

Groupe de travail Réseau
Request for Comments : 2250
 RFC rendue obsolète : 2038
 Catégorie : En cours de normalisation
 Traduction Claude Brière de L'Isle

D. Hoffman & G. Fernando
 Sun Microsystems, Inc.
 V. Goyal, Precept Software, Inc.
 M. Civanlar, AT&T Labs - Research
 janvier 1998

Format de charge utile RTP pour la vidéo MPEG1/MPEG2

Statut de ce mémoire

Le présent document spécifie un protocole Internet en cours de normalisation pour la communauté de l'Internet, et appelle à des discussions et des suggestions pour son amélioration. Prière de se reporter à l'édition actuelle du STD 1 "Normes des protocoles officiels de l'Internet" pour connaître 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 (1998). Tous droits réservés

Résumé

Le présent mémoire décrit un schéma de mise en paquet pour les flux MPEG vidéo et audio. Le schéma proposé peut être utilisé pour transporter un tel flux vidéo ou audio sur les protocoles de transport pris en charge par RTP. Deux approches sont décrites. La première est conçue pour assurer une interopérabilité maximale avec les environnements de systèmes MPEG. La seconde est conçue pour assurer une compatibilité maximale avec les autres flux de supports encapsulés dans RTP et le futur travail de l'IETF sur le contrôle de conférence.

Le présent mémoire est une révision de la RFC2038, protocole en cours de normalisation de l'Internet. Dans cette révision, les mécanismes de résilience à la perte de paquet du paragraphe 3.4 ont été étendus pour inclure des informations supplémentaires d'en-tête d'image exigées par MPEG2. On a ajouté un nouveau paragraphe sur les considérations pour la sécurité sur ce type de charge utile.

Table des Matières

1. Introduction.....	1
2. Encapsulation de système MPEG et de flux de transport.....	2
2.1 Utilisation de l'en-tête RTP.....	3
3. Encapsulation des flux élémentaires MPEG.....	3
3.1 Flux élémentaires MPEG vidéo.....	3
3.2 Flux élémentaires MPEG audio.....	4
3.3 En-tête RTP fixe pour encapsulation ES MPEG.....	4
3.4 En-tête MPEG spécifique de la vidéo.....	4
3.5 En-tête MPEG spécifique de l'audio.....	6
4. Considérations pour la sécurité.....	6
Appendice 1 Récupération d'erreur et stratégies de resynchronisation.....	6
Références.....	8
Adresse des auteurs.....	8
Déclaration complète de droits de reproduction.....	8

1. Introduction

Le groupe de travail ISO/CEI JTC1/SC29 (aussi appelé comité MPEG) a défini la norme MPEG1 (ISO/CEI 11172) [ISO11172] et la norme MPEG2 (ISO/CEI3818) [ISO13818]. Le présent mémoire décrit un schéma de mise en paquet pour transporter les flux MPEG vidéo et audio en utilisant le protocole de transport en temps réel 'RTP, *Real-time Transport Protocol*), version 2 [RFC1889], [RFC1890].

La spécification MPEG1 est définie en trois parties : système, vidéo et audio. Elle est principalement conçue pour les applications fondées sur le CD-ROM, et est optimisée pour des taux de données combinés d'approximativement 1,5 Mbit/s. Les portions vidéo et audio de la spécification décrivent le format de base des flux vidéo ou audio. Ces formats définissent les flux élémentaires (ES, *Elementary Stream*). La spécification du système MPEG1 définit une encapsulation de l'ES qui contient des horodatages de présentation (PTS, *Presentation Time Stamp*), des horodatages de décodage (DTS, *Decoding Time Stamp*) et des références d'horloge système, et effectue le multiplexage des ES vidéo et audio MPEG1 compressés

avec les données d'utilisateur.

La spécification MPEG2 est structurée de façon similaire. Cependant, elle n'a pas été restreinte aux seules applications sur CD-ROM. La spécification du système MPEG2 définit deux formats de flux de système : le flux de transport MPEG2 (MTS, *MPEG2 Transport Stream*) et le flux de programme MPEG2 (MPS, *MPEG2 Program Stream*). Le MTS est construit pour communiquer avec, ou pour mémoriser, un ou plusieurs programmes de données MPEG2 compressées et aussi d'autres données dans des environnements relativement enclins à l'erreur. Le MPS est conçu pour des environnements relativement libres d'erreur.

