

T R - I E T F - R F C 3 9 8 4

H.264 ビデオのための  
RTP ペイロードフォーマット  
に関する技術レポート

Technical Report for  
RTP Payload Format for H.264 Video

第 1 版

2006 年 2 月 28 日制定

社団法人

情報通信技術委員会

THE TELECOMMUNICATION TECHNOLOGY COMMITTEE



本書におけるRFC 3984原英文以外の部分については、(社)情報通信技術委員会が著作権を保有しています。

内容の一部又は全部を(社)情報通信技術委員会の許諾を得ることなく複製、転載、改変、転用及びネットワーク上での送信、配布を行うことを禁止します。

## 著作權事項

Copyright (C) The Internet Society (2005). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

## 目次

<参考>	5
1. はじめに	7
1.1. H.264 コーデック	7
1.2. パラメータセットの概念	8
1.3. ネットワーク抽象化レイヤユニットタイプ	8
2. 規則	9
3. 範囲	9
4. 定義と略語	9
4.1. 定義	9
4.2. 略語	11
5. RTP ペイロードフォーマット	11
5.1 RTP ヘッダの使い方	11
5.2 RTP ペイロードフォーマットの共通構造	13
5.3 NAL ユニットオクテットの使い方	15
5.4 パケット化モード	16
5.5. 復号順序番号(DON)	17
5.6. シングル NAL ユニットパケット	19
5.7. 集合パケット	20
5.7.1. シングルタイム集合パケット	22
5.7.2. マルチタイム集合パケット(MTAP)	25
5.8. 分割ユニット (FU)	29
6. パケット化規則	33
6.1. 共通パケット化規則	33
6.2. シングル NAL ユニットモード	34
6.3. 非インタリーブモード	34
6.4. インタリーブモード	34
7. 逆パケット化処理 (Informative)	34
7.1. シングル NAL ユニットと非インタリーブモード	34
7.2. インタリーブモード	35
7.2.1. 逆インタリーブバッファのサイズ	35
7.2.2. 逆インタリーブ処理	35
7.3. 追加の逆パケット化ガイドライン	36
8. ペイロードフォーマットパラメータ	36
8.1. MIME 登録	37
8.2. SDP パラメータ	47
8.2.1 MIME パラメータの SDP へのマッピング	47
8.2.2 SDP オファー/アンサーモデルの使用	48
8.2.3. 宣言型セッション記述 (Declarative Session Descriptions) における使用方法	52
8.3. 例	53
8.4. パラメータセットの考慮点	56
9. セキュリティの考察	57

10. 輻輳制御.....	58
11. IANA の考慮点.....	59
12. 付録:アプリケーション例.....	59
12.1. ITU-T 勧告 H.241 付属資料 A に従ったテレビ電話.....	59
12.2. テレビ電話、スライス・データパーティションなし、NAL ユニットアグリゲーションなし.....	59
12.3. テレビ電話：NAL ユニット集合を用いたインタリーブパケット化.....	60
12.4. データパーティションを行うテレビ電話.....	60
12.5. FU と前方誤り訂正によるテレビ電話またはストリーミング.....	60
12.6. 低ビットレートのストリーミング.....	62
12.7. ビデオストリーミングにおける強固なパケットスケジューリング.....	63
13. 復号順序番号に対する論理的根拠.....	63
13.1. 序論.....	63
13.2. 複数枚ピクチャのスライスインタリーブの例.....	64
13.3. 強固なパケットスケジューリングの例.....	65
13.4. 冗長符号化スライスの強固な伝送スケジューリング.....	69
13.5. その他の設計可能性に関する備考.....	69
14. 謝辞.....	70
15. References.....	70
15.1. Normative References.....	70
15.2. Informative References.....	70

## <参考>

### 1. 概要

本技術レポートは、IETF より発行されている RFC3984(RTP Payload Format for H.264 Video) [1] に対し作成されたものである。TTC メディア符号化専門委員会マルチメディアシステム SWG では、日本国内における H.264 ビデオ伝送の普及を目的とし、RFC3984[1]に対する日本語訳の検討を行った。その結果を本技術レポートとして制定するものである。

#### (1) 国際標準等との関連

本技術レポートは IETF 標準 RFC3984[1]に基づき日本語訳を行ったものである。

#### (2) 上記国際標準等に対する追加項目等

なし。

#### (3) 上記国際標準等に対する変更事項

なし。

#### (4) 参照した国際標準との章立て構成の相違

なし。(RFC3984[1]の章立てと同様)

#### (5) 改版の履歴

版数	制定日	改版内容
第1版	2006年2月28日	初版制定

#### (6) 工業所有権

なし。(本書の配布は無制限である)

#### (7) その他(参照している標準等)

[1] S. Wenger, M. M. Hannuksela, T. Stockhammer, M. Westerlund, D. Singer, "RTP Payload Format for H.264 Video", RFC3984, February 2005, Internet Engineering Task Force

その他 RFC3984[1] 内で参照している勧告、標準については本文 15.1 節(Normative References)及び 15.2 節(Informative References)にて記述している。

## 2. 対象となる RFC の前文

Network Working Group

Request for Comments: 3984

Category: Standards Track

S. Wenger

M. M. Hannuksela

T. Stockhammer

M. Westerlund

D. Singer

February 2005

### RTP Payload Format for H.264 Video

#### Status of This Memo

This document specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" (STD 1) for the standardization state and status of this protocol. Distribution of this memo is unlimited.

#### Copyright Notice

Copyright (C) The Internet Society (2005).

#### Abstract

This memo describes an RTP Payload format for the ITU-T Recommendation H.264 video codec and the technically identical ISO/IEC International Standard 14496-10 video codec. The RTP payload format allows for packetization of one or more Network Abstraction Layer Units (NALUs), produced by an H.264 video encoder, in each RTP payload. The payload format has wide applicability, as it supports applications from simple low bit-rate conversational usage, to Internet video streaming with interleaved transmission, to high bit-rate video-on-demand.

## 1. はじめに

### 1.1. H.264 コーデック

この文書は、ITU-T 勧告 H.264[1]と ISO/IEC 国際標準 14496 パート 10[2]（両方ともアドバンストビデオ符号化、または AVC としても知られている）として知られているビデオ符号化の RTP ペイロード仕様を示したものである。勧告 H.264 は 2003 年 5 月に ITU-T で標準化され、ドラフト仕様書は公文書[8]として入手できる。（訳注：ITU-T で正式に公開されている）この文書内で、H.264 の表記はコーデックや標準として使われるが、この文書は同様に ISO/IEC の対になる符号化標準にも適用できる。

H.264 ビデオコーデックは非常に幅広い適用範囲を有し、低ビットレートのインターネットストリーミングアプリケーションから HDTV 放送やデジタル映画アプリケーションのようなほとんど損失のない符号化までのデジタル圧縮ビデオ形式のすべてを網羅している。現在の技術と比較して、H.264 の総合的な性能は 50%以上のビットレート削減であるといわれている。例えば、デジタル衛星 TV の品質は 1.5Mbit/s で実現できるといわれており、現在の 3.5Mbit/s[9]あたりの MPEG2 ビデオと同じ品質である

コーデックの仕様書[1]そのものは、ビデオ符号化レイヤ（VCL）とネットワーク抽象レイヤ（NAL）との間で概念的にはっきりと区別している。VCL はコーデックの信号処理機能と変換、量子化、動き予測補償などの機構とループフィルターを含んでいる。今日の大多数のビデオコーデックの一般的なコンセプトは、動き補償フレーム間予測と残差信号の符号化変換を利用したマクロブロックベースコーダである。VCL エンコーダはスライスを出力し、スライスヘッダ（スライス内の最初のマクロブロックの空間アドレス、初期量子化パラメータや同様の情報を含んでいる）は、マクロブロックの整数個のマクロブロックデータとスライスヘッダの情報を含んでいるビット列である。スライス内のマクロブロックは、異なったマクロブロックの割り当てが明示されなければ、いわゆるフレキシブルマクロブロック順序シンタックスを利用することにより、スキャン順序で配置される。イン-ピクチャ予測はスライス内のみで用いられる。更なる情報は[9]を見よ。

ネットワーク抽象レイヤ（NAL）エンコーダは VCL エンコーダのスライス出力をネットワーク抽象レイヤユニット（NAL ユニット）にカプセル化しており、それらはパケットネットワーク上またはパケット指向多重化環境での伝送に適している。H.264 の付録 B では、NAL ユニットのバイトストリーム指向のネットワーク経路で伝送するためのカプセル化の過程が定義されている。但し、付録 B は、本標準の範囲には含まれない。内部で、NAL は NAL ユニットを利用している。NAL ユニットは 1 バイトのヘッダとペイロードバイトストリングで構成されている。ヘッダは NAL ユニットタイプ、（潜在的な）ビットエラーの存在や NAL ユニットペイロード内のシンタックス違反、復号過程における NAL ユニットの相対的な重要性に関する情報を示している。この RTP ペイロード仕様書は、NAL ユニットペイロード内のビットストリングを意識しないで、デザインされている。

H.264 の主な特徴の 1 つは、伝送時間、符号化時間、そしてスライスやピクチャのサンプリングまたはプレゼンテーション時間が完全に切り離されているということである。H.264 で定められている符号化過程は時間のことは気にせず、そして H.264 のシンタックスはスキップしたフレーム数のような情報（初期のビデオ圧縮標準では Temporal Reference の形で一般的な）を伝えない。また、多くのピクチャに影響をあたえる、それゆえに本質的に時間の要素を含まない NAL ユニットも存在する。この理由により、RTP タイムスタンプの取り扱いは、サンプリングまたはプレゼンテーション時間が定義されていない、または伝送時間が不明



であるなどの NAL ユニットに対し、特別の配慮を必要とする。

### 1.2. パラメータセットの概念

H.264 のデザイン概念の非常に基本的なことの 1 つは、RFC2429[10]のヘッダ複製または MPEG-4 のヘッダ拡張符号 (HEC) [11]のような仕組みを不要にして自己完結パケットを生成することである。これは、メディアストリームから、2 つ以上のスライスに関連する情報を切り離すことによって成し遂げられる。このより高次元のメタ情報は、スライスパケットを含んだ RTP パケットストリームとは別に信頼性が高くあらかじめ送信されているべきである。(このためのアウトバンド伝送チャンネルが用意されていないアプリケーションに対してはインバンドで送る仕組みが用意されている。)

より高いレベルのパラメータの組み合わせは、パラメータセットと呼ばれる。H.264 仕様書は、2 種類のパラメータセットを含んでいる。シーケンスパラメータセットとピクチャパラメータセットである。アクティブシーケンスパラメータセットは符号化ビデオシーケンスを通じて不変であり、ピクチャパラメータは符号化ピクチャ内において不変である。シーケンスとピクチャパラメータセット構造はピクチャサイズ、オプションの符号化モードの使用やマクロブロックとスライスグループの対応マップを含んでいる。

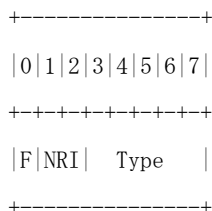
パラメータセットとスライスパケットストリームを同時に更新することなくピクチャパラメータ (ピクチャサイズのような) を更新するために、エンコーダとデコーダは 2 つ以上のシーケンスとピクチャパラメータセットのリストを維持することができる。各スライスヘッダはシーケンスとピクチャパラメータセットが利用されたことを示すコードを含んでいる。

この機構は、パケットストリームからのパラメータセットの伝送切り離し、外部の手段 (例えば、能力交換の副次効果として) や (信頼できるまたは信頼できない) コントロールプロトコルを通じたそれらの伝送を可能とする。さらにアプリケーションデザイン仕様書で固定されていて、パラメータを全く送らないですませることすら可能である。

### 1.3. ネットワーク抽象化レイヤユニットタイプ

NAL 設計のチュートリアル情報は[12]、[13]そして[14]でみることができる。

すべての NAL ユニットは 1 つの NAL ユニットタイプオクテットを構成要素とし、NAL ユニットタイプオクテットは、この RTP ペイロードフォーマットのペイロードヘッダとしても役に立つ。NAL ユニットのペイロードは、その直後に続く。



NAL ユニットタイプオクテットのシンタクスとセマンティクスは[1]に明示されている。しかし NAL ユニットタイプオクテットの必須の特性は以下に要約されている。NAL ユニットタイプオクテットは以下のフォーマットである。

H.264 仕様書に明示されている、NAL ユニットタイプオクテットの要素のセマンティックスは簡単に以下に記述してある。

F:1 ビット

forbidden\_zero\_bit。H.264 仕様書では、1 はシンタックス違反として宣言している。

NRI:2 ビット

nal\_ref\_idc。00 の値は NAL ユニットの内容が、ピクチャ間予測のための参照ピクチャを再構築するには使われないことを示している。そのような NAL ユニットは、リファレンスピクチャの完全性のリスクなしに廃棄することができる。値が 00 以上は、NAL ユニットの復号がリファレンスピクチャの完全性を維持する必要があるということを示している。

Type:5 ビット

nal\_unit\_types。このコンポーネントは、[1]の表 7-1 とこの文書内のあとに定義されている NAL ユニットペイロードタイプを示している。NAL ユニットタイプとセマンティックスの現在のすべての定義については、[1]の 7.4.1 節を参照せよ。

この文書は、5.2 節で新しい NAL ユニットタイプを紹介している。この文書で定義されている NAL ユニットタイプは[1]で明示されていない。その上さらに、この仕様書は、5.3 節で記述されている F と NRI のセマンティックスを拡張している。

## 2. 規則

"MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", "OPTIONAL" のキーワードは、このドキュメントにおいて、RFC2119[3]、BCP14 に記述されているように解釈される。

この仕様書は、ビットフィールドを扱うとき、ビットのセットやクリアの概念を用いる。ビットのセットはビットの値を 1 (ON) に割り当てると同等である。ビットのクリアはビットの値を 0 (OFF) に割り当てると同等である。

## 3. 範囲

このペイロード仕様書は RTP を利用して H.264 の NAL ユニットストリームそのものを運ぶためのみに利用され、H.264 の付録 B で議論されているビットストリームフォーマットを運ぶためではない。それと同様に、この仕様書の最初のアプリケーションは、会話を行うマルチメディア分野、テレビ電話やテレビ会議である。しかしペイロードフォーマットは IP を介したインターネットストリーミングや TV のようなほかのアプリケーションも包含している。

## 4. 定義と略語

### 4.1. 定義

この文書は[1]の定義を利用している。以下の語句は[1]に定義されており、便宜上ここでは概要を述べている。

アクセスユニット：

主符号化ピクチャを常に含む 1 組の NAL ユニット。主符号化ピクチャに加えてアクセスユニットは 1 つ以上の冗長符号化ピクチャまたは符号化ピクチャのスライスやスライスデータパーティションを含んでいない他の NAL ユニットを含んでもよい。アクセスユニットの復号により常に一つの復号ピクチャが生成される。

符号化ビデオシーケンス：

復号順序で、ある IDR のアクセスユニットではじまり、それに続く 0 個以上のすべての非 IDR アクセスユニットからなる。ただし次の後続する IDR アクセスユニットを含まない。

