

Groupe de travail Réseau
Request for Comments : 4425
 Catégorie : Sur la voie de la normalisation

A. Klemets, Microsoft
 février 2006
 Traduction Claude Brière de L'Isle

Format de charge utile RTP pour codec vidéo 1 (VC-1)

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 mémoire spécifie un format de charge utile RTP pour encapsuler des flux binaires compressés de codec vidéo (VC-1, *Video Codec 1*) comme défini par la norme de la Société des ingénieurs d'images animées et de télévision (SMPTE) SMPTE 421M. SMPTE est le principal organisme de normalisation de l'industrie des images animées, et la norme SMPTE 421M définit un format de flux binaire vidéo compressé et un processus de décodage pour la télévision.

Table des matières

1. Introduction.....	2
1.1 Conventions utilisées dans le présent document.....	2
2. Définitions et abréviations.....	2
3. Vue d'ensemble de VC-1.....	3
3.1 Modèle de mise en couches de flux binaire VC-1.....	4
3.2 Unités de données de flux binaire dans le profil évolué.....	4
3.3 Paramètres d'initialisation du décodeur.....	4
3.4 Ordre des trames.....	5
4. Encapsulation de flux binaires au format VC-1 dans RTP.....	5
4.1 Unités d'accès.....	5
4.2 Fragmentation des trames VC-1.....	6
4.3 Considérations d'horodatage.....	7
4.4 Points d'accès aléatoire.....	8
4.5 Suppression des paramètres de HRD.....	8
4.6 Répétition de l'en-tête de couche de séquence.....	8
4.7 Signalisation des paramètres de type de support.....	9
4.8 Paramètres de type de support de "mode=1".....	9
4.9 Paramètres de type de support de "mode=3".....	9
5. Syntaxe du format de charge utile RTP.....	9
5.1 Usage de l'en-tête RTP.....	9
5.2 Syntaxe de l'en-tête AU.....	10
5.3 Syntaxe du champ de contrôle AU.....	11
6. Paramètres du format de charge utile RTP.....	11
6.1 Enregistrement de type de support.....	11
6.2 Transposition des paramètres de type de support en SDP.....	15
6.3 Usage avec modèle SDP d'offre/réponse.....	15
6.4 Usage dans les descriptions de session déclaratives.....	16
7. Considérations sur la sécurité.....	16
8. Contrôle d'encombrement.....	17
9. Considérations relatives à l'IANA.....	17
10. Références.....	18
10.1 Références normatives.....	18
10.2 Références pour information.....	18
Remerciements.....	18
Adresse de l'auteur.....	19
Déclaration complète de droits de reproduction.....	19

1. Introduction

Le présent mémoire spécifie un format de charge utile RTP pour le codec vidéo standard de codage de vidéo, aussi appelé VC-1. La spécification du format de flux binaire VC-1 et le processus de décodage est publié par la Société des ingénieurs d'images animées et de télévision (SMPTE, *Society of Motion Picture and Television Engineers*) dans [SMPTE421M].

VC-1 a une large applicabilité, car il convient pour les applications de flux directs Internet à bas débit jusqu'aux applications de diffusion de télévision haute définition (HDTV, *High Definition Television*) et les applications de cinéma numérique avec un codage presque sans pertes. Les performances globales de VC-1 sont telles que des économies de débit binaire de plus de 50 % sont rapportées [WMV9] comparées avec MPEG-2. Voir les détails dans [WMV9] sur la façon de comparer VC-1 avec les autres codecs, comme MPEG-4 et H.264/AVC. (Dans [WMV9], VC-1 est appelé par son ancien nom, VC-9.)

VC-1 est largement utilisé pour télécharger et passer en direct des films sur l'Internet, sous la forme de Windows Media Video 9 [WMV9], parce que le codec WMV-9 est conforme à la norme VC-1. VC-1 a aussi été récemment adopté comme format de compression obligatoire pour les formats de DVD haute définition HD DVD et Blu-ray.

SMPTE 421M définit la syntaxe du flux binaire VC-1 et spécifie les contraintes qui doivent être satisfaites par les flux binaires conformes à VC-1. SMPTE 421M spécifie aussi le processus complet exigé pour décoder le flux binaire. Cependant, il ne spécifie pas l'algorithme de compression VC-1, permettant ainsi différentes façons de mettre en œuvre un codeur VC-1.

La syntaxe du flux binaire VC-1 a trois profils. Chaque profil a des éléments de syntaxe spécifiques du flux binaire et des algorithmes qui lui sont associés. Selon l'application dans laquelle VC-1 est utilisé, certains profils peuvent mieux convenir que d'autres. Par exemple, le profil simple est conçu pour les flux directs à bas débit de l'Internet et pour l'exécution sur des appareils qui peuvent seulement traiter un décodage de faible complexité. Le profil évolué est conçu pour les applications de diffusion, comme la télévision numérique, HD DVD, ou HDTV. Le profil évolué est le seul profil VC-1 qui prend en charge les trames vidéos entrelacées et les pixels non carrés.

La Section 2 définit les abréviations utilisées dans ce document. La Section 3 donne une vue d'ensemble plus détaillée de VC-1. Les Sections 4 et 5 définissent le format de charge utile RTP pour VC-1, et la section 6 définit le type de support et les paramètres SDP pour VC-1. Voir à la Section 7 les considérations de sécurité, et à la Section 8 les exigences pour le contrôle d'encombrement.

1.1 Conventions utilisées dans le présent document

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] et indiquent les niveaux d'exigence pour les mises en œuvre RTP conformes..

2. Définitions et abréviations

Le présent document utilise les définitions de [SMPTE421M]. Pour des raisons pratiques, les termes suivants de SMPTE 421M sont redonnés ici :

Accès aléatoire : un point d'accès aléatoire dans le flux binaire est défini par la garantie suivante : si le décodage commence à ce point, toutes les trames nécessaires pour l'affichage après ce point n'auront pas de dépendance de décodage à des données précédant ce point, et elles sont aussi présentes dans la séquence de décodage après ce point. Un point d'accès aléatoire est aussi appelé un point d'entrée.

BDU encapsulée (EBDU, *Encapsulated BDU*) : BDU qui a été encapsulée en utilisant le mécanisme d'encapsulation décrit à l'Annexe E de [SMPTE421M], pour empêcher l'émulation du préfixe de code de début dans le flux binaire.

Codes de début (SC, *start code*) : codes uniques de 32 bits qui sont incorporés dans le flux binaire codé et qui identifient le début d'une BDU. Les codes de début consistent en trois octets de préfixe de code de début (SCP, *Start Code Prefix*) unique, et un suffixe de code de début (SCS, *Start Code Suffix*) d'un octet.

Entrelacement : propriété de trames où les lignes alternées de la trame représentent différentes instances dans le temps. Dans une trame entrelacée, un des champs est destiné à être affiché en premier.

Image : pour la vidéo progressive, une image est identique à une trame, tandis qu'en vidéo entrelacée, une image peut se référer à une trame, ou au champ supérieur ou inférieur de la trame selon le contexte.

Image B : image qui est codée en utilisant la prédiction de mouvement compensé à partir des champs ou trames de référence passées et/ou futures. Une image B ne peut pas être utilisée pour prédire une autre image.

Image BI : une image B qui est codée en utilisant les informations provenant seulement d'elle-même. Une image BI ne peut pas être utilisée pour prédire une autre image.

Image I : image codée en utilisant des information provenant seulement d'elle-même.

Image P : image qui est codée en utilisant la prédiction de mouvement compensé provenant de champs ou trames de référence passés.

Niveau : ensemble défini de contraintes sur les valeurs qui peuvent être prises par les paramètres (comme le débit binaire et la taille de mémoire tampon) au sein d'un profil particulier. Un profil peut contenir un ou plusieurs niveaux.

Point d'entrée : point dans le flux binaire qui offre un accès aléatoire.

Profil : sous ensemble défini de la syntaxe de VC-1 avec un ensemble spécifique d'outils de codage, algorithmes, et syntaxe qui lui sont associés. Il y a trois profils VC-1 : simple, principal, et évolué.

Progressive : propriété des trames où tous les échantillons de la trame représentent la même instance de temps.

Séquence : représentation codée d'une série d'une ou plusieurs images. Dans le profil évolué VC-1, une séquence consiste en une série d'un ou plusieurs segments de point d'entrée, où chaque segment de point d'entrée consiste en une série d'une ou plusieurs images, et où la première image dans chaque segment de point d'entrée fournit un accès aléatoire. Dans les profils VC-1 simple et principal, la première image dans chaque séquence est une image I.

Trame : une trame contient des lignes d'informations spatiales d'un signal vidéo. Pour une vidéo progressive, ces lignes contiennent des échantillons commençant à un instant et continuant à travers des lignes successives jusqu'au fond de la trame. Pour la vidéo entrelacée, une trame consiste en deux champs, un champ supérieur et un champ inférieur. Un de ces champs va commencer une période de champ plus tard que l'autre.

Tranche : série consécutive de rangées de macroblocs dans une image, qui sont codées comme une seule unité.

Unité de données de flux binaire (BDU, *Bit-stream data unit*) : unité des données compressées qui peut être analysée (c'est-à-dire, décodée selon la syntaxe) indépendamment des autres informations au même niveau hiérarchique. Une BDU peut être, par exemple, un en-tête de couche de séquence, un en-tête de point d'entrée, une trame, ou une tranche.

3. Vue d'ensemble de VC-1

La syntaxe du flux binaire VC-1 consiste en trois profils : simple, principal, et évolué. Le profil simple est destiné aux applications à faible débit binaire et à faible complexité, comme l'exécution du support sur une tablette. Le débit binaire maximum supporté par le profil simple est 384 kbit/s. Le profil principal vise les applications à haut débit, comme les flux directs et la TV sur IP. Le profil principal prend en charge les images B, qui fournissent une efficacité de compression améliorée au prix d'une complexité accrue.

Certaines caractéristiques qui peuvent être utilisées pour réaliser une bonne efficacité de compression, comme des pixels non carrés et la prise en charge d'images entrelacées, sont seulement incluses dans le profil évolué. Le débit binaire maximum supporté par le profil évolué est 135 Mbit/s, le rendant convenable pour le codage presque sans pertes des signaux de HDTV.

Seul le profil évolué prend en charge le transport de données d'utilisateur (méta-données) dans la bande avec le flux binaire compressé. Les données d'utilisateur peuvent être utilisées par exemple pour la prise en charge du sous titrage.

Des trois profils, seul le profil évolué permet que la configuration des paramètres du codec, comme le ratio d'aspect d'image, soit changée au moyen de la signalisation dans la bande dans le flux binaire compressé.

Pour chacun des profils, un certain nombre de "niveaux" ont été définis. À la différence d'un "profil", qui implique un certain ensemble de caractéristiques ou éléments de syntaxe, un "niveau" est un ensemble de contraintes sur les valeurs des paramètres d'un profil, comme le débit binaire ou la taille de la mémoire tampon. Le profil simple de VC-1 a deux niveaux, le profil principal en a trois, et le profil évolué cinq. Voir à l'Annexe D de [SMPTE421M] la liste détaillée des profils et niveaux.

3.1 Modèle de mise en couches de flux binaire VC-1

Le flux binaire VC-1 est défini comme une hiérarchie de couches. Ceci est conceptuellement similaire à la notion d'une pile de protocoles des protocoles de réseautage. La couche la plus externe est appelée la couche de séquence. Les autres couches sont point d'entrée, image, tranche, macrobloc, et bloc.

Dans les profils simple et principal, une séquence dans la couche de séquence consiste en une série d'une ou plusieurs images codées. Dans le profil évolué, une séquence consiste en un ou plusieurs segments de point d'entrée, où chaque segment de point d'entrée consiste en une série d'une ou plusieurs images, et où la première image dans chaque segment de point d'entrée fournit un accès aléatoire. Une image est décomposée en macroblocs. Une tranche comprend une ou plusieurs rangées contiguës de macroblocs.

Les couches point d'entrée et tranche sont seulement présentes dans le profil évolué. Dans le profil évolué, le début de chaque segment de couche de point d'entrée indique un point d'accès aléatoire. Dans les profils simple et principal, chaque image I est un point d'accès aléatoire.

Chaque image peut être codée comme une image I, image P, image sautée, image BI, ou image B. Ces termes sont définis à la Section 2 de ce document et au paragraphe 4.12 de [SMPTE421M].

3.2 Unités de données de flux binaire dans le profil évolué

Dans le profil évolué, chaque image et tranche est considérée comme une unité de données de flux binaire (BDU, *Bit-stream Data Unit*). Une BDU est toujours alignée sur l'octet et est définie comme une unité qui peut être analysée (c'est-à-dire, décodée selon la syntaxe) indépendamment des autres informations dans la même couche.

Le début d'une BDU est signalé par un identifiant appelé code de début (SC, *Start Code*). Les en-têtes de couche séquence et les en-têtes de point d'entrée sont aussi des BDU et peuvent donc être facilement identifiés par leur code de début. Voir à l'Annexe E de [SMPTE421M] la liste complète des codes de début. Les blocs et macroblocs ne sont pas des BDU et n'ont donc pas de code de début et ne sont pas nécessairement alignés sur l'octet.

Le code de début consiste en quatre octets. Les trois premiers octets sont 0x00, 0x00 et 0x01. Le quatrième octet est appelé le suffixe de code de début (SCS, *Start Code Suffix*) et il est utilisé pour indiquer le type de BDU qui suit le code de début. Par exemple, le SCS d'un en-tête de couche séquence (0x0F) est différent du SCS sur un en-tête de point d'entrée (0x0E). Le code de début est toujours aligné sur l'octet et est transmis dans l'ordre des octets du réseau.

Pour prévenir l'émulation accidentelle du code de début dans le flux binaire codé, SMPTE 421M définit un mécanisme d'encapsulation qui utilise le bourrage d'octets. Une BDU qui a été encapsulée par ce mécanisme est appelée une BDU encapsulée (EBDU, *Encapsulated BDU*).

3.3 Paramètres d'initialisation du décodeur

Dans le profil évolué de VC-1, l'en-tête de couche séquence contient les paramètres qui sont nécessaires pour initialiser le décodeur VC-1.

Les paramètres s'appliquent à tous les segments de point d'entrée jusqu'à la prochaine occurrence d'un en-tête de couche séquence dans le flux binaire codé.

Les paramètres dans l'en-tête de couche séquence incluent le niveau de profil évolué, les dimensions maximum des trames codées, le ratio d'aspect, les informations d'entrelacement, le débit de trame et jusqu'à 31 réglages de paramètre de dispositif

4. Encapsulation de flux binaires au format VC-1 dans RTP

4.1 Unités d'accès

Chaque paquet RTP contient un nombre entier d'unités de données d'application (ADU, *Application Data Unit*). Pour le flux binaire de format VC-1, une ADU est équivalente à une unité d'accès (AU, *Access Unit*). Une AU est définie comme l'en-tête d'AU (défini au paragraphe 5.2) suivi par une charge utile de longueur variable, avec les règles et contraintes décrites aux paragraphes 4.1 et 4.2. La Figure 2 montre la disposition d'un paquet RTP avec plusieurs AU.

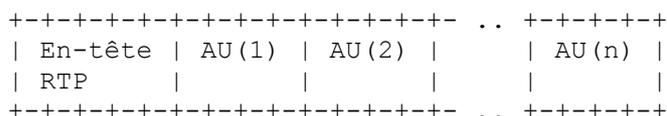


Figure 2. Structure de paquet RTP

Chaque AU DOIT commencer par l'en-tête AU défini au paragraphe 5.2. La charge utile AU DOIT contenir les données appartenant à exactement une trame VC-1. Cela signifie que les données provenant de trames VC-1 différentes vont toujours être dans des AU différentes. Cependant, il est possible qu'une seule trame VC-1 soit fragmentée sur plusieurs AU (voir le paragraphe 4.2).

Dans le cas de vidéo entrelacée, une trame VC-1 consiste en deux champs qui peuvent être codés comme des images séparées. Les deux images appartiennent quand même à la même trame VC-1.

Les règles suivantes s'appliquent au contenu de chaque charge utile AU quand le profil VC-1 évolué est utilisé:

- La charge utile AU DOIT contenir les données de flux binaire VC-1 en format EBDU (c'est-à-dire, le flux binaire doit utiliser le mode d'encapsulation avec bourrage d'octet défini à l'Annexe E de [SMPTE421M].)
- La charge utile AU PEUT contenir plusieurs EBDU, par exemple, un en-tête de couche séquence, un en-tête de point d'entrée, un en-tête de trame (image), un en-tête de champ, et plusieurs tranches et les données d'utilisateur associées. Cependant, toutes les tranches et leurs macroblocs correspondants DOIVENT appartenir à la même trame vidéo.
- La charge utile AU DOIT commencer à une limite de EBDU, sauf quand la charge utile AU contient une trame fragmentée, auquel cas les règles du paragraphe 4.2 s'appliquent.

Quand les profils VC-1 simple ou principal sont utilisés, la charge utile AU DOIT commencer au début d'une trame, sauf quand la charge utile AU contient une trame fragmentée. Le paragraphe 4.2 décrit comment traiter les trames fragmentées.

Les AU DOIVENT être alignées sur l'octet. Si les données dans une AU (les EBDU dans le cas du profil évolué et la trame dans le cas du profil simple et du profil principal) ne se terminent pas sur une limite d'octet, jusqu'à sept bits de bourrage de valeur zéro DOIVENT être ajoutés pour réaliser l'alignement sur l'octet.

4.2 Fragmentation des trames VC-1

Chaque charge utile AU DEVRAIT contenir une trame VC-1 complète. Cependant, si cela causerait le dépassement par le paquet RTP de la taille de MTU, la trame DEVRAIT être fragmentée en plusieurs AU pour éviter la fragmentation au niveau IP. Quand une AU contient une trame fragmentée, cela DOIT être indiqué en réglant le champ FRAG dans l'en-tête AU comme défini au paragraphe 5.3.

Les charges utiles AU qui ne contiennent pas de trame fragmentée ou qui contiennent le premier fragment d'une trame DOIVENT commencer à une limite d'une EBDU si le profil évolué est utilisé. Dans ce cas, pour les profils simple et principal, la charge utile AU DOIT commencer au début d'une trame.

Si le profil évolué est utilisé, les charge utiles AU qui contiennent un fragment d'une trame autre que le premier fragment DEVRAIT commencer au début d'une limite de EBDU, comme le début d'une tranche.

Cependant, les tranches sont seulement définies pour le profil évolué, et ne sont pas toujours utilisées. Les blocs et macroblocs ne sont pas des BDU (ils n'ont pas de code de début) et ne sont pas alignés sur l'octet. Donc, il peut n'être pas toujours possible de continuer une trame fragmentée à une frontière d'EBDU. On peut déterminer si une charge utile AU

commence à une limite d'EBDU en inspectant les trois premiers octets de la charge utile AU. La charge utile AU commence à une limite de EBDU si les trois premiers octets sont identiques au préfixe de code de début (c'est-à-dire, 0x00, 0x00, 0x01).

Dans le cas des profils simple et principal, comme les blocs et macroblocs ne sont pas alignés sur l'octet, la limite de fragmentation peut être choisie arbitrairement.

Si un paquet RTP contient une AU avec le dernier fragment d'une trame, des AU supplémentaires NE DEVRAIENT PAS être incluses dans le paquet RTP.

Si le champ PTS Delta est présent dans l'en-tête AU, chaque fragment d'une trame DOIT avoir le même instant de présentation. Si le champ DTS Delta est présent dans l'en-tête AU, chaque fragment d'une trame DOIT avoir le même instant de décodage.

4.3 Considérations d'horodatage

Les trames vidéo VC-1 DOIVENT être transmises dans l'ordre codé. Un ordre codé implique qu'aucune trame ne dépende des trames suivantes, comme expliqué au paragraphe 3.4. Quand une trame vidéo consiste en une seule image, l'instant de présentation de la trame est identique à l'instant de présentation de l'image. Quand le mode de codage VC-1 entrelacé est utilisé, les trames peuvent contenir deux images, une pour chaque champ. Dans ce cas, l'instant de présentation d'une trame est l'instant de présentation du champ qui est affiché en premier.

Le champ Horodatage RTP DOIT être réglé à l'instant de présentation de la trame vidéo contenue dans la première AU dans le paquet RTP. L'instant de présentation peut être utilisé comme champ d'horodatage dans l'en-tête RTP parce qu'il diffère de l'instant d'échantillonnage de la trame seulement d'un décalage arbitraire constant.

Si la trame vidéo dans une AU a un instant de présentation qui diffère du champ Horodatage RTP, l'instant de présentation DOIT alors être spécifié en utilisant le champ PTS Delta dans l'en-tête AU. Comme le champ Horodatage RTP doit être identique à l'instant de présentation de la première trame vidéo, cela ne peut arriver que si un paquet RTP contient plusieurs AU. La syntaxe du champ PTS Delta est définie au paragraphe 5.2.

L'instant de décodage d'une trame VC-1 est toujours à accroissement monotone quand les trames vidéo sont transmises dans l'ordre codé. Si des images ni B ni BI ne sont présentes dans le flux binaire codé, alors l'instant de décodage d'une trame DEVRA être égal à l'instant de présentation de la trame. Une image BI est un cas particulier d'image B, et dans le reste de ce paragraphe les termes d'image B et de trame B s'appliquent aussi respectivement aux images et trames BI.

Si des images B peuvent être présentes dans le flux binaire codé, les instants de décodage des trames sont alors déterminés comme suit :

- trames B : l'instant de décodage DEVRA être égal à l'instant de présentation de la trame B.
- première trame non B dans l'ordre codé : l'instant de décodage DEVRA être au moins d'une période de trame de moins que l'instant de décodage de la prochaine trame dans l'ordre codé. Une période de trame est définie comme l'inverse du débit de trame utilisé dans le flux binaire codé (par exemple, 100 millisecondes si le débit de trame est de 10 trames par seconde). Pour un flux binaire avec un débit de trame variable, le débit de trame maximum DEVRA déterminer la période de trame. Si le maximum de trame n'est pas spécifié, le débit de trame maximum permis par le profil et niveau DEVRA être utilisé.
- trames non B (autres que la première trame dans l'ordre codé) : l'instant de décodage DEVRA être égal à l'instant de présentation de la précédente trame non B dans l'ordre codé.

Par exemple, considérons la Figure 1 du paragraphe 3.4. Pour déterminer l'instant de décodage de la première trame, I0, on doit d'abord déterminer l'instant de décodage de la prochaine trame, P1. Parce que P1 est une trame non B, son instant de décodage est égal à l'instant de présentation de I0, qui est 3 unités de temps. Donc, l'instant de décodage de I0 doit être au moins une période de trame de moins que 3. Dans cet exemple, la période de trame est 1, parce que une trame est affichée à chaque unité de temps. Par conséquent, l'instant de décodage de I0 est choisi comme deux unités de temps. L'instant de décodage de la troisième trame dans l'ordre codé, P4, est 4, parce que il doit être égal à l'instant de présentation de la précédente trame non B dans l'ordre codé, P1. Par ailleurs, l'instant de décodage de la trame B2 est de 5 unités de temps, qui est identique à son instant de présentation.

Si l'instant de décodage d'une trame vidéo diffère de son instant de présentation, l'instant de décodage DOIT alors être spécifié en utilisant le champ DTS Delta dans l'en-tête AU. La syntaxe du champ DTS Delta est définie au paragraphe 5.2.

Il n'est pas exigé des receveurs qu'ils utilisent le champ DTS Delta. Cependant, les utilisations possibles incluent la gestion de mémoire tampon et la régulation des trames avant le décodage. Si des paquets RTP sont perdus, il est possible d'utiliser le champ DTS Delta pour déterminer si la séquence des paquets RTP perdus contenait des trames de référence ou seulement des trames B. Cela peut être fait en comparant l'instant de décodage et celui de présentation de la première trame reçue après la séquence perdue à l'instant de présentation de la dernière trame de référence reçue avant la séquence perdue.

Savoir si le flux va contenir des images B peut aider le receveur à allouer plus efficacement des ressources et peut réduire le délai, car l'absence d'images B dans le flux implique qu'aucune réorganisation des trames ne va être nécessaire entre le processus de décodage et l'affichage des trames décodées. Cela peut être important pour des applications interactives.

Le receveur DEVRA supposer que le flux binaire codé peut contenir des images B dans les cas suivants :

- profil évolué : si la valeur du paramètre de type de support "bpic" défini au paragraphe 6.1 est 1, ou si le paramètre "bpic" n'est pas spécifié.
- profil principal : si le champ MAXBFRAMES dans le paramètre d'initialisation du décodeur STRUCT_C a une valeur non zéro. STRUCT_C est porté dans le paramètre de type de support "config", qui est défini au paragraphe 6.1.

Le profil simple n'utilise pas d'image B

4.4 Points d'accès aléatoire

L'en-tête de point d'entrée contient des informations qui sont nécessaires au décodeur pour décoder les trames dans ce segment de point d'entrée. Cela signifie que en cas de paquets RTP perdus, le décodeur peut être inutilisable pour décoder les trames jusqu'à la réception du prochain en-tête de point d'entrée.

La première trame après un en-tête de point d'entrée est un point d'accès aléatoire dans le flux binaire codé. Les profils simple et principal n'ont pas d'en-tête de point d'entrée, donc pour ces profils, chaque image I est un point d'accès aléatoire.

Pour permettre au receveur RTP de détecter qu'un paquet RTP qui a été perdu contenait un point d'accès aléatoire, ce format de charge utile RTP définit un champ appelé "compte RA". Ce champ est présent dans chaque AU, et sa valeur est incrémentée (modulo 256) pour chaque point d'accès aléatoire. Pour plus de détails, voir la définition de "compte RA" au paragraphe 5.2.

Pour faciliter la détermination de si une AU contient un point d'accès aléatoire, ce format de charge utile RTP définit aussi un bit appelé le fanion "RA" dans le champ de contrôle AU. Ce bit est réglé à 1 seulement sur les AU qui contiennent un point d'accès aléatoire. Le bit RA est défini au paragraphe 5.3.

4.5 Suppression des paramètres de HRD

L'en-tête de couche séquence du profil évolué peut inclure jusqu'à 31 réglages de paramètre de dispositif à fuite pour le décodeur de référence hypothétique (HRD, *Hypothetical Reference Decoder*). Chaque réglage de paramètre de dispositif à fuite spécifie un possible débit binaire de transmission de crête (HRD_RATE) et une capacité de mémoire tampon de décodeur (HRD_BUFFER). Voir au paragraphe 3.3 un exposé supplémentaire sur le HRD.

Si le débit de transmission de crête réel est connu de l'envoyeur RTP, il PEUT supprimer les réglages de paramètre de dispositif à fuite sauf celui correspondant au débit de transmission de crête réel.

Pour chaque réglage de paramètre de dispositif à fuite dans l'en-tête de couche séquence, il y a aussi un paramètre dans l'en-tête de point d'entrée qui spécifie le remplissage initial (HRD_FULL) du dispositif à fuite.