On cherche à réaliser l'interopérabilité parmi quatre types de systèmes d'extrémité dans la spécification qui suit. Les quatre types sont :

1. Unité d'interfonctionnement en transmission (TIU, *Transmitting Interworking Unit*)

Elle reçoit les informations MPEG d'un système MTS natif pour les distribuer sur les réseaux des paquets en utilisant une couche système fondée sur RTP natif (telle qu'un inter réseau fondé sur IP). Exemples : coder en temps réel, liaison satellite MTS à l'Internet, serveur vidéo avec matériau de source codé en MTS.

2. Unité d'interfonctionnement en réception (RIU, *Receiving Interworking Unit*)

Elle reçoit les informations MPEG en temps réel d'un réseau fondé sur RTP pour les transmettre à un environnement MTS natif. Exemples : d'un serveur vidéo fondé sur Internet à un complexe de distribution par câble fondé sur MTS.

3. Système d'extrémité Internet de transmission (TAES, *Transmitting Internet End-System*)

Il transmet les informations MPEG générées ou mémorisées au sein du système d'extrémité Internet lui-même, ou reçues de réseaux informatiques fondés sur l'Internet. Exemple : serveur vidéo.

4. Système d'extrémité Internet de réception (RAES, *Receiving Internet End-System*)

Il reçoit les informations MPEG sur un internet fondé sur RTP pour la consommation par le système d'extrémité internet ou transmission au réseau informatique traditionnel. Exemple : ordinateur de bureau ou station de travail pour voir des vidéos d'entraînement.

Chacun des deux types d'émetteur doit fonctionner avec chacun des deux types de receveur. Comme il est probable que le TAES, et certain que le RAES, seront fondés sur les ordinateurs existants et prévus connectés à l'Internet, il est tout à fait souhaitable que le protocole interopérable soit fondé sur RTP.

À cause de la gamme d'applications qui peuvent employer des flux MPEG, on propose de définir deux formats de charge utile.

La communauté MPEG est très intéressée par l'utilisation d'un des codages du système MPEG, et donc, la Section 2 propose des encapsulations de flux de système MPEG1 et de flux de transport et de programme MPEG2 avec RTP. Ce profil prend en charge toute la sémantique du système MPEG et offre l'interopérabilité de base entre les quatre types de systèmes d'extrémité.

Lorsque il ne fonctionne qu'entre des systèmes d'extrémité fondés sur internet (c'est-à-dire TAES et RAES) on désire un format de charge utile qui fournit une plus grande compatibilité avec l'architecture de l'Internet, renvoyant certains des problèmes du système aux autres protocoles définis dans la communauté de l'Internet (comme le groupe de travail MMUSIC). La Section 3 propose une encapsulation de données vidéo et audio compressées (qu'on appelle dans la documentation MPEG des "flux élémentaires" (ES)) conforme aussi bien à MPEG1 qu'à MPEG2. Ici, aucun des standard systèmes de MPEG1 ou MPEG2 n'est utilisé. Les ES sont directement encapsulés avec RTP.

Tout au long de la présente spécification, on fait une utilisation extensive de la terminologie MPEG. Le lecteur devrait consulter les principales références MPEG pour avoir les descriptions définitives de cette terminologie.

2. Encapsulation de système MPEG et de flux de transport

Chaque paquet RTP va contenir un horodatage déduit de la référence d'horloge à 90 kHz de l'envoyeur. Cette horloge est synchronisée à la référence d'horloge programme (PCR, *Program Clock Reference*) ou à la référence d'horloge système (SCR, *System Clock Reference*) du flux de système et représente l'heure de transmission cible du premier octet de la charge utile du paquet. L'horodatage RTP ne sera pas passé au décodeur MPEG. Cette utilisation de l'horodatage est un peu différente de ce qui est normalement le cas dans RTP, en ce qu'il n'est pas considéré comme étant l'horodatage de l'affichage ou de la présentation du support. Le principal objet de l'horodatage RTP sera d'estimer et réduire toute gigue induite par le réseau et de synchroniser le décalage temporel relatif entre l'émetteur et le receveur.

Pour les flux de transport MPEG2 la charge utile RTP va contenir un nombre entier de paquets de transport MPEG. Pour éviter les inefficacités aux systèmes d'extrémité, les données provenant de multiples petits paquets MTS (normalement fixés à une taille de 188 octets) sont agrégées en un seul paquet RTP. Le nombre de paquets de transport contenus est calculé en divisant la longueur de la charge utile RTP par la longueur d'un paquet MTS (188).