IDR アクセスユニット：

主符号化ピクチャが IDR ピクチャであるアクセスユニット。

IDR ピクチャ：

復号処理において「リセット」を起動する、I、SI スライスタイプのスライスのみを含んだ符号化ピクチャ。IDR ピクチャ復号後、復号順序で後続のすべての符号化ピクチャは IDR ピクチャの前に復号されたどんなピクチャからの INTER 予測なしで復号が可能である。

主符号化ピクチャ：

H. 264 に適合するビットストリームの復号処理で使用される、ピクチャの符号化表現。主符号化ピクチャは、ピクチャのすべてのマクロブロックを含む。

冗長符号化ピクチャ：

ピクチャまたはピクチャの一部の符号化表現。冗長符号化ピクチャの内容は、H. 264 に適合したビットストリームの復号化処理で使用されてはならない。冗長符号化ピクチャの内容は、エラーやロスを含んだビットストリームの復号化処理で利用してもよい。

VCL NAL ユニット：

符号化スライスと符号化データパーティション NAL ユニットの示す集合的な語句。

加えて以下の定義を適用する。

符号化順序番号 (DON)：

ペイロード構造のフィールド、または NAL ユニット符号化順序を示す導出変数。DON の値は 0 以上 65535 以下のすべての値を含んだ範囲にある。最大値に達した後は、DON の値は 0 にもどる。

NAL ユニット復号順序：

[1]の 7.4.1.2 節で与えられる NAL ユニット順序の制約に一致する NAL ユニット順序。

伝送順序：

昇順 RTP シーケンス番号 (モジュロ算術) によるパケットの順序。集合パケットの中で、NAL ユニット伝送順序は、パケット内の NAL ユニットの出現順序と同じである。

メディア・アウェア・ネットワーク要素 (MANE)：

RTP ペイロードヘッダの一部や RTP ペイロードを構文解析してメディアの内容に反応できるミドルボックスやアプリケーションレイヤゲートウェイのようなネットワーク要素。

注：

MANE の概念は、MANE はシグナリングを意識しているという点において (例えば、メディアストリームのペイロード種別マッピングに関して学習するような) または SRTP が機能しているとき信頼できるものでなければならないという点において、一般のルータやゲートウェイの範囲を超える。MANE を利用

する利点は、メディア符号化の必要に関連してパケットが落ちることを許しているということである。例えば、もし MANE があるリンクの混雑のためにパケットを落とさなければならないとき、パケットの欠落がユーザの体験に最小限の影響となるようにし、混雑を削減し、そして／または遅延を最小にするためにそれらを削除するパケットを特定することができる。

## 4.2 略語

DON:	復号順序番号(Decoding Order Number)
DONB:	復号順序番号ベース(Decoding Order Number Base)
DOND:	復号順序番号差分(Decoding Order Number Difference)
FEC:	前方誤り訂正(Forward Error Correction)
FU:	分割ユニット(Fragmentation Unit)
IDR:	瞬時復号更新(Instantaneous Decoding Refresh)
IEC:	国際電気標準会議(International Electrotechnical Commission)
ISO:	国際標準化機構(International Organization for Standardization)
ITU-T:	国際電気通信連合－電気通信標準化部門(International Telecommunication Union, Telecommunication Standardization Sector)
MANE:	メディア・アウェア・ネットワーク要素(Media Aware Network Element)
MTAP:	マルチタイム集合パケット(Multi-Time Aggregation Packet)
MTAP16:	16ビットタイムスタンプオフセット MTAP(MTAP with 16-bit timestamp offset)
MTAP24:	24ビットタイムスタンプオフセット MTAP(MTAP with 24-bit timestamp offset)
NAL:	ネットワーク抽象化レイヤ(Network Abstraction Layer)
NALU:	NAL ユニット(NAL Unit)
SEI:	付加拡張情報(Supplemental Enhancement Information)
STAP:	シングルタイム集合パケット(Single-Time Aggregation Packet)
STAP-A:	STAP A 型(STAP type A)
STAP-B:	STAP B 型(STAP type B)
TS:	タイムスタンプ(Timestamp)
VCL:	ビデオ符号化レイヤ(Video Coding Layer)

## 5. RTP ペイロードフォーマット

### 5.1 RTP ヘッダの使い方

RTP ヘッダのフォーマットは RFC 3550 [4] で定義され、便宜上このフォーマットを図 1 に再記載する。このペイロードフォーマットは、RFC 3550 に従った形式でヘッダのフィールドを使用する。

1 つの NAL ユニットが RTP パケット単位でカプセル化される時、推奨される RTP ペイロードフォーマットは 5.6 節で定義される。集合パケットや分割ユニット用の RTP ペイロード（といくつかの RTP ヘッダビットの設定）は、それぞれ 5.7 節と 5.8 節で定義される。

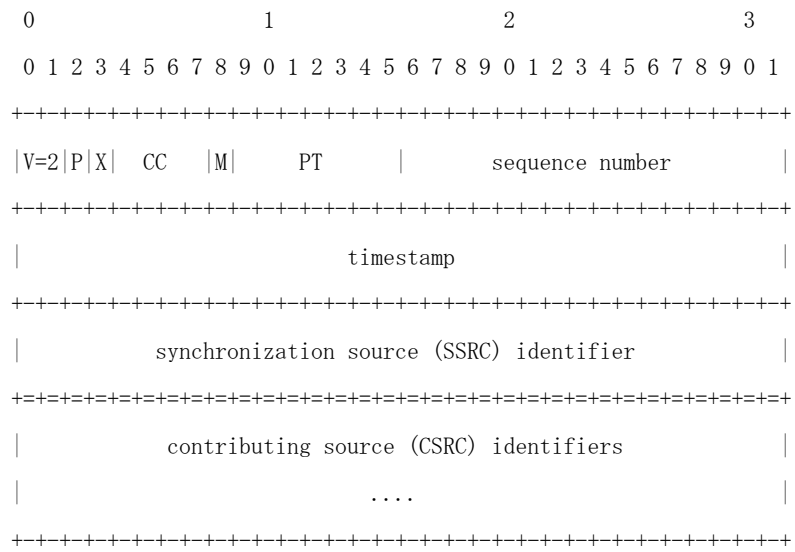


図1 RFC 3550 に従った RTP ヘッダ (RFC 3984)

この RTP ペイロードフォーマットに従って設定される RTP ヘッダ情報は次の通り。

マーカビット (M) : 1 ビット

RTP タイムスタンプによって示されるアクセスユニットの最後の packets について設定すること。これはビデオフォーマットにおける M ビットの標準的な使い方に沿っており、効率的な初期 (プレイアウト) バッファ処理を可能にする。集合 packets (STAP や MTAP) について、RTP ヘッダにあるマーカビットに設定しなければならない値は、RTP packets 自身で転送されたなら、集合 packets の最終 NAL ユニットの最初にあるであろうマーカビットである。デコーダは、アクセスユニットにある最終 packets の最初に表れる目印としてこのビットを使ってもよいものの、この特性に頼ってはいけぬ。

注: M ビットは 1 個だけである。複数の NAL ユニットの集合 packets と関連付けられる。このように、もしゲートウェイが集合 packets をいくつかの packets に再 packets 化した場合、それらの packets の M ビットを確実に設定することはできない。

ペイロードタイプ (PT) : 7 ビット

この新しい packets フォーマットについての RTP ペイロードタイプの割り当ては、この文書の範囲外であり、ここでは定義されない。ペイロードタイプの割り当ては、使われたプロファイルを通してか、動的な方法で行われなければならない。

シーケンス番号 (SN) : 16 ビット

RFC 3550 に従って設定され使用される。シングル NAL ユニットの非インターリーブ packets 化モードについては、シーケンス番号は NAL ユニットの復号順序を決定するために使われる。

タイムスタンプ : 32 ビット

RTP タイムスタンプは、コンテンツにあるサンプリングされたタイムスタンプで設定される。90kHz クロックレートを適用しなければならない。

もし NAL ユニットがそれ自身のタイミング特性を持たないのならば（例えば、パラメータセットや SEI NAL ユニット）、[1]の 7.4.1.2 節に従って、RTP タイムスタンプは NAL ユニットを含むアクセスユニットの主符号化ピクチャの RTP タイムスタンプで設定される。

MTAP についての RTP タイムスタンプの設定は、5.7.2 節で定義される。

受信側は、アクセスユニットに含まれる表示タイムスタンプが 1 個だけの場合、どんなピクチャタイミング SEI メッセージも無視すべきである。その代わりに、受信側は表示処理の同期について RTP タイムスタンプを使用すべきである。

RTP 送信側は、複数フィールドの表示を想定しないピクチャについてピクチャタイミング SEI メッセージを送出すべきでない。

もし 1 つのアクセスユニットが、ピクチャタイミング SEI メッセージで運ばれる 1 つ以上の表示タイムスタンプを有するならば、SEI メッセージにある情報は RTP タイムスタンプに関連するように扱うべきである。最も早いイベントが RTP タイムスタンプで与えられた時間で発生し、SEI メッセージピクチャタイミング値の差異でもたらされた遅延で派生的なイベントが発生する。アクセスユニットの SEI メッセージで運ばれる表示タイムスタンプを、 $tSEI1$ 、 $tSEI2$ 、...、 $tSEIn$  とし、ここで  $tSEI1$  は全タイムスタンプの最も早いものであるとする。SEI メッセージのタイムスケールを 90kHz タイムスケールに調整する関数を  $tmadjst()$  とする。RTP タイムスタンプを  $TS$  とする。すると、 $tSEI1$  で割り当てられたイベントについての表示時間は  $TS$  となる。 $tSEIx$  で割り当てられたイベントについての表示時間は、ここで  $x$  は  $[2...n]$  であるとする、 $TS + tmadjst(tSEIx - tSEI1)$  となる。

注：符号化フレームのフィールド表示は、3:2 プルダウンとして知られている処理において共通に必要である。これは、符号化フレームから成るフィルムコンテンツが、インターレーススキャンを使用したディスプレイへの表示で使用される。ピクチャタイミング SEI メッセージは同じ符号化ピクチャについて複数のタイムスタンプの転送が可能で、それゆえ 3:2 プルダウン処理が完全に制御される。符号化フレームあたり 1 つだけのタイムスタンプが RTP タイムスタンプにおいて運ばれることができるため、ピクチャタイミング SEI メッセージのメカニズムが必要である。

注：H. 264 は復号順序が表示順序と異なることを許容するので、RTP タイムスタンプの値は RTP シーケンス番号の関数として昇順にならなくてもよい。その上、RTCP レポートで報告される到着パケット間ジッタの値が、ネットワークパフォーマンスの信頼できる指標にならない場合もある。なぜなら、到着パケット間ジッタについての演算規則（RFC 3550 の 6.4.1 節）はパケットの RTP タイムスタンプが直接、転送時間に比例することを仮定しているためである。

## 5.2. RTP ペイロードフォーマットの共通構造

ペイロードフォーマットは、3 つの異なった基本ペイロード構造を定義する。受信側は RTP ペイロードの最初のバイトによってペイロード構造を特定することが可能である。それは、RTP ペイロードヘッダとして、ある場合においてはペイロードの最初のバイトとして機能する。このバイトは常に NAL ユニットヘッダとして構造化される。NAL ユニットタイプフィールドはどの構造が存在するかを示す。可能な構造は次の通

り。

シングル NAL ユニットパケット：ペイロードにシングル NAL ユニットだけを含む。NAL ヘッダタイプフィールドはオリジナルの NAL ユニットタイプに等しい。すなわち、1 から 23 までの範囲にある。5.6 節において定義される。

集合パケット：複数の NAL ユニートを 1 つの RTP ペイロードに集約するために使われるパケットタイプ。このパケットは、4 つのバージョンがあり、シングルタイム集合パケットタイプ A (STAP-A)、シングルタイム集合パケットタイプ B (STAP-B)、マルチタイム集合パケット (MTAP) の 16 ビットオフセット版 (MTAP16)、マルチタイム集合パケット (MTAP) の 24 ビットオフセット版 (MTAP24) がある。STAP-A、STAP-B、MTAP16、MTAP24 で割り当てられた NAL ユニット番号はそれぞれ 24、25、26、27 である。5.7 節において定義される。

分割ユニット：シングル NAL ユニットが複数の RTP パケットに渡って断片化するために使用される。2 つのバージョン、FU-A と FU-B が存在し、NAL ユニットタイプ番号としてはそれぞれ 28、29 として示される。5.8 節において定義される。

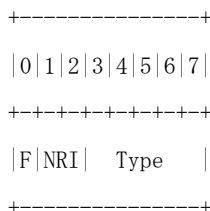
表 1 NAL ユニットタイプとそれらのペイロード構造の概要 (RFC 3984)

種別	パケット	種別名	節
0	未定義		-
1-23	NAL ユニット	H. 264 によるシングル NAL ユニットパケット	5.6 節
24	STAP-A	シングルタイム集合パケット	5.7.1 節
25	STAP-B	シングルタイム集合パケット	5.7.1 節
26	MTAP16	マルチタイム集合パケット	5.7.2 節
27	MTAP24	マルチタイム集合パケット	5.7.2 節
28	FU-A	分割ユニット	5.8 節
29	FU-B	分割ユニット	5.8 節
30-31	未定義		-

注：この定義は、シングル NAL ユニットパケットと分割ユニットにおいてカプセル化される NAL ユニットのサイズを制限しない。どんな集合パケットにおいてもカプセル化される NAL ユニットの最大サイズは 65535 バイトである。

### 5.3. NAL ユニットオクテットの使い方

NAL ユニットオクテットの構造とセマンティクスは 1.3 節で紹介された。便宜上、NAL ユニットタイプオクテットのフォーマットを以下のように再記載する。



本節はこの仕様に従って F と NRI のセマンティクスを定義する。

#### F : 1 ビット

forbidden\_zero\_bit を示す。値ゼロは、NAL ユニットタイプオクテットとペイロードがビットエラーか他のシンタックス違反を含むべきでないことを示す。値 1 は、NAL ユニットタイプオクテットとペイロードがビットエラーか他のシンタックス違反を含んでも良いことを示す。

MANE は、NAL ユニットにある検出されたビットエラーを示すために F ビットを設定すべきである。H.264 の仕様では、F ビットが 0 に等しいことが求められる。F ビットがセットされる時、デコーダはビットエラーか他のどんなシンタックス違反がペイロードまたは NAL ユニットタイプオクテットにおいて出現しているかもしれないことを知らせている。F ビットが 1 に等しい NAL ユニットに反応する簡易なデコーダは、そのような NAL ユニットの廃棄し、廃棄した NAL ユニットで消失したデータを隠蔽することになる。

#### NRI : 2 ビット

nal\_ref\_idc を示す。値 00 と非ゼロの値のセマンティクスは H.264 の仕様から未変更で残る。言い換えると、値 00 は NAL ユニットの内容がピクチャ間予測のための再構築参照ピクチャに使用されないことを示す。そのような NAL ユニットは参照ピクチャの状態を危うくすることなしに廃棄することが可能である。00 以上の値は、参照ピクチャの完全性を維持するために NAL ユニットの復号が要求されることを示す。

上記の仕様に追加して、この RTP ペイロードの仕様に従うと、00 より大きい NRI 値は、符号化器で決定される、相対的な転送の優先度を示す。MANE は、重要でない NAL ユニットよりもより重要な NAL ユニットを守るために、この情報の使用が可能である。最も高い転送の優先度は 11 で、次に 10 が続き、次に 01、最後に最も低い 00 が続く。

注：NRI のどんな非ゼロ値でも、H.264 デコーダにおいて同様に扱われる。それゆえ、NAL ユニットのデコーダに渡す時、受信側は NRI 値を操作する必要はない。

H.264 エンコーダは、nal\_unit\_type の値が 1 から 12 までの範囲内である時、H.264 の仕様 (7.4.1 節) に従って NRI 値を設定しなくてはならない。とりわけ H.264 の仕様は、6、9、10、11、12 に等しい nal\_unit\_type を持つ全ての NAL ユニットについて、NRI 値が 0 に等しくなくてはならないことを要求する。



7 または 8 (それぞれシーケンスパラメータセットまたはピクチャパラメータセットを示す) に等しい `nal_unit_type` を持つ NAL ユニットについては、H.264 エンコーダは NRI 値を 11 (バイナリフォーマット) に設定すべきである。5 (IDR ピクチャに属する符号化スライスを示す) に等しい `nal_unit_type` を持つ、最初に符号化されたピクチャの符号化スライス NAL ユニットについては、H.264 エンコーダは NRI 値を 11 (バイナリフォーマット) に設定すべきである。

NRI 値に対する残っている `nal_unit_type` の割り当てについては、次の例を使用してもよく、特定の環境において効率的であることが示された[13]。アプリケーションや使用中の H.264 付属資料 A に示されたプロファイルに依存して、他の割り当ての方が望ましい場合もある。

注：データパーティショニングは、特定のプロファイルにおいて利用可能ではない。例えば、メインプロファイルやベースラインプロファイルにおいてである。その結果として、NAL ユニットタイプの 2、3 または 4 が発生できるのは、ビデオビットストリームがデータパーティショニングを可能にするプロファイルの場合だけであって、メインプロファイルまたはベースラインプロファイルのストリームの場合ではない。

表 2：符号化スライス／主符号化参照ピクチャの  
符号化スライスデータパーティションに関する NRI 値の例 (RFC 3984)

NAL ユニットタイプ	NAL ユニットの内容	NRI (バイナリ)
1	非 IDR 符号化スライス	10
2	符号化スライスデータパーティション A	10
3	符号化スライスデータパーティション B	01
4	符号化スライスデータパーティション C	01

注：以前に述べたこととして、非参照ピクチャの NRI 値は、H.264 で指定された 00 である。

H.264 エンコーダは、符号化スライスと冗長符号化参照ピクチャの符号化スライス NAL ユニットについて、NRI 値を 01 (バイナリフォーマット) に設定すべきである。

NAL ユニットタイプ 24 から 29 までの NRI 値の定義は、本文書の 5.7 節と 5.8 節で与えられる。

NRI の推奨値が、13 から 23 までの範囲において `nal_unit_type` を持つ NAL ユニットに対しては与えられない。なぜなら、これらの値は ITU-T または ISO/IEC で予約されているからである。NRI の推奨値が、0 に等しいか 30 から 31 までの範囲において `nal_unit_type` を持つ NAL ユニットに対しては与えられない。これらの値のセマンティクスは本文書で定義されないからである。

#### 5.4 パケット化モード

この文書では三つのパケット化モードの仕様を定めている

- シングル NAL ユニットモード
- 非インタリーブモード
- インタリーブモード

シングル NAL ユニットモードは、H.241 に従う会話型システムをターゲットにしている。[15] (12.1 節参照) 非インタリーブモードは ITU-T 勧告 H.241 に従っていない会話型システムをターゲットにしている。非インタリーブモードでは、NAL ユニットは NAL ユニット復号順序で伝送される。インタリーブモードは、超低遅延を要求しないシステムをターゲットにしている。インタリーブモードは NAL ユニットデコード順以外での NAL ユニットの伝送を許容している。

パケット化モードはオプションのパケット化モード MIME パラメータの値、またはそのほかの手段で通知することができる。使用されたパケット化モードは、どの NAL ユニットタイプが RTP ペイロード中に許可されているかを管理している。いくつかの NAL ユニットタイプ値(表 3 で定義されていない)は将来の拡張のために予約されている。これらのタイプの NAL ユニットは送信者によって送られるべきではなく、また、受信者はそれを無視しなければならない。たとえば、パケットタイプが NAL ユニットでタイプ番号が 1-23 のものはシングル NAL ユニットモードと非インタリーブモードでは許可されているが、インタリーブモードでは許可されていない。パケット化モードはさらに詳細に 6 節で説明している。

表 3. 各パケット化モードについて許可されている NAL ユニットタイプのまとめ

(yes = 許可, no = 不許可, ig = 無視)

Type	Packet	Single NAL Unit Mode	Non-Interleaved Mode	Interleaved Mode
0	undefined	ig	ig	ig
1-23	NAL unit	yes	yes	no
24	STAP-A	no	yes	no
25	STAP-B	no	no	yes
26	MTAP16	no	no	yes
27	MTAP24	no	no	yes
28	FU-A	no	yes	yes
29	FU-B	no	no	yes
30-31	undefined	ig	ig	ig

### 5.5. 復号順序番号(DON)

インタリーブされたパケット化モードにおいて、NAL ユニットの伝送順序は NAL ユニットの復号順序と異なっていることが許可されている。復号順序番号(DON)は NAL ユニットの復号順序を示すもので、ペイロード構造中のフィールドや派生変数である。復号順序以外の伝送の場合や DON を使用した場合の原理や例は 13 節で与えられている。

伝送順序と復号順序の組み合わせは、次に示すオプションの `sprop-interleaving-depth` MIME パラメータにより制御される。オプションの `sprop-interleaving-depth` MIME パラメータは 0 に等しいか(明示的にあるいはデフォルトで)、それらの復号順序以外での NAL ユニットの伝送がそのほかの手段で許可されていないとき、NAL ユニットの伝送順序は NAL ユニットの復号順序に従わなければならない。オプションの `sprop-interleaving-depth` MIME パラメータの値が 0 より大きいとき、それらの復号順序以外での NAL ユニットの伝送

がそのほかの手段で許可されているとき、

- o MTAP16 や MTAP24 での NAL ユニットの順序は NAL ユニット復号順序になることは要求されない、  
そして

- o 二つの連続したパケット中で STAP-Bs, MTAPs, や FUs の逆カプセル化をすることにより生成される NAL ユニット順序は、NAL ユニット復号順序になることを要求されない。

シングル NAL ユニットパケットのための RTP ペイロード構造、STAP-A、や FU-A は DON を含んでいない。STAP-B や FU-B 構造は DON を含んでいて、そして MTAPs の構造は 5.7.2 節で記載されるように DON の導出を可能にする。

注：インタリーブモードにおいて FU-A が発生したとき、それはいつも DON を設定した FU-B に続く。

注：送信機がパケット毎にシングル NAL ユニットのキャプセル化し、そしてそれらの復号順序以外でパケットを送信するならば、STAP-B パケットタイプを使用することができる。

シングル NAL ユニットパケット化モードにおいて、RTP のシーケンス番号によって決定される NAL ユニットの伝送順序は NAL ユニット復号順序と同じでなければならない。

非インタリーブパケット化モードにおいて、シングル NAL ユニットパケット、STAP-As や FU-As での NAL ユニットの伝送順序はそれらの NAL ユニットの復号順序と同じでなければならない。STAP 中の NAL ユニットの伝送順序は NAL ユニット復号順序で出現しなければならない。

したがって、復号順序は、まず STAP 内では暗黙の番号から提供され、次に STAPs や FUs や単一 NAL ユニットパケットの間の順序については RTP シーケンス番号から提供される。

STAP-B、MTAP や FU-B に始まる分割ユニットの連続により伝送される、NAL ユニットのための DON 値の通知方法は、5.7.1,5.7.2 や 5.8 節でそれぞれに定められる。

伝送順序での最初の NAL ユニットの DON 値は任意の値に設定してもよい。

DON 値は、0 から 65535 までの範囲にある。最大値になった後 DON 値は 0 に戻る。

任意の STAP-B, MTAP または FU-B で始まる分割ユニットに含まれる二つの NAL ユニットの復号順序は次のように決定される。DON(i) を伝送順序中で i 番目のインデックスを持つ NAL ユニットの復号順序番号とする。関数  $don\_diff(m,n)$  は次のように指定される。

もし  $DON(m) == DON(n)$  ならば  $don\_diff(m,n) = 0$

もし  $DON(m) < DON(n)$  かつ  $DON(m) - DON(n) < 32768$  ならば  
 $don\_diff(m,n) = DON(n) - DON(m)$

もし  $DON(m) > DON(n)$  かつ  $DON(m) - DON(n) \geq 32768$  ならば  
 $don\_diff(m,n) = 65536 - DON(m) + DON(n)$

もし  $DON(m) < DON(n)$  かつ  $DON(n) - DON(m) \geq 32768$  ならば  
 $don\_diff(m,n) = - (DON(m) + 65536 - DON(n))$

もし  $DON(m) > DON(n)$  かつ  $DON(m) - DON(n) < 32768$  ならば

$don\_diff(m,n) = - (DON(m) - DON(n))$

$don\_diff(m,n)$ の正の値は、復号順序においてインデックス  $n$  の伝送順序を持つ NAL ユニットのインデックス  $m$  の伝送順序の NAL ユニットに続くことを示している。 $don\_diff(m,n)$ が 0 に等しいとき、二つの NAL ユニットの復号順序はどちらでも良い。 $don\_diff(m,n)$ が負の値の時は、復号順序においてはインデックス  $m$  の伝送順序をもつ NAL ユニットより先行したインデックス  $n$  の伝送順序を持つ NAL ユニットがあることを示す。

DON の値に関連するフィールド(DON,DONB,DOND 5.7 節参照)は、上記に指定した DON の値により決定される復号順序が NAL ユニットの復号順序に合致するようになっていなければならない。もし NAL ユニットの復号順序において二つの NAL ユニットの復号順序が入れ替えられていて、新しい順序が NAL ユニットの復号順序に従っていないなら NAL ユニットは同じ DON の値を持っていてはならない。もし NAL ユニットストリーム中に二つの連続した NAL ユニットの順序が入れ替えられていて、新しい順序がまだ NAL ユニットの復号順序に従っているならば同じ NAL ユニットは DON 値を有しているかもしれない。たとえば、任意のスライス順序が使用しているビデオ符号化プロファイルによって許可されているとき、符号化された映像のすべての符号化されたスライス NAL ユニットは同じ DON の値を持つことを許されている。その結果、同じ DON 値をもつ NAL ユニットは任意の順序で復号することができる。そして、異なる値を持つ二つの NAL ユニットは、上記で指定された順序でデコーダに渡されるべきである。NAL ユニットの復号順序において二つの連続した NAL ユニットが異なる DON 値を持つとき、復号順序について二つ目の NAL ユニットの DON 値は最初の DON 値に 1 加算したものであるべきである。

7 節で NAL ユニットの復号順序を回復する逆カプセル化処理の例を示す。

注: エラーフリーの伝送においても、受信機は NAL ユニットの復号順序において二つの連続した NAL ユニットに対する DON 値の差分絶対値が 1 に等しいことを期待するべきではない。DON の値と NAL ユニットの関連付けの時点で、すべての NAL ユニットが受信機に渡されるかどうかかわからないため、1 の増分は必要ではない。たとえば、パケットをフォワードしているネットワークでビットレートが不足している場合、ゲートウェイは非参照ピクチャまたは SEI NAL ユニットの符号化されたスライス NAL ユニットをフォワードしないかもしれない。別の例では、生中継は時々、コマercialのように事前に符号化されたコンテンツによって割り込まれることがある。事前に符号化された最初のイントラピクチャは、受信側ですぐに表示できることを保証するために予め転送される。最初のイントラピクチャを転送するとき、コマercialの発信元は、符号順序に従って、事前に符号化したクリップの最初のイントラピクチャを送信する前にどれくらいの数の NAL ユニットが符号化されたかを正確には知らない。したがって、転送される時、事前に符号化されたクリップの最初のイントラピクチャの NAL ユニットのための DON 値は推定されなければならない。DON 値のギャップが起こるかもしれない。

## 5.6. シングル NAL ユニットパケット

ここで定義するシングル NAL ユニットパケットは[1]で定義されたタイプのただひとつだけの NAL ユニットを含まなければならない。これは、シングル NAL ユニットパケットの中では、集合パケットも断片化されたユニットも両方とも使用することができないことを意味している。RTP シーケンス番号順序でシングル

NAL ユニットパケットを逆カプセル化して得られる NAL ユニットストリームは、NAL ユニット復号順序に従わなければならない。シングル NAL ユニットパケットの構成を図 2 に示す。

注：NAL ユニットの最初のバイトは RTP ペイロードヘッダと共用している。

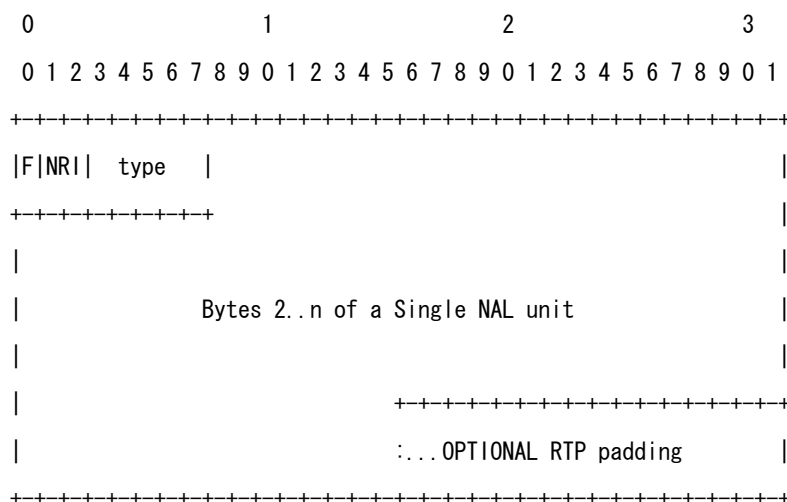


図 2. シングル NAL ユニットパケットのための RTP ペイロードフォーマット

### 5.7. 集合パケット

集合パケットは、本ペイロード仕様における NAL ユニートを集合させる仕組みである。この仕組みは、ターゲットとなる 2 種類のネットワークの間で MTU サイズが劇的に異なる場合に対応するために導入された。その 2 種類とは、有線 IP ネットワーク(MTU サイズは、Ethernet の MTU サイズ(およそ 1500 バイト)に制限されていることが多い)と、IP/非 IP(例：ITU-T H.324/M)ベースの無線通信システム(254 バイト以下の伝送ユニットサイズにすることが望ましいとされる)である。この 2 つの世界の間でのメディア変換を避けるために、またパケット化によるオーバーヘッドを極力避けるために、NAL ユニートの集合の仕組みが導入されている。

本仕様では、次の 2 種類の集合パケットが定義されている。

- ・シングルタイム集合パケット(Single-time aggregation packet, STAP)：単一の NAL ユニット(NALU)時刻で NAL ユニートを集合する。次の 2 種類の STAP が定義される：DON(復号順序番号)のない STAP(STAP-A)、ならびに DON を持つ STAP(STAP-B)。
- ・マルチタイム集合パケット(Multi-time aggregation packet, MTAP)：NALU 時刻が異なる可能性のある NAL ユニートを集合する。NAL ユニートのタイムスタンプオフセット長が異なる 2 個の MTAP が定義される。

NALU 時刻という用語は、該当する NAL ユニットが RTP パケットで伝送されるならばその RTP タイムスタンプが持つであろう値として定義される。

集合パケットとして伝送される各 NAL ユニットの、集合ユニットにカプセル化されている。4 種類の集合ユニットとそれぞれの特性については、後述を参照のこと。

図 3 は、集合パケットの RTP ペイロードフォーマットの構造を示している。

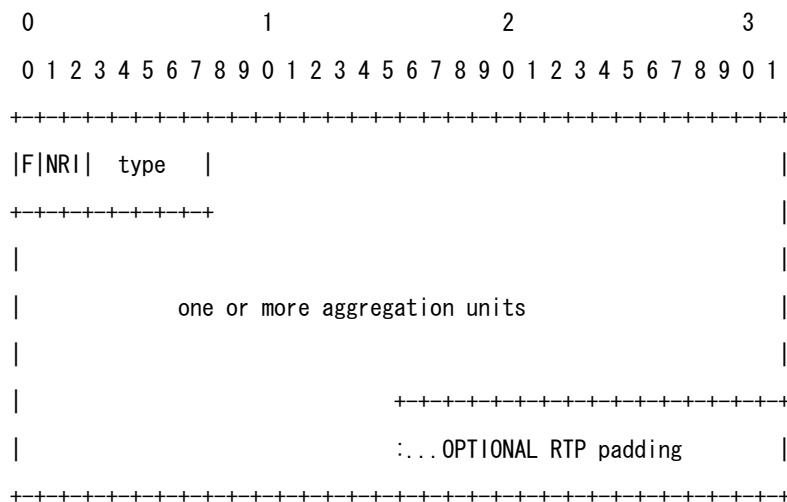


図 3. 集合パケットの RTP ペイロードフォーマット

MTAP と STAP で共通して、次のパケット化ルールを持つ。RTP タイムスタンプは、集合されるすべての NAL ユニットの NALU 時刻のうち、もっとも早い時刻のものをセットする。NAL ユニットタイプのバイトフィールドは、表 4 に示す適切な値をセットする必要がある。集合された NAL ユニットのすべての F bit が 0 にセットされている場合は、この F bit は 0 となり、それ以外の場合は 1 となる。NRI の値は、集合パケットとして伝送されるすべての NAL ユニットのうちの最大値をセットする。

表 4. STAP および MTAP のフィールド

Type フィールド	パケット	タイムスタンプ オフセットフィールド長 (bit)	DON 関連フィールド (DON, DONB, DOND) が存在するか否か
24	STAP-A	0	no
25	STAP-B	0	yes
26	MTAP16	16	yes
27	MTAP24	24	yes

集合パケットが伝送される RTP ヘッダのマーカービットは、集合パケットの最後の NAL ユニット自身が伝送されるとしたら持つであろうマーカービット値をセットする。

集合パケットのペイロードは、1 つ以上の集合ユニットから成り立っている。4 種類の集合ユニットについ

ては、5.7.1 節/5.7.2 節を参照のこと。集合パケットは、必要に応じて任意数の集合ユニットを伝送することが可能である。しかし、集合パケットの総データサイズは、当然ながら 1 個の IP パケット内に収まる必要があり、最終的には IP パケットが MTU サイズより小さくなるような選択をしなければならない。1 個の集合パケットは、5.8 節で示すようなフラグメンテーションユニットを含んではいけない。集合パケットはネストしてはいけない。すなわち、1 個の集合パケットが他の集合パケットを含んではならない。

### 5.7.1. シングルタイム集合パケット

シングルタイム集合パケット(STAP)は、集合される全ての NAL ユニットが同じ NALU 時刻を共有するような場合に用いることを推奨する。STAP-A のペイロードは、図 4 で示す通り、DON を含まず、1 個以上のシングルタイム集合ユニットを含んでいる。STAP-B のペイロードは、図 5 で示す通り、(ネットワークバイトオーダーで)16 ビット符号なしの復号順序番号(DON)と、その後続く 1 個以上の単一時間集合ユニットより成り立っている。

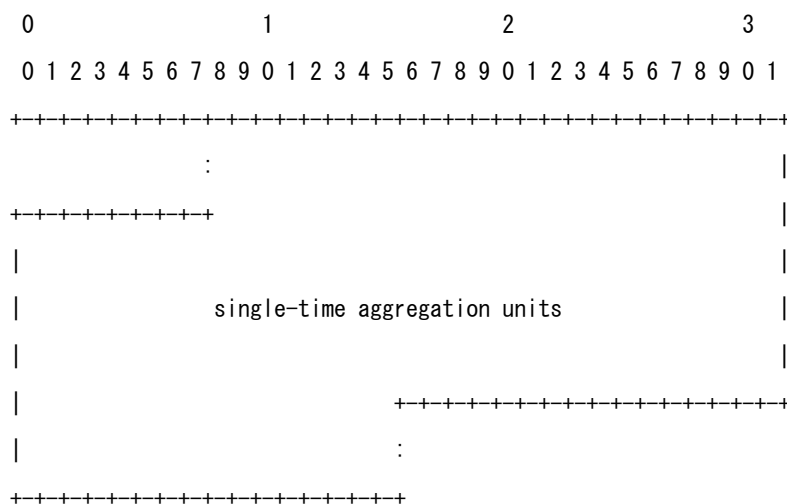


図 4. STAP-A のペイロードフォーマット

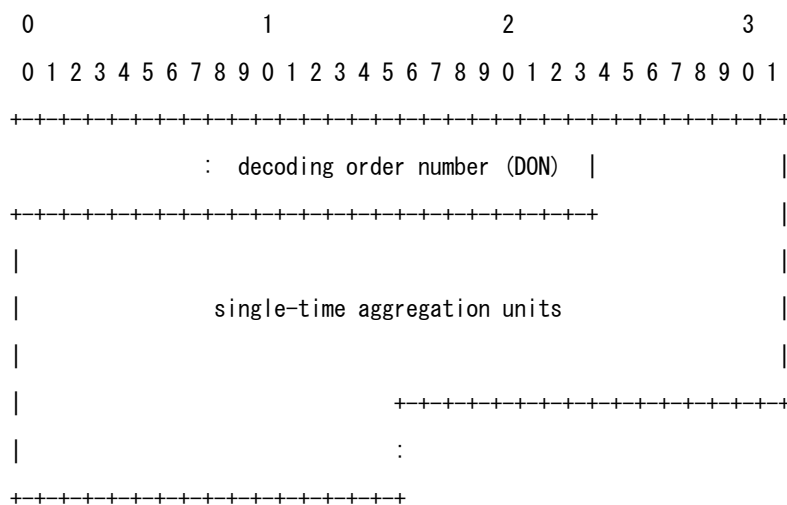


図 5. STAP-B のペイロードフォーマット

DON フィールドは、STAP-B 内の、伝送順序で最初の NAL ユニットの DON 値を表している。STAP-B 内の後続する NAL ユニットについては、DON 値は、

$$(STAP-B \text{ 内の直前の NAL ユニットの DON 値}+1) \% 65536$$

で表される。ただし、'%' はモジュロ演算を示す。

シングルタイム集合ユニットは、(ネットワークバイトオーダーで)16 ビット符号なしのサイズ情報と、NAL ユニットタイプのバイトを含む NAL ユニット自身より成り立っている。サイズ情報は、バイト単位での NAL ユニットの大きさを示す。シングルタイム集合ユニットは、RTP ペイロード内でバイトアライメントが取られているが、必ずしも 32 ビットワード単位でのアライメントは取られているとは限らない。図 6 は、シングルタイム集合ユニットの構造を示すものである。

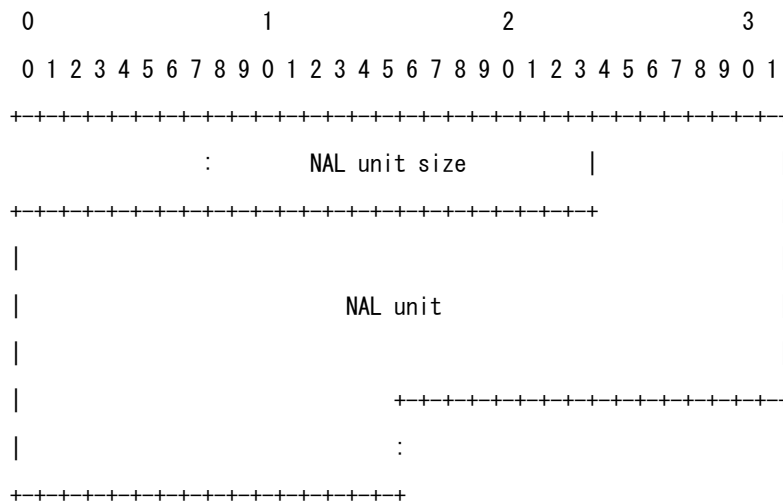


図 6. シングルタイム集合ユニットの構造



図 7 は、1 個の STAP-A を含む RTP パケットの例を表している。この STAP は、2 個のシングルタイム集合ユニットを持ち、各々のユニットは図内では 1、2 とラベル付けしてある。

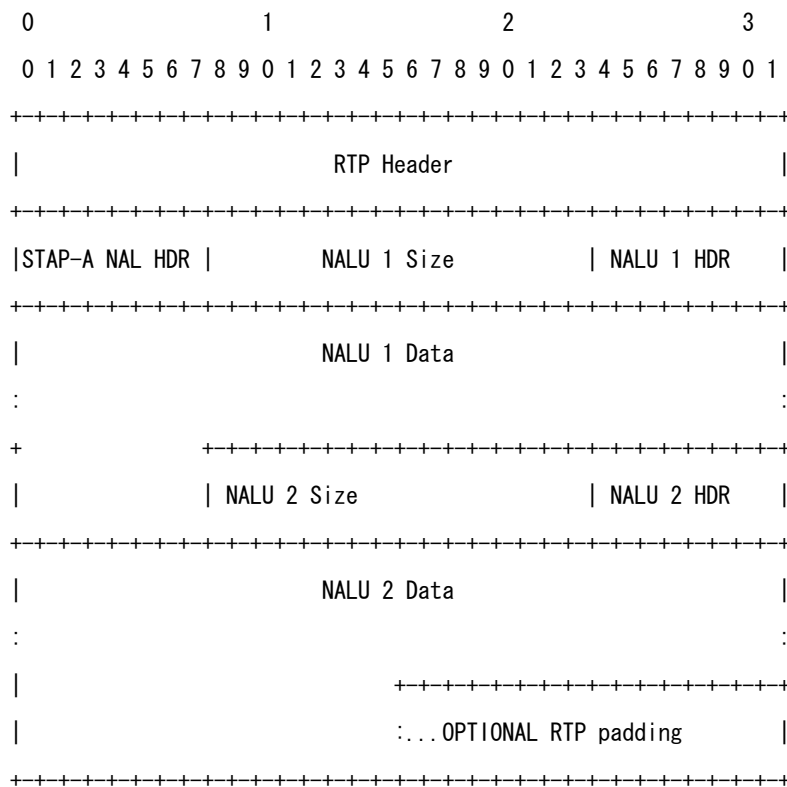


図 7. 1 個の STAP-A と 2 個のシングルタイム集合ユニットを含む RTP パケットの例

図 8 は、1 個の STAP-B を含む RTP パケットの例を表している。この STAP は、2 個のシングルタイム集合ユニットを持ち、各々のユニットは図内では 1、2 とラベル付けしてある。

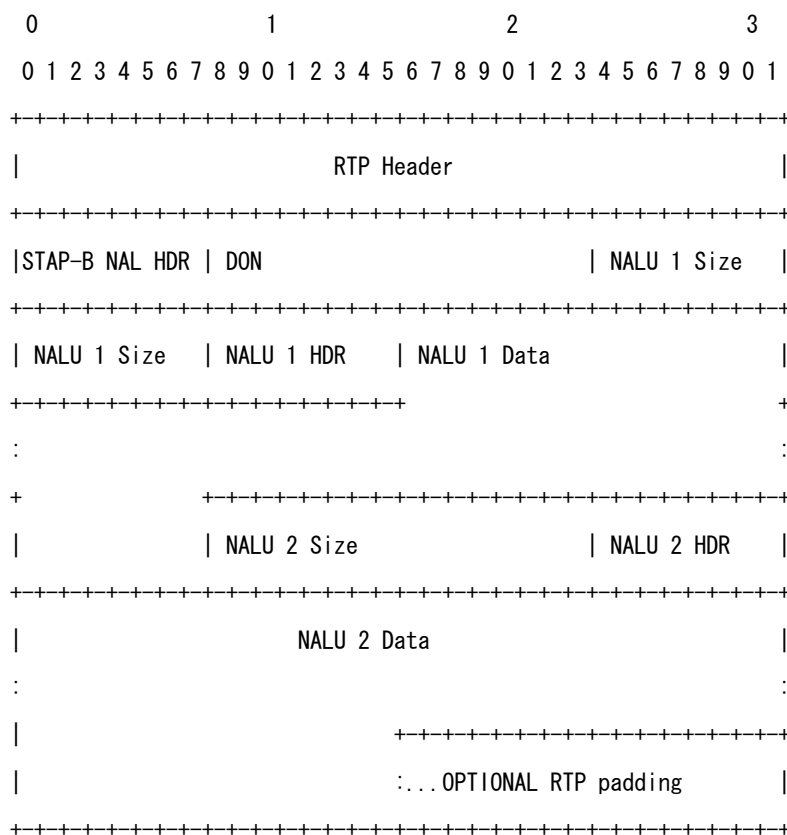


図 8. 1 個の STAP-B と 2 個のシングルタイム集合ユニットを含む RTP パケットの例

### 5.7.2. マルチタイム集合パケット(MTAP)

MTAP の NAL ユニットペイロードは、図 9 で示されるように、16 ビット長符号無し の復号順序番号ベース (DONB) (ネットワークバイトオーダーで)、及び、一つ以上のマルチタイム集合ユニットから構成される。

DONB は、MTAP の NAL ユニット中の NAL ユニット復号順序における最初の NAL ユニットの DON 値を含まなければならない。

注：NAL ユニット復号順序における最初の NAL ユニットは、必ずしも MTAP 中にカプセル化された NAL ユニットの順序において、最初の NAL ユニットであるとは限らない。

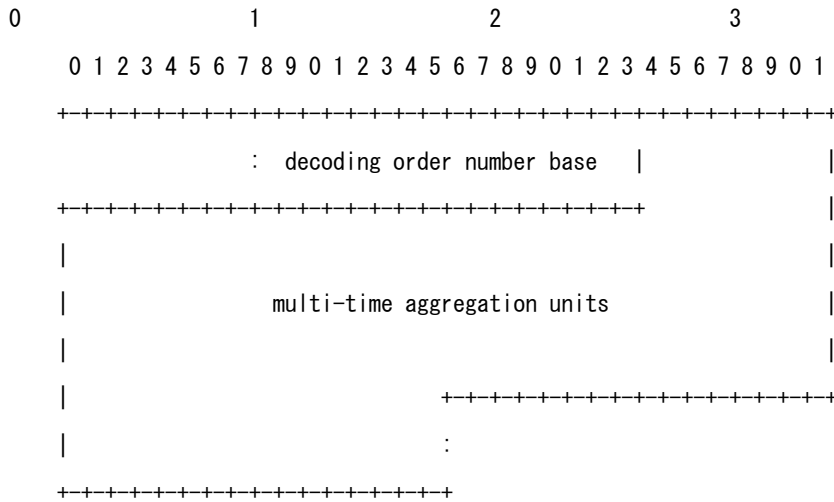


図 9. MTAP のための NAL ユニットペイロードフォーマット

この仕様では 2 つの異なるマルチタイム集合ユニットが定義される。

それらの両方とも、(ネットワークバイトオーダーで)次の NAL ユニットの 16 ビット長符号無しの情報、8 ビット長符号無し番号差分(DOND)、この NAL ユニットの(ネットワークバイトオーダーで)n ビット長のタイムスタンプオフセット(TS オフセット)から構成される。ここで、n は、16 ビット、または、24 ビットである。

異なる MTAP タイプ (MTAP16、及び、MTAP24)での選択は、アプリケーション依存である。

より大きなタイムスタンプオフセットは、MTAP の柔軟性をより高くするが、オーバーヘッドもまた、高くなる。

MTAP16、MTAP24 のためのそれぞれのマルチタイム集合ユニットの構造を、図 10、図 11 に示す。

パケット内の集合ユニットの開始位置、または、終了位置が 32 ビットワード境界である必要はない。

次の NAL ユニットの DON は、( DONB + DOND ) % 65536 に等しい。% はモジュロ演算を意味する。

本文書では、MTAP 中の NAL ユニットがどのような順序であるかは規定しない。しかし、(ほとんどの場合)、NAL ユニット復号順序が使われるべきである。

タイムスタンプオフセットフィールドは、次の式の値に等しい値に設定すべきである。

NALU 時間がパケットの RTP タイムスタンプより大きい、または、同じ場合、タイムスタンプオフセットは、

$$(NAL \text{ ユニットの NALU 時間}) - (\text{パケットの RTP タイムスタンプ})$$

と等しい。

NALU 時間がパケットの RTP タイムスタンプより小さい場合、タイムスタンプオフセットは、

$$(NALU \text{ 時間}) + (2^{32} - \text{パケットの RTP タイムスタンプ})$$

と等しい。

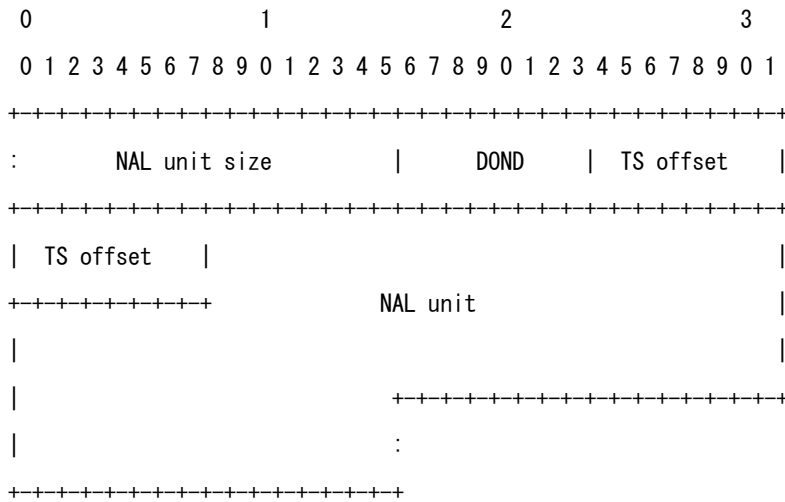


図 10. MTAP16 のためのマルチタイム集合ユニット

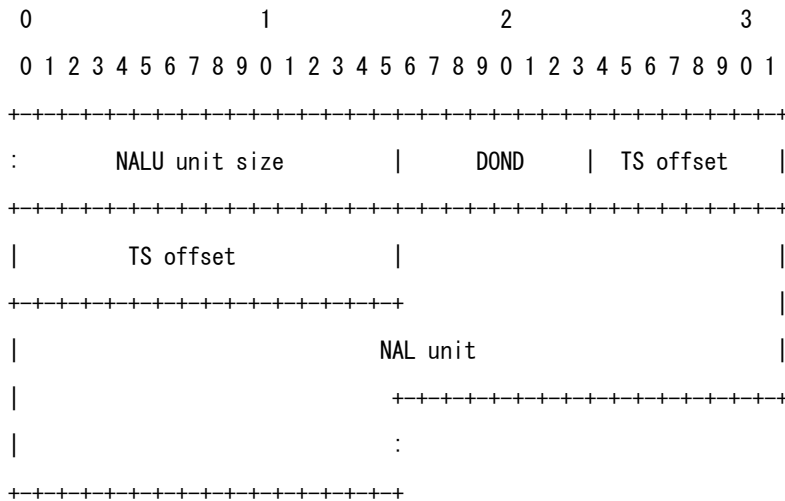


図 11. MTAP24 のためのマルチタイム集合ユニット

MTAP における「最も早期の」マルチタイム集合ユニットのためのタイムスタンプオフセットは、ゼロでなければならない。従って、MTAP そのものの RTP タイムスタンプは、最も早期の NALU 時間と同じである。

注：「最も早期の」マルチタイム集合ユニットは、集合ユニットが シングル NAL ユニットパッケージに入れられた場合、MTAP の全ての集合ユニットの中で最も小さな拡張 RTP タイムスタンプを持つものである。

拡張タイムスタンプは、32 ビットを超える大きさを持ち、タイムスタンプフィールドのラップアラウンドのカウントが可能である。そのため、タイムスタンプが巡回した場合に最も小さな一つの値を決める事ができる。

そのような「最も早期の」集合ユニットは、MTAP にカプセル化された集合ユニットの順番で最初のものでなくてもよい。

「最も早期の」NAL ユニットは、NAL ユニット復号順序における最初の NAL ユニットと同じである必要

はない。

図 12 は、MTAP16 タイプのマルチタイム集合パッケージを含んでいる RTP パッケージの例を示しており、マルチタイム集合パッケージには、図中に 1、2 としてラベル表示されている二つのマルチタイム集合ユニットを含んでいる。

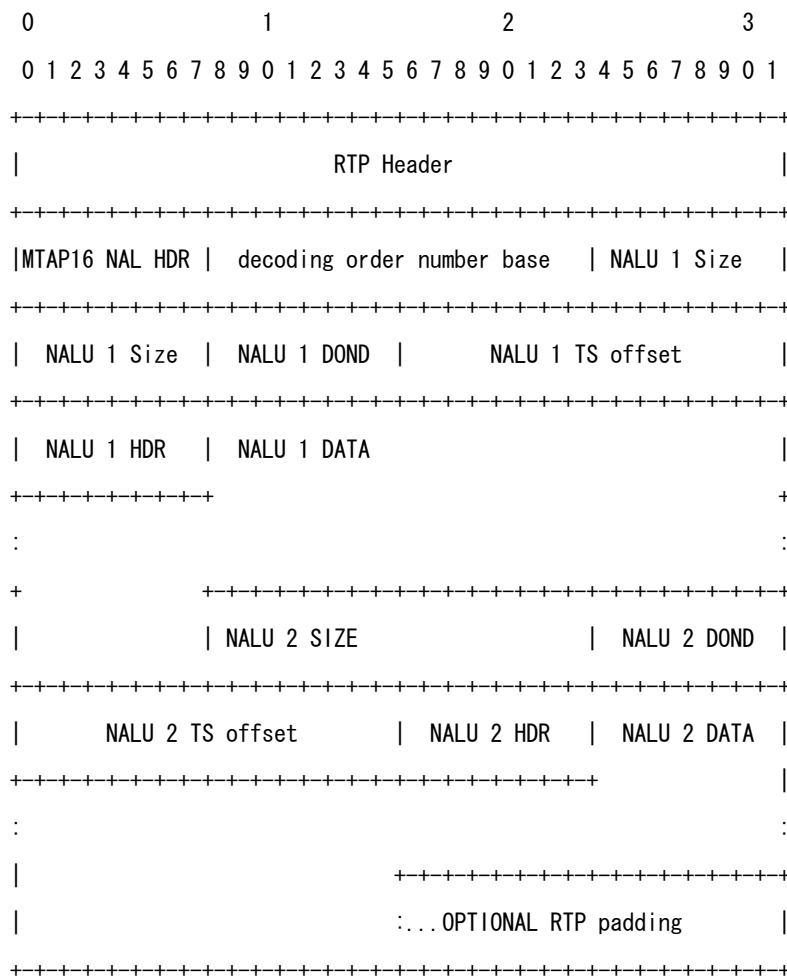


図 12. MTAP16 タイプのマルチタイム集合パッケージ (2つのマルチタイム集合ユニットから成る) を含む RTP パッケージ

図 13 は、MTAP24 タイプのマルチタイム集合パッケージを含んでいる RTP パケットの例を示しており、マルチタイム集合パッケージには、図中に 1、2 としてラベル表示されている二つのマルチタイム集合ユニットを含んでいる。

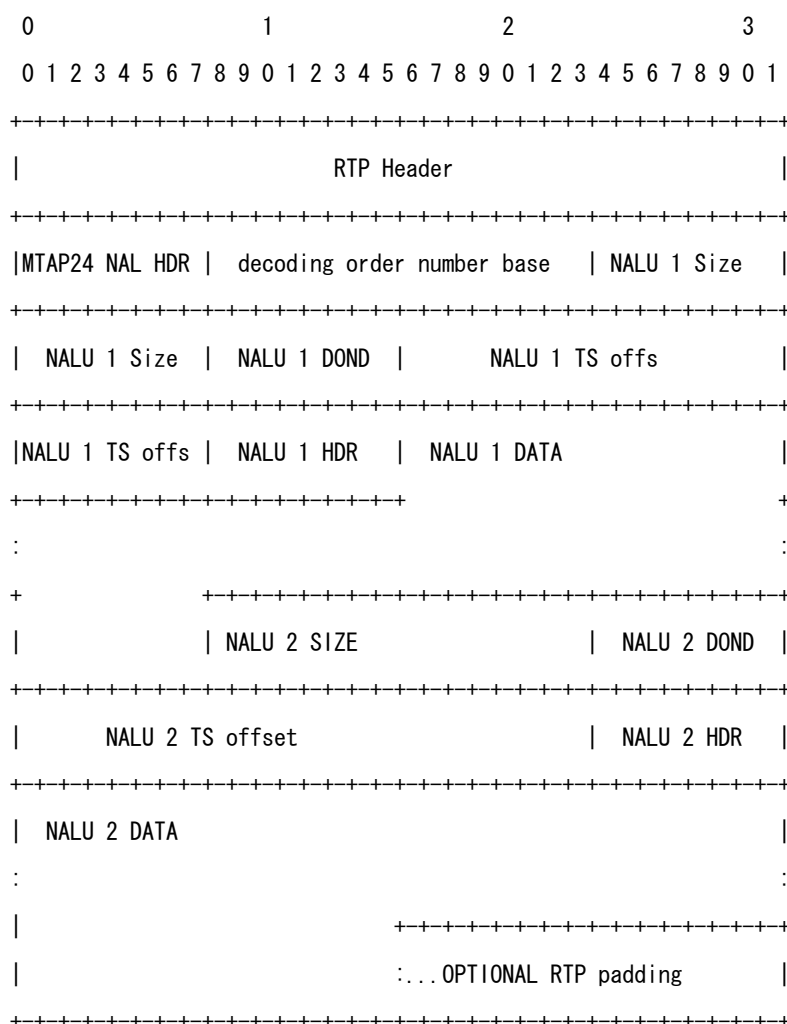


図 13. MTAP24 タイプのマルチタイム集合パッケージ (2 つのマルチタイム集合ユニットから成る) を含む RTP パッケージ

### 5.8. 分割ユニット (FU)

このペイロードタイプは、1 つの NAL ユニットのいくつかの RTP パッケージに細分化することを許可している。より低いレイヤでの分割 (例えば IP レイヤによる) に頼る代わりにアプリケーションレイヤで細分化することにより以下の有利な点がある。

- ペイロードフォーマットは、IPv4 ネットワーク上で、録画ビデオ、特に High Definition フォーマットで現れるような 64kbyte よりも大きい NAL ユニットの送信する能力がある。H.264 では、ピクチャあたりのスライス数の制限があり、それは結果としてピクチャあたりの NAL ユニットの制限となり、大きい NAL ユニットの制限となる。

一分割メカニズムは、シングルピクチャの細分化を可能とし、12.5 節で表現される一般的前方誤り訂正が適用できる。

分割はシングル NAL ユニットのみに定義されており、集合パケットには定義されていない。NAL ユニットのフラグメントは、その NAL ユニットの整数個の連続オクテットより構成されている。NAL ユニットの各々のオクテットは、その NAL ユニットの厳密に 1 つのフラグメントに帰属しなければならない。同じ NAL ユニットのフラグメントは、連続した順序で、昇順の RTP シーケンス番号をつけて送信しなければならない（最初と最後のフラグメント間では同じ RTP ストリームの別の RTP パケットを送信してはいけない）。同様に NAL ユニットのフラグメントは、RTP シーケンス番号順に再集合させるべきである。

NAL ユニットが分割され、分割ユニット (FU) 内で送信された場合、分割 NAL ユニットと呼ぶ。STAP や MTAP は分割してはいけない。FU は入れ子にしてはいけない、すなわち FU に他の FU を含めてはいけない。

FU を伝送している RTP パケットの RTP タイムスタンプは、分割された NAL ユニットの NALU タイムにセットされる。

図 14 は FU-A の RTP ペイロードフォーマットを表している。FU-A は 1 オクテットの分割ユニットインジケータ、1 オクテットの分割ユニットヘッダ、そして分割ユニットペイロードから成り立っている。

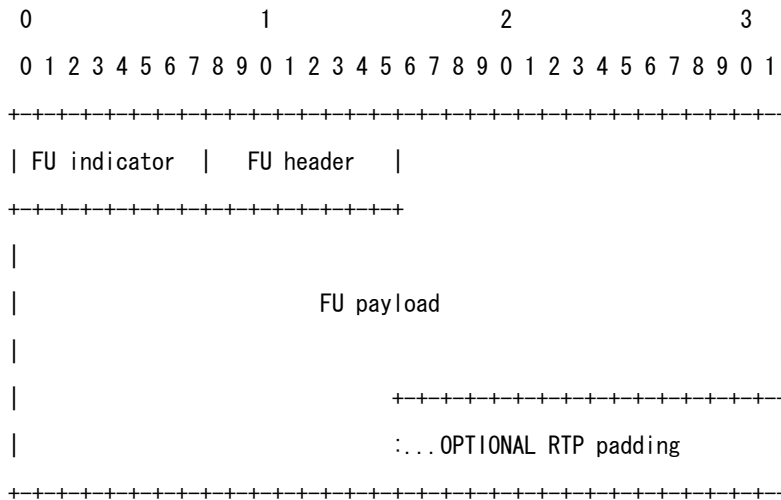


図 14 FU-A における RTP ペイロードフォーマット

図 15 は FU-B の RTP ペイロードフォーマットを表している。FU-B は 1 オクテットの分割ユニットインジケータ、1 オクテットの分割ユニットヘッダ、復号順序番号 (DON) (ネットワークバイトオーダーで表現される)、分割ユニットペイロードから構成されている。言い換えれば、追加の DON 領域をのぞいて FU-B の構造は FU-A と同様である。

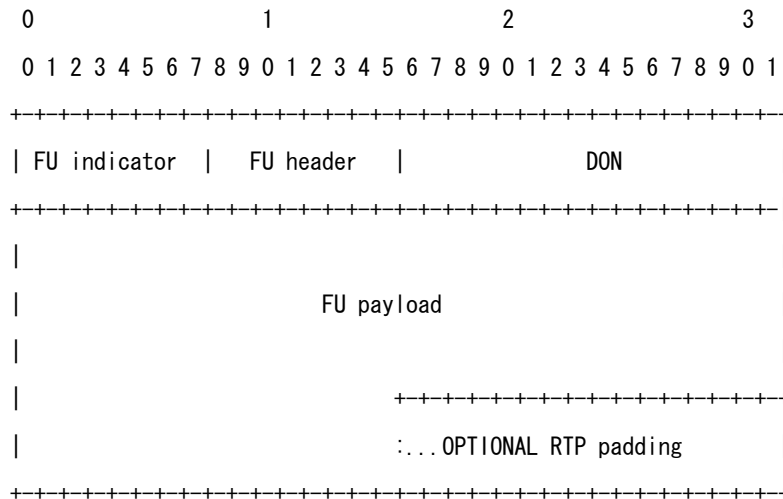
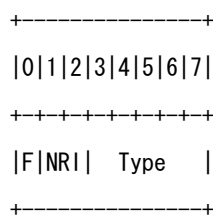


図 15 FU-B の RTP ペイロードフォーマット

FU-B 型の NAL ユニットの断片化は、断片化した NAL ユニットの最初の分割ユニットのインタリーブパッケージ化モードを利用しなければならない。FU-B 型の NAL ユニットのほかの場合では利用してはいけない。言い換えれば、インタリーブパッケージ化モードでは、分割した各 NALU は最初のフラグメントとして FU-B があり、その後 1 つ以上の FU-A のフラグメントが続く。

FU インジケータオクテットは以下のようなフォーマットである。



FU インジケータオクテットの **Type** フィールドの値が 28、29 の場合は、それぞれ FU-A と FU-B を示している。F ビットの利用は 5.3 節に記述してある。NRI フィールド値は、分割 NAL ユニットの NRI フィールドの値にしたがって設定しなければならない。



FU ヘッダは以下のフォーマットになっている。

```
+-----+
|0|1|2|3|4|5|6|7|
+-----+
|S|E|R| Type |
+-----+
```

#### S : 1 ビット

1 に設定されたとき、Start ビットは分割 NAL ユニットのスタートを示す。それに続く FU ペイロードが分割 NAL ユニットペイロードのスタートでなければ、Start ビットは 0 にセットされる。

#### E : 1 ビット

1 に設定されたとき、End ビットは分割 NAL ユニットの終わりを示す、すなわちペイロードの最後のバイトは分割 NAL ユニットの最後のバイトでもある。それに続く FU ペイロードが分割 NAL ユニットの最後のフラグメントでなければ、End ビットは 0 に設定される。

#### R : 1 ビット

予約ビットは 0 と等しくなければならず、受信器は無視しなければならない。

#### Type : 5 ビット

[1] の表 7-1 で定義された NAL ユニットペイロードタイプ。

FU-B 内の DON 値は 5.5 節に記述されたものとして選択される。

注：FU-B 内の DON フィールドは、ゲートウェイが入力 NAL ユニットの復号順序に再編することなく、分割 NAL ユニットの FU-B に分割可能にする。

分割 NAL ユニットの 1 つの FU で送信してはいけない、すなわち Start ビットと End ビットは、同じ FU ヘッダで両方ともセットしてはいけない。

FU ペイロードは分割 NAL ユニットのペイロードのフラグメントから成り立っている。一連の FU の分割ユニットペイロードが連続的につながっていたならば、分割 NAL ユニットのペイロードが再構築できるようにするためである。

分割 NAL ユニットの NAL ユニットタイプオクテットのようなのは、分割ユニットペイロードには含まれていない。むしろ分割 NAL ユニットの NAL ユニットタイプオクテットの情報は、分割ユニットの FU インジケータオクテットの F や NRI フィールドや FU ヘッダの Type フィールドの中で運ばれる。FU ペイロードは、任意数のオクテットあってもよいし、なくともよい。

注：空の FU は、ほとんど損失のない環境で、あるクラスの送信側の遅延を減らすために許されている。これらの送信者は、NALU が完全に生成される前、それゆえに NALU サイズがわかる前に、NALU フラグメントをパケット化するのが特徴である。長さ 0 の NALU フラグメントが許されないとしたら、送信者は、現在のフラグメントが送信される前にそれに続くフラグメントの少なくとも 1 ビットを生成しなければならない。時々いくつかのマクロブロックは 0 ビットで占有されるという H.264 の特徴のために、このことは望ましくなく、遅延を増加させることになる。しかしながら、長さ 0 の

NALU の使用（可能性）は、その送信のために付加的なパケットを要し、NALU 損失リスクの増加することに注意すべきである。

分割ユニットが損失した時、受信器は、伝達順序が同じ分割 NAL ユニットと一致するような、次に続くすべての分割ユニットは捨てるべきである。

エンドポイントや MANE の受信者は、たとえその NAL ユニットのフラグメント  $n$  が受信されていなくとも、NAL ユニットの最初の  $n-1$  フラグメントを集めて（不完全な）NAL ユニットにしてもよい。

この場合、NAL ユニットの `forbidden_zero_bit` はシンタックス違反を示すために、1 に設定しなければならない。

## 6. パケット化規則

パケット化モードは 5.2 節で紹介されている。2 つ以上のパケット化モードに共通のパケット化規則は 6.1 節に定義される。シングル NAL ユニットモードや非インタリーブモード、インタリーブモードについてのパケット化規則はそれぞれ、6.2 節、6.3 節、6.4 節で定義される。

### 6.1. 共通パケット化規則

すべての送信機は、使用中のパケット化モードに関わらず次のパケット化規則に従わなければならない。

- 同じ符号化ピクチャに属している符号化スライスNALユニットまたは符号化スライスデータパーティションNALユニット（従って同じRTPタイムスタンプ値を共有している）は、[1]で定義される適用プロファイルによって許されるどのような順序で送出されても良い。しかしながら、遅延に厳格なシステムについては、遅延を最小にするために元々の符号化順序で送出されるべきである。符号化順序はスキャン順序ではないが、NALパケットがRTPスタックに利用可能となる順序であることに注意すること。
- パラメータセットは、8.4 節で示される規則と推奨手順に従って処理される。
- MANE（メディア・アウェア・ネットワーク要素）は、シーケンスパラメータセットNALユニットまたはピクチャパラメータセットNALユニットを除いたNALユニットを複製してはいけない。本文書も H.264 もどちらも複製NALユニットを見分ける手段は提供されない。シーケンスパラメータセットNALユニットやピクチャパラメータセットNALユニットは、正しい受信をより確実にするために複製しても良い。しかし、どんな複製も、有効なシーケンスパラメータセットやピクチャパラメータセットの内容に影響を与えてはいけない。複製は、アプリケーションレイヤ上で行われるべきであり、（同一のシーケンス番号の）RTPパケットの複製によって行われるべきではない。

非インタリーブモードやインタリーブモードを使用する送信機は、次のパケット化ルールを実施しなければならない。

- MANEは、RTPトランスレータにおいて、シングルNALユニットパケットを1つの集合パケットに変換するか、集合パケットをいくつかのシングルNALユニットに変換するか、それらを組み合わせた変換をしてもよい。RTPトランスレータは、少なくとも次のパラメータを考慮に入れるべきである。
  - パスMTUサイズ
  - 不均一な誤り保護機構（例えば、特にシーケンスパラメータセットNALユニットやピクチャパラメータセットNALユニット、符号スライスデータパーティションNALユニットのための、RFC2733 [18]に従ったパケットベースのFECによる）
  - システムの許容遅延
  - 受信機のバッファリング能力

注：RTP トランスレータは、RTCP を RFC3550 で処理することが求められる。

## 6.2. シングル NAL ユニットモード

このモードが使用されるのは、オプションで MIME パラメータの `packetization-mode` の値が 0 に等しいか、`packetization-mode` が存在しないか、外部手段によって他のパケット化モードが通知されていない時である。全ての受信機はこのモードをサポートしなければならない。このモードは、主として H.241 [15]を使ったシステム（12.1 節を参照）と互換性のある低遅延アプリケーションを意図したものである。このモードではシングル NAL ユニットだけを使用しても良い。STAP や MTAP、FU は使用してはいけない。シングル NAL ユニットパケットの転送順序は、NAL ユニット復号順序に従わなければならない。

## 6.3. 非インタリーブモード

このモードが使用されるのは、オプションで MIME パラメータの `packetization-mode` の値が 1 に等しいか、外部手段によってモードが有効になる時である。このモードをサポートするべきである。このモードは、主として、低遅延アプリケーション用である。シングル NAL ユニットパケット、STAP-A、FU-A だけはこのモードを使用してもよい。STAP-B、MTAP、FU-B は使用してはいけない。NAL ユニットの転送順序は、NAL ユニット復号順序に従わなければならない。

## 6.4. インタリーブモード

このモードが使用されるのは、オプションで MIME パラメータの `packetization-mode` の値が 2 に等しいか、外部手段によってこのモードが有効になった時である。一部の受信機がこのモードをサポートする場合がある。STAP-B や MTAP、FU-A、FU-B は使用しても良い。STAP-A とシングル NAL ユニットだけのパケットは使用してはいけない。パケットと NAL ユニットの転送順序は、5.5 節で示されるように制限される。

## 7. 逆パケット化処理 (Informative)

逆パケット化処理は実装依存である。それゆえ、以下の記述を適切な実装の一例として見るべきである。他の手法も同様に使用しても良い。記述されたアルゴリズムについても最適化の余地があるであろう。7.1 節はシングル NAL ユニットと非インタリーブパケット化モード用の逆パケット化処理を表すが、7.2 節はインタリーブモード用の処理を記述している。7.3 節は高度な受信機用の追加の逆カプセル化指針を含む。

バッファ管理に関連する全ての標準的な RTP の機構は適用される。特に、（RTP シーケンス番号や RTP タイムスタンプによって示される）重複や過度に遅延した RTP パケットは廃棄される。復号の正確な時刻を決定するために、適切なストリーム間同期を可能にするための意図的な遅延のような要因は計算に入れなければならない。

### 7.1. シングル NAL ユニットと非インタリーブモード

受信機は、送信遅延ジッタを補償する受信バッファを含む。受信機は、到達パケットを受信順序で受信バッファに蓄積する。パケットは、RTP シーケンス番号順序で逆カプセル化される。もし逆カプセル化パケットがシングル NAL ユニットパケットである場合、パケットに含まれる NAL ユニットはデコーダに直接渡される。もし逆カプセル化パケットが STAP-A である場合、パケットに含まれる NAL ユニットはパケットにおいてカプセル化された順序でデコーダに渡される。もし逆カプセル化パケットが FU-A である場合、分割化された NAL ユニットの全てのフラグメントは結合され、デコーダに渡される。

注：もし、デコーダが任意スライス順序をサポートする場合、ピクチャの符号化スライスは受信順序や送信順序に関わらず、どんな順序でもデコーダに渡すことができる。

## 7.2. インタリーブモード

これらの逆パケット化規則の背景にある一般的な概念は、NAL ユニットの送信順序から NAL ユニット復号順序に順序替えするということである。

受信機は受信バッファを含み、その受信バッファは、送信遅延ジッタの補償や、パケットを送信順序から NAL ユニット復号順序への順序替えのために使われる。本節では、受信機の動作は送信遅延ジッタがないという仮定で記述されている。送信遅延ジッタの補償のためにも使われる実際の受信バッファと区別するために、本節ではこれ以降、受信バッファを逆インタリーブバッファと呼ぶことにする。受信機も送信遅延バッファを用意すべきである。すなわち、送信遅延ジッタバッファリングと逆インタリーブバッファリングのために独立したバッファを用意するか、受信バッファを送信遅延ジッタバッファリングと逆インタリーブバッファリング両方のために使うとかである。その上、受信機はバッファリング処理において送信遅延ジッタを考慮に入れるべきである。例えば、復号や再生の開始前に追加の初期バッファリングによってである。本節は次の事項をまとめたものである。

- 7.2.1 節は逆インタリーブバッファのサイズの計算方法を示す。
- 7.2.2 節は受信NALユニットをNALユニット復号順序に再編するという受信処理を示したものである。

### 7.2.1. 逆インタリーブバッファのサイズ

SDP オファー／アンサーモデルまたは他の能力交換手順をセッション設定で使用する時、受信ストリームの特性は受信能力を超えないようにすべきである。SDP オファー／アンサーモデルにおいて受信機は、MIME パラメータの `deint-buf-cap` を使って逆インタリーブバッファを割り当てる自身の能力を示すことができる。送信機は、MIME パラメータの `sprop-deint-buf-req` を使って逆インタリーブバッファサイズの要求を示す。それゆえ推奨は、逆インタリーブバッファサイズをバイト単位で、MIME パラメータの `sprop-deint-buf-req` の値以上に設定することである。MIME パラメータの `deint-buf-cap` や `sprop-deint-buf-req` に関する詳細情報については 8.1 節を参照し、SDP オファー／アンサーモデルにおけるそれらの使用については 8.2.2 節を参照のこと。

宣言型セッション記述 (declarative session description) をセッション設定で使用する時、MIME パラメータの `sprop-deint-buf-req` は逆インタリーブバッファサイズの要求を通知する。それゆえ推奨は、逆インタリーブバッファサイズをバイト単位で、MIME パラメータの `sprop-deint-buf-req` の値以上に設定することである。

### 7.2.2. 逆インタリーブ処理

受信機では、初期バッファリングと再生中バッファリングという 2 つのバッファリング状態が存在する。初期バッファリングは RTP セッションが初期化されたときに発生する。初期バッファリング後、復号や再生が開始され、再生中バッファリングモードが使用される。

バッファリング状態に関わらず、受信機は、受信順序で、次のようにして逆インタリーブバッファに、到達 NAL ユニットの格納する。集合パケットの NAL ユニットの、それぞれ逆インタリーブバッファに格納される。全ての NAL ユニットの対して、DON 値が計算され格納される。

受信処理は、次の関数や定数を使って以下のように記述される。

- 関数 `AbsDON` は 8.1 節で示される。
- 関数 `don_diff` は 5.5 節で示される。
- 定数 `N` は、オプションで MIME 種別パラメータの `sprop-interleaving-depth` の値 (8.1 節参照) を 1 でインクリメントした値である。

次の条件のうちの1つが成立するまで、初期バッファリングは動作し続ける。

- 逆インタリーブバッファにN個のVCL NALユニットがある。
- もしsprop-max-don-diffが存在する場合、don\_diff(m,n)はsprop-max-don-diffの値よりも大きい。nは受信NALユニットの間で最も大きいAbsDON値を持つNALユニットに対応する。mは受信NALユニットの間で最も小さいAbsDON値を持つNALユニットに対応する。
- 初期バッファリングが、オプションでMIMEパラメータのsprop-init-buf-timeの値に等しいか大きい値の期間で動作し続けた。

逆インタリーブバッファから取り出されるNALユニットは次のように決定される。

- もし逆インタリーブバッファに少なくともN個のVCL NALユニットを含む場合、NALユニットは逆インタリーブバッファから取り出され、バッファがN-1個のVCL NALユニットを含むまで以下に示される順序でデコーダに渡される。
- もしsprop-max-don-diffが存在する場合、don\_diff(m,n)がsprop-max-don-diffよりも大きいmの値を持つ全てのNALユニットは逆インタリーブバッファから取り出され、以下に示される順序でデコーダに渡される。ここでnは、受信NALユニットの間でAbsDONの最大値のNALユニットに対応する。

NALユニットをデコーダに渡す順序は次のように示される。

- RTPセッションの開始で0に初期化される変数をPDONとする。
- DON値と関連付けられるそれぞれのNALユニットについて、DON間隔は次のように計算される。NALユニットのDON値がPDON値よりも大きい場合、DON間隔はDON-PDONに等しい。言い換えれば、DON間隔は65535 - PDON + DON + 1に等しい。
- NALユニットは、DON間隔の昇順でデコーダに渡される。もし、いくつかのNALユニットが同一のDON間隔を共有する場合、どんな順序でもそれらをデコーダに渡すことができる。
- 所望の数のNALユニットがデコーダに渡された時、PDON値はデコーダに渡された最後のNALユニットのDON値に設定される。

### 7.3. 追加の逆パケット化ガイドライン

以下の追加の逆パケット化ルールをH.264の逆パケット化操作に使用してもよい。

- 高度なRTP受信器(たとえば、ゲートウェイのような)は、符号化スライスデータパーティションA(DPAs)のロスを認識する。もしDPAのロスが発見されると、ゲートウェイは、対応する符号化スライスデータパーティションBとCがH.264デコーダにとって意味のない情報になるので、それらを送信しないことを決定する。この方法で、MANEは、複雑なビットストリームを分析することなく不要なパケットを廃棄することによりネットワークの負荷を軽減させることができる。
- 高度なRTP受信器(たとえば、ゲートウェイのような)は、FUsのロスも認識してもよい。もしFUのロスが発見されると、ゲートウェイは、それに続く断片化された同じNALユニットのFUsはH.264デコーダにとって意味のないデータになるので、それらを送信しないことを決定する。この方法で、MANEは、複雑なビットストリームを分析することなく不要なパケットを廃棄することによりネットワークの負荷を軽減させることができる。
- 不要なパケットやNALUを廃棄できる高度な受信器は、まず、NALユニットタイプオクテットのNRIフィールドの値が0である全てのパケット/NALUを廃棄しなければならない。これはユーザの知覚上の影響を最小にし、参照ピクチャを完全に維持する。もし更に多くのパケットが廃棄されると、数値的に小さなNRIを持つパケットが数値的に大きなNRIの値を持つパケットより先に廃棄されるべきである。しかしながら、0よりも大きなNRIを持つパケットを廃棄することは、デコーダでドリフトを発生させる可能性が高く、回避されるべきである。

## 8. ペイロードフォーマットパラメータ

本節では、ペイロードフォーマットのオプションの特徴と、ビットストリームのいくつかの特徴を選択使用するためのパラメータについて記述する。パラメータはここでは、ITU-T H.264 | ISO/IEC 14496-10 コーデックのためにMIMEサブタイプ登録の一部として記述される。セッション記述プロトコル(SDP)[5]にパラメータをマッピングすることは、SDPを使用したアプリケーションのために提供される。同等のパラメータは、

MIME や SDP を使用しない制御プロトコルを使う別の箇所に定義することができる。  
数種のパラメータが、送信されるであろうストリームのプロパティを持って受信器に提供される。これらのパラメータの名称はストリームプロパティでは"sprop"で始まる。数種のこれら"sprop"パラメータは、他のペイロードやコーデック構成パラメータにより制限される。たとえば、sprop-parameter-sets パラメータは profile-level-id パラメータにより制限される。メディア送信器は受信器よりも、全ての"sprop"パラメータを選択する。"sprop"パラメータのこの珍しい特性は数種のシグナリングプロトコル概念と互換性がないかもしれない。その場合、これらのパラメータの使用は避けなければならない。

## 8.1. MIME 登録

ITU-T H.264 | ISO/IEC 14496-10 コーデックのための MIME サブタイプは、IETF ツリーから確保されている。受信器は記述されていないパラメータを無視せねばならない。

メディアタイプ名: video

メディアサブタイプ名: H264

要求されるパラメータ: なし

オプションパラメータ:

profile-level-id:

[1] で記述されたシーケンスパラメータセット NAL ユニットの中の以下の 3 バイトの base16[6] (16 進数) 表現

- 1) profile-idc
- 2) profile-iop として参照される 1 バイト。

その内訳は、

```
constraint_set0_flag,  
constraint_set1_flag,  
constraint_set2_flag,  
reserved_zero_5bits ビット  
(MSB First)
```

- 3) level-idc

注: [1] では、reserved\_zero\_5bits は 0 である必要がある。しかし、その将来的な値は ITU-T または ISO/IEC で決定される。

もし、profile-level-id パラメータが NAL ユニットストリームのプロパティを示すために使用されるならば、それはデコーダがストリームを復号するとき、[1] に従うためにサポートしなければならないプロファイルとレベルが示されている。profile-iop バイトは NAL ユニットストリームが、以下に示すプロファイルのすべての制約にも従うかどうかを示している。もし profile-iop のビット 7 (最上位ビット)、ビット 6、またはビット 5 が 1 に等しければ、その NAL ユニットストリームにおいて、Baseline プロファイル、Main プロファイル、拡張プロファイルのすべての制約に従う。

もし profile-level-id パラメータが能力交換または、セッション確立手順に使用されるなら、コーデックがサポートし、通知されたプロファイルのためにサポートされる最高位のプロファイルを示す。profile-iop バイトはコーデックが、profile-iop バイトと共に通知され、アルゴリ

ズムの特徴の通常のサブセットおよびコーデックがサポートする profile\_idc によって示されるプロファイルの制約によってのみ追加の制約があるかどうかを示す。たとえば、もしコーデックが Baseline プロファイルとレベル 2.1 およびそれ以下での Main プロファイル符号化ツールの通常のサブセットのみをサポートするならば、profile-level-id は 42E015 になる。ここで 42 は Baseline プロファイルを表し、E0 は、すべてのプロファイルのための通常のサブセットのみがサポートされることを表し、15 はレベル 2.1 を表す。

注：能力交換およびセッション確立手順は、サポートされたコーデックのプロファイルのための能力を提示する方法を提供するべきである。たとえば、SDP オファー/アンサーモデルの 1 対 N のコーデック選択は(10.2 節の[7])を使用することができる。

もし profile-level-id が提供されないのであれば、Level1 の追加制約のない Baseline プロファイルが暗示されなければならない。

max-mbps, max-fs, max-cpb, max-dpb, and max-br:

これらのパラメータは受信器実装の能力を通知するのに使用してもよい。これらのパラメータは、いかなる他の目的にも使用してはならない。profile-level-id パラメータは、これらのパラメータのいくつかを含む同じ受信器能力記述の中に存在していなければならない。profile-level-id パラメータの値で伝達されたレベルは受信器がサポートする全ての能力でなければならない。

max-mbps, max-fs, max-cpb, max-dpb と max-br は、以下で記述されるように、通知されるレベルの要求能力を拡大する受信器の能力を示すのに使用してもよい。

(max-mbps, max-fs, max-cpb, max-dpb, max-br)セットから二つ以上のパラメータが存在するとき、受信器はすべての通知された能力を同時にサポートしなければならない。たとえば、もし max-mbps と max-br が存在するなら、フレームレートとビットレート両方の拡張を有する通知レベルがサポートされる。即ち、受信器は、次のような NAL ユニットストリームの全てを復号できる。マクロブロック処理レート  $\leq$  max-mbps、ビットレート  $\leq$  max-br、符号化ピクチャバッファサイズは以下に示す max-br パラメータのセマンティックス内から導かれ、他のパラメータは profile-level-id パラメータの値で指定されたレベルに従う。

もし受信器がレベル A の特性すべてをサポートできるとしても、受信器は、profile-level-id パラメータの値の中で記述されたレベルと比較して、ここでレベル A と記載された、より高い要求に合致するような、max-mbps, max-fs, max-cpb, max-dpb と max-br の値を通知してはならない。

注：オプションの MIME タイプパラメータが 1 つの NAL ユニットストリームのプロパティを通知するとき、max-mbps, max-fs, max-cpb, max-dpb および max-br は存在しない。そして profile-level-id の値は NAL ユニットストリームが常に指定されたプロファイルとレベルに完全に従うようなものでなければならない。

max-mbps:

max-mbps の値は整数で、1 秒間あたりのマクロブロック数の単位で表した最大マクロブロック処理レートを示している。max-mbps パラメータは受信器が profile-level-id パラメータの値で伝達される通知レベルによって要求されるよりも、より高いビデオ復号能力を通知する。max-mbps

が通知されるとき、受信器は、通知されたレベルの[1]の表 A-1 にある MaxMBPS 値が max-mbps の値で置き換えられる例外を除き、通知されたレベルに適合する NAL ユニットストリームを復号できなければならない。max-mbps の値は[1]の表 A-1 で与えられるレベルの MaxMBPS の値よりも大きいか等しくなければならない。送信器は、通知されたレベルで示されているよりも速いピクチャレートで得られたサイズのピクチャを送るためにこの知識を使用してもよい。

max-fs:

max-fs の値はマクロブロックユニット数が単位の最大フレームサイズを整数で示している。max-fs パラメータは受信器が profile-level-id パラメータ値内で伝達された通知レベルによって要求されるサイズより大きいサイズの画像を復号できる能力があることを通知する。

max-fs が通知されたとき、その通知レベルのための[1]の表 A-1 中の MaxFS 値を max-fs の値に置き換えることを除いて、受信器はその通知レベルに従う NAL ユニットストリームを復号することができなければならない。max-fs の値は[1]の表 A-1 中で与えられたレベルのため MaxFS の値よりも大きいか等しくなければならない。送信器は、その通知レベルで示されているよりも、比例的に低いフレームレートでより大きいピクチャを送信するためにこの知識を使ってもよい。

max-cpb:

max-cpb の値は VCL HRD パラメータ ([1]の A. 3. 1 の i 項参照)の 1000 ビットの単位での、および NAL HRD パラメータ ([1]の A. 3. 1 の j 項参照)の 1200 ビット単位での最大符号化ピクチャバッファサイズを示す整数である。max-cpb パラメータは、受信器が profile-level-id パラメータ値で伝達され通知レベルによって要求される符号化ピクチャバッファメモリの最小量より多くのメモリを持っていることを通知する。max-cpb が通知されるとき、受信器は通知レベルの[1]の表 A-1 中の MaxCPB を max-cpb の値に置き換えることを除いて、通知レベルに適合する NAL ユニットストリームを復号することができなければならない。max-cpb の値は[1]の表 A-1 で与えられたレベルの MaxCPB の値よりも大きいか等しくなければならない。送信器は、この[1]の表 A-1 中の MaxCPB によって達成されることができるとも、より大きなビットレート変動を有する符号化ビデオストリームを作り出すためにこの知識を使用してもよい。

注: 符号化ピクチャバッファは H. 264 の仮想参照デコーダで使用されている。仮想参照デコーダの使用は、その生成ビットストリームがその標準に適合しているかを検証するよう、および出力ビットレートを制御するよう H. 264 エンコーダにおいて推奨されている。したがって、符号化ピクチャバッファは、逆インタリーブバッファとジッタ吸収バッファを含む受信器における他の潜在的バッファとは概念的に独立している。符号化ピクチャバッファは H. 264 の Annex C で記述されているようにデコーダで実現される必要はない、しかし、標準に準拠しているデコーダはそれらが標準に準拠したビットストリームを復号できる限りどのようなバッファ構造を持つこともできる。したがって、実際には、ビデオデコーダの入力バッファは、受信器の逆インタリーブバッファおよびジッタ吸収バッファと統合させることができる。

max-dpb:

max-dpb 値は、最大復号ピクチャサイズを 1024 バイトを 1 単位として示す整数値である。max-dpb パラメータは、profile-level-id パラメータ値にて伝送された通知レベルで必要とされる復



号ピクチャバッファメモリの最小量よりも多くのメモリを受信端末が保持していることを示すものである。max-dpb が通知された場合、受信端末は通知レベルに従った NAL ユニットストリームを復号可能でなければならない。ただし、文献[1]の Table A-1 にある通知レベル用の MaxDPB 値が max-dpb 値に置き換えられる場合を除く。結果として、max-dpb を通知する受信端末は、以下の数の分だけの復号フレーム、相互補完するフィールドペア、あるいはペアとなっていない複数フィールドを、復号ピクチャバッファで蓄積できなければならない：

$$\text{Min}(1024 * \text{max-dpb} / (\text{PicWidthInMbs} * \text{FrameHeightInMbs} * 256 * \text{ChromaFormatFactor}), 16)$$

PicWidthInMbs、FrameHeightInMbs、ChromaFormatFactor は、[1]で定義される。

max-dpb 値は、[1]の Table A-1 にあるレベル用の MaxDPB 値以上でなければならない。送信端末は、圧縮率を向上した符号化ビデオストリームを構築する際に、この知識を用いる可能性がある。

注：本パラメータは、シグナリングゲートウェイの設計を容易にする目的で、主に ITU-T 勧告 H.245 における類似のコードポイントを補完するために追加された。復号ピクチャバッファは、再構築されたサンプルを蓄積するものであり、ビデオデコーダのみに固有のものである。復号ピクチャバッファと、RTP に用いられるバッファ、特に逆インターリーブバッファやジッタ吸収用バッファとの間には、何の関連性も存在しない。

max-br:

max-br 値は、ビデオの最大ビットレートを示す整数で、VCL HRD パラメータ ([1]の A.3.1 item i 参照)においては 1000bps が 1 単位となり、NAL HRD パラメータ ([1]の A.3.1 item j 参照)においては 1200bps が 1 単位となる。

max-br パラメータは、受信端末のビデオデコーダが、profile-level-id パラメータ値にて伝送された通知レベルで必要とされるものより高いビットレートで復号可能であることを通知するものである。max-br 値は、[1]の Table A-1 で与えられるレベルに対する MaxBR 値以上でなければならない。

max-br が通知された場合、受信端末のビデオコーデックは、profile-level-id パラメータで伝送された通知レベルに従った NAL ユニットストリームを復号可能でなければならない。ただし、レベルによって特定される制限による以下の例外の場合を除く。

- max-br 値で、([1]の Table A-1 で)通知されたレベルの MaxBR 値を置き換える。
- max-cpb パラメータが存在しない場合、以下の式で[1]の Table A-1 の MaxCPB 値を置き換える。  
(通知されたレベルの MaxCPB) × max-br / (通知されたレベルの MaxBR)

例えば、受信端末がレベル 1.2 に対する許容量を max-br=1550 で通知した場合、これは VCL HRD

パラメータにおいては最大 1550kbps、NAL HRD パラメータにおいては最大 1860kbps、さらに CPB サイズが 4036458 ビット (1550000 / 384000 \* 1000 \* 1000)であることを示している。

max-br 値は、[1]の Table A-1 で与えられる通知レベルの MaxBR 値以上でなければならない。

送信端末は、ビデオ品質を改善する目的で、H.264 の Annex A のレベル定義で許可されたよりも高いビットレートのビデオを送信する際に、この知識を用いる可能性がある。

注：本パラメータは、シグナリングゲートウェイの設計を容易にする目的で、主に ITU-T 勧告 H.245 における類似のコードポイントを補完するために追加された。本パラメータ値は、ネットワークがあらゆる時間においてそのビットレートでの伝送処理が可能である、ということを前提にしたものではない。特に、混雑制御の制約のもとでは、通知されたビットレートでの伝送が可能である、という結論にはならない。

redundant-pic-cap:

本パラメータは、受信端末の実装能力を通知する。値が 0 の場合、受信端末が正しく復号されなかった主符号化ピクチャを修正するのに冗長符号化ピクチャを利用するつもりがないことを、本パラメータは示している。値が 0 の場合は、受信端末は冗長スライスを用いることができないため、送信端末は帯域の節約のため冗長スライスの送信は避けるべきである。値が 1 の場合、受信端末は主符号化ピクチャの(少なくとも部分的に)破壊された領域をカバーする任意の冗長スライスを復号することができる。そのため、送信端末は冗長スライスの送信を行っても良い。本パラメータが存在しない場合は、redundant-pic-cap 値としては 0 を用いなければならない。存在する場合は、redundant-pic-cap 値は 0 か 1 のどちらかでなければならない。

profile-level-id パラメータが redundant-pic-cap パラメータと同一の(capability signaling)に存在し、かつ profile-level-id で示されるプロファイルが冗長符号化ピクチャの使用を禁止している(例：メインプロファイル)といったような場合、redundant-pic-cap は 0 でなければならない。受信端末が redundant-pic-cap=0 を示す場合、受信されたストリームは冗長符号化ピクチャを含んではならない。

注：redundant-pic-cap が 0 であったとしても、デコーダは、仮に冗長符号化ピクチャが許可されているようなプロファイル(ベースライン、拡張)をサポートするものとした場合でも、冗長符号化ピクチャを無視することは可能である。

注：redundant-pic-cap が 1 であったとしても、受信端末は、冗長スライスによる復号の置換・補完を行うのに別のエラー隠蔽手法を選択しても良い。

sprop-parameter-sets:

このパラメータは、あらゆるシーケンスやピクチャパラメータセット NAL ユニットの伝達するために使用しても良い。(ここで、初期パラメータセット NAL ユニットの参照される事を考慮すると)それは、復号順序において他の NAL ユニットの先行していなければならない。そのパラメータは、あらゆる能力交換手続きにおいてコーデック能力を示すために使用してはな

らない。

パラメータ値は、[1] の 7.3.2.1 節、及び、7.3.2.2 節において規定された初期パラメータセット NAL ユニットの base64 [6] 表現である。

パラメータセットは、復号順序に従って伝達され、そして、パラメータセット NAL ユニットのフレーミングは、起こらない。

コンマは、リストにおいて連続する 2 つのパラメータセットを分割するために使用する。

パラメータセット NAL ユニットにおけるバイト数は、典型的には 10 バイトより小さいが、ピクチャパラメータセット NAL ユニットは、数 100 バイトを含むことが有り得る事に注意すること。

注：複数のペイロードタイプがそれぞれの sprop-parameter-sets パラメータを伴って、SDP オフラー/アンサーモデルにおいて提供される場合、受信器は、それらのパラメータセットが矛盾するストレージロケーションを使用しない（すなわち、パラメータセット識別子が同値である）とは決め付けられない。

従って、受信器は、全ての sprop-parameter-sets を二重にバッファリングし、それらを特定のペイロードタイプを復号するデコーダインスタンスにて利用可能にするべきである。

#### parameter-add:

このパラメータは、このパラメータの受信器が sprop-parameter-sets MIME パラメータを使うシグナリング応答においてパラメータセットを加えることを許可されるか否かを伝えるために使用しても良い。

このパラメータの値は、0 か 1 のいずれかである。

0 は、偽である。すなわち、パラメータセットを加えることが許可されない。

1 は、真である。すなわち、パラメータセットを加えることが許可されている。

もし、パラメータが存在しない場合は、その値は、1 でなければならない。

#### packetization-mode:

このパラメータは、RTP ペイロードタイプのプロパティ、または、受信器実装の能力を伝える。

1 つのコンフィグレーションのみが、示され得る。従って、1 つ以上の packetization-mode をサポートする能力が宣言された場合、複数のコンフィグレーション(RTP ペイロードタイプ)を、使用しなければならない。

packetization-mode の値が 0 に等しい場合、もしくは、packetization-mode が存在しない場合は、本文書の 6.2 節にて定義された シングル NAL モードを使用しなければならない。

このモードは、ITU-T 勧告 H.241 [15] (12.1 節を参照)を用いる標準において使用する。

packetization-mode の値が 1 に等しい場合、本文書の 6.3 節において定義された非インタリーブモードを使用しなければならない。

packetization-mode の値が 2 に等しい場合、本文書の 6.4 節において定義されたインタリーブモードを使用しなければならない。

packetization-mode の値は、0~2 の範囲に含まれる整数でなければならない。

#### sprop-interleaving-depth:

packetization-mode が存在しない場合、または、packetization-mode の値が 0 か 1 に等しい場

合、このパラメータは、存在してはならない。

packetization-mode の値が 2 に等しい場合、このパラメータは、存在しなくてはならない。

このパラメータは、NAL ユニットストリームのプロパティを伝える。

それは、NAL ユニットストリーム中の任意の VCL NAL ユニットに対し、伝送順序では先行し、復号順序では後続する VCL NAL ユニットの最大数を指定する。

従って、NAL ユニット復号化順序リカバリのためのバッファサイズが VCL NAL ユニットで数えて、少なくとも (sprop-interleaving-depth + 1) の値である場合、受信器は、NAL ユニット復号順序を再構成できる事を保証する。

sprop-interleaving-depth の値は、0~32767 の範囲に含まれる整数でなければならない。

sprop-deint-buf-req:

packetization-mode が存在しない場合、または、packetization-mode の値が 0 か 1 に等しい場合、このパラメータは、存在してはならない。

packetization-mode の値が 2 に等しい場合、このパラメータは、存在しなくてはならない。

sprop-deint-buf-req は、NAL ユニットストリームのための逆インタリーブバッファの要求サイズを伝える。

パラメータの値は、本文書の 7.2 節において規定されたような逆インタリーブバッファに要求された最大バッファ使用量(バイト単位)以上でなければならない。

逆インタリーブバッファサイズが、バイト単位で、少なくとも sprop-deint-buf-req の値である場合、受信器は、インタリーブされた NAL ユニットの NAL ユニット復号順序に逆インタリーブできる事を保証する。

sprop-deint-buf-req の値は、0~4294967295 の範囲に含まれる整数でなくてはならない。

注：sprop-deint-buf-req は、逆インタリーブバッファのみの要求サイズを示している。

ネットワークジッタが発生する可能性がある場合、適切なサイズのジッタバッファが同様に用意されるべきである。

deint-buf-cap:

このパラメータは、受信器実装の能力を伝え、その受信器が NAL ユニット復号順序を再構成する事を可能とする逆インタリーブバッファ空間の総量をバイト単位で示す。

受信器は、sprop-deint-buf-req パラメータの値が、このパラメータの値以下のあらゆるストリームを扱う事が可能である。

このパラメータが存在しない場合、0 の値を deint-buf-cap として使用しなければならない。

deint-buf-cap の値は、0~4294967295 の範囲に含まれる整数でなければならない。

注：deint-buf-cap は、受信器の逆インタリーブバッファのみの最大可能サイズを示している。

ネットワークジッタが発生する可能性がある場合、適切なサイズのジッタバッファが同様に用意されるべきである。

sprop-init-buf-time:

このパラメータは、NAL ユニットストリームのプロパティを伝えるために使用しても良い。  
packetization-mode の値が 0 か 1 に等しい場合、このパラメータは、存在してはならない。

このパラメータは、受信器が伝送順序から NAL ユニット復号順序を回復するためにバッファリングする初期バッファリング時間を伝える。

このパラメータは、信頼性のある瞬時の伝送であり、伝送と復号が同じ時系列で、最初のパケットが到着した時に復号を開始すると仮定すると、( ある NAL ユニットの伝送時間 - その NAL ユニットの復号時間 ) の最大値である。

sprop-init-buf-time の値の規定例は以下の通りである。

NAL ユニットストリームは、以下のようなインタリーブされた順序で送信されており、値は復号時間を表し、伝送順序は、左から右である。

0 2 1 3 5 4 6 8 7 ...

NAL ユニットの安定した伝送速度を仮定すると、伝送時間は、

0 1 2 3 4 5 6 7 8 ...

である。

伝送時間から復号時間を列方向で引き算すると、以下の一連の結果となる。

0 -1 1 0 -1 1 0 -1 1 ...

従って、NAL ユニット伝送時間の間隔については、この例では、sprop-init-buf-time の値は、1 である。

このパラメータは、90kHz のクロックチックで、非負の 10 進の整数表現として符号化される。

パラメータが存在しない場合、最初のバッファリング時間値は、定義されない。

そうでなければ、sprop-init-buf-time の値は、0~4294967295 の範囲に含まれる整数である。

伝えられた sprop-init-buf-time に加えて、受信器は、ミキサ、トランスレータ、ゲートウェイ、プロキシ、トラフィックシェーパ、その他のネットワーク要素に起因する遅延ジッタのためのバッファリングを含めた伝送遅延ジッタバッファリングを考慮すべきである。

sprop-max-don-diff:

このパラメータは、NAL ユニットストリームのプロパティを伝えるために使用しても良い。  
このパラメータは、送信器、または、受信器、または、コーデックの能力を伝えるために使用してはならない。

packetization-mode の値が 0 か 1 に等しい場合、このパラメータは、存在してはならない。

sprop-max-don-diff は、0~32767 の範囲に含まれる整数である。

sprop-max-don-diff が存在しない場合、パラメータの値は、規定しない。

sprop-max-don-diff は、以下の通りに計算される。

$j > i$  となる全ての  $i$  について

$$\text{sprop-max-don-diff} = \max \{ \text{AbsDON}(i) - \text{AbsDON}(j) \}$$

ここで、 $i$ 、及び、 $j$  は、伝送順序における NAL ユニットのインデックスを示しており、AbsDON は、65535 の後に 0 にラップアラウンドしない NAL ユニットの復号順序番号を示している。

すなわち、AbsDON は、以下のように計算される。 $m$ 、及び、 $n$  は、伝送順序における連続した NAL ユニットとする。

伝送順序において最初の NAL ユニット(インデックスが 0)では、

$$\text{AbsDON}(0) = \text{DON}(0)$$

である。

他の NAL ユニットでは、AbsDON は、以下の通りに計算される。

If  $\text{DON}(m) == \text{DON}(n)$ ,

$$\text{AbsDON}(n) = \text{AbsDON}(m)$$

If  $(\text{DON}(m) < \text{DON}(n) \text{ and } \text{DON}(n) - \text{DON}(m) < 32768)$ ,

$$\text{AbsDON}(n) = \text{AbsDON}(m) + \text{DON}(n) - \text{DON}(m)$$

If  $(\text{DON}(m) > \text{DON}(n) \text{ and } \text{DON}(m) - \text{DON}(n) \geq 32768)$ ,

$$\text{AbsDON}(n) = \text{AbsDON}(m) + 65536 - \text{DON}(m) + \text{DON}(n)$$

If  $(\text{DON}(m) < \text{DON}(n) \text{ and } \text{DON}(n) - \text{DON}(m) \geq 32768)$ ,

$$\text{AbsDON}(n) = \text{AbsDON}(m) - (\text{DON}(m) + 65536 - \text{DON}(n))$$

If  $(\text{DON}(m) > \text{DON}(n) \text{ and } \text{DON}(m) - \text{DON}(n) < 32768)$ ,

$$\text{AbsDON}(n) = \text{AbsDON}(m) - (\text{DON}(m) - \text{DON}(n))$$

ここで、DON( $i$ )は、伝送順序におけるインデックス  $i$  を持つ NAL ユニットの復号順序番号である。  
復号順序番号は、本文書の 5.5 節において規定される。

注：受信器は、受信器のバッファにおいてどの NAL ユニットのデコーダに渡して良いかを決定するために sprop-max-don-diff を使用しても良い。

#### max-rcmd-nalu-size:

このパラメータは、受信器の能力を通知するために利用してもよい。

このパラメータはほかの目的のために利用してはならない。このパラメータの値は、受信器が効率よく処理することができる最大の NALU バイトサイズを示す。

このパラメータ値は推奨であり、厳格な上限値ではない。送信器は、より大きな NALU をつくってもよいが、この操作は制限値を守った NALU より高いコストを要することに注意しなければならない。

max-rcmd-nalu-size の値は、0 以上 4292967295 以下の整数でなければならない。もし、このパラメータが明記されなかったら、NALU サイズの制限が存在しない。

送信器は、送信器と受信器間で利用できる MTU サイズを考慮しなければならない、そしてこの目的のために、MTU 検出 (MTU discovery) を実行しなければならない。

このパラメータは、例えば H. 223 の送信データユニットよりも小さい NALU のほうが効率的となる IP-H. 223 テレビ電話ゲートウェイに動機づけられている。

ゲートウェイは IP を終端してもよい、従って MTU 検出 (MTU discovery) は通常ゲートウェイの向こう側には及ばない。

注：このパラメータの必要な値以下の設定は、悪い影響を与える可能性がある。

#### 符号化の考察:

この型は RTP (RFC3550) によって送信するためのみの定義である。

H. 264/AVC ビデオのファイルフォーマットは[29]に定義されている。この定義は、3GPP マルチメディアファイルフォーマット (MIME タイプ video/3gpp) [30]や MP4 ファイルフォーマット (MIME タイプ video/mp4) のようなそのほかのