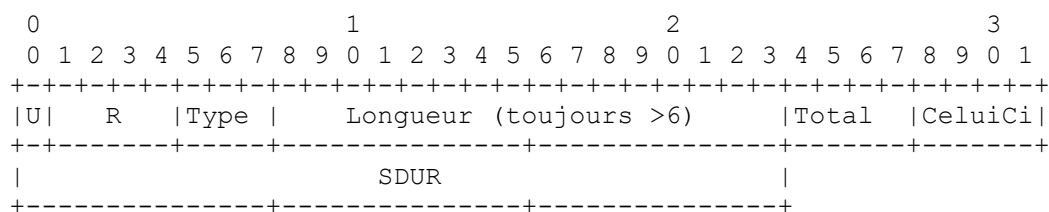
Si un échantillon de texte contenant des modificateurs est fragmenté, cet en-tête DOIT être utilisé pour transporter le premier fragment ou, si possible, les modificateurs complets.

En détails :

- o Les champs U, R, et Type sont définis au paragraphe 4.1.1.
- o Longueur indique la longueur du contenu du modificateur. Sa valeur est obtenue à la fragmentation. De plus, le champ Longueur DOIT être supérieur à six (0x0006). Autrement, l'unité DOIT être éliminée.
- o Les champs Total/CeluiCi ont la même signification que pour le Type 2. Pour les unités de Type 3 qui contiennent le dernier (de queue) fragment du modificateur, la valeur de Total DOIT être égale à celle de CeluiCi (Total=CeluiCi). De plus, Total=CeluiCi DOIT être supérieur à un, parce que le nombre total de fragments d'un échantillon de texte est logiquement toujours supérieur à un. Autrement, si Total est différent de CeluiCi dans une unité de Type 3, cela signifie que l'unité contient le premier fragment des modificateurs.
- o Le champ SDUR a la même définition pour le Type 1. Comme les fragments sont toujours transportés dans leurs propres paquets RTP, ce champ n'est nécessaire que pour savoir pendant combien de temps le fragment est valide. Cela peut, par exemple, être utilisé pour déterminer combien de temps il devrait être conservé dans la mémoire tampon d'affichage.

Noter que les champs SLEN et SIDX ne sont pas présents dans les en-têtes d'unités de Type 3. C'est parce que a) ces fragments ne contiennent pas de chaîne de texte et b) ces types de fragments sont appliqués sur des fragments de chaîne de texte, qui contiennent déjà ces informations.

#### 4.1.5 En-tête de type 4



**Figure 7. Format d'en-tête de type 4**

Ce type d'en-tête est placé avant les fragments de modificateur, autres que le premier.

Les champs U, R, et Type sont utilisés comme au paragraphe 4.1.1.

Longueur indique comme pour le type 3 la longueur du contenu de modificateur et DEVRA aussi être obtenu à la

fragmentation. Le champ Longueur DOIT être supérieur à six (0x0006). Autrement, l'unité DOIT être éliminée.

Total/CeluiCi est utilisé comme dans le Type 2.

Le champ SDUR est défini comme dans le Type 1. Le raisonnement sur l'absence de SLEN et SIDX est le même que dans les unités de Type 3.

4.1.6. En-tête de type 5

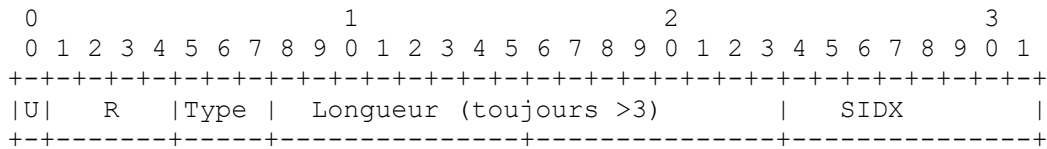


Figure 8. Format d'en-tête de type 5

Ce type d'en-tête est utilisé pour transporter des descriptions d'échantillons dynamiques. Chaque description d'échantillon DOIT avoir son propre en-tête de type 5.

Les champs U, R, et Type sont utilisés comme au paragraphe 4.1.1.

Le champ Longueur indique la longueur de la description d'échantillon, plus trois unités comptant pour le champ SIDX et Longueur lui-même. Donc, ce champ DOIT être supérieur à trois (0x0003). Autrement, l'unité DOIT être éliminée.

Si l'échantillon est écoulé d'un fichier 3GP, la longueur du contenu de la description d'échantillon (c'est-à-dire, ce qui vient après le SIDX dans l'unité elle-même) est obtenu du fichier (voir au paragraphe 4.3).

Le champ SIDX contient une valeur de SIDX dynamique allouée à la description d'échantillon portée comme un contenu d'échantillon de cette unité. Comme seules les descriptions d'échantillons dynamiques sont portées en utilisant le type 5, les valeurs de SIDX possibles sont dans l'intervalle (clos) [0,127].

Les envoyeurs PEUVENT utiliser des unités de Type 5. Tous les receveurs DOIVENT mettre en œuvre la prise en charge des unités de Type 5, car cela n'ajoute qu'une complexité minimale et peut augmenter la robustesse de la session de flux.

Le paragraphe qui suit spécifie comment les valeurs de SIDX sont calculées.

4.2 Mise en mémoire tampon des descriptions d'échantillon

La mise en mémoire tampon des descriptions d'échantillons relève de la mise en œuvre du codec de texte synchronisé de client. Pour qu'il fonctionne correctement, ce format de charge utile exige que :

- o les descriptions d'échantillons statiques DOIVENT être mises en mémoire tampon chez le client, au moins, pour la durée de la session ;
- o si des descriptions d'échantillons dynamiques sont utilisées, leur mise en mémoire tampon et la mise à jour des valeurs de SIDX DOIT suivre le mécanisme décrit au paragraphe suivant.

4.2.1 Mécanisme dynamique d'enveloppement SIDX

L'utilisation de descriptions d'échantillons dynamiques par les envoyeurs est FACULTATIVE. Cependant, si elles sont utilisées, les envoyeurs DOIVENT mettre en œuvre ce mécanisme. Les receveurs DOIVENT toujours le mettre en œuvre.

Les valeurs de SIDX dynamique restent actives durant toute la session (si elles sont utilisées juste une fois) ou à des intervalles différents (si utilisées plus d'une fois).

Note : dans la suite, SIDX s'entend de SIDX dynamique.

Pour choisir le mécanisme de débordement, on a utilisé le raisonnement suivant : il y a 128 valeurs possibles de SIDX, [0..127]. Si on choisit de permettre un maximum de 127 SIDX, tout paquet réordonné avec une nouvelle description d'échantillon va faire échouer le mécanisme. Par exemple, si le dernier paquet reçu est SIDX=5, toutes les 127 valeurs

excepté SIDX=6 vont être "actives". Maintenant, si un paquet réordonné arrive avec une nouvelle description, SIDX=9, il va être éliminé à tort, parce que le SIDX=9 est, à ce moment, marqué comme "actif" et les descriptions d'échantillons actives ne doivent pas être réécrites. Donc, un "intervalle de garde" est introduit. Cet intervalle de garde réduit le nombre de SIDX actifs à 64 à tout moment. Bien que la plupart des applications de texte synchronisé aient probablement besoin de moins de 64 descriptions d'échantillons durant une session (au total) un mécanisme de débordement pour traiter le besoin de plus est décrit ici.

Pour cela, on utilise une fenêtre glissante de 64 valeurs de SIDX actives. Les valeurs dans la fenêtre sont "actives" ; toutes les autres sont marquées "inactives". Une valeur de SIDX devient active si au moins une description d'échantillon identifiée par ce SIDX a été reçue. Comme les descriptions d'échantillons PEUVENT être envoyées de façon redondante, il est possible qu'un client reçoive plusieurs fois un certain SIDX. Cependant, les descriptions d'échantillons actives NE DEVRONT PAS être écrasées : le receveur DEVRA ignorer les descriptions d'échantillons redondantes et il DOIT utiliser la copie qui est déjà en antémémoire. L'intervalle de garde de (64) valeurs inactives assure que l'association correcte de SIDX <-> description d'échantillon est toujours utilisée.

Note : Comme pour la valeur de "l'intervalle de garde" elle-même, 64 comme 128/2 a été considéré assez simple tout en satisfaisant au nombre maximum attendu de descriptions d'échantillons. Par ailleurs, il n'y a aucune autre raison au choix de 64 plutôt qu'une autre valeur.

L'algorithme suivant est utilisé pour mettre en mémoire tampon les descriptions d'échantillons dynamiques et pour conserver les valeurs de SIDX dynamiques :

Soit X le dernier SIDX reçu qui a mis à jour la gamme des descriptions d'échantillons actives. Soit Y une valeur dans la gamme admise pour les SIDX dynamiques : [0,127], et différent de X. Soit Z le SIDX de la dernière description d'échantillon reçue. Alors,

1. Initialiser toutes les valeurs de SIDX dynamiques comme inactives. Pour les contenus mémorisés, lire l'indice de description d'échantillon dans la boîte d'échantillons à tronçonner (stsc, *Sample to Chunk*) pour cet échantillon. Pour les flux en direct, la première valeur PEUT être zéro ou toute autre valeur dans l'intervalle ci-dessus. Passer à l'étape 2.
2. D'abord, une description d'échantillon dans la bande avec SIDX=Z est reçue et mémorisée; régler X=Z. Passer à l'étape 3.
3. Tout SIDX dans l'intervalle [X+1 modulo(128), X+64 modulo(128)] est marqué comme inactif, et toute description d'échantillon correspondante est supprimée. Tout SIDX dans l'intervalle [X+65 modulo(128), X] est réglé à actif. Passer à l'étape 4 (état d'attente).
4. Attendre la prochaine description d'échantillon. Une fois le client initialisé, l'intervalle des valeurs de SIDX actives DOIT changer chaque fois qu'est reçue une description d'échantillon avec une valeur de SIDX dans l'ensemble inactif. C'est à dire qu'à réception d'une description d'échantillon avec SIDX=Z, on fait ce qui suit :
  - a. Si Z est dans l'intervalle (clos) [X+1 modulo(128), X+64 modulo(128)] régler X=Z, mémoriser la description d'échantillon, et passer à l'étape 3.
  - b. Autrement, Z doit être dans l'intervalle [X+65 modulo(128), X], donc :
    - i. Si SIDX=Z n'est pas mémorisé, mémoriser la description d'échantillon. Passer au début de l'étape 4 (état d'attente).
    - ii. Autrement, passer au début de l'étape 4 (état d'attente).

Note : Il est permis que toute valeur de SIDX=X soit envoyée dans l'intervalle [0,127]. Par exemple, si [64..127] est l'ensemble actuellement actif et si SIDX=0 est envoyé, une nouvelle description d'échantillon est définie (0) et une ancienne est supprimée (64) ; donc [65..127] et [0] sont actives. De même, on peut maintenant envoyer SIDX=64, inversant donc les ensembles actif et inactif.

Exemple :

Si X=4, tout SIDX dans l'intervalle [5,68] est inactif. Les valeurs de SIDX actives sont dans l'intervalle complémentaire [69,127] plus [0,4]. Par exemple, si le client reçoit un SIDX=6, l'intervalle actif est maintenant différent : [0,6] plus [71,127]. Si le SIDX reçu est dans l'intervalle actif en cours, aucun changement NE DEVRA être appliqué.

### 4.3 Trouver les valeurs d'en-tête de charge utile dans les fichiers 3GP

Pour l'écoulement des contenus de texte synchronisé, certaines valeurs dans les boîtes contenues dans un fichier 3GP sont transposées en des champs de cet en-tête de charge utile. Ce paragraphe explique où trouver ces valeurs.

De plus, pour la durée et les indices de description d'échantillon, des mécanismes d'extension sont fournis. Tous les envoyeurs DOIVENT mettre en œuvre les mécanismes d'extension décrits ici.

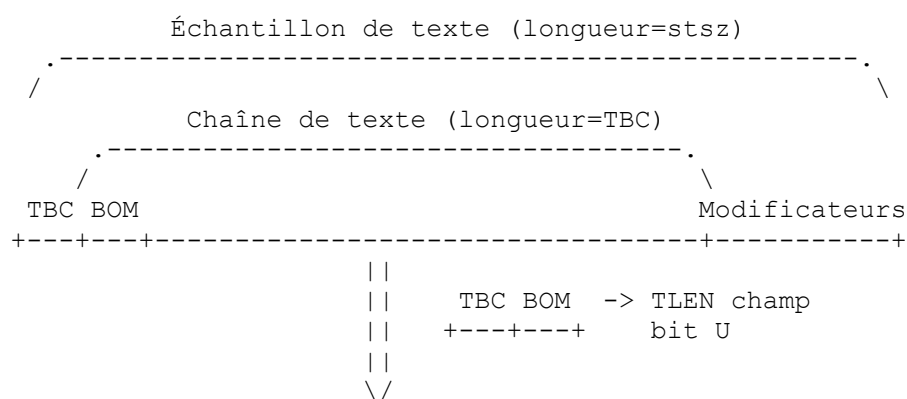
Si le fichier est écoulé à partir d'un fichier 3GP, les lignes directrices suivantes DEVRONT être suivies.