Pour les flux de programme MPEG2 et les flux systèmes MPEG1, il n'y a pas de restrictions de mise en paquet ; ces flux sont traités comme un flux d'octets mis en paquets.

2.1 Utilisation de l'en-tête RTP

Les champs de l'en-tête RTP sont utilisés comme suit, :

Type de charge utile : des types distincts de charge utile devraient être alloués aux flux système MPEG1, aux flux de programme MPEG2 et aux flux de transport MPEG2. Voir dans la [RFC1890] les allocations de type de charge utile.

Bit M : réglé à 1 chaque fois que l'horodatage est discontinu (comme ce qui peut arriver lorsque un envoyeur passe d'une source de données à une autre). Cela permet au receveur et à tout mélangeur ou traducteur RTP intermédiaire qui sont synchronisés au flux d'ignorer la différence entre cet horodatage et tout horodatage antérieur dans son détecteur de phase d'horloge.

Horodatage : horodatage de 32 bits à 90 kHz représentant le temps de transmission cible pour le premier octet du paquet.

3. Encapsulation des flux élémentaires MPEG

Les types ES suivants peuvent être encapsulés directement en RTP :

- (a) MPEG1 Vidéo (ISO/IEC 11172-2)
- (b) MPEG2 Vidéo (ISO/IEC 13818-2)
- (c) MPEG1 Audio (ISO/IEC 11172-3)
- (d) MPEG2 Audio (ISO/IEC 13818-3)

Un type de charge utile RTP distinct est alloué respectivement à la vidéo MPEG1/MPEG2 et à l'audio MPEG1/MPEG2. Il n'est pas nécessaire de fournir d'autres indications sur le fait que les données sont MPEG1 ou MPEG2 dans les en-têtes RTP ou spécifiques de MPEG de cette encapsulation, car cette information est disponible dans les en-têtes ES.

Des horodatages de présentation (PTS) de 32 bits avec une précision de 90 kHz devront être portés dans l'en-tête fixe RTP. Tous les paquets qui constituent une trame audio ou vidéo devront avoir le même horodatage.

3.1 Flux élémentaires MPEG vidéo

La vidéo MPEG1 peut être distinguée de la vidéo MPEG2 à l'en-tête de séquence vidéo, c'est-à-dire que pour la vidéo MPEG2 un `sequence_header()` est suivi par `sequence_extension()`. Les profil et niveau particuliers de la vidéo MPEG2 (MAIN_Profile@MAIN_Level, HIGH_Profile@HIGH_Level, etc) sont déterminés par le champ `profile_and_level_indicator` de l'en-tête `sequence_extension` de la vidéo MPEG2.

La sémantique du flux binaire MPEG a été conçue pour des environnements relativement libres d'erreur, et il y a une dépendance significative (à la fois temporelle et spatiale) au sein du flux qui fait que la perte de quelques données rend inutiles les autres données non corrompues. Le format tel que défini dans cette encapsulation utilise les informations de tramage de couche application plus des informations supplémentaires dans l'en-tête RTP spécifique du flux pour permettre certains mécanismes de récupération. L'Appendice 1 suggère plusieurs stratégies de récupération fondées sur les propriétés de cette encapsulation.

Comme les images MPEG peuvent être grandes, elles vont normalement être fragmentées en paquets d'une taille inférieure à la MTU de LAN/WAN typique. Les règles de fragmentation suivantes s'appliquent :

1. Le `Video_Sequence_Header` MPEG, lorsque il est présent, va toujours être au début d'une charge utile RTP.
2. Un `GOP_header` MPEG, lorsque présent, sera toujours au début de la charge utile RTP, ou va suivre un `Video_Sequence_Header`.
3. Un `Picture_Header` MPEG, lorsque présent, va toujours être au début d'une charge utile RTP, ou va suivre un `GOP_header`.

Chaque en-tête ES doit être entièrement contenu dans le paquet. Par conséquent, une taille minimum de charge utile RTP de 261 octets doit être prise en charge pour contenir le plus grand en-tête seul défini dans l'ES (c'est-à-dire, l'en-tête `extension_data()` qui contient la `quant_matrix_extension()`). Autrement, il n'y a pas de restriction sur les endroits où peuvent apparaître les en-têtes au sein de la charge utile du paquet.