Si l'envoyeur RTP a supprimé tous les réglages de paramètre de dispositif à fuite de l'en-tête de couche séquence, alors pour tout réglage de paramètre de dispositif à fuite supprimé, il DOIT aussi supprimer le paramètre correspondant HRD_FULL dans l'en-tête de point d'entrée.

Supprimer les réglages de paramètre de dispositif à fuite, comme décrit ci-dessus, peut significativement réduire la taille des en-têtes de couche séquence et des en-têtes de point d'entrée.

4.6 Répétition de l'en-tête de couche de séquence

Pour améliorer la robustesse à la perte de paquets RTP, il est RECOMMANDÉ que si l'en-tête de couche séquence change, il devrait être répété fréquemment dans le flux binaire. Dans ce cas, il est RECOMMANDÉ que le nombre de paramètres Dispositif à fuite dans l'en-tête de couche séquence et les en-têtes de point d'entrée soit réduit à un, comme décrit au paragraphe 4.5. Cela va aider à réduire les frais généraux causés par la répétition de l'en-tête de couche séquence.

Toutes les données dans le flux binaire VC-1, y compris les copies répétées de l'en-tête de séquence lui-même, doivent être comptées lors du calcul du paramètre Dispositif à fuite pour le HRD. Voir au paragraphe 3.3 la discussion sur le HRD.

Si la valeur de TFCNTRFLAG dans l'en-tête de couche séquence est 1, chaque en-tête d'image contient un champ Compteur de trame (TFCNTR). Chaque fois que l'en-tête de couche séquence est inséré dans le flux binaire, la valeur de ce compteur DOIT être réinitialisée.

Pour permettre au receveur RTP de détecter qu'un paquet RTP qui a été perdu contenait un nouvel en-tête de couche séquence, le champ de contrôle AU définit un bit appelé le fanion "SL". Ce bit est établi quand un en-tête de couche séquence est transmis, mais seulement si cet en-tête est différent du plus récent en-tête de couche séquence transmis. Le bit SL est défini au paragraphe 5.3.

4.7 Signalisation des paramètres de type de support

Quand ce format de charge utile RTP est utilisé avec SDP, les paramètres d'initialisation de décodeur décrits au paragraphe 3.3 DOIVENT être signalés dans SDP en utilisant les paramètres de type de support spécifiés au paragraphe 6.1. Le paragraphe 6.2 spécifie comment transposer les paramètres de type de support en SDP [RFC3264], le paragraphe 6.3 définit les règles spécifiques du modèle d'offre/réponse SDP, et le paragraphe 6.4 définit les règles quand SDP est utilisé en style déclaratif.

Quand les profils simple ou principal sont utilisés, il n'est pas possible de changer les paramètres d'initialisation de décodeur à travers le flux binaire codé. Tout changement des paramètres d'initialisation de décodeur va devoir être fait par des moyens hors bande, par exemple, par un re-invite SIP [RFC2543] ou un moyen similaire qui porte une description de session mise à jour.

Quand le profil évolué est utilisé, les paramètres d'initialisation de décodeur PEUVENT être changés en insérant un nouvel en-tête de couche séquence ou en-tête de point d'entrée dans le flux binaire codé.

L'en-tête de couche séquence spécifie le niveau VC-1, la taille maximum des trames codées et facultativement aussi le débit maximum de trames. Les paramètres de type de support "niveau", "largeur", "hauteur", et "débit de trames" spécifient les limites supérieures pour ces paramètres. Donc, l'en-tête de couche séquence PEUT spécifier des valeurs qui sont inférieures aux valeurs des paramètres de type de support "niveau", "largeur", "hauteur", ou "débit de trames", mais l'en-tête de couche séquence NE DOIT excéder les valeurs d'aucun de ces paramètres de type de support.

4.8 Paramètres de type de support de "mode=1"

Dans certaines applications qui utilisent le profil évolué, l'en-tête de couche séquence ne change jamais. Cela PEUT être signalé avec le paramètre de type de support "mode=1". (Le paramètre "mode" est défini au paragraphe 6.1.) Le paramètre "mode=1" sert "d'indication" au receveur RTP que tous les en-têtes de couche séquence dans le flux binaire vont être identiques. Si "mode=1" est signalé et si un en-tête de couche séquence est présent dans le flux binaire codé, il DOIT alors être identique à l'en-tête de couche séquence spécifié par le paramètre de type de support "config".

Comme l'en-tête de couche séquence ne change jamais en "mode=1", l'envoyeur RTP PEUT le supprimer du flux binaire. Noter cependant, que si la valeur de TFCNTRFLAG dans l'en-tête de couche séquence est 1, chaque en-tête d'image contient un champ Compteur de trames (TFCNTR). Ce champ est réinitialisé chaque fois que l'en-tête de couche séquence se produit dans le flux binaire. Si l'envoyeur RTP choisit de supprimer l'en-tête de couche séquence, il DOIT alors s'assurer que le flux binaire résultant est toujours conforme à la spécification VC-1 (par exemple, en ajustant le champ TFCNTR, si nécessaire.)

4.9 Paramètres de type de support de "mode=3"

Dans certaines applications qui utilisent le profil évolué, les deux en-têtes de couche séquence et de point d'entrée ne

changent jamais. Cela PEUT être signalé avec le paramètre de type de support "mode=3". Les mêmes règles s'appliquent au "mode=3" et au "mode=1", décrit au paragraphe 4.8. De plus, si "mode=3" est signalé, l'envoyeur RTP PEUT alors "compresser" le flux binaire codé en n'incluant pas d'en-tête de couche séquence ni d'en-tête de point d'entrée dans les paquets RTP.

Le receveur RTP DOIT "décompresser" le flux binaire codé en ré-insérant les en-têtes de point d'entrée avant de livrer le flux binaire codé au décodeur VC-1. L'en-tête de couche séquence n'a pas besoin d'être décompressé par le receveur, car il ne change jamais.

Si "mode=3" est signalé et si le receveur RTP reçoit une AU complète ou le premier fragment d'une AU, et si le bit RA est réglé à 1 mais si l'AU ne commence pas par un en-tête de point d'entrée, cela indique alors que l'en-tête de point d'entrée a été "comprimé". Dans ce cas, le receveur RTP DOIT insérer un en-tête de point d'entrée au début de l'AU. Quand il insère l'en-tête de point d'entrée, le receveur RTP DOIT utiliser celui qui a été spécifié par le paramètre de type de support "config".

5. Syntaxe du format de charge utile RTP

5.1 Usage de l'en-tête RTP

Le format de l'en-tête RTP est spécifié dans la [RFC3550] et est reproduit dans la Figure 3 pour l'agrément du lecteur.

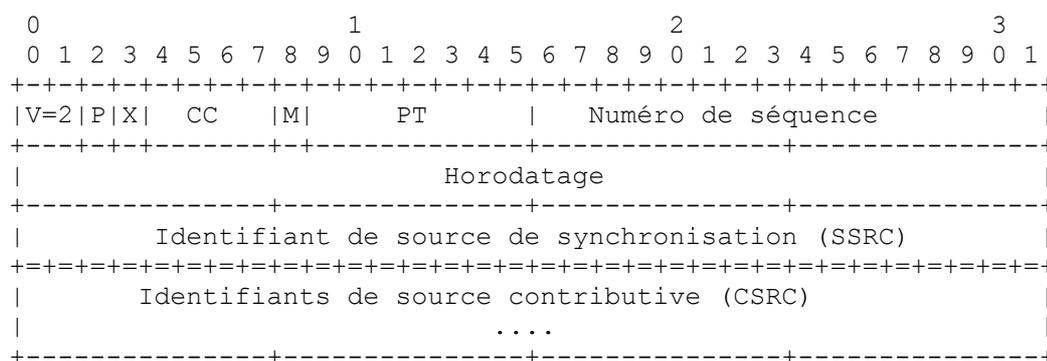


Figure 3. En-tête RTP selon la RFC 3550

Les champs de l'en-tête fixe RTP ont leur signification usuelle, définie dans la RFC 3550 et par le profil RTP utilisé, avec les notes supplémentaires suivantes :

Bit Marqueur (M) : 1 bit. Ce bit est réglé à 1 si le paquet RTP contient une AU contenant une trame VC-1 complète ou le dernier fragment d'une trame VC-1.

Type de charge utile (PT, *Payload type*) : 7 bits. Le présent document n'alloue pas de type de charge utile RTP pour ce format de charge utile RTP. L'allocation d'un type de charge utile doit être effectué soit par le profil RTP utilisé, soit de façon dynamique.

Numéro de séquence : 16 bits. Le receveur RTP peut utiliser le champ Numéro de séquence pour récupérer l'ordre codé des trames VC-1. Un décodeur VC-1 normal va exiger que les trames VC-1 soient livrées dans l'ordre codé. Quand des trames VC-1 ont été fragmentées entre des paquets RTP, le receveur RTP peut utiliser le champ Numéro de séquence pour d'assurer qu'aucun fragment ne manque.

Horodatage : 32 bits. L'horodatage RTP est réglé à l'instant de présentation de la trame VC-1 dans la première AU. Un débit d'horloge de 90 kHz DOIT être utilisé.

5.2 Syntaxe de l'en-tête AU

L'en-tête AU consiste en un champ de contrôle AU d'un octet, le champ Compte de RA, et trois champs facultatifs. Tous les champs DOIVENT être écrits dans l'ordre des octets du réseau. La structure de l'en-tête AU est illustrée à la Figure 4.