Note : Tous les champs dans les objets (boîtes) d'un fichier 3GP sont dans l'ordre des octets du réseau.

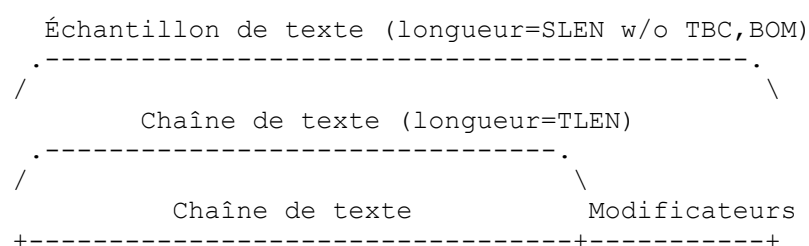
Informations obtenues de la boîte de tableau d'échantillons (stbl, *Sample Table Box*) :

- o Descriptions d'échantillon et Longueur de description d'échantillon : la boîte de description d'échantillon (stsd, *Sample Description*) (à l'intérieur de stbl) contient les descriptions d'échantillons. Pour les supports de texte synchronisé, chaque élément de stsd est une entrée d'échantillon de texte synchronisé (type "tx3g"). Les 32 bits (non signés) du champ "Taille" dans la boîte stsd représentent la longueur (en octets) de la description d'échantillon, telle que portée dans les unités de type 5. Par ailleurs, le champ Longueur des unités de type 5 est restreint à 16 bits. Donc, si la valeur de "taille" est supérieure à  $(2^{16}-1-3)$ [octets], la description d'échantillon NE DEVRA PAS être écoulée avec ce format de charge utile. Aucun mécanisme d'extension n'est défini dans ce cas, car la fragmentation de descriptions d'échantillons n'est pas définie (les descriptions d'échantillons font normalement environ 200 octets). Note : le trois (3) représente le champ d'en-tête de type 5 inclus dans la valeur de Longueur.
- o SDUR provenant de la boîte d'heure de décodage d'échantillon (stts, *Decoding Time to Sample Box*). Les 32 bits (non signés) du champ "delta d'échantillon" sont utilisés pour calculer la SDUR. Cependant, comme le champ SDUR fait seulement 3 octets, les échantillons de texte de valeur de durée supérieure à  $(2^{24}-1)/(\text{débit d'horloge d'horodatage})$  [secondes] ne peuvent pas être écoulés directement. La solution est simple : les copies de l'échantillon de texte correspondant DEVRONT être envoyées. Ainsi, les valeurs d'horodatage et de durée DEVRONT être ajustées afin qu'un affichage continu soit garanti comme si juste un échantillon avait été envoyé. C'est-à-dire qu'un échantillon avec l'horodatage TS et la durée SDUR peut être envoyé comme deux échantillons ayant les horodatages TS1 et TS2 et les durées SDUR1 et SDUR2, de telle sorte que  $TS1=TS$ ,  $TS2=TS1+SDUR1$ , et  $SDUR=SDUR1+SDUR2$ .
- o Longueur d'échantillon de texte à partir de la boîte de taille d'échantillon (stsz, *Sample Size Box*). Les 32 bits (non signés) de "taille d'échantillon" ou "taille d'entrée" (l'un d'eux, selon que la taille d'échantillon est fixe ou variable) indique la longueur (en octets) de l'échantillon de texte 3GP. Pour obtenir la longueur (réelle) de l'échantillon de texte écoulé, les longueurs du compte d'octets de la chaîne de texte (2 octets) et, dans le cas de chaînes UTF-16, la longueur de la BOM (aussi 2 octets) DEVRONT être déduites. Ceci est illustré à la Figure 9.

Échantillon de texte conforme à 3GPP TS 26.245



Échantillon de texte conforme à ce format de charge utile





Légende :

TBC (*Text string Byte Count*) : compte d'octet de la chaîne de texte

BOM (*Byte Order Mark*) : marque d'ordre des octets

### Figure 9. Composition d'échantillon de texte

De plus, comme le champ Longueur dans l'en-tête d'unité de type 1 fait 16 bits, de plus grande tailles d'échantillon de texte que  $(2^{16}-1-8)$  [octets] NE DEVRONT PAS être écoulées. Aussi, dans ce cas, aucun mécanisme d'extension n'est défini. C'est parce que ce maximum est considéré suffisant pour les applications de flux visées. (Note : le huit (8) est pour les champs d'en-tête de type 1 inclus dans la valeur de Longueur).

- o SIDX provenant de la boîte d'échantillon à tronquer (stsc, *Sample to Chunk Box*) : la boîte stsc est utilisée pour trouver les échantillons et leurs descriptions d'échantillons correspondantes. Elles sont référencées par un "indice de description d'échantillon", un entier (non signé) de 32 bits. Si possible, ces indices peuvent être directement transposés en le champ SIDX. Cependant, il y a plusieurs cas où ceci n'est pas
  - a) Le nombre total d'indices utilisés est supérieur au nombre d'indices disponibles, c'est-à-dire, si les descriptions d'échantillons statiques sont plus de 127 ou si les dynamiques sont plus de 64.
  - b) Les gammes originales de valeurs de SIDX ne tiennent pas dans les gammes admises pour les valeurs statiques (129 à 254) ou dynamiques (0 à 127).

Donc, quand on alloue des valeurs de SIDX aux descriptions d'échantillons, les lignes directrices suivantes sont fournies :

- o Les descriptions d'échantillons statiques peuvent simplement être allouées à des valeurs consécutives dans la gamme de 129 à 254 (intervalle clos). Cette gamme devrait suffire pour les descriptions d'échantillons statiques.
- o Pour les descriptions d'échantillons dynamiques :
  - a) Les flux qui utilisent moins de 64 descriptions d'échantillons dynamiques DEVRAIENT utiliser des valeurs consécutives pour les SIDX partout dans la gamme de 0 à 127 (intervalle clos).
  - b) Pour les flux avec plus de 64 descriptions d'échantillons, les valeurs de SIDX DOIVENT être allouées dans l'ordre d'utilisation, et si une description d'échantillon doit être utilisée après qu'elle a été réglé à inactive, elle va devoir être envoyée à nouveau et recevoir une nouvelle valeur de SIDX (conformément à l'algorithme du paragraphe 4.2.1).

Informations obtenues de la boîte de données de support :

- o Chaînes de texte, TLEN, bit U, et modificateurs provenant de la boîte de données de support (mdat). Les chaînes de texte, le compte d'octets de 16 bits de chaîne de texte, la marque d'ordre des octets (BOM, indiquant le codage UTF) et les boîtes de modificateur peuvent y être trouvés.

Pour les unités de type 1, la valeur de TLEN est extraite du compte d'octets de chaîne de texte qui précède la chaîne de texte dans l'échantillon de texte, comme mémorisé dans le fichier 3GP. Si le codage UTF-16 est utilisé, deux (2) octets de plus doivent être précédemment déduits de ce compte d'octets, afin d'exclure la BOM. Voir la Figure 9.

#### 4.4 Fragmentation des échantillons de texte synchronisé

Ce paragraphe explique pourquoi les échantillons de texte peuvent devoir être fragmentés et discute les approches possibles pour le faire. Une solution est proposée avec les règles et recommandations pour fragmenter et transporter les échantillons de texte.

Les applications de texte synchronisé 3GPP sont supposées fonctionner à de faibles débits binaires. Ce fait, ainsi que la petite taille des échantillons de texte synchronisé (normalement une ou deux centaines d'octets) rend la fragmentation des échantillons de texte un événement rare. Les échantillons devraient généralement tenir dans la taille de MTU du chemin de réseau utilisé.

Néanmoins, certaines chaînes de texte (par exemple, le générique de fin d'un film) et certaines boîtes de modificateurs (c'est-à-dire, pour des hyperliens, pour du karaoke, ou pour des styles) peuvent devenir plus grandes. Ceci peut aussi s'appliquer pour de futures boîtes de modificateur. Dans ces cas, la première option à considérer est si il est possible d'ajuster le codage (par exemple, la taille de l'échantillon) afin d'éviter la fragmentation. Si cela est possible, il faut le

préférer à la fragmentation et cela DEVRAIT être fait.

Autrement, si ce n'est pas possible ou si d'autres contraintes l'empêchent, la fragmentation PEUT être utilisée, et les lignes directrices de base données dans ce document DOIVENT être suivies :

- o Il est RECOMMANDÉ que les échantillons de texte soient fragmentés aussi rarement que possible, c'est-à-dire, que le plus petit nombre possible de fragments soient créés à partir d'un échantillon de texte.
- o Si il y a du débit binaire et de l'espace libre dans la charge utile, les descriptions d'échantillons (si il y en a) DEVRAIENT être agrégées.
- o Les chaînes de texte DOIVENT se partager aux frontières de caractères ; voir l'en-tête de type 2. Autrement, il n'est pas possible d'afficher le contenu du texte d'un fragment si un fragment précédent a été perdu. Par conséquent, la fragmentation de chaîne de texte exige la connaissance du format de codage UTF-8/UTF-16 pour déterminer les limites de caractères.
- o À la différence des chaînes de texte, les boîtes de modificateur NE SONT PAS obligées d'être partagées à des limites particulières. Cependant, il est RECOMMANDÉ que cela soit fait chaque fois que possible. Cela diminue les effets de la perte de paquet. Ce format de charge utile n'assure pas que des modificateurs partiellement reçus s'appliquent aux chaînes de texte. Si seulement une partie des modificateurs est reçue, il revient à l'application de déterminer comment traiter cela, c'est-à-dire, de les utiliser ou non.

Note : S'assurer que les modificateurs partiellement reçus peuvent être appliqués aux chaînes de texte dans tous les cas (pour tous les types de modificateurs et pour toutes les constellations de perte de fragment) ferait peser des exigences supplémentaires sur le format de charge utile. En particulier, cela exigerait que : a) les envoyeurs comprennent la sémantique des boîtes de modificateur et b) des en-têtes de fragment spécifiques pour chaque boîte de modificateur soient définies, en plus des formats de charge utile définis ci-dessous. Comprendre la sémantique des modificateurs signifie de savoir, par exemple, où commence et se termine chaque modificateur, quels fragments de texte sont affectés, quels modificateurs peuvent être ou non partagés, ou ce qu'indiquent les champs. Ceci est nécessaire pour être capable de partager les modificateurs d'une façon telle que chaque fragment puisse être appliqué indépendamment des pertes de paquets précédentes. Cela exigerait une entité de fragmentation plus intelligente et des en-têtes plus complexes. Étant donné la faible probabilité de fragmentation et le désir de garder un faible niveau d'exigences, il ne semble pas raisonnable de spécifier de tels en-têtes spécifiques de boîte de modificateur.

- o Les fragments de modificateur et de chaîne de texte DEVRAIENT être protégés contre les pertes de paquets, c'est-à-dire, utiliser la FEC [RFC2733], la retransmission [RFC4588], la répétition (Section 5), ou une technique équivalente. Cela minimise les effets de la perte de paquet.
- o Une exigence supplémentaire de la fragmentation d'un échantillon de texte est que le début des modificateurs DOIT être indiqué en utilisant l'en-tête de charge utile défini à cette fin, c'est-à-dire, une unité de type 3 DOIT être utilisée (voir le paragraphe 4.1.4). Cela permet à un receveur de détecter le début des modificateurs pour autant qu'il n'y ait pas deux pertes de paquets consécutives, ou plus.
- o Finalement, les descriptions d'échantillons NE DEVRONT PAS être fragmentées parce qu'elles contiennent des informations importantes qui peuvent affecter plusieurs échantillons de texte.

#### 4.5 Réassemblage des échantillons de texte chez le receveur

Les en-têtes de charge utile définis dans le présent document permettent le réassemblage des fragments d'échantillons de texte. À cette fin sont utilisés l'horodatage RTP standard, le champ de durée (SDUR), et les champs Total/CeluiCi dans les en-têtes de charge utile.

Les unités qui appartiennent au même échantillon de texte DOIVENT avoir le même horodatage. Les unités de type 5 ne se conforment pas à cette règle car elles ne font pas partie d'un échantillon de texte particulier.

Le processus de collecte des différents fragments (unités) d'un échantillon de texte est le suivant :

1. Chercher les unités qui ont la même valeur d'horodatage, c'est-à-dire, les unités qui appartiennent au même échantillon de texte ou descriptions d'échantillons qui devront devenir disponibles à cet instant. Si plusieurs unités du même échantillon sont répétées, une seule d'entre elles DEVRA être utilisée. Les unités répétées sont celles qui ont le même

horodatage et les mêmes valeurs pour Total/CeluiCi.