Dans MPEG, chaque image est constituée d'une ou plusieurs "tranches," et une tranche est destinée à être l'unité de récupération de perte ou de corruption de données. Un décodeur conforme à MPEG va normalement avancer au début de la prochaine tranche chaque fois que se rencontre une erreur dans le flux. Les bits de début et de fin de tranche MPEG sont fournis dans l'en-tête d'encapsulation pour faciliter cela.

Le début d'une tranche doit être soit les premières données dans un paquet (après tous les en-têtes ES MPEG) soit suivre après un certain nombre entier de tranches dans un paquet. Cette exigence assure que le début de la tranche suivante, après une qui a un paquet manquant, peut être trouvé sans exiger que le receveur examine le contenu du paquet. Les tranches peuvent être fragmentées à travers des paquets pour autant que les règles ci-dessus soient satisfaites.

Une mise en œuvre fondée sur cette encapsulation suppose que le `Video_Sequence_Header` est répété périodiquement dans le flux binaire MPEG. En pratique (bien que ce ne soit pas exigé par la norme MPEG) ceci est utilisé pour permettre la commutation de canal et de recevoir et commencer à décoder un flux binaire MPEG relayé en continu à des points arbitraires du flux de support. Il est suggéré quand on rejoue d'un flux MPEG à partir d'un format de fichier (où le `Video_Sequence_Header` peut n'être représenté qu'au début du flux) que le premier `Video_Sequence_Header` (précédé par un indicateur de fin de flux) soit sauvegardé par l'empaqueteur pour l'injecter périodiquement dans le flux réseau.

3.2 Flux élémentaires MPEG audio

MPEG1 audio peut être distingué de MPEG2 audio à partir de l'en-tête MPEG `ancillary_data()`. Pour l'audio MPEG1 ou MPEG2, des horodatages de présentation distincts peuvent être présents pour des trames qui correspondent soit à 384 échantillons pour la couche I, soit à 1152 échantillons pour la couche II ou la couche III. Le nombre réel d'octets requis pour représenter ce nombre d'échantillons va varier selon les paramètres de codage.

Plusieurs trames audio peuvent être encapsulées au sein d'un paquet RTP. Dans ce cas, un nombre entier de trames audio doit être contenu dans le paquet et l'en-tête de fragmentation défini au paragraphe 3.5 devra être réglé à 0.

Aussi, si des paquets relativement courts doivent être utilisés, une trame peut être si grande qu'elle peut s'étendre sur plusieurs paquets RTP. Par exemple, pour la couche II MPEG audio échantillonnée à un taux de 44,1 kHz, chaque trame représenterait un intervalle de temps de 26,1 ms. À ce taux d'échantillonnage, si le taux binaire compressé est de 384 kbit/s (c'est-à-dire, 48 koctet/s) la taille moyenne de trame audio sera de 1,25 koctet. Si les paquets devaient faire 500 octets, chaque trame audio s'étendrait sur trois paquets RTP.

L'en-tête Indicateur de fragmentation audio (voir au paragraphe 3.5) devra être présent pour qu'un type de charge utile MPEG1/2 assure cette fragmentation.

3.3 En-tête RTP fixe pour encapsulation d'ES MPEG

Les champs d'en-tête RTP sont utilisés comme suit :

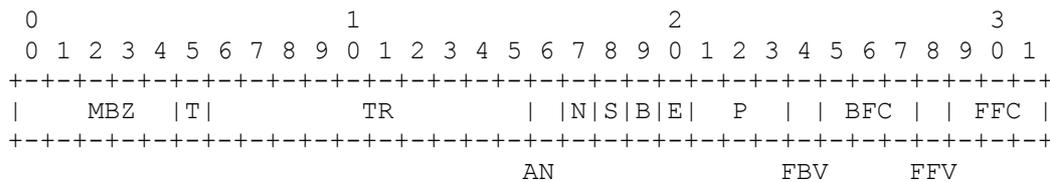
Type de charge utile : des types de charge utile distincts devraient être alloués aux flux vidéo élémentaires et aux flux audio élémentaires. Voir dans la [RFC1890] les allocations de type de charge utile.