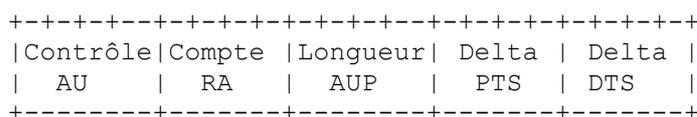


Figure 4. Structure de l'en-tête AU

Contrôle AU : 8 bits. L'usage du champ de contrôle AU est défini au paragraphe 5.3.

Compte RA : 8 bits. Compteur de point d'accès aléatoire. Ce champ est un compteur binaire modulo 256. La valeur de ce champ DOIT être incrémentée de 1 chaque fois qu'un AU est transmis lorsque le bit RA dans le champ de contrôle AU est réglé à 1. La valeur initiale de ce champ est indéfinie et PEUT être choisie au hasard.

Longueur AUP : 16 bits. Longueur de la charge utile de l'unité d'accès. Spécifie la taille, en octets, de la charge utile de l'unité d'accès. Le champ n'inclut pas la taille de l'en-tête AU lui-même. Le champ DOIT être inclus dans chaque en-tête AU dans un paquet RTP, sauf pour le dernier en-tête AU dans le paquet. Si ce champ n'est pas inclus, la charge utile de l'unité d'accès DEVRA être supposée s'étendre jusqu'à la fin de la charge utile RTP.

Delta PTS : 32 bits. Différence d'heure de présentation. Spécifie l'heure de présentation de la trame comme un décalage (delta) de complément à 2 par rapport au champ Horodatage dans l'en-tête RTP de ce paquet RTP. Le champ Delta PTS DOIT utiliser le même débit d'horloge que le champ Horodatage dans l'en-tête RTP.

Ce champ NE DEVRAIT PAS être inclus dans le premier en-tête AU du paquet RTP, parce que le champ Horodatage RTP spécifie l'heure de présentation de la trame dans la première AU. Si ce champ n'est pas inclus, l'heure de présentation de la trame DEVRA être supposé être spécifiée par le champ Horodatage dans l'en-tête RTP.

Delta DTS : 32 bits. Différence d'heure de décodage. Spécifie l'instant de décodage de la trame comme un décalage (delta) de complément à deux entre l'heure de présentation et l'instant de décodage. Noter que si l'heure de présentation est supérieure à l'instant de décodage, il en résultera une valeur pour ce champ Delta DTS qui sera supérieure à zéro. Le champ Delta DTS DOIT utiliser le même débit d'horloge que le champ Horodatage dans l'en-tête RTP. Si ce champ n'est pas inclus, l'instant de décodage de la trame DEVRA être supposé identique à l'instant de présentation de la trame.

5.3 Syntaxe du champ de contrôle AU

La structure du champ de contrôle AU de 8 bits est montré à la Figure 5.

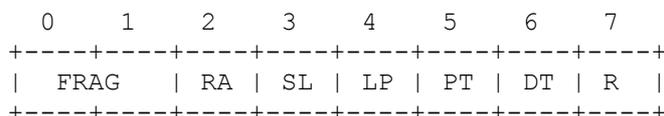


Figure 5. Syntaxe du champ de contrôle AU.

FRAG : 2 bits. Informations de fragmentation. Ce champ indique si la charge utile AU contient une trame complète ou un fragment d'une trame. Il DOIT être réglé comme suit :

- 0 : la charge utile AU contient un fragment d'une trame autre que le premier ou le dernier fragment.
- 1 : la charge utile AU contient le premier fragment d'une trame.
- 2 : la charge utile AU contient le dernier fragment d'une trame.
- 3 : la charge utile AU contient une trame complète (non fragmentée).

RA : 1 bit. Indicateur de point d'accès aléatoire. Ce bit DOIT être réglé à 1 si l'AU contient une trame qui est un point d'accès aléatoire. Dans le cas des profils simple et principal, toute image I est un point d'accès aléatoire.

Dans le cas du profil évolué, la première trame après un en-tête de point d'entrée est un point d'accès aléatoire.

Si les en-têtes de point d'entrée ne sont pas transmis à chaque point d'accès aléatoire, cela DOIT être indiqué en utilisant le paramètre de type de support "mode=3".

SL : 1 bit. Compteur de couche séquence. Ce bit DOIT être permuté, c'est-à-dire, changé de 0 à 1 ou de 1 à 0, si l'AU contient un en-tête de couche séquence et si il est différent du plus récent en-tête de couche séquence transmis. Autrement, la valeur de ce bit doit être identique à la valeur du bit SL dans l'AU précédente.

La valeur initiale de ce bit est indéfinie et PEUT être choisie au hasard.

Le bit DOIT être 0 pour les flux binaires des profils simple et principal ou si l'en-tête de couche séquence ne change jamais.

LP : 1 bit. Longueur présente. Ce bit DOIT être réglé à 1 si l'en-tête AU inclut le champ Longueur AUP.

PT : 1 bit. Delta PTS présent. Ce bit DOIT être réglé à 1 si l'en-tête AU inclut le champ Delta PTS.

DT : 1 bit. Delta DTS présent. Ce bit DOIT être réglé à 1 si l'en-tête AU inclut le champ Delta DTS.

R : 1 bit. Réserve. Ce bit DOIT être réglé à 0 et DOIT être ignoré par les receveurs.

6. Paramètres du format de charge utile RTP

6.1 Enregistrement de type de support

Cet enregistrement utilise le gabarit défini dans la [RFC4288] et suit la [RFC3555].

Nom de type : video

Nom de sous type : vc1

Paramètres exigés :

profil : la valeur est un entier identifiant le profil VC-1. Les valeurs suivantes sont définies :

0 : profil simple

1 : profil principal

3 : profil évolué

Si le paramètre profil est utilisé pour indiquer les propriétés d'un flux binaire codé, il indique le profil VC-1 qu'un décodeur doit prendre en charge quand il décode le flux binaire.

Si le paramètre profil est utilisé pour échanger les capacités ou dans une procédure d'établissement de session, il indique le profil VC-1 que le codec prend en charge.

niveau : la valeur est un entier qui spécifie le niveau du profil VC-1.

Pour le profil évolué, les valeurs valides sont de 0 à 4, qui correspondent respectivement aux niveaux L0 à L4. Pour les profils simple et principal, les valeurs suivantes sont définies :

1 : niveau bas

2 : niveau moyen

3 : niveau haut (seulement valide pour le profil principal)

Si le paramètre Niveau est utilisé pour indiquer les propriétés d'un flux binaire codé, il indique le plus haut niveau du profil VC-1 qu'un décodeur doit prendre en charge quand il décode le flux binaire. Noter que la prise en charge d'un niveau implique celle de tous les niveaux numériquement inférieurs de ce profil.

Si le paramètre Niveau est utilisé pour l'échange de capacités ou dans une procédure d'établissement de session, il indique le plus haut niveau du profil VC-1 que le codec prend en charge. Voir au paragraphe 6.3 de la RFC 4425 les règles spécifiques de la façon dont ce paramètre est utilisé avec le modèle d'offre/réponse SDP.

Paramètres facultatifs :

config : la valeur est une représentation en base16 [RFC3548] (hexadécimal) d'une chaîne d'octets qui exprime les paramètres d'initialisation du décodeur. Les paramètres d'initialisation du décodeur sont transposés en la chaîne d'octets base16 sur la base du MSB en premier. Le premier bit des paramètres d'initialisation de décodeur DOIT être situé au MSB du premier octet. Si les paramètres d'initialisation de décodeur ne sont pas un multiple de 8 bits, jusqu'à 7 bits de bourrage de valeur zéro DOIVENT être ajoutés dans le dernier octet pour réaliser l'alignement sur l'octet.

Pour les profils simple et principal, les paramètres d'initialisation du décodeur sont STRUCT_C, comme défini à l'Annexe J de [SMPTE421M].

Pour le profil évolué, les paramètres d'initialisation du décodeur sont un en-tête de couche séquence directement suivi par un en-tête de point d'entrée. Les deux en-têtes DOIVENT être en format EBDU, ce qui signifie qu'ils doivent inclure leur code de début et doivent utiliser la méthode d'encapsulation définie à l'Annexe E de [SMPTE421M].

width (*largeur*) : la valeur est un entier supérieur à zéro, qui spécifie la taille horizontale maximum des trames codées, en échantillons luma (pixels dans l'image luma).

Pour les profils simple et principal, la valeur DEVRA être identique à la taille horizontale réelle des trames codées.

Pour le profil évolué, la valeur DEVRA être supérieure ou égale à la plus grande taille horizontale des trames codées. Si ce paramètre n'est pas spécifié, il prend par défaut la taille horizontale maximum permise par le profil et niveau spécifiés.

height (hauteur) : la valeur est un entier supérieur à zéro, qui spécifie la taille maximum verticale des trames codées, en échantillons luma (pixels dans une image luma codée progressivement).

Pour les profils simple et principal, la valeur DEVRA être identique à la taille verticale réelle des trames codées.

Pour le profil évolué, la valeur DEVRA être supérieure ou égale à la plus grande taille verticale des trames codées.

Si ce paramètre n'est pas spécifié, il prend par défaut la taille verticale maximum permise par le profil et niveau spécifiés.

bitrate (débit binaire) : la valeur est un entier supérieur à zéro, qui spécifie le débit de transmission de crête du flux binaire codé en bits par seconde. Le nombre n'inclut pas les frais généraux causés par l'encapsulation RTP, c'est-à-dire, il n'inclut pas les en-têtes AU, ni aucun en-tête RTP, UDP, ou IP.

Si ce paramètre n'est pas spécifié, il prend par défaut le débit binaire maximum permis par le profil et niveau spécifiés.

Voir les valeurs pour "RMax" à l'Annexe D de [SMPTE421M].

buffer (mémoire tampon) : la valeur est un entier qui spécifie la taille du dispositif à fuite, B, en millisecondes, exigé pour contenir un flux transmis au débit de transmission spécifié par le paramètre bitrate. Ce paramètre est défini dans le modèle de décodeur hypothétique de référence pour VC-1, à l'Annexe C de [SMPTE421M].

Noter que ce paramètre se rapporte seulement au flux binaire du codec, et ne tient pas compte du temps de mise en mémoire tampon qui peut être requis pour compenser la gigue dans le réseau.

Si ce paramètre n'est pas spécifié, il prend par défaut la taille maximum de mémoire tampon permise par le profil et niveau spécifiés. Voir les valeurs pour "BMax" et "RMax" à l'Annexe D de [SMPTE421M].

framerate (débit de trames) : la valeur est un entier supérieur à zéro, qui spécifie le nombre maximum de trames par seconde dans le flux binaire codé, multiplié par 1000 et arrondi à la plus proche valeur d'entier. Par exemple, 30000/1001 (approximativement 29,97) trames par seconde est représenté comme 29 970.

Ce paramètre peut être utilisé pour contrôler l'allocation de ressources chez le receveur. Par exemple, un receveur peut choisir d'effectuer un post-traitement supplémentaire sur les trames décodées seulement si le débit de trames est supposé être faible. Le paramètre NE DOIT PAS être utilisé pour réguler le processus de rendu, car le débit de trames réel peut différer de la valeur spécifiée.

Si le paramètre n'est pas spécifié, il prend par défaut le débit de trames maximum permis par le profil et niveau spécifiés.

bpic (image B) : ce paramètre signale que des images B et BI peuvent être présentes quand le profil évolué est utilisé. Si ce paramètre est présent, et si des images B ou BI peuvent être présentes dans le flux binaire codé, ce paramètre DOIT être égal à 1.

Une valeur de 0 indique que des images B et BI NE DEVRONT PAS être présentes dans le flux binaire codé, même si l'en-tête de couche séquence change. L'inclusion de ce paramètre avec une valeur de 0 est RECOMMANDÉE, si des images B ni BI ne sont pas incluses dans le flux binaire codé.

Ce paramètre NE DOIT PAS être utilisé avec les profils simple et principal. Pour le profil principal, la présence d'images B et BI est indiquée par le champ MAXBFRAMES dans le paramètre STRUCT_C d'initialisation du décodeur.

Pour le profil évolué, si ce paramètre n'est pas spécifié, une valeur de 1 DEVRA être supposée.

mode : la valeur est un entier qui spécifie l'utilisation de l'en-tête de couche séquence et de l'en-tête de point d'entrée. Ce paramètre est seulement défini pour le profil évolué. Les valeurs suivantes sont définies :

0 : l'en-tête de couche séquence et l'en-tête de point d'entrée peuvent tous deux changer, et les en-têtes changés vont être inclus dans les paquets RTP.

1 : l'en-tête de couche séquence spécifié dans le paramètre config ne change jamais. Les règles du paragraphe 4.8 de la RFC 4425 DOIVENT être suivies.

3 : l'en-tête de couche séquence et l'en-tête de point d'entrée spécifiés dans le paramètre config ne changent jamais. Les règles du paragraphe 4.9 de la RFC 4425 DOIVENT être suivies.

Si le paramètre mode n'est pas spécifié, une valeur de 0 DEVRA être supposée. Le paramètre mode DEVRAIT être spécifié si les modes 1 ou 3 s'appliquent au flux binaire VC-1.

max-width, max-height, max-bitrate, max-buffer, max-framerate : ces paramètres sont définis comme étant utilisés dans la procédure d'échange de capacité. Ces paramètres ne signalent pas des propriétés du flux binaire codé, mais plutôt des limites supérieures ou des valeurs préférées pour les paramètres "width", "height", "bitrate", "buffer", et "framerate". Le

paragraphe 6.3 de la RFC 4425 donne des règles spécifiques sur la façon dont ces paramètres sont utilisés avec le modèle d'offre/réponse SDP.

Les receveurs qui signalent la prise en charge d'un certain profil et niveau DOIVENT accepter les valeurs maximum pour ces paramètres pour ce profil et niveau. Par exemple, un receveur qui indique la prise en charge du profil principal, niveau bas, doit accepter une largeur de 352 échantillons luma et une hauteur de 288 échantillons luma, même si cela exige d'adapter l'image pour qu'elle tienne la résolution d'un appareil de plus petit affichage.

Un receveur PEUT utiliser tout paramètre max-width, max-height, max-bitrate, max-buffer, et max-framerate pour indiquer les capacités préférées. Par exemple, un receveur peut choisir de spécifier des valeurs pour max-width et max-height qui correspondent à la résolution de l'affichage de son appareil, car un flux binaire codé en utilisant ces paramètres n'aura pas besoin d'être réadapté.

Si un des paramètres max-width, max-height, max-bitrate, max-buffer, et max-framerate signale une capacité qui est inférieure aux capacités requises du profil et niveau signalés, alors le paramètre DEVRA être interprété comme valeur préférée pour cette capacité.

Chacun de ces paramètres PEUT aussi être utilisé pour signaler des capacités qui excèdent les capacités requises du profil et niveau signalés. Dans ce cas, le paramètre DEVRA être interprété comme la valeur maximum qui peut être acceptée pour cette capacité.

Quand plus d'un paramètre de l'ensemble (max-width, max-height, max-bitrate, max-buffer, et max-framerate) est présent, toutes les capacités signalées DOIVENT être prises en charge simultanément.

Un envoyeur ou receveur NE DOIT PAS utiliser ces paramètres pour signaler des capacités qui satisfont les exigences d'un niveau supérieur du profil VC-1 à celui spécifié dans le paramètre "niveau", même si l'envoyeur ou receveur peut prendre en charge toutes les propriétés du niveau supérieur, sauf si spécifier un niveau supérieur n'est pas permis à cause d'autres restrictions. Un exemple d'une telle restriction, dans le modèle d'offre/réponse SDP, est que la valeur du paramètre Niveau qui peut être utilisée dans une réponse est limitée par ce qui était spécifié dans l'offre.

max-width (largeur maximum) : la valeur est un entier supérieur à zéro, qui spécifie une taille horizontale pour les trames codées, en échantillons luma (pixels dans l'image luma). Si la valeur est inférieure à la taille horizontale maximum permise par le profil et niveau, alors la valeur spécifie la taille horizontale préférée. Autrement, elle spécifie la taille horizontale maximum qui est prise en charge.

Si ce paramètre n'est pas spécifié, il prend par défaut la taille horizontale maximum permise par le profil et niveau spécifiés.

max-height (hauteur maximum) : la valeur est un entier supérieur à zéro, qui spécifie une taille verticale pour les trames codées, en échantillons luma (pixels dans une image luma codée progressivement). Si la valeur est inférieure à la taille maximum verticale permise par le profil et niveau, alors la valeur spécifie la taille verticale préférée. Autrement, elle spécifie la taille maximum verticale qui est prise en charge.

Si ce paramètre n'est pas spécifié, il prend par défaut la taille maximum verticale permise par le profil et niveau spécifiés.

max-bitrate : la valeur est un entier supérieur à zéro, qui spécifie un débit de transmission de crête pour le flux binaire codé en bits par seconde. Le nombre n'inclut pas les frais généraux causés par l'encapsulation RTP, c'est-à-dire, il n'inclut pas d'en-têtes AU, ni aucun en-tête RTP, UDP, ou IP.

Si la valeur est inférieure au débit binaire maximum permis par le profil et niveau, alors la valeur spécifie le débit binaire préféré. Autrement, elle spécifie le débit binaire maximum qui est accepté.

Si ce paramètre n'est pas spécifié, il prend par défaut le débit binaire maximum permis par le profil et niveau spécifiés. Voir les valeurs pour "RMax" à l'Annexe D de [SMPTE421M].

max-buffer : la valeur est un entier qui spécifie une taille de dispositif à fuite, B, en millisecondes, exigée pour contenir un flux transmis au débit de transmission spécifié par le paramètre max-bitrate. Ce paramètre est défini dans le modèle de décodeur de référence hypothétique pour VC-1, à l'Annexe C de [SMPTE421M].

Noter que ce paramètre se rapporte seulement au flux binaire du codec et ne tient pas compte du temps de mise en mémoire tampon qui peut être requis pour compenser la gigue dans le réseau.

Si la valeur est inférieure à la taille maximum de dispositif à fuite permise par le paramètre max-bitrate et le profil et niveau, alors la valeur spécifie la taille préférée de dispositif à fuite. Autrement, elle spécifie le maximum de taille de dispositif à fuite qui est supportée pour le débit binaire spécifié par le paramètre max-bitrate.

Si ce paramètre n'est pas spécifié, il prend par défaut la taille maximum de mémoire tampon permise par le profil et

niveau spécifiés. Voir les valeurs pour "BMax" et "RMax" à l'Annexe D de [SMPTE421M].

max-framerate : la valeur est un entier supérieur à zéro, qui spécifie un nombre de trames par seconde pour le flux binaire codé. La valeur est le débit de trames multiplié par 1000 et arrondi à la plus proche valeur d'entier. Par exemple, 30000/1001 (approximativement 29,97) trames par seconde est représenté comme 29 970.

Si la valeur est inférieure au débit maximum de trames permis par le profil et niveau, alors la valeur spécifie le débit de trames préféré. Autrement, elle spécifie le débit maximum de trames qui est supporté.

Si le paramètre n'est pas spécifié, il prend par défaut le débit maximum de trames permis par le profil et niveau spécifiés.

Considérations de codage : ce type de support est tramé et contient des données binaires.

Considérations de sécurité : voir la Section 7 de la RFC 4425.

Considérations d'interopérabilité : aucune.

Spécification publiée : RFC 4425.

Applications qui utilisent ce type de support : flux directs multimédia et outils de conférence.

Informations supplémentaires : aucune.

Personne & adresse de messagerie à contacter pour plus d'informations : Anders Klemets <anderskl@microsoft.com> ;
groupe de travail IETF AVT.

Usage prévu : COMMUN

Restrictions d'usage : ce type de support dépend du tramage RTP ; donc, il est seulement défini pour le transfert via RTP [RFC3550].

Auteur : Anders Klemets

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

6.2 Transposition des paramètres de type de support en SDP

Les informations portées dans la spécification de type de support ont une transposition spécifique en les champs du protocole de description de session (SDP) [RFC2327]. Si SDP est utilisé pour spécifier les sessions en utilisant ce format de charge utile, la transposition est comme suit :

- o Le nom du support dans la ligne "m=" de SDP DOIT être video (le nom du type).
- o Le nom du codage dans la ligne "a=rtpmap" de SDP DOIT être vc1 (le nom du sous type).
- o Le débit d'horloge dans la ligne "a=rtpmap" DOIT être 90000.
- o Les paramètres EXIGÉS "profil" et "niveau" DOIVENT être inclus dans la ligne "a=fmtp" de SDP.
Ces paramètres sont exprimés sous la forme d'une liste de paires de paramètre=valeur séparées par des points-virgules.
- o Les paramètres FACULTATIFS "config", "width", "height", "bitrate", "buffer", "framerate", "bpic", "mode", "max-width", "max-height", "max-bitrate", "max-buffer", et "max-framerate", quand ils sont présents, DOIVENT être inclus dans la ligne "a=fmtp" de SDP.
Ces paramètres sont exprimés sous la forme d'une liste de paires de paramètre=valeur séparées par des points-virgules.

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 DOIT être ignoré. Par exemple, les paramètres définis dans des spécifications ultérieures PEUVENT être copiés dans le SDP et DOIVENT être ignorés par les receveurs qui ne les comprennent pas.

6.3 Usage avec le modèle SDP d'offre/réponse

Quand VC-1 est offert sur RTP en utilisant SDP dans un modèle d'offre/réponse [RFC3264] pour la négociation en usage d'envoi individuel, les règles et limitations suivantes s'appliquent :

- o Le paramètre "profil" DOIT être utilisé de façon symétrique, c'est-à-dire que celui qui répond DOIT soit conserver le paramètre, soit supprimer complètement le format de support (type de charge utile) si le profil VC-1 offert n'est pas pris en charge.
- o Le paramètre "niveau" spécifie le plus haut niveau du profil VC-1 supporté par le codec. Celui qui répond NE DOIT PAS spécifier un niveau numériquement supérieur dans la réponse à celui spécifié dans l'offre. Celui qui répond PEUT spécifier un niveau inférieur à celui spécifié dans l'offre, c'est-à-dire que le paramètre Niveau peut être "dégradé".
Si l'offre spécifie l'attribut de direction sendrecv ou sendonly et si la réponse dégrade le paramètre Niveau, cela peut exiger qu'une nouvelle offre spécifie un paramètre "config" mis à jour. Si le paramètre "config" ne peut pas être utilisé avec le niveau spécifié dans la réponse, l'offreur DOIT alors initier un autre tour d'offre/réponse ou ne pas utiliser ce format de support (type de charge utile).
- o Les paramètres "config", "bpic", "width", "height", "framerate", "bitrate", "buffer", et "mode", décrivent les propriétés du flux binaire VC-1 que l'offreur ou celui qui répond envoie pour cette configuration de format de support. Dans le cas d'usage en envoi individuel, et quand l'attribut de direction dans l'offre ou la réponse est recvonly, l'interprétation de ces paramètres est indéfinie et ils NE DOIVENT PAS être utilisés.
- o Les paramètres "config", "width", "height", "bitrate", et "buffer" DOIVENT être spécifiés quand l'attribut de direction est sendrecv ou sendonly.
- o Les paramètres "max-width", "max-height", "max-framerate", "max-bitrate", et "max-buffer" PEUVENT être spécifiés dans une offre ou une réponse, et leur interprétation est comme suit :
Quand l'attribut de direction est sendonly, les paramètres décrivent les limites du flux binaire VC-1 que l'expéditeur est capable de produire pour le profil et niveau donnés, et pour tout niveau inférieur du même profil.
Quand l'attribut de direction est recvonly ou sendrecv, les paramètres décrivent les propriétés de la mise en œuvre de receveur. Si la valeur d'une propriété est inférieure à celle permise par le niveau du profil VC-1, elle DEVRA alors être interprétée comme la valeur préférée et le flux binaire VC-1 de l'expéditeur NE DEVRAIT PAS l'excéder. Si la valeur d'une propriété est supérieure à celle qui est permise par le niveau du profil VC-1, elle DEVRA alors être interprétée comme la limite supérieure de la valeur que le receveur accepte pour le profil et niveau donnés, et pour tout niveau inférieur du même profil.
Par exemple, si une offre recvonly ou sendrecv spécifie "profil=0;niveau=1;max-bitrate=48000", alors 48 kbit/s est simplement un débit binaire suggéré, parce que toutes les mises en œuvre de receveur de profil simple, niveau bas, sont obligées de prendre en charge les débits binaires jusqu'à 96 kbit/s. En supposant que l'offre est acceptée, la réponse devrait spécifier "bitrate=48000", mais toute valeur jusqu'à 96000 est permise. Mais si l'offre spécifie "max-bitrate=200000", cela signifie que la mise en œuvre de receveur prend en charge un maximum de 200 kbit/s pour le profil et niveau donnés (ou un niveau inférieur). Dans ce cas, il est permis à celui qui répond d'envoyer un paramètre bitrate de jusqu'à 200 000.
- o Si un offreur souhaite avoir des capacités non symétriques entre l'envoi et la réception, par exemple, utiliser des niveaux différents dans chaque direction, l'offreur doit alors offrir des sessions RTP différentes. Cela peut être fait en spécifiant des lignes de support différentes déclarées respectivement comme "recvonly" et "sendonly".

Pour des flux livrés en diffusion groupée, les règles suivantes s'appliquent en plus :

- o Le paramètre "niveau" spécifie le plus haut niveau du profil VC-1 utilisé par les participants à la session de diffusion groupée. La valeur de ce paramètre NE DOIT PAS être changée par celui qui répond. Donc, un type de charge utile peut être soit accepté inchangé, soit supprimé.
- o Les paramètres "config", "bpic", "width", "height", "framerate", "bitrate", "buffer", et "mode", spécifient les propriétés du flux binaire VC-1 qui vont être envoyées et/ou reçues dans la session de diffusion groupée. Les paramètres PEUVENT être spécifiés, même si l'attribut de direction est recvonly.
Les valeurs de ces paramètres NE DOIVENT PAS être changées par celui qui répond. Donc, un type de charge utile peut être soit accepté inchangé, soit supprimé.
- o Les valeurs des paramètres "max-width", "max-height", "max-framerate", "max-bitrate", et "max-buffer" DOIVENT

être acceptées par le receveur pour tous les flux déclarés comme sendrecv ou recvonly. Autrement, une des actions suivantes DOIT être effectuée : le format de support est supprimé, ou la session est rejetée.

6.4 Usage dans les descriptions de session déclaratives

Quand VC-1 est offert sur RTP en utilisant SDP dans le style déclaratif, comme dans RTSP [RFC2326] ou SAP [RFC2974], les règles et limitations suivantes s'appliquent :

- o Les paramètres "profil" et "niveau" indiquent seulement les propriétés du flux binaire codé. Ils n'impliquent pas de limite aux capacités acceptées par l'envoyeur.
- o Les paramètres "config", "width", "height", "bitrate", et "buffer" DOIVENT être spécifiés.
- o Les paramètres "max-width", "max-height", "max-framerate", "max-bitrate", et "max-buffer" NE DOIVENT PAS être utilisés.

Un exemple de représentation de support dans SDP est comme suit (profil simple, niveau moyen) :

```
m=video 49170 RTP/AVP 98
a=rtpmap:98 vc1/90000
a=fmtp:98 profile=0;level=2;width=352;height=288;framerate=15000;
bitrate=384000;buffer=2000;config=4e291800
```

7. Considérations sur la sécurité

Les paquets RTP qui utilisent le format de charge utile défini dans cette spécification sont soumis aux considérations de sécurité discutées dans la spécification RTP [RFC2327], et dans tout profil RTP approprié. Cela implique que la confidentialité du flux de supports est réalisée par le chiffrement ; par exemple, par l'application de SRTP [RFC3711].

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 du côté du receveur. L'attaquant peut injecter dans le flux des paquets RTP pathologiques qui sont complexes à décoder et qui causent la surcharge du receveur. VC-1 est particulièrement vulnérable à de telles attaques, parce que il est possible à un attaquant de générer des paquets RTP contenant des trames qui affectent le processus de décodage de nombreuses futures trames. Donc, l'usage de l'authentification de l'origine des données et de la protection de l'intégrité des données d'au moins le paquet RTP est RECOMMANDÉE ; par exemple, avec SRTP [RFC3711].

Noter que le mécanisme approprié pour assurer la confidentialité et l'intégrité des paquets RTP et de leurs charges utiles dépend de l'application et des protocoles de transport et de signalisation employés. Donc, bien que SRTP soit donné en exemple ci-dessus, d'autres choix possibles existent.

Le flux binaire VC-1 peut porter des données d'utilisateur, comme des informations de sous-titrage et des méta-données de contenu. La spécification VC-1 ne définit pas comment interpréter les données d'utilisateur. Les identifiants pour les données d'utilisateur doivent être enregistrés avec SMPTE. Il est concevable que les types de données d'utilisateur soient définies comme incluant un contenu de programmation, comme des scripts ou commandes qui vont être exécutés par le receveur. Selon le type de données d'utilisateur, il serait possible à un envoyeur de générer des données d'utilisateur d'une manière non conforme pour faire échouer le receveur ou le rendre temporairement indisponible. Les envoyeurs qui transportent des flux binaires VC-1 DEVRAIENT s'assurer que les données d'utilisateur sont conformes à la spécification enregistrée par SMPTE (voir l'Annexe F de [SMPTE421M]). Les receveurs DEVRAIENT empêcher les dysfonctionnements en cas de données d'utilisateur non conformes.

Il est important de noter que le flux VC-1 peut avoir des exigences de bande passante très élevées (jusqu'à 135 Mbit/s pour la vidéo haute définition). Cela cause un potentiel de déni de service si la transmission est faite sur de nombreux chemins Internet. Donc, les utilisateurs de ce format de charge utile DOIVENT se conformer aux exigences du contrôle d'encombrement décrites à la Section 8.

8. Contrôle d'encombrement

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

Si un service au mieux est utilisé, les utilisateurs de ce format de charge utile DOIVENT surveiller la perte de paquet pour s'assurer que la perte de paquet est dans des limites acceptables. La perte de paquets est considérée acceptable si un flux TCP à travers le même chemin du réseau et subissant les mêmes conditions de réseau réaliserait un débit moyen, mesuré sur une échelle de temps raisonnable, qui n'est pas inférieur à ce que le flux RTP réalise. Cette condition peut être satisfaite en mettant en œuvre des mécanismes de contrôle d'encombrement pour adapter le débit de transmission ou en s'arrangeant pour qu'un receveur quitte la session si le taux de perte est d'un niveau inacceptable.

L'adaptation du débit binaire nécessaire pour respecter le principe du contrôle d'encombrement est facilement réalisable quand le codage en temps réel est utilisé. Quand un contenu pré-codé est transmis, l'adaptation de la bande passante exige une ou plusieurs des conditions suivantes :

- La disponibilité de plus d'une représentation codée du même contenu à des débits binaires différents. La commutation entre les différentes représentations peut normalement être effectuée dans la même session RTP en changeant de flux aux limites de point d'accès aléatoire.
- L'existence de trames de non référence (par exemple, des trames B) dans le flux binaire. Des trames de non référence peuvent être éliminées par l'émetteur avant l'encapsulation dans RTP.

C'est seulement quand il est exigé que des paramètres non dégradables (comme le paramètre VC-1 "profile") soient changés qu'il devient nécessaire de terminer et redémarrer le flux de supports. Cela peut être accompli en utilisant un type de charge utile RTP différent.

Sans considération de la méthode utilisée pour l'adaptation de la bande passante, le flux binaire résultant DOIT être conforme à la spécification VC-1 [SMPT421M]. Par exemple, si des trames non de référence sont éliminées, alors l'élément de syntaxe FRMCNT (trames de profils simple et principal seulement) et l'élément de syntaxe facultatif TFCNTR (trames de profil évolué seulement) doivent être incrémentées comme si aucune trame n'avait été éliminée. Parce que l'élément de syntaxe TFCNTR compte les trames dans l'ordre d'affichage, qui est différent de l'ordre de transmission (l'ordre codé) cela exige que l'émetteur "regarde devant" ou mette en mémoire tampon un certain nombre de trames.

Daas un autre exemple, quand on commute entre différentes représentations du même contenu, il peut être nécessaire de signaler une discontinuité en modifiant le champ FRMCNT, ou si le profil évolué est utilisé, en établissant le fanion BROKEN_LINK dans l'en-tête de point d'entrée à 1.

Ce format de charge utile peut aussi être utilisé dans des réseaux qui fournissent des garanties de qualité de service. Si un service amélioré est utilisé, les receveurs DEVRAIENT surveiller la perte de paquets pour s'assurer que le service demandé est bien fourni. Si il ne l'est pas, ils DEVRAIENT alors supposer qu'ils reçoivent un service au mieux et se comporter en conséquence.

9. Considérations relatives à l'IANA

L'IANA a enregistré le type de support "video/vc1" et le format de charge utile RTP associé dans le registre Types de support du registre des types MIME de format de charge utile RTP, comme spécifié au paragraphe 6.1.

10. Références

10.1 Références normatives

[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](#))

- [RFC3264] J. Rosenberg et H. Schulzrinne, "[Modèle d'offre/réponse](#) avec le protocole de description de session (SDP)", juin 2002. (*P.S.* ; *MàJ par* [RFC8843](#))
- [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](#), [RFC8860](#))
- [RFC3555] S. Casner et P. Hoschka, "Enregistrement de type MIME pour les types de charge utile RTP", juillet 2003. (*P.S.* ; *Rendue obsolète par les RFC* 4855, 4856)
- [RFC4288] N. Freed et J. Klensin, "Spécifications du [type de support et procédures d'enregistrement](#)", [BCP 13](#), décembre 2005.
- [SMPTE421M] Society of Motion Picture and Television Engineers, "VC-1 Compressed Video Bitstream Format and Decoding Process", SMPTE 421M.

10.2 Références pour information

- [H.264/AVC] Ribas-Corbera, J., Chou, P.A., and S.L. Regunathan, "A generalized hypothetical reference decoder for H.264/AVC", IEEE Transactions on Circuits and Systems for Video Technology, août 2003.
- [RFC2326] H. Schulzrinne, A. Rao et R. Lanphier, "Protocole de [flux directs en temps réel](#) (RTSP)", avril 1998. (*Remplacée par* [RFC7826](#))
- [RFC2543] M. Handley, H. Schulzrinne, E. Schooler, J. Rosenberg, "SIP : protocole d'initialisation de session", mars 1999. (*Obsolète, voir* [RFC3261](#), [RFC3262](#), [RFC3263](#), [RFC3264](#), [RFC3265](#)) (*P.S.*)
- [RFC2974] M. Handley, C. Perkins, E. Whelan, "Protocole d'annonce de session (SAP)", octobre 2000. (*Expérimentale*)
- [RFC3551] H. Schulzrinne et S. Casner, "[Profil RTP pour conférences audio](#) et vidéo avec contrôle minimal", STD 65, juillet 2003. (*MàJ par* [RFC8860](#))
- [RFC3711] M. Baugher et autres, "Protocole de [transport sécurisé en temps réel](#) (SRTP)", mars 2004. (*P.S.*)
- [WMV9] Srinivasan, S., Hsu, P., Holcomb, T., Mukerjee, K., Regunathan, S.L., Lin, B., Liang, J., Lee, M., and J. Ribas-Corbera, "Windows Media Video 9: overview and applications", Signal Processing: Image Communication, Volume 19, Issue 9, octobre 2004.

Remerciements

Merci à Regis Crinon, Miska Hannuksela, Colin Perkins, Shankar Regunathan, Gary Sullivan, Stephan Wenger, et Magnus Westerlund qui ont fourni des commentaires détaillés sur ce document.

Adresse de l'auteur

Anders Klemets
Microsoft Corp.
1 Microsoft Way
Redmond, WA 98052
USA
mél : Anders.Klemets@microsoft.com

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, 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.

Remerciement

Le financement de la fonction d'édition des RFC est fourni par l'activité de soutien administratif (IASA) de l'IETF.