Noter que, comme mentionné au paragraphe 4.1.1, le receveur DEVRA ignorer les unités qui ont une valeur de type non reconnue. Cependant, les champs d'en-tête RTP et le reste des unités (si il en est) dans la charge utile sont toujours utiles.

2. Vérifier dans cet ensemble si il manque des unités dans cet échantillon de texte. Cela se fait en utilisant les champs Total et CeluiCi ; le champ Total indique combien de fragments ont été créés à partir de l'échantillon de texte, et le champ CeluiCi indique la position de ce fragment dans l'échantillon de texte. Par suite de cette opération, deux résultats sont possibles :
  - a. Aucun fragment ne manque. Le champ CeluiCi DEVRA alors être utilisé pour ordonner les fragments et assembler l'échantillon de texte avant de le transmettre à l'application de décodage. Une attention particulière DEVRA être apportée au réassemblage de la chaîne de texte comme indiqué au point 4 ci-dessous.
  - b. Un ou plusieurs fragments manquent : Vérifier si ce fragment appartient à la chaîne de texte ou aux modificateurs. Les unités de type 2 identifient les fragments de chaîne de texte, et celles de type 3 et 4 identifient les fragments de modificateur :
    - i. Si le ou les fragments manquants appartiennent à la chaîne de texte et si les modificateurs ont été reçus complets, les caractères de texte reçus peuvent, au moins, être affichés comme texte en clair. Certains modificateurs peuvent n'être appliqués qu'autant qu'il est possible d'identifier le nombre de caractères, par exemple, si seulement le dernier fragment de la chaîne de texte est perdu. Ce peut être le cas pour les modificateurs qui définissent des styles de fonte spécifiques ("styl"), des caractères en surbrillance ("hlit"), des caractéristiques de karaoke ("krok"), et des caractères clignotants ("blnk"). D'autres modificateurs comme "dlay" ou "tbox" peuvent être appliqués sans connaissance du nombre de caractères. C'est l'affaire d'une application de décider d'appliquer ou non les modificateurs.
    - ii. Si le fragment manquant appartient aux modificateurs et si les chaînes de texte ont été reçues complètes, les modificateurs incomplets peuvent être utilisés. La chaîne de texte DEVRAIT au moins être affichée comme texte en clair. Comme mentionné au paragraphe 4.4, les modificateurs peuvent être partagés sans respect de frontières significatives. Donc, il peut n'être pas toujours possible d'utiliser des modificateurs reçus partiellement. Cependant, pour éviter cela, il est RECOMMANDÉ que les modificateurs se coupent à des frontières de signification.
    - iii. Une troisième possibilité est qu'il ne soit pas possible de discerner si des modificateurs ou des chaînes de texte ont été reçues complètes. Par exemple, si l'unité de type 3 d'un échantillon plus le paquet suivant ou précédant est perdu, il n'y a aucun moyen pour le receveur RTP de savoir si un des paquets perdus ou les deux appartient aux modificateurs ou si des chaînes de texte sont aussi manquantes. La répétition, la FEC, la retransmission, ou d'autres mécanismes de protection comme au paragraphe 4.6 sont RECOMMANDÉS pour éviter cette situation.
    - iv. Finalement, si il est sûr que ni les chaînes de texte ni les modificateurs n'ont été reçus complets, les chaînes de texte et les modificateurs peuvent alors être rendus partiellement ou peuvent être éliminés. C'est un choix de l'application.
3. Les descriptions d'échantillon peuvent être directement associées à des échantillons de texte réassemblés via l'indice de description d'échantillon (SIDX).
4. Le réassemblage des chaînes de texte : comme les chaînes de texte transportées dans des paquets RTP NE DOIVENT PAS inclure de marque d'ordre d'octets (BOM) le receveur DOIT lui faire précéder la chaîne UTF-16 réassemblée avant de la traiter au décodeur de texte synchronisé (voir la Figure 9). La valeur de la BOM est 0xFEFF parce que seule la mise en série des chaînes UTF-16 est prise en charge par ce format de charge utile.

#### 4.6. Charges utiles agrégées

Les unités DEVRAIENT être agrégées pour éviter les frais généraux, chaque fois que possible. Les charges utiles agrégées DOIVENT se conformer à une des configurations ordonnées suivantes :

1. Zéro, une ou plusieurs descriptions d'échantillons (Type 5) suivies par zéro, un ou plusieurs échantillons de texte complets (unités de type 1). Au moins une unité de l'un ou l'autre type DOIT être présente.
2. Zéro, une ou plusieurs descriptions d'échantillons suivies par zéro ou un fragment de modificateur, de type 3 ou 4. Au moins une unité DOIT être présente.
3. Zéro, une ou plusieurs descriptions d'échantillons, suivies par zéro ou un fragment de chaîne de texte (Type 2) suivi par

zéro ou une unité de type 3. Si une unité de type 2 et une unité de type 3 sont présentes, elles DOIVENT alors appartenir au même échantillon de texte. Au moins une unité DOIT être présente.

Quelques observations :

- o Des agrégats différents de ceux mentionnés ci-dessus NE DEVRONT PAS être utilisés.
- o Les descriptions d'échantillon DOIVENT être placées dans la charge utile agrégée avant l'occurrence de toute unité qui n'est pas de type 5.
- o La réception correcte des unités de type 5 est importante car leur contenu peut être référencé par plusieurs autres unités dans le flux.

Les receveurs sont incapables d'utiliser les échantillons de texte tant que leurs descriptions d'échantillons correspondantes ne sont pas reçues. Par conséquent, un expéditeur DEVRAIT envoyer plusieurs copies d'une description d'échantillon pour assurer la fiabilité (voir la Section 5). Les receveurs PEUVENT utiliser des messages de retour spécifiques de la charge utile [RFC4585] pour dire à l'expéditeur qu'ils ont reçu une description d'échantillon particulière.

- o Concernant le calcul de l'horodatage : en général, les règles de calcul de l'horodatage des unités dans une charge utile agrégée dépendent du type d'unité. Sur la base des constellations possibles pour les charges utiles agrégées, comme ci-dessus, on a :
  - les descriptions d'échantillon DOIVENT recevoir l'horodatage RTP du paquet dans lequel elles sont incluses.
 Noter que pour les unités de type 5, l'horodatage ne représente en fait pas l'instant où elles sont exécutées, mais plutôt l'instant auquel elles deviennent utilisables.
  - Pour la première configuration : la première unité de type 1 reçoit l'horodatage RTP. L'horodatage de toute unité de type 1 suivante DOIT être obtenu en ajoutant la durée et l'horodatage de l'échantillon de l'unité précédente de type 1.
  - Pour la seconde et la troisième configurations, toutes les unités, de type 2, 3, et 4, DOIVENT recevoir l'horodatage RTP.

Se reporter aux exemples détaillés sur le calcul d'horodatage ci-dessous.

- o Comme pour la configuration 3 ci-dessus, une charge utile PEUT contenir plusieurs fragments d'un (et un seul) échantillon de texte. Si elle le fait, alors exactement une unité de type 2 suivie par exactement une unité de type 3 est permise dans la même charge utile. Ceci est conforme au paragraphe 2.4 de la [RFC3640] qui interdit explicitement de combiner des fragments de différents échantillons dans la même charge utile RTP. Noter que, dans ce cas particulier, aucun calcul d'horodatage n'est nécessaire. C'est-à-dire que l'horodatage RTP des deux unités est égal à l'horodatage de l'en-tête RTP du paquet.
- o Finalement, on note que l'utilisation d'échantillons de texte vides permet l'agrégation d'unités de type 1 non consécutives dans la même charge utile. Deux échantillons de texte, avec les horodatages TS1 et TS3 et les durées SDUR1 et SDUR3, sont non consécutifs si  $TS1 + SDUR1 < TS3$ . Une solution à cela est d'inclure une unité de type 1 avec la durée SDUR2 entre eux, de sorte que  $TS2 + SDUR2 = TS1 + SDUR1 + SDUR2 = TS3$ .

Des exemples de charges utiles agrégées sont illustrés à la Figure 10. (Note : la figure n'est pas à l'échelle.)

```

  N/A    TS1    TS2    TS3
+-----+-----+-----+-----+
|Type5| |Type1| |Type1| |Type1|
+-----+-----+-----+-----+
  N/A    sdur1  sdur2  sdur3

```

```

  N/A    TS4
+-----+-----+
|Type5| |Type 1|
+-----+-----+
  N/A    sdur4

```

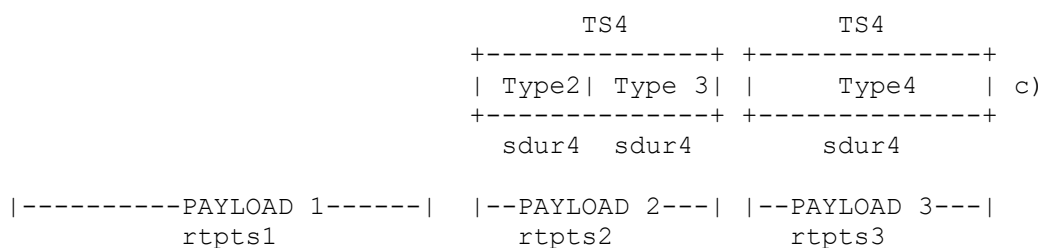
a)

```

          TS4          TS4    TS4
+-----+-----+ +-----+-----+
|   Type2   | |Type2| |Type 3|
+-----+-----+ +-----+-----+
          sdur4          sdur4  sdur4

```

b)



Légende :

TSx = échantillon de texte x

rtptsy = horodatage RTP standard pour PAYLOAD y

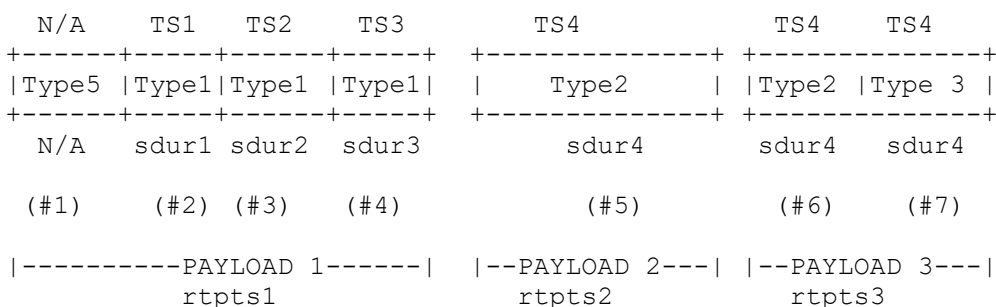
sdurx = durée de l'échantillon de texte x

N/A = non applicable

**Figure 10. Exemple de charges utiles agrégées**

Dans la Figure 10, quatre échantillons de texte (TS1 à TS4) sont envoyés en utilisant trois paquets RTP. Ces configurations ont été choisies pour montrer comment les en-têtes de type 5 sont utilisés. De plus, trois différentes possibilités pour le dernier échantillon de texte, TS4, sont décrites : a), b), et c).

Dans la Figure 11, l'option b) de la Figure 10 est choisie pour illustrer comment l'horodatage est trouvé pour chaque unité.



**Figure 11. Charges utiles choisies de la Figure 10**

En supposant que TSx signifie l'échantillon de texte x, rtptsy représente l'horodatage RTP standard pour PAYLOAD y et sdurx, la durée de l'échantillon de texte x, l'horodatage pour l'unité #z, ts(#z) peut être trouvé comme la somme de rtptsy et la somme cumulative des durées des unités précédentes dans cette charge utile (sauf dans le cas de PAYLOAD 3 conformément à la règle 3 ci-dessus). Donc, on a :

1. pour les unités dans la première charge utile agrégée, PAYLOAD 1 :

$$ts(\#1) = rtpts1$$

$$ts(\#2) = rtpts1$$

$$ts(\#3) = rtpts1 + sdur1$$

$$ts(\#4) = rtpts1 + sdur1 + sdur2$$

Noter que l'unité de type 5 et la première unité de type 1 ont toutes deux l'horodatage RTP.