Bit M : pour la vidéo, réglé à 1 sur un paquet contenant un code de fin de trame MPEG, à 0 autrement. Pour l'audio, réglé à 1 sur le premier paquet d'une "salve de parole", 0 autrement.

horodatage : horodatage de 32 bits à 90 kHz représentant l'heure de présentation de l'image MPEG ou de la trame audio. Le même pour tous les paquets qui constituent une image ou une trame audio. Peut n'être pas à accroissement monotone dans les flux de vidéo si des images B sont présentes dans le flux. Pour des paquets qui ne contiennent que une séquence vidéo et/ou un en-tête GOP, l'horodatage est montré sur la figure qui suit.

3.4 En-tête MPEG spécifique de la vidéo

Cet en-tête devra être joint à chaque paquet RTP après l'en-tête fixe RTP.



MBZ : non utilisé. Doit être réglé à zéro dans la présente spécification. Cet espace est réservé pour une utilisation future.

T : l'extension d'en-tête spécifique de MPEG-2 est présente (1 bit). Réglé à 1 quand l'extension d'en-tête spécifique de la vidéo MPEG-2 (voir au paragraphe 3.4.1) suit cet en-tête. Cette extension peut être nécessaire pour une meilleure résilience à l'erreur ; mais son inclusion dans un paquet RTP est facultative. (Voir l'Appendice 1.)

TR : référence temporelle (*Temporal-Reference*) (10 bits) de l'image actuelle au sein du GOP actuel. La gamme des valeurs va de 0 à 1023 et est constante pour tous les paquets RTP d'une certaine image.

AN : active le bit N pour la résilience à l'erreur (1 bit). Réglé à 1 lorsque le bit suivant (N) est utilisé pour signaler des changements des informations d'en-tête de l'image pour les charges utiles MPEG-2. Il doit être réglé à 0 pour les charges utiles MPEG-1 ou lorsque le bit N n'est pas utilisé.

N : en-tête de nouvelle image (1 bit). Utilisé pour les charges utiles MPEG-2 lorsque le bit (AN) précédent est réglé à 1. Autrement, il doit être réglé à zéro. Réglé à 1 quand les informations contenues dans les en-têtes d'image transmis précédemment ne peuvent pas être utilisés pour reconstruire un en-tête pour l'image actuelle. Cela arrive lorsque l'image en cours est codée avec un ensemble différent de paramètres que les images précédentes de même type. Le bit N doit être constant pour tous les paquets RTP qui appartiennent à la même image de sorte que la réception de tout paquet provenant d'une image permet de détecter si les informations nécessaires pour la reconstruction étaient contenues dans cette image (N = 1) ou dans une image précédente (N = 0).

S : en-tête Séquence présent (1 bit). Normalement à 0 et réglé à 1 à l'occurrence de chaque en-tête de séquence MPEG. Utilisé pour détecter la présence de l'en-tête de séquence dans un paquet RTP.

B : début de tranche (BS, *Beginning-of-Slice*) (1 bit). Établi lorsque le début de la charge utile du paquet est un code de début de tranche, ou lorsque un code de début de tranche est précédé de seulement un ou plusieurs de Video_Sequence_Header, GOP_header et/ou Picture_Header.

E : fin de tranche (ES, *End-of-Slice*) (1 bit). Établi lorsque le dernier octet de la charge utile est la fin d'une tranche MPEG.

P : type d'image (3 bits). I (1), P (2), B (3) ou D (4). Cette valeur est constante pour chaque paquet RTP d'une certaine image. La valeur 000B est interdite et 101B à 111B sont réservées pour prendre en charge de futures extensions à la spécification d'ES MPEG.

FBV : full_pel_backward_vector

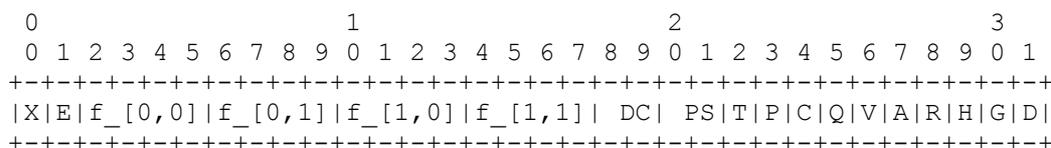
BFC : backward_f_code

FFV : full_pel_forward_vector

FFC : forward_f_code

Obtenues du plus récent en-tête d'image, elles sont constantes pour chaque paquet RTP d'une certaine image. Pour les trames I aucune de ces valeurs n'est présente dans l'en-tête d'image et elles doivent être réglées à zéro dans l'en-tête RTP. Pour les trames P, seules les deux dernières valeurs sont présentes et FBV et BFC doivent être réglées à zéro dans l'en-tête RTP. Pour les trames B, les quatre valeurs sont toutes présentes.

3.4.1 Extension d'en-tête MPEG-2 spécifique de la vidéo



X : non utilisé (1 bit). Doit être réglé à zéro dans la présente spécification. Cet espace est réservé pour un usage futur.

E : Extensions présentes (1 bit). Réglée à 1, cette extension d'en-tête, incluant l'extension d'affichage composite lorsque

D = 1, sera suivie par une ou plusieurs des extensions suivantes : extension de matrice quant, extension d'affichage d'image, extension d'adaptation temporelle d'image, extension d'adaptation spatiale de l'image et extension de droit de reproduction.

Le premier octet de ces données d'extension donne la longueur des extensions dans des mots de 32 bits incluant le champ Longueur lui-même. Des octets de bourrage à zéro sont utilisés à la fin si nécessaire pour aligner les extensions sur des limites de 32 bits.

Comme elles peuvent n'être pas vitales pour le décodage d'une image, l'inclusion d'une de ces extensions dans un paquet RTP est facultative même lorsque l'extension d'en-tête spécifique de la vidéo MPEG-2 est incluse dans le paquet (T = 1) (voir l'Appendice 1). Si elles sont présentes, elles devraient être copiées des extensions correspondantes suivant la plus récente extension de codage d'image MPEG-2 et elles restent constantes pour chaque paquet RTP d'une certaine image.

Le code de début d'extension (32 bits) et l'identifiant de code de début d'extension (4 bits) sont inclus. Donc, les extensions sont auto identifiantes.

F_[0,0] : horizontal vers l'avant, f_code (4 bits)

f_[0,1] : vertical vers l'avant, f_code (4 bits)

f_[1,0] : horizontal vers l'arrière, f_code (4 bits)

f_[1,1] : vertical vers l'arrière, f_code (4 bits)

DC : précision intra DC (2 bits)

PS : structure d'image (2 bits)

T : premier champ supérieur (1 bit)

P : frame_predicted_frame_dct (1 bit)

C : vecteurs de dissimulation du mouvement (1 bit)

Q : type q_scale (1 bit)

V : format intra_vlc (1 bit)

A : examen de remplacement (1 bit)

R : répétition du premier champ (1 bit)

H : type chroma_420 (1 bit)

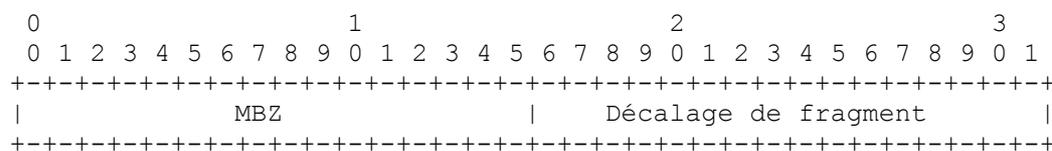
G : trame progressive (1 bit)

D : fanion d'affichage composite (1 bit). Réglé à 1, les 32 bits qui suivent celui-ci contiennent 12 zéros suivis par 20 bits d'informations d'affichage composite.

Ces valeurs sont copiées de la plus récente extension de codage d'image et sont constantes pour chaque paquet RTP d'une certaine image. Leur signification est expliquée dans la norme MPEG-2.

3.5 En-tête MPEG spécifique de l'audio

Cet en-tête devra être joint à chaque paquet RTP au début de la charge utile et après tous en-têtes RTP pour un type de charge utile MPEG1/2 audio.



Décalage de fragment : décalage en octets dans la trame audio pour les données de ce paquet.

4. Considérations pour la sécurité

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

Une menace potentielle de déni de service existe pour les codages de données qui utilisent des techniques de compression qui ont une charge de calcul non uniforme à l'extrémité receveuse. L'attaquant peut injecter des datagrammes pathologiques

dans le flux, qui sont complexes à décoder et causent la surcharge chez le receveur. Cependant, ce codage n'exhibe aucune non uniformité significative.

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

Une revue de la sécurité de ce format de charge utile n'a trouvé aucune considération de sécurité supplémentaire au delà de celles de la spécification RTP.

Appendice 1 Récupération d'erreur et stratégies de resynchronisation

Les stratégies suivantes de récupération d'erreur et de resynchronisation sont destinées à n'être que des lignes directrices. Un receveur conforme est libre d'employer (ou non) des stratégies de remplacement.

Lors du décodage initial d'un flux élémentaire MPEG encapsulé dans RTP, le receveur peut éliminer tous les paquets jusqu'à ce que le bit en-tête de séquence présent soit réglé à 1. À ce moment, des informations d'état suffisantes sont contenues dans le flux pour permettre le traitement par un décodeur MPEG.

La perte de paquets contenant l'en-tête GOP et/ou le Picture_Header est détectée par un changement inattendu des valeurs de référence temporelle et de type d'image. Considérons l'exemple suivant de séquence GOP :

Dans l'ordre de l'affichage : 0B 1B 2I 3B 4B 5P 6B 7B 8P GOP_HDR 0B ...

Dans l'ordre du flux : 2I 0B 1B 5P 3B 4B 8P 6B 7B GOP_HDR 2I ...

Considérons aussi deux compteurs :

ref_pic_temp (Référence temporelle (I,P) de l'image de référence)

dep_pic_temp (Référence temporelle (B) dépendante de l'image)

À chaque début de GOP, on règle ces compteurs à la valeur de la référence temporelle du type d'image correspondant. Pour notre exemple de séquence de GOP, ref_pic_temp = 2 et dep_pic_temp = 0. On continue d'incrémenter LES DEUX compteurs de un à chaque image suivante. Ref_pic_temp devrait correspondre aux références temporelles des trames I et P, et dep_pic_temp devrait correspondre aux références temporelles des trames B.

dep_pic_temp:	-	0	1	2	3	4	5	6	7		8	9		
Ordre du flux :	2I	0B	1B	5P	3B	4B	8P	6B	7B	GOP_H	2I	0B	1B	...
ref_pic_temp:	2	3	4	5	6	7	8	9	10	^	11			
	-----											^		
	Concordance										Aband	Discordance		
	dans ref_pic_temp													

La perte d'un en-tête GOP peut être détectée en faisant correspondre le compteur approprié (sur la base du type d'image) à la valeur de référence temporelle. Une discordance indique la perte d'un en-tête de GOP. Si on le désire, un en-tête de GOP peut être reconstruit en utilisant un code de temps "nul", répétant le fanion closed_gop des en-têtes GOP précédents, et réglant le fanion broken_link (*liaison interrompue*) à 1.

La perte d'un en-tête d'image peut aussi être détectée par une discordance dans la référence temporelle contenue dans le paquet RTP d'après le compteur approprié dep_pic_temp ou ref_pic_temp chez le receveur.

Pour les charges utiles MPEG-1, après examen du prochain début de tranche, l'en-tête d'image est reconstruit à partir des champs P, TR, FBV, BFC, FFV et FFC contenus dans ce paquet, et à partir de valeurs par défaut qui dépendent du flux.

Pour MPEG-2, des informations supplémentaires sont nécessaires pour la reconstruction. Ces informations sont fournies par l'extension d'en-tête spécifique de la vidéo MPEG-2 contenue dans ce paquet si le bit T est réglé à 1, ou par l'en-tête d'image pour l'image en cours qui peut être disponible à partir des paquets précédents appartenant à la même image. La stratégie de l'émetteur pour l'inclusion de l'extension d'en-tête spécifique de la vidéo MPEG-2 peut dépendre d'un certain

nombre de facteurs. Cet en-tête peut n'être pas nécessaire quand :

1. les informations ont été transmises un nombre de fois suffisant dans des paquets précédents pour assurer la réception avec la probabilité désirée, ou
2. les informations sont transmises sur un canal fiable distinct, ou
3. les taux de perte attendus sont assez bas pour que les trames manquantes ne soient pas un problème, ou
4. préserver la bande passante est plus important que la résilience à l'erreur, etc.

Si $T=1$ et $E=0$, il peut y avoir des extensions présentes dans le flux binaire original de la vidéo qui ne sont pas incluses dans le paquet en cours. L'émetteur peut choisir de ne pas inclure d'extension dans un paquet lorsque elles ne sont pas nécessaires pour le décodage, ou si un des cas énumérés ci-dessus pour ne pas inclure l'extension d'en-tête spécifique de la vidéo MPEG-2 dans un paquet s'applique seulement aux données d'extension.

Si $N=0$, alors l'en-tête d'image provenant d'une image précédente du même type (I, P ou B) peut être utilisée tant qu'au moins un paquet a été reçu pour chaque image concernée du même type et que le bit N était 0 pour chacune de ces images. Cela peut impliquer :

1. de sauvegarder les informations d'en-tête d'image pertinentes qui peuvent être obtenues de l'extension d'en-tête spécifique de la vidéo MPEG-2 ou directement du flux binaire vidéo pour chaque type d'image,
2. de conserver des indicateurs de validité pour ces informations sauvegardées sur la base des bits N reçus et des paquets perdus, et,
3. de mettre à jour les données chaque fois qu'est reçu un paquet avec $N=1$.

Si les informations nécessaires ne sont pas disponibles à partir d'une de ces sources, il est conseillé de supprimer les données jusqu'à ce qu'arrive un code de début de nouvelle image.

Chaque fois qu'un paquet RTP est perdu (comme l'indique un trou dans les numéros de séquence RTP) le receveur peut éliminer tous les paquets jusqu'à ce que le bit Début de tranche soit établi. À ce moment, des informations d'état suffisantes sont contenues dans le flux pour permettre le traitement par un décodeur MPEG qui va commencer à la prochaine limite de tranche (éventuellement après reconstruction de l'en-tête de GOP et/ou l'en-tête d'image comme décrit ci-dessus).

Références

- [ISO11172] Norme internationale ISO/CEI 11172, "Codage des images animées et du son associé pour les supports de mémorisations numériques jusqu'à environ 1,5 Mbit/s", novembre 1993.
- [ISO13818] Norme internationale ISO/CEI 13818, "Codage générique des images animées et des informations de son associées", novembre 1994.
- [RFC1112] S. Deering, "Extensions d'hôte pour [diffusion groupée sur IP](#)", STD 5, août 1989. (*Mise à jour par la RFC2236*)
- [RFC1889] H. Schulzrinne, S. Casner, R. Frederick et V. Jacobson, "RTP : protocole de transport pour applications en temps réel", janvier 1996. (*Obsolète, voir [RFC3550](#) STD64*)
- [RFC1890] H. Schulzrinne, "Profil RTP pour conférences audio et vidéo avec contrôle minimal", janvier 1996. (*Obsolète, voir [RFC3551](#) (P.S.)*)

Adresse des auteurs

Gerard Fernando
Sun Microsystems, Inc.
Mail-stop UMPK14-305
2550 Garcia Avenue
Mountain View, California 94043-1100
USA
téléphone : +1 415-786-6373
mél : gerard.fernando@eng.sun.com

Vivek Goyal
Precept Software, Inc.
1072 Arastradero Rd,
Palo Alto, CA 94304
USA
téléphone : +1 415-845-5200
mél : goyal@precept.com

Don Hoffman
Sun Microsystems, Inc.
Mail-stop UMPK14-305
2550 Garcia Avenue
Mountain View, California 94043-1100
USA
téléphone : +1 503-297-1580
mél : don.hoffman@eng.sun.com

M. Reha Civanlar
AT&T Labs - Research
100 Schutlz Drive, 3-213
Red Bank, NJ 07701-7033
USA
téléphone: +1 732-345-3305
mél : civanlar@research.att.com

Déclaration complète de droits de reproduction

Copyright (C) The Internet Society (1998). Tous droits réservés.

Le présent document et ses traductions peuvent être copiés et fournis aux tiers, et les travaux dérivés qui les commentent ou les expliquent ou aident à leur mise en œuvre peuvent être préparés, copiés, publiés et distribués, en tout ou partie, sans restriction d'aucune sorte, pourvu que la déclaration de droits de reproduction ci-dessus et le présent paragraphe soient inclus dans toutes telles copies et travaux dérivés. Cependant, le présent document lui-même ne peut être modifié d'aucune façon, en particulier en retirant la notice de droits de reproduction ou les références à la Internet Society ou aux autres organisations Internet, excepté autant qu'il est nécessaire pour le besoin du développement des normes Internet, auquel cas les procédures de droits de reproduction définies dans les procédures des normes Internet doivent être suivies, ou pour les besoins de la traduction dans d'autres langues que l'anglais.

Les permissions limitées accordées ci-dessus sont perpétuelles et ne seront pas révoquées par la Internet Society ou ses successeurs ou ayant droits.

Le présent document et les informations y contenues sont fournies 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 ci encloses ne violent aucun droit ou aucune garantie implicite de commercialisation ou d'aptitude à un objet particulier.

Remerciement

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