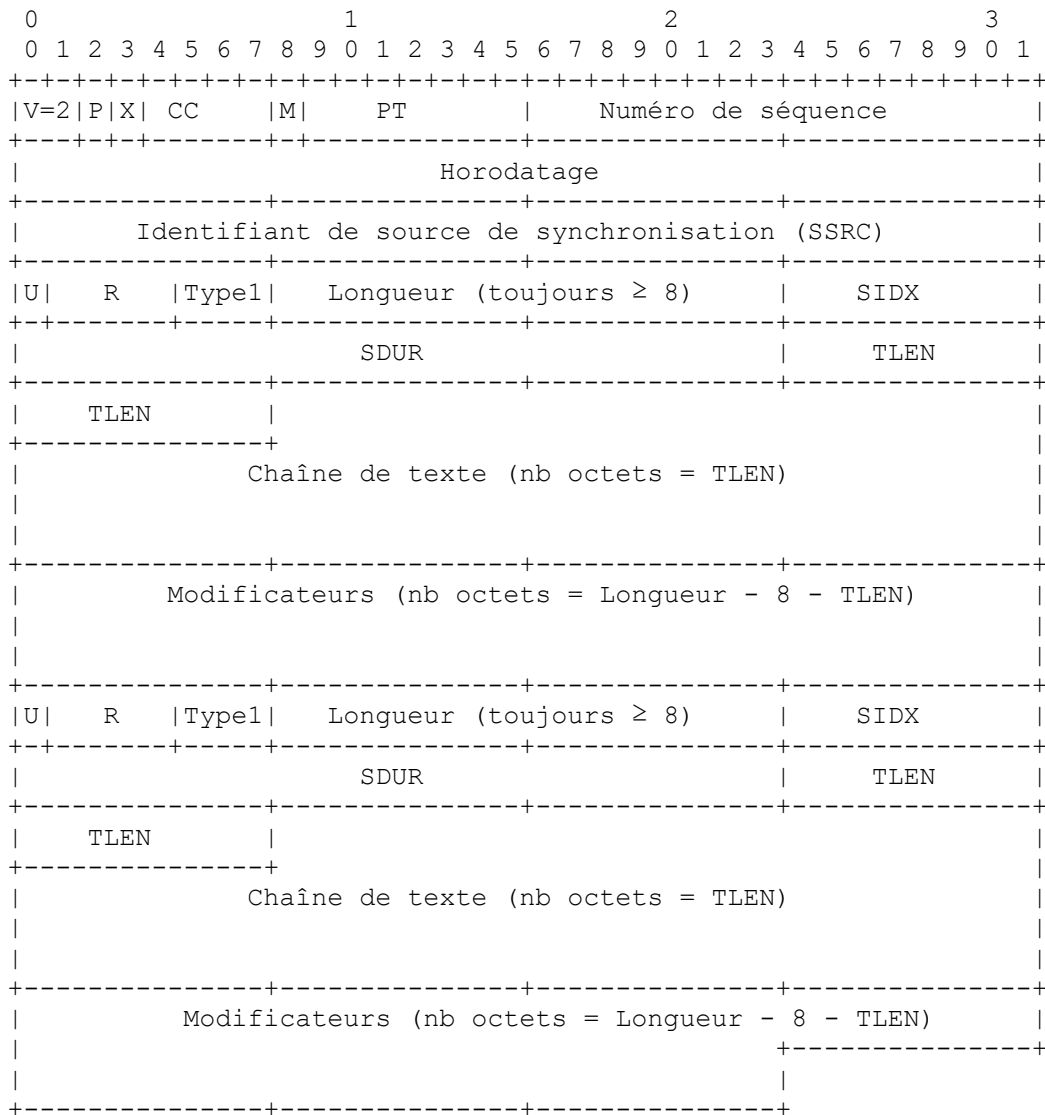
2. pour PAYLOAD 2 : ts(#5) = rtpts2

3. pour PAYLOAD 3 : ts(#6) = ts(#7) = rtpts2 = rtpts3

Selon la configuration 3 ci-dessus, les unités de type2 et 3 devront appartenir au même échantillon. Donc, rtpts3 doit être égal à rtpts2. Pour la même raison, la valeur de SDUR ne sera pas utilisée pour calculer l'horodatage de la prochaine unité.

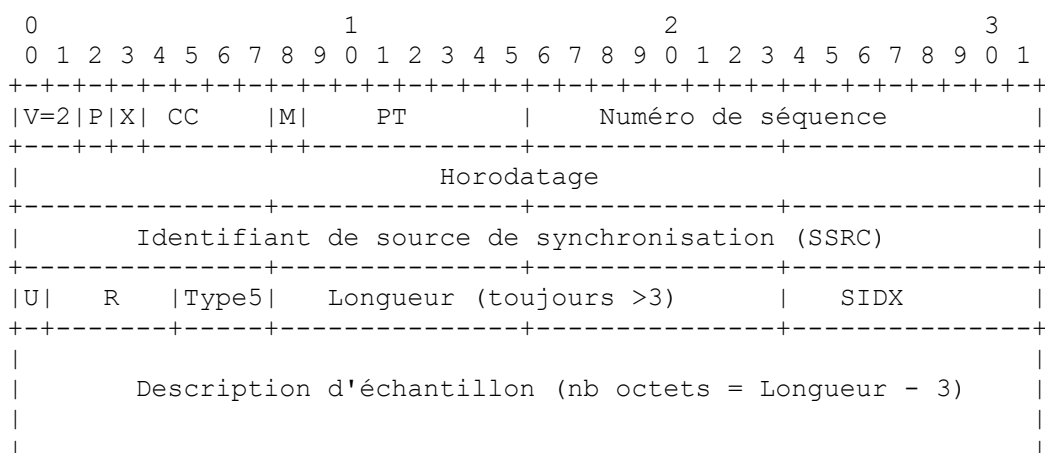
#### 4.7 Exemples de charges utiles

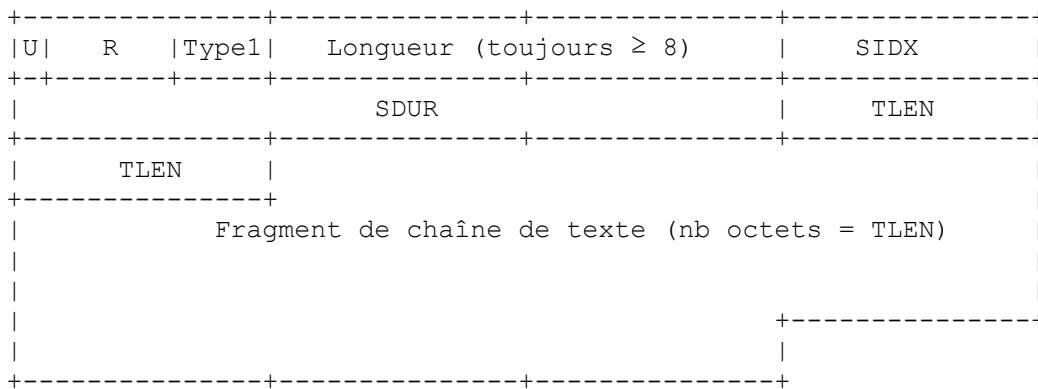
Des exemples de charges utiles utilisant les en-têtes définis sont montrés ci-dessous :



**Figure 12. Charge utile portant deux unités de type 1**

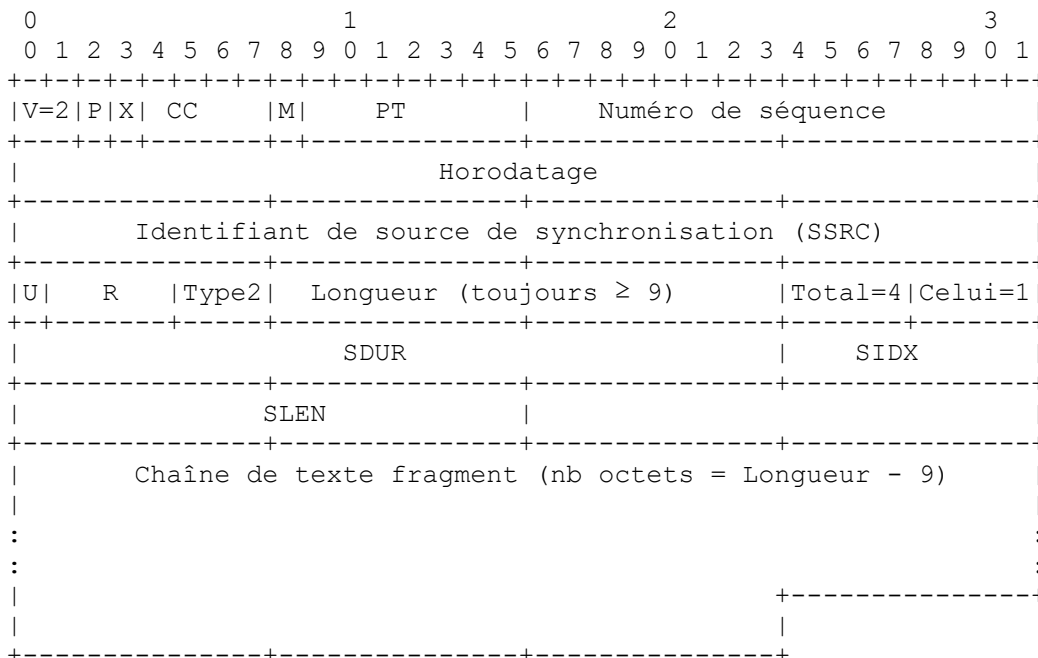
À la Figure 12, un paquet RTP portant deux unités de type 1 est décrit. On peut voir comment les champs Longueur et TLEN peuvent être utilisés pour trouver le début de la prochaine unité (Longueur), le début des modificateurs (TLEN), et la longueur des modificateurs (Longueur - TLEN).





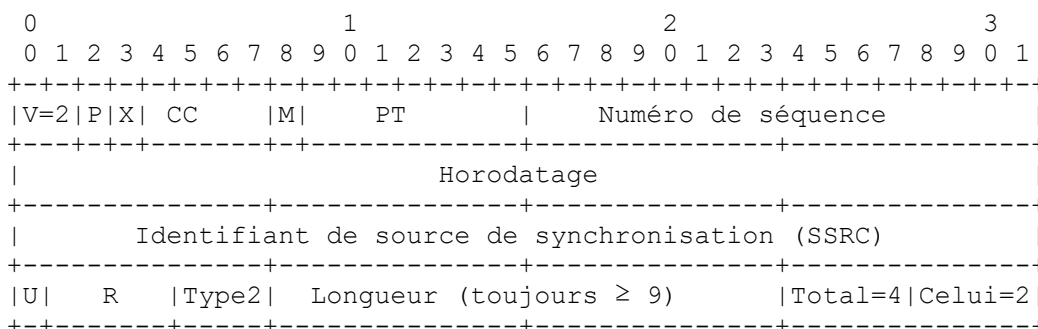
**Figure 13. Paquet RTP portant une unité de Type 5 et une unité de Type 1**

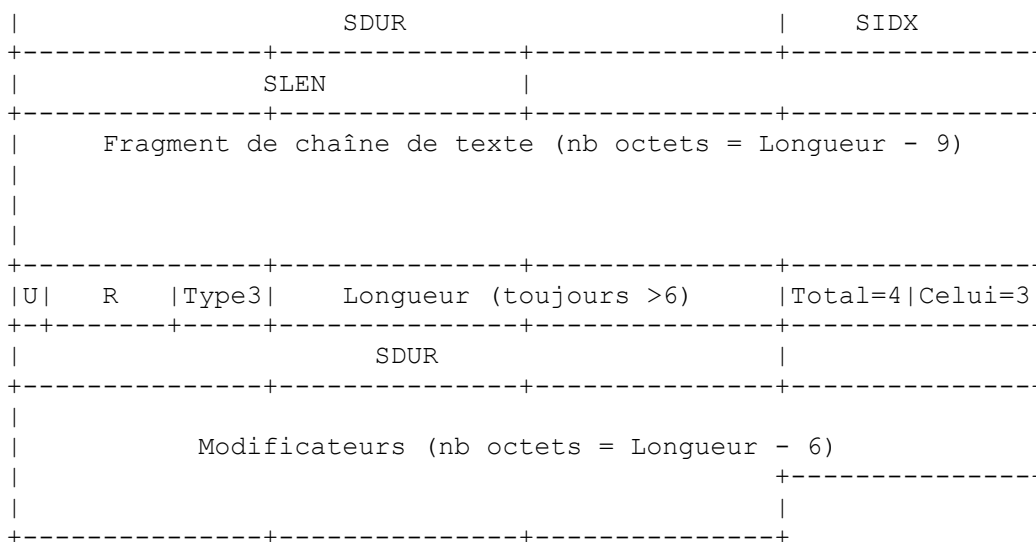
Dans la Figure 13, une description d'échantillon et une unité de Type 1 sont agrégées. L'unité de Type 1 se trouve ne contenir que des chaînes de texte et est petite, de sorte qu'une unité supplémentaire de Type 5 est incluse pour tirer parti des bits disponibles dans le paquet.



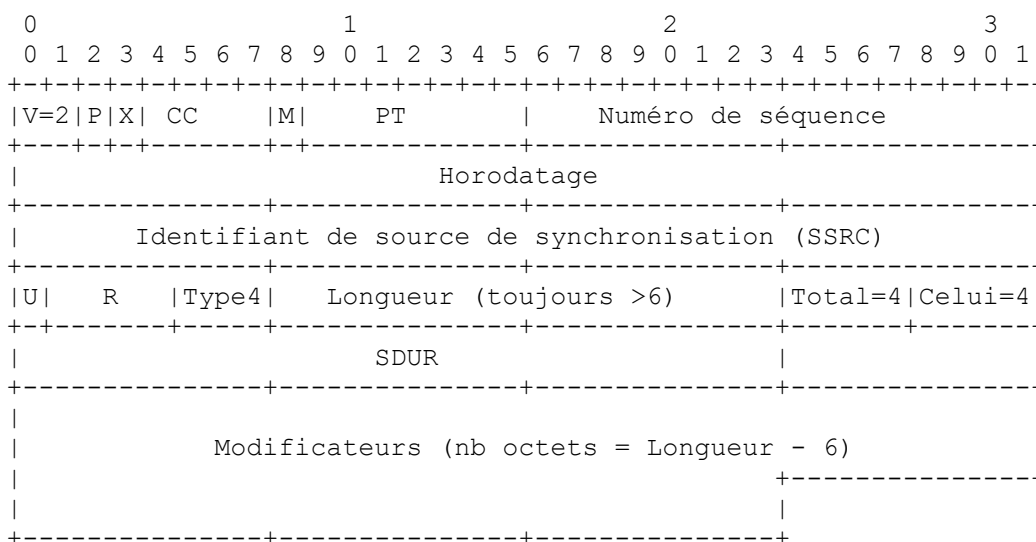
**Figure 14. Charge utile avec le premier fragment de chaîne de texte d'un échantillon**

Dans les Figures 14, 15, et 16, un échantillon de texte est partagé en trois paquets RTP. Dans la Figure 14, la chaîne de texte est grosse et prend toute la longueur du paquet. Dans la Figure 15 est représentée la seule possibilité pour porter deux fragments du même échantillon de texte (voir la configuration 3 au paragraphe 4.6). Le dernier paquet, montré à la Figure 16, porte le dernier fragment de modificateur, une unité de type 4.





**Figure 15. Paquet RTP portant une unité de Type 2 et une unité de Type 3**



**Figure 16. Paquet RTP portant le dernier fragment de modificateurs (Type 4)**

#### 4.8 Relations avec la RFC 3640

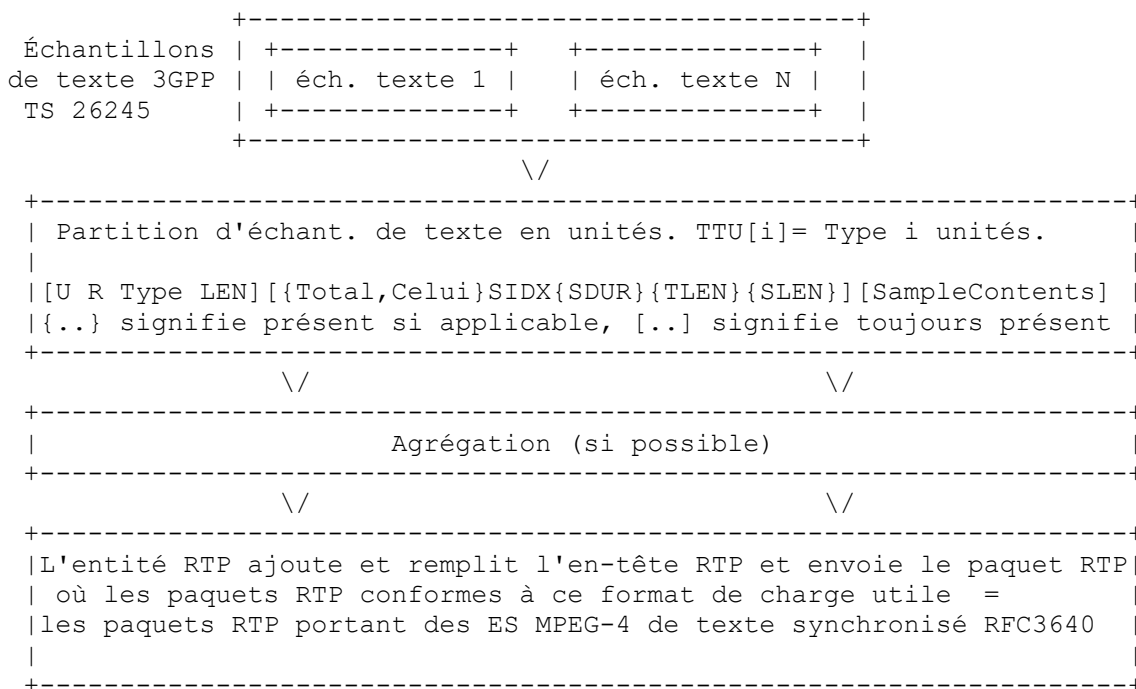
La [RFC3640] définit un format de charge utile pour le transport de tout flux élémentaire MPEG-4 non multiplexé. Un des divers types de flux élémentaires MPEG-4 est le flux MPEG-4 de texte synchronisé, spécifié dans MPEG-4 partie 17 [ISO14496-17], aussi connu comme ISO/CEI 14496-17. Les flux MPEG-4 de texte synchronisé sont capables de porter les données de texte synchronisé 3GPP, comme spécifié dans le TS 26.245 du 3GPP [PSS-TTF].

Les flux MPEG-4 de texte synchronisé sont intentionnellement construits de façon à garantir l'interopérabilité entre la RFC 3640 et le présent format de charge utile. Cela signifie que la construction des paquets RTP portant du texte synchronisé est la même. C'est-à-dire que le flux élémentaire MPEG-4 de texte synchronisé selon la norme ISO/CEI 14496-17 est identique aux charges utiles (agrégées) construites en utilisant ce format de charge utile.

La Figure 17 illustre le processus de construction d'un paquet RTP contenant du texte synchronisé. Comme on peut le voir dans le bloc de partition, les unités (transport) utilisées dans ce format de charge utile sont identiques aux unités de texte synchronisé (TTU, *Timed Text Unit*) définies dans la norme ISO/CEI 14496-17. De même, les règles pour l'agrégation de charge utile selon le paragraphe 4.6 sont identiques à celles définies dans ISO/CEI 14496-17 et sont conformes à la RFC 3640. Par suite, un paquet RTP qui utilise ce format de charge utile est identique à un paquet RTP qui utilise la RFC 3640 pour convoyer des TTU selon la norme ISO/CEI 14496-17. En particulier, MPEG-4 Partie 17 spécifie que quand on utilise la RFC 3640 pour transporter des flux de texte synchronisé, la valeur du paramètre "streamType" est réglée à



0x0D, et la valeur de "objectTypeIndication" dans "config" prend la valeur 0x08.



**Figure 17. Relation avec la RFC 3640**

Note : l'utilisation de la RFC 3640 pour le transport de données ISO/CEI 14496-17 n'exige aucun nouveau paramètre SDP ou aucune nouvelle définition de mode.

#### 4.9 Relations avec la RFC 2793

La [RFC2793] et sa révision, la [RFC4103], spécifient un protocole pour permettre une conversation de texte. Les applications normales de ce format de charge utile sont des terminaux de communication de texte et des outils de conférence de texte. Le contenu d'une session de texte est spécifié dans la Recommandation UIT-T T.140 [T.140]. Le texte de T.140 est codé en UTF-8 comme spécifié dans [T.140] sans autre tramage. Le bloc T140 contient un ou plusieurs éléments de code T.140 comme spécifié dans T.140. Les éléments de code sont des séquences de contrôle comme "Nouvelle ligne", "Interruption", "Terminaison de chaîne", ou "Début de chaîne". La plupart des éléments de code T.140 sont des caractères [ISO10646-1] seuls, mais certains sont des séquences de plusieurs caractères. Chaque caractère est codé en UTF-8 [RFC3629] sur un ou plusieurs octets.

Ce format de charge utile peut aussi être utilisé pour des applications de conversation (même pour la messagerie instantanée). Cependant, ce n'est pas sa cible principale. La caractéristique qui différencie le format de support de texte synchronisé 3GPP est qu'il permet la décoration du texte. Ceci est particulièrement utile dans les présentations multimédia, le karaoke, les annonces commerciales, les nouveaux tiqueurs, les chaînes de texte cliquables, et les légendes. Le contenu de texte T.140 utilisé dans la RFC 2793 ne permet pas l'utilisation des décorations de texte.

De plus, le format de charge utile RTP de texte conversationnel recommande une méthode pour inclure du texte redondant provenant de paquets déjà transmis afin de réduire le risque de perte de texte causé par des pertes de paquets. Ainsi les charges utiles vont inclure une copie redondante de la dernière charge utile envoyée. Ce format de charge utile ne décrit pas une telle méthode, mais elle est aussi applicable ici. Comme expliqué à la Section 5, la redondance de paquet DEVRAIT être utilisée chaque fois que possible. Les lignes directrices d'agrégation du paragraphe 4.6 permettent les charges utiles redondantes.

## 5. Transport résilient

À part les lignes directrices de base de la fragmentation décrites à la section précédente, la plus simple option pour un transport résilient à la perte de paquet est la répétition du paquet. Ce mécanisme peut consister en un mécanisme strict de

répétition fondé sur la fenêtre, ou simplement un mécanisme de répétition dans un sens plus large, où nouveaux et anciens paquets sont mêlés, par exemple.

Un serveur PEUT décider d'utiliser la répétition comme une mesure de résilience à la perte de paquet. Ainsi, un serveur PEUT envoyer les mêmes charges utiles RTP ou juste certaines des unités des charges utiles.

Comme pour le cas des charges utiles complètes, les unités répétées une seule fois DOIVENT correspondre exactement aux mêmes unités envoyées dans la première transmission ; c'est-à-dire, si une fragmentation est nécessaire, elle DEVRA n'être effectuée qu'une seule fois pour chaque échantillon de texte. C'est seulement alors qu'un receveur peut utiliser les unités déjà reçues et les unités répétées pour reconstruire l'échantillon de texte original. Comme l'horodatage RTP est utilisé pour regrouper les fragments d'un échantillon, il faut veiller à préserver l'ordonnance des unités lors de la construction de nouveaux paquets RTP.

Par exemple, si un échantillon de texte était à l'origine envoyé comme un seul échantillon de texte non fragmenté (une unité de type 1) une répétition de cet échantillon DOIT être envoyée aussi comme un seul échantillon de texte non fragmenté en une unité. De même, si l'échantillon de texte original était fragmenté et étalé sur plusieurs paquets RTP (disons, un total de 3 unités) les fragments répétés DEVRONT alors aussi avoir les mêmes limites d'octets et utiliser les mêmes en-têtes et octets d'unités par fragment.

Avec la répétition, les unités répétées se résolvent en le même horodatage que leur original. Lorsque des unités redondantes sont disponibles, une seule d'entre elles DEVRA être utilisée.

Concernant les champs d'en-tête RTP :

- o Si la charge utile RTP entière est répétée, tous les champs spécifiques de charge utile dans l'en-tête RTP (les champs M, TS et PT) DOIVENT garder leurs valeurs d'origine excepté le numéro de séquence, qui DOIT être incrémenté pour se conformer à RTP (les champs Total/CeluiCi permettent de ré-assembler les fragments avec des numéros de séquence différents).
- o Dans les paquets qui contiennent des unités répétées une seule fois, les règles générales de la Section 3 pour allouer des valeurs aux champs d'en-tête RTP s'appliquent. Garder la valeur de l'horodatage RTP pour préserver l'ordonnancement des unités est particulièrement pertinent ici.

À part la répétition, d'autres mécanismes comme la FEC [RFC2733], la retransmission [RFC4588], ou des techniques similaires pourraient être utilisées pour faire face aux pertes de paquets.

## 6. Contrôle d'encombrement

Le contrôle d'encombrement pour RTP DEVRA être mis en œuvre conformément à RTP [RFC3550] et au profil RTP applicable, par exemple, RTP/AVP [RFC3551].

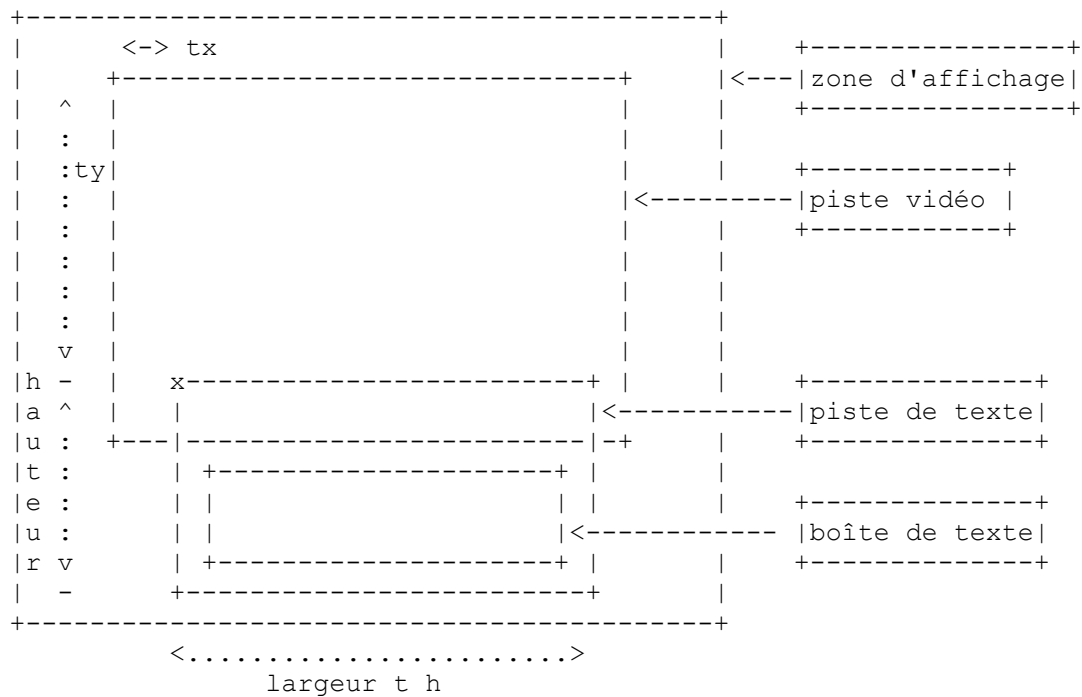
Lors de l'utilisation de ce format de charge utile, deux facteurs principaux peuvent affecter le contrôle d'encombrement :

- o l'utilisation de l'agrégation (d'unités) peut rendre le format de charge utile plus efficace en bande passante, en évitant une surcharge d'en-têtes et en réduisant donc le débit binaire utilisé ;
- o l'utilisation de mécanismes de transport résilients : bien que les applications de texte synchronisé fonctionnent normalement à de faibles débits binaires, l'augmentation due au transport résilient devra être considérée pour le mécanisme de contrôle d'encombrement. Cela s'applique à tous les mécanismes mais plus particulièrement aux moins efficaces comme la répétition.

## 7. Description d'un scénario

### 7.1 Position et composition d'un rendu de texte

Afin d'établir une session de texte synchronisé, sans considération de ce que le flux est mémorisé dans un fichier 3GP ou écoulé en direct, des informations sur la disposition initiale sont nécessaires aux homologues communicants.



**Figure 18. Illustration de la position et composition d'un rendu de texte**

Les paramètres utilisés pour la négociation de la position et la taille de la piste de texte dans la zone d'affichage sont montrés à la Figure 18. Ce sont la "largeur" et la "hauteur" de la piste de texte, ses valeurs de translation, "tx" et "ty", et sa "couche" ou proximité de l'utilisateur.

En même temps, l'envoyeur du flux a besoin de savoir les capacités du receveur. Dans ce cas, les valeurs maximum admissibles pour la hauteur et la largeur de la piste de texte : "max-h" et "max-w", pour le flux que le receveur devra afficher.

Ces informations de disposition DOIVENT être convoyées dans une forme fiable avant le début de la session, par exemple, durant l'annonce de session ou dans un échange offre/réponse (O/R). Un exemple de transport fiable peut être le canal hors bande utilisé pour SDP. Les Sections 8 et 9 donnent des détails sur la transposition de ces paramètres en descriptions SDP et leur usage en O/R.

Pour le contenu mémorisé, les valeurs de disposition qui expriment les propriétés de flux DOIVENT être obtenues de la boîte d'en-tête de piste. Voir au paragraphe 7.3.

Pour les flux de direct, les valeurs appropriées négociées durant l'établissement de session devront être utilisées.

## 7.2 Usage de SMIL

Les attributs contenus dans les boîtes d'en-tête de piste d'un fichier 3GP spécifient seulement les relations spatiales des pistes au sein du fichier 3GP.

Si plusieurs fichiers 3GP sont envoyés, ils requièrent une synchronisation spatiale. Par exemple, pour un flux de texte et de vidéo, les positions des pistes de texte et de vidéo de la Figure 18 doivent être déterminées. À cette fin, le langage SMIL [SMIL] PEUT être utilisé.

SMIL alloue des régions de l'affichage à chacun de ces fichiers et place les pistes dans ces régions. Généralement, dans SMIL, la position d'une piste (ou flux) est exprimée par rapport à une autre piste. Ceci est différent du fichier 3GP, où le coin supérieur gauche est la référence pour tous les décalages de translation. Donc, c'est seulement si la position dans SMIL est relative à l'origine de la piste de vidéo que ce décalage de translation a la même valeur que (tx, ty) dans le fichier 3GP.

Noter aussi que les informations d'en-tête de la piste originale ne sont utilisées pour chaque piste qu'au sein de sa région, comme allouée par SMIL. Donc, même si la description de scène SMIL est utilisée, les éléments d'information d'en-tête de

piste DEVRAIENT être envoyées de toutes façons, car elles représentent les propriétés intrinsèques du support. Voir les détails dans le profil de langage SMIL du 3GPP dans [PSS-SMIL].

### 7.3 Trouver les valeurs de présentation dans un fichier 3GP

Dans un fichier 3GP, au sein de la boîte d'en-tête de piste (tkhd, *Track Header Box*) :

- o tx, ty : ces valeurs spécifient le décalage de translation de la piste (texte) par rapport au coin supérieur gauche de la piste vidéo, si elle est présente. Ce sont l'avant dernière et l'antépénultième valeurs dans la matrice de l'unité ; les valeurs sont à virgule fixe 16.16, restreinte à être des entiers (signés) (c'est-à-dire que les 16 bits de poids faible de chaque valeur doivent être des zéros). Donc, seuls les 16 bits premiers bits sont utilisés pour obtenir la valeur des paramètres de type de support.
- o largeur, hauteur : ils ont le même nom dans la boîte tkhd. Tous les 32 bits (non signé) sont significatifs.
- o couche : tous les 16 bits (signés) sont utilisés.

## 8. Type de support texte synchronisé 3GPP

Le sous type de support pour le codec de texte synchronisé 3GPP est alloué à partir de l'arborescence standard. Le type de support de niveau supérieur sous lequel ce format de charge utile est enregistré est "vidéo". Cet enregistrement est fait en utilisant le gabarit défini dans la [RFC4288] et suivant la [RFC3555].

Le receveur DOIT ignorer tout paramètre non reconnu.

Type de support : vidéo

Sous type de support : 3gpp-tt

Paramètres exigés :

débit : voir la Section 3 de la RFC 4396.

sver : le paramètre "sver" contient une liste des versions rétro-compatibles prises en charge de la spécification du format de texte synchronisé (3GPP TS 26.245) que l'envoyeur accepte de recevoir (et qui sont les mêmes qu'il voudrait envoyer).

La première valeur est la valeur préférée à recevoir (ou préférée à envoyer). La première valeur PEUT être suivie par une liste des versions, séparées par des virgules, qui DEVRAIENT être utilisées comme remplacements. L'ordre est significatif, la première étant la préférée et la dernière celle qui l'est le moins. Chaque entrée a le format  $Z_i(x_i*256+y_i)$ , où "Z<sub>i</sub>" est le numéro de la livraison et "x<sub>i</sub>" et "y<sub>i</sub>" sont tirés de la version de la spécification 3GPP (c'est-à-dire, vZ<sub>i</sub>.x<sub>i</sub>.y<sub>i</sub>). Par exemple, pour 3GPP TS 26.245 v6.0.0,  $Z_i(x_i*256+y_i)=6(0)$ , la valeur de version est "60". (Noter que "60" est l'enchaînement des valeurs  $Z_i=6$  et  $(x_i*256+y_i)=0$  et non leur produit.)

Si aucune valeur "sver" n'est disponible, par exemple, lors de l'écoulement d'un fichier 3GP, la valeur par défaut "60", correspondant à la version de la livraison 6 du 3GPP TS 26.245, DEVRA être utilisée.

Paramètres facultatifs :

tx : ce paramètre indique le décalage de translation horizontale en pixels de la piste de texte par rapport à l'origine de la piste vidéo. Cette valeur est la représentation en décimal d'un entier signé de 16 bits. Voir dans TS 3GPP 26.245 une illustration de ce paramètre.

ty : ce paramètre indique le décalage de translation verticale en pixels de la piste de texte par rapport à l'origine de la piste vidéo. Cette valeur est la représentation en décimal d'un entier signé de 16 bits. Voir dans TS 3GPP 26.245 une illustration de ce paramètre.

couche : ce paramètre indique la proximité de la piste de texte pour celui qui la voit. Des valeurs négatives signifient plus proche de celui qui voit. Ce paramètre n'a pas d'unités. Cette valeur est la représentation en décimal d'un entier signé de 16 bits.

tx3g : ce paramètre DOIT être utilisé pour porter des descriptions d'échantillons hors bande. Il contient une liste d'entrées séparées par des virgules codées en base64. Les entrées de cette liste PEUVENT suivre un ordre quelconque et la liste NE DEVRA PAS être vide. Chaque entrée est le résultat d'un codage base64 sur l'enchaînement de la valeur de SIDX (statique) comme entier non signé de 8 bits et de la description d'échantillon (statique) pour ce SIDX, dans cet ordre. Le format d'une entrée de description d'échantillon peut être trouvé dans 3GPP TS 26.245 livraison 6 et ultérieures. Tous

les serveurs et clients DOIVENT comprendre ce paramètre et DOIVENT être capables d'utiliser la ou les descriptions d'échantillon qu'il contient. Se référer à la [RFC3548] pour les détails du codage base64.

largeur : ce paramètre indique la largeur en pixels de la piste de texte ou de la zone de texte envoyé. Cette valeur est la représentation en décimal d'un entier non signé de 32 bits. Voir dans TS 3GPP 26.245 une illustration de ce paramètre.

hauteur : ce paramètre indique la hauteur en pixels de la piste de texte envoyé. Cette valeur est la représentation en décimal d'un entier non signé de 32 bits. Voir dans TS 3GPP 26.245 une illustration de ce paramètre.

max-w : ce paramètre indique les capacités d'affichage. C'est la valeur maximum de "largeur" que l'expéditeur de ce paramètre accepte. Cette valeur est la représentation en décimal d'un entier de 32 bits non signé.

max-h : ce paramètre indique les capacités d'affichage. c'est la valeur maximum de "hauteur" que l'expéditeur de ce paramètre accepte. Cette valeur est la représentation en décimal d'un entier de 32 bits non signé.

Considérations de codage : ce type de support est tramé (voir au paragraphe 4.8 de la [RFC4288]) et contient partiellement des données binaires.

Restrictions d'usage : ce type de support dépend du tramage RTP, et n'est donc défini que pour le transfert via RTP [RFC3550]. Le transport dans d'autres protocoles de tramage n'est pas défini pour l'instant.

Considérations de sécurité : se référer à la Section 11 de la RFC 4396.

Considérations d'interopérabilité : le format de support de texte synchronisé 3GPP et sa mémorisation de fichier sont spécifiés dans la livraison 6 de 3GPP TS 26.245, "Service transparent de bout en bout de flux de paquets commutés (PSS) ; Format de texte synchronisé (Livraison 6)". Noter aussi que le 3GPP peut dans de futures livraisons spécifier des extensions ou des mises à jour au format de support de texte synchronisé d'une façon rétro compatible, par exemple, de nouvelles boîtes de modificateur ou des extensions aux descriptions d'échantillons. Le format de charge utile défini dans la RFC 4396 permet de telles extensions. Pour les futures livraisons du 3GPP du format de texte synchronisé, le paramètre "sver" est utilisé pour identifier la spécification exacte utilisée. Le format de mémorisation défini pour le format de texte synchronisé 3GPP est le format de fichier 3GPP (3GP) [PSS-FF]. Les fichiers 3GP peuvent être transférés en utilisant le type de support vidéo/3gpp enregistré par la [RFC3839]. Le format de fichier 3GPP est un fichier conteneur qui peut contenir, par exemple, de l'audio et vidéo qui peuvent être synchronisés avec le texte synchronisé 3GPP.

Spécification publiée : RFC 4396

Applications qui utilisent ce type de support : applications de flux multimédia.

Informations supplémentaires : le format de support de texte synchronisé 3GPP est spécifié dans le TS 26.245 du 3GPP, "Service transparent de bout en bout de flux de paquets commutés (PSS) ; Format de texte synchronisé (Livraison 6)". Ce document et les futures extensions au format de texte synchronisé 3GPP sont disponibles au public à <http://www.3gpp.org>.

Numéro magique : aucun

Extension de fichier : aucune.

Code de type de fichier Macintosh : aucun.

Adresse personnelle & de messagerie à contacter pour d'autres informations :

Jose Rey, [jose.rey@eu.panasonic.com](mailto:jose.rey@eu.panasonic.com)

Yoshinori Matsui, [matsui.yoshinori@jp.panasonic.com](mailto:matsui.yoshinori@jp.panasonic.com)

Groupe de travail Transport audio/vidéo.

Utilisation prévue : COMMUNE

Auteurs : Jose Rey ; Yoshinori Matsui

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

## 9. Usage de SDP

### 9.1 Transposition en SDP

Les informations portées dans la spécification de type de support ont une transposition spécifique dans les champs de SDP [RFC2327]. Si SDP est utilisé pour spécifier des sessions qui utilisent ce format de charge utile, la transposition est faite comme suit :

- o Le type de support ("vidéo") passe dans le "m=" de SDP comme le nom de support.

m=video <numéro d'accès> RTP/<profil RTP> <type de charge utile dynamique>

- o Le sous type de support ("3gpp-tt") et le débit d'horloge d'horodatage "rate" (la valeur RECOMMANDÉE de 1000 Hz ou une autre) passe dans la ligne SDP "a=rtpmap" comme, respectivement, le nom et le débit de codage :

a=rtpmap:<type de charge utile> 3gpp-tt/1000

- o Le paramètre EXIGÉ "sver" passe dans l'attribut SDP "a=fmtp" en le copiant directement de la chaîne de type de support comme une paire paramètre/valeur séparés par un point virgule .

- o Les paramètres FACULTATIFS "tx", "ty", "couche", "tx3g", "largeur", "hauteur", "max-w" et "max-h" passent dans l'attribut SDP "a=fmtp" en les copiant directement de la chaîne de type de support comme des paires paramètre/valeur séparées par un point virgule :

a=fmtp:<type de charge utile dynamique> <nom de paramètre>=<valeur>[,<valeur>][; <nom de paramètre>=<valeur>]

- o Tout paramètre inconnu de l'appareil qui utilise SDP DEVRA être ignoré. Par exemple, les paramètres ajoutés au format de support dans des spécifications ultérieures PEUVENT être copiés dans SDP et DEVRONT être ignorés par les receveurs qui ne les comprennent pas.

### 9.2 Usage de paramètre dans le modèle SDP d'offre/réponse

Dans ce paragraphe, on explique la signification des paramètres SDP définis dans ce document au sein du contexte d'offre/réponse [RFC3264].

En envoi individuel, expéditeur et receveur négocient normalement le flux, c'est-à-dire, quelles valeurs de codecs et de paramètres seront utilisés dans la session. Ceci est aussi possible dans une moindre mesure en diffusion groupée.

De plus, la signification des paramètres PEUT varier selon la direction utilisée. Dans les paragraphes qui suivent, une "offre de <directionality>" signifie une offre qui contient un flux réglé à <directionality>. <directionality> peut prendre les valeurs sendrecv, sendonly, et recvonly. Des considérations similaires s'appliquent aux réponses. Par exemple, une réponse à une offre "sendonly" est une réponse "recvonly".

#### 9.2.1 Usage en envoi individuel

Les types de paramètres suivants sont utilisés dans ce format de charge utile :

1. Paramètres déclaratifs : l'offreur et celui qui répond déclarent les valeurs qu'ils vont utiliser pour le flux entrant (sendrecv/recvonly) ou sortant (sendonly). L'offreur et celui qui répond PEUVENT utiliser des valeurs différentes.
  - a. "tx", "ty", et "couche" : Ces paramètres décrivent où est placé la piste de texte reçu. Selon la direction :
    - i. Ils DOIVENT apparaître dans toutes les offres et réponses "sendrecv" et dans toutes les offres et réponses "recvonly" (s'appliquant donc au flux entrant). Dans le cas des offres et réponses "sendrecv" et des offres "recvonly", ces valeurs DEVRAIENT être utilisées par l'expéditeur du flux sauf si il a une préférence particulière, et dans ce cas, il DOIT s'assurer que ces valeurs différentes ne corrompent pas la présentation. Pour les réponses "recvonly", le répondant PEUT accepter les valeurs proposées pour le flux entrant (dans une offre "sendonly", voir en ii. ci-dessous) ou répondre avec des valeurs différentes. L'offreur DOIT utiliser les valeurs retournées.
    - ii. Ils PEUVENT apparaître dans des offres "sendonly" et DOIVENT apparaître dans les réponses "sendonly". Dans les offres "sendonly", ils spécifient les valeurs que l'offreur propose d'envoyer (voir l'exemple du paragraphe 9.3). Dans

les réponses "sendonly", ces valeurs DEVRAIENT être copiées de l'offre "recvonly" correspondante à l'acceptation du flux, sauf si une préférence particulière du receveur du flux existe, comme expliqué au point précédent.

2. Paramètres décrivant les capacités d'affichage, "max-h" et "max-w", qui indiquent les dimensions maximum de la piste de texte (zone d'affichage de texte) pour les valeurs "tx" et "ty" du flux entrant (voir la Figure 18). "max-h" et "max-w" DOIVENT être inclus dans toutes les offres et réponses lorsque "tx" et "ty" se réfèrent au flux entrant, excluant donc les offres et réponses "sendonly" (voir l'exemple du paragraphe 9.3) où ils NE DEVRONT PAS être présents.
3. Paramètres décrivant les propriétés du flux envoyé, c'est-à-dire, l'expéditeur du flux décide de leurs valeurs :
  - a. "largeur" et "hauteur" spécifient les dimensions de la piste de texte. Ils DEVRONT toujours être présents dans les offres et réponses "sendrecv" et "sendonly". Pour les réponses "recvonly", le répondant DOIT inclure les valeurs de paramètre offertes (si il en est) sans en rien changer dans la réponse qui accepte le flux.
  - b. "tx3g" contient les descriptions d'échantillons statiques. Il PEUT être présent dans les seules offres et réponses "sendrecv" et "sendonly". Ce paramètre s'applique aux flux qu'envoient les offreurs ou ceux qui répondent.
4. Paramètres négociables, qui DOIVENT faire l'objet d'un accord. C'est le cas de "sver". Ce paramètre DOIT être présent dans chaque offre et réponse. Celui qui répond DEVRA choisir une réponse acceptée dans la liste de l'offreur, ou autrement il DOIT supprimer le flux ou rejeter la session.
5. Paramètres symétriques : "rate", débit d'horloge d'horodatage, appartient à cette classe. Les paramètres symétriques DOIVENT être retournés en écho exactement comme dans la réponse. Autrement, le flux DOIT être retiré ou la session rejetée.

Le tableau suivant résume toutes les options :

Type de paramètre	Direction/ O ou A	sendrecv	recvonly	sendonly
Déclaratif	tx, ty, layer	M/M	M/M	m/M
Capacités d'affichage	max-h, max-w	M/M	M/M	-/-
Propriétés du flux	hauteur, largeur, tx3g	M/M m/m	-/ (M) -/-	M/M m/m
Négociable	sver	M/M	M/M	M/M
Symétrique	rate	M/M	M/M	M/M

**Table 1. Usage de paramètres dans l'offre/réponse en envoi individuel**

Légende :

- o M signifie DOIT être présent.
- o m signifie PEUT être présent (comme des valeurs proposées).
- o (M) ou (m) signifie DOIT ou PEUT, si applicable.
- o un tiret ("-") signifie que le paramètre NE DOIT PAS être présent.

Autres observations concernant l'usage des paramètres :

- o Valeurs de translation et transparence : dans les offres "sendonly", "tx", "ty", et "layer" indiquent les valeurs proposées. C'est utile pour les sessions composées visuellement où les différents flux occupent des parties différentes de l'affichage, par exemple, un flux de vidéo et les sous titres. Ce sont juste des valeurs suggérées ; l'homologue qui rend le texte décide en fin de compte où placer la piste de texte.

- o Dimensions de la piste (zone) de texte, "hauteur" et "largeur" : dans le cas des offres "sendonly", un répondant qui accepte l'offre DOIT être prêt à rendre le flux en utilisant ces valeurs. Si une de ces conditions n'est pas satisfaite, le flux DOIT être supprimé ou la session rejetée.
- o Capacités d'affichage, "max-h" et "max-w" : un répondant qui envoie un flux DEVRA s'assurer que les valeurs de "hauteur" et "largeur" dans la réponse, sont compatibles avec les capacités signalées par l'offreur.
- o Traitement de version via "sver" : l'idée est que l'offreur et le répondant communiquent en utilisant la même version. Ceci se fait en laissant le répondant choisir sur une liste de versions prises en charge, "sver". Pour les flux "recvonly", la première valeur de la liste est la version préférée à recevoir. Par conséquent, pour les flux "sendonly" (et "sendrecv") la première valeur est celle préférée pour l'envoi (et la réception). Le répondant DOIT choisir une valeur et la retourner dans la réponse. À réception de la réponse, l'offreur DEVRA être prêt à envoyer ("sendonly" et "sendrecv") et recevoir ("recvonly" et "sendrecv") un flux utilisant cette version. Si aucune des versions de la liste n'est prise en charge, le flux DOIT être supprimé ou la session rejetée. Noter que, si des versions non compatibles sont offertes, cela DEVRA alors être fait en utilisant des types différents de charge utile.

### 9.2.2 Usage de diffusion groupée

En diffusion groupée, l'usage des paramètres est similaire au cas d'envoi individuel, sauf comme suit :

- o les paramètres "tx", "ty", et "layer" dans les offres en diffusion groupée n'ont de signification que pour les flux "sendrecv" et "recvonly". Afin que tous les clients aient la même vision de la session, ils DOIVENT être utilisés de façon symétrique.
- o pour "hauteur", "largeur", et "tx3g" (pour "sendrecv" et "sendonly") les offres en diffusion groupée spécifient quelles valeurs de ces paramètres les participants DOIVENT utiliser pour envoyer. Donc, si le flux est accepté, le répondant DOIT aussi les inclure exactement tels quels dans la réponse (et aussi "tx3g", si présent).
- o les paramètres de capacité, "max-h" et "max-w", NE DEVRONT PAS être utilisés en diffusion groupée. Si la piste de texte offerte devait changer en taille, une nouvelle offre DEVRA être utilisée à la place.
- o Concernant le traitement de version : dans le cas des offres en diffusion groupée, un répondant PEUT accepter une offre en diffusion groupée tant que une des versions de la liste de "sver" est prise en charge. Donc, si le flux est accepté, le répondant DOIT choisir sa version préférée, mais, à la différence de l'envoi individuel, l'offreur NE DEVRA PAS changer le flux offert de cette version choisie parce que il peut y avoir d'autres participants à la session qui ne prennent pas en charge les extensions plus récentes. Par conséquent, les différents participants à la session peuvent finir par utiliser des versions de format de supports différents rétro compatibles. Il est RECOMMANDÉ que l'offre en diffusion groupée contienne un nombre limité de versions, afin que tous les participants aient la même vue de la session. C'est de la responsabilité du créateur de la session. Si aucune des versions offertes n'est acceptée, le flux DEVRA être supprimé ou la session rejetée. Aussi dans ce cas, si des versions de remplacement non compatibles sont offertes, cela DEVRA être fait en utilisant des types de charge utile différents.

### 9.3 Exemples d'offre/réponse

Dans ces exemples d'O/R en envoi individuel, les longues lignes sont coupées. Les descriptions d'échantillons statiques sont raccourcies pour que l'exemple soit plus clair.

Pour "sendrecv" :

O -> R

m=video <accès> RTP/AVP 98

a=rtpmap:98 3gpp-tt/1000

a=fmtp:98 tx=100; ty=100; layer=0; hauteur=80; largeur=100; max-h=120; max-w=160; sver=6256,60; tx3g=81...

a=sendrecv

R -> O

m=video <accès> RTP/AVP 98..

a=rtpmap:98 3gpp-tt/1000

a=fmtp:98 tx=100; ty=95; layer=0; hauteur=90; largeur=100; max-h=100; max-w=160; sver=60; tx3g=82...

a=sendrecv



Dans cet exemple, l'offreur dit au répondant où il va placer le flux reçu et ce que sont la hauteur et la largeur maximum admissibles pour le flux qu'il va recevoir. Aussi, il dit au répondant les dimensions de la piste de texte pour le flux envoyé et quelle description d'échantillon il devra utiliser. Il offre deux versions, 6256 et 60. Le répondant répond avec un ensemble équivalent de paramètres pour le flux qu'il reçoit. Dans ce cas, les "max-h" et "max-w" du répondant sont compatibles avec les "hauteur" et "largeur" de l'offreur. Autrement, le répondant aurait du supprimer ce flux, et l'offreur aurait du produire une nouvelle offre prenant en compte les capacités du répondant. Ceci n'est possible que si plusieurs types de charge utile sont présents dans l'offre initiale afin qu'au moins une d'elles corresponde aux capacités du répondant telles qu'exprimées par le "max-h" et "max-w" dans la réponse négative. Noter aussi que les dimensions de la boîte de texte du répondant tiennent dans les valeurs maximales signalées dans l'offre. Finalement, le répondant choisit d'utiliser la version 60 du format de texte synchronisé.

Pour "recvonly" :

Offreur -> Répondant

m=video <accès> RTP/AVP 98

a=rtpmap:98 3gpp-tt/1000

a=fmtp:98 tx=100; ty=100; layer=0; max-h=120; max-w=160; sver=6256,60

a=recvonly

R -> O

m=video <accès> RTP/AVP 98..

a=rtpmap:98 3gpp-tt/1000

a=fmtp:98 tx=100; ty=100; layer=0; hauteur=90; largeur=100; sver=60; tx3g=82...

a=sendonly

Dans ce cas, l'offre est différente de celle du cas précédent : elle n'inclut pas les propriétés "hauteur", "largeur", et "tx3g" du flux. Le répondant copie les valeurs "tx", "ty", et "layer", en accusant donc réception. "max-h" et "max-w" ne sont pas présents dans la réponse parce que les paramètres "tx" et "ty" (et "layer") ne s'appliquent pas dans ce cas particulier au flux reçu, mais au flux envoyé. Aussi, si l'offreur et le répondant avaient de très différentes tailles d'affichage, il ne serait pas possible d'exprimer les capacités du répondant. Dans l'exemple ci-dessus et pour un répondant avec un affichage de 50x50, les valeurs de translation sont déjà hors gamme.

Pour "sendonly" :

O -> R

m=video <accès> RTP/AVP 98

a=rtpmap:98 3gpp-tt/1000

a=fmtp:98 tx=100; ty=100; layer=0; hauteur=80; largeur=100; sver=6256,60; tx3g=81...

a=sendonly

R -> O

m=video <accès> RTP/AVP 98..

a=rtpmap:98 3gpp-tt/1000

a=fmtp:98 tx=100; ty=100; layer=0; hauteur=80; largeur=100; max-h=100; max-w=160; sver=60

a=recvonly

Noter que "max-h" et "max-w" ne sont pas présents dans l'offre. Aussi, avec cette réponse, le répondant va accepter l'offre telle qu'elle (faisant donc écho à "tx", "ty", "hauteur", "largeur", et "layer") et de plus informe l'offreur sur ses capacités : "max-h" et "max-w".

Une autre réponse possible pour ce cas serait :

R -> O

m=video <accès> RTP/AVP 98..

a=rtpmap:98 3gpp-tt/1000

a=fmtp:98 tx=120; ty=105; layer=0; max-h=95; max-w=150; sver=60

a=recvonly

Dans ce cas, le répondant n'accepte pas les valeurs offertes. L'offreur DOIT utiliser ces valeurs ou supprimer le flux.

#### 9.4 Usage de paramètres en dehors de offre/réponse

SDP peut aussi être employé en dehors du contexte Offre/Réponse, par exemple pour des sessions multimédia qui sont annoncées avec le protocole d'annonce de session (SAP, *Session Announcement Protocol*) [RFC2974] ou qui s'écoulent avec le protocole de flux en temps réel (RTSP, *Real Time Streaming Protocol*) [RFC2326].

Dans ce cas, le receveur d'une description de session est obligé de prendre en charge les paramètres et les valeurs données pour les flux, ou autrement il DOIT rejeter la session. Il est de la responsabilité de l'envoyeur (ou créateur) des descriptions de session de définir les paramètres de session afin que la probabilité d'échec de l'établissement de la session soit minimisée. Ceci sort du domaine d'application de ce document.

## 10. Considérations relatives à l'IANA

L'IANA a enregistré le nom de sous type de support "3gpp-tt" pour le type de support "vidéo" comme spécifié à la Section 8 du présent document.

## 11. Considérations sur la sécurité

Les paquets RTP qui utilisent le format de charge utile défini dans la présente spécification sont soumis aux considérations sur la sécurité discutées dans la spécification RTP [RFC3550] et tout profil RTP applicable, par exemple, AVP [RFC3551].

En particulier, un attaquant peut invalider l'ensemble courant des descriptions d'échantillons actives chez le client au moyen de la répétition d'un paquet avec une vieille description d'échantillon, c'est-à-dire, une attaque en répétition. Cela voudrait dire que l'affichage du texte serait corrompu, si il s'affiche. Une autre forme d'attaque peut consister en l'envoi de fragments redondants, dont les limites ne correspondent pas aux limites exactes des originaux (comme indiqué par le champ Longueur) ou des fragments qui portent des longueurs d'échantillon différentes (SLEN). Ceci peut causer la défaillance du décodeur.

Ces types d'attaques peuvent être facilement évités en utilisant l'authentification de la source et la protection de l'intégrité. De plus, les homologues dans une session de texte synchronisé peuvent désirer conserver une certaine confidentialité à leur communication. Ce format de charge utile ne fournit aucun mécanisme pour réaliser cela. La protection de la confidentialité, de l'intégrité, et l'authentification doit être résolue par un mécanisme externe à ce format de charge utile, par exemple, SRTP [RFC3711].

## 12. Références normatives

### 12.1 Références normatives

[ISO14496-12] ISO/CEI 14496-12, "Technologie de l'information - Codage des objets audio-visuels - Partie 12 : format de fichier de support de base ISO", 2004.

[PSS-TTF] W3C, "Transparent end-to-end packet switched streaming service (PSS); Timed Text Format (Release 6)", TS 26.245 v 6.0.0, juin 2004.

[RFC2119] S. Bradner, "[Mots clés à utiliser](#) dans les RFC pour indiquer les niveaux d'exigence", BCP 14, mars 1997. (MàJ par [RFC8174](#))

[RFC2327] M. Handley et V. Jacobson, "SDP : [Protocole de description de session](#)", avril 1998. (*Obsolète; voir [RFC4566](#)*)

[RFC3548] S. Josefsson, "Codages de données Base16, Base32, et Base64", juillet 2003. (*Obsolète, voir [4648](#)*) (*Info*)

[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](#))

## 12.2 Références pour information

- [ISO10646-1] ISO/CEI 10646-1, "Jeu de caractères universel codé sur plusieurs octets", (1993).
- [ISO14496-17] ISO/CEI FCD 14496-17, "Technologie de l'information - Codage des objets audio-visuels - Partie 17 : format de flux de texte", juin 2004.
- [PSS6] W3C, "Transparent end-to-end packet switched streaming service (PSS); Protocols et codecs (Release 6)", TS 26.234 v 6.1.0, septembre 2004.
- [PSS-FF] W3C, "Transparent end-to-end packet switched streaming service (PSS); 3GPP file format (3GP) (Release 6)", TS 26.244 V6.3. mars 2005.
- [PSS-SMIL] W3C, "Transparent end-to-end Packet-switched Streaming Service (PSS); 3GPP SMIL language profile, (Release 6)", TS 26.246 v 6.0.0, juin 2004.
- [RFC2326] H. Schulzrinne, A. Rao et R. Lanphier, "Protocole de [flux directs en temps réel](#) (RTSP)", avril 1998. (*Remplacée par [RFC7826](#)*)
- [RFC2354] C. Perkins et O. Hodson, "Options pour [réparer un support de direct](#)", juin 1998. (*Information*)
- [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.)
- [RFC2781] P. Hoffman et F. Yergeau, "[UTF-16](#), un codage de la norme ISO 10646", février 2000.
- [RFC2793] G. Hellstrom, "Charge utile RTP pour conversation textuelle", mai 2000. (*Obsolète, voir [RFC4103](#)*) (P.S.)
- [RFC2974] M. Handley, C. Perkins, E. Whelan, "Protocole d'annonce de session (SAP)", octobre 2000. (*Expérimentale*)
- [RFC3264] J. Rosenberg et H. Schulzrinne, "[Modèle d'offre/réponse](#) avec le protocole de description de session (SDP)", juin 2002.
- [RFC3551] H. Schulzrinne et S. Casner, "[Profil RTP pour conférences audio](#) et vidéo avec contrôle minimal", STD 65, juillet 2003.
- [RFC3555] S. Casner et P. Hoschka, "Enregistrement de type MIME pour les types de charge utile RTP", juillet 2003.
- [RFC3611] T. Friedman, R. Caceres et A. Clark, éditeurs, "[Rapports étendus du protocole](#) de contrôle de RTP (RTCP XR)", novembre 2003. (P.S.)
- [RFC3629] F. Yergeau, "[UTF-8, un format de transformation](#) de la norme ISO 10646", STD 63, novembre 2003.
- [RFC3640] J. van der Meer et autres, "[Format de charge utile RTP pour le transport de flux élémentaires MPEG-4](#)", novembre 2003. (P.S., *MàJ par [RFC5691](#)*)
- [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.)
- [RFC4103] G. Hellstrom, P. Jones, "[Charge utile RTP pour conversation textuelle](#)", juin 2005. (*Remplace [RFC2793](#)*) (P.S.)
- [RFC4288] N. Freed et J. Klensin, "Spécifications du [type de support et procédures d'enregistrement](#)", [BCP 13](#), décembre 2005.
- [RFC4585] J. Ott et autres, "Profil RTP étendu pour rétroaction fondée sur le protocole de contrôle de transport en temps réel (RTCP) (RTP/AVPF)", juillet 2006. (P.S., *MàJ par [RFC8108](#)*)

[RFC4588] J. Rey et autres, "Format de charge utile de retransmission RTP", juillet 2006. (*P.S.*)

[SMIL] W3C, "Synchronised Multimedia Integration Language (SMIL 2.0)", août 2001.

[T.140] Recommandation UIT-T T.140 (1998) "Protocole de conversation de texte pour application multimédia", avec amendement 1, (2000).

### 13. Bases de la structure de fichier 3GP

Cette Section donne une vue d'ensemble grossière de la structure du fichier 3GP, qui suit le format de base de support ISO [ISO14496-12].

Chaque fichier 3GP consiste en "boîtes". En général, un fichier 3GP contient la boîte Type de fichier (ftyp), la boîte Image (moov), et la boîte Données de support (mdat). La boîte Type de fichier identifie le type et les propriétés du fichier 3GP lui-même. La boîte Image et la boîte Données de support servant de conteneurs, incluent leurs propres boîtes pour chaque support. Les boîtes commencent par un en-tête, qui indique la taille et le type (ces champs sont appelés, "taille" et "type"). De plus, chaque type de boîte peut inclure un certain nombre de boîtes.

Dans ce qui suit, seules sont mentionnées les boîtes qui sont utiles pour les besoins de ce format de charge utile.

La boîte Image (moov) contient une ou plusieurs boîtes Piste (trak) qui incluent des informations sur chaque piste. Une boîte Piste contient, entre autres choses, la boîte En-tête de piste (tkhd), la boîte En-tête de support (mdhd), et la boîte Informations sur le support (minf).

La boîte En-tête de piste spécifie les caractéristiques d'une seule piste, où une piste est, dans ce cas, le texte écoulé durant une session. Exactement une boîte En-tête de piste est présente pour une piste. Elle contient les informations sur la piste, comme la disposition spatiale (largeur et hauteur) la matrice de transformation vidéo, et le numéro de couche. Comme ces éléments d'informations sont essentiels et statiques (c'est-à-dire, constants) pour la durée de la session, ils doivent être envoyés avant la transmission de tout échantillon de texte.

La boîte En-tête de support contient "l'échelle des temps" ou le nombre d'unités de temps qui passent en une seconde, c'est-à-dire, les cycles par seconde ou Hertz. La boîte Informations de support inclut la boîte Tableau d'échantillons (stbl) qui contient tous les indices de temps et de données des échantillons du support dans la piste. En utilisant cette boîte, il est possible de localiser les échantillons dans le temps et de déterminer leur type, taille, conteneur, et décalage au sein de ce conteneur. Dans la boîte Tableau d'échantillons, on peut trouver la boîte Description d'échantillons (stsd, pour trouver les descriptions d'échantillons) la boîte Heure de décodage d'échantillon (stts, pour trouver la durée de l'échantillon) la boîte Taille d'échantillon (stsz) et la boîte Tronçons d'échantillon (stsc, pour trouver l'indice de description d'échantillon).

Finalement, la boîte Données de support contient les données du support elles-mêmes. Dans les pistes de texte synchronisé, cette boîte contient les échantillons de texte. Son équivalent en audio et vidéo sont les trames audio et vidéo, respectivement. L'échantillon de texte consiste en la longueur de texte, la chaîne de texte, et une ou plusieurs boîtes de modificateur. La longueur de texte est la taille du texte en octets.

La chaîne de texte est du texte en clair à rendre. La boîte Modificateur est de l'information à rendre en plus du texte, comme la couleur, la fonte, etc.

### 14. Remerciements

Les auteurs tiennent à remercier Dave Singer, Jan van der Meer, Magnus Westerlund, et Colin Perkins de leurs commentaires et suggestions sur ce document.

Les auteurs souhaitent aussi remercier Markus Gebhard pour l'éditeur ASCII gratuit et publiquement disponible JavE (utilisé pour les dessins ASCII de ce document) et Henrik Levkowitz pour le service web Idnits.

## Adresse des auteurs

Jose Rey  
Panasonic R&D Center Germany GmbH  
Monzastr. 4c  
D-63225 Langen, Germany  
mél : [jose.rey@eu.panasonic.com](mailto:jose.rey@eu.panasonic.com)  
téléphone : +49-6103-766-134  
Fax : +49-6103-766-166

Yoshinori Matsui  
Matsushita Electric Industrial Co., LTD.  
1006 Kadoma  
Kadoma-shi, Osaka, Japan  
mél : [matsui.yoshinori@jp.panasonic.com](mailto:matsui.yoshinori@jp.panasonic.com)  
téléphone : +81 6 6900 9689  
Fax : +81 6 6900 9699

## 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](http://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](mailto: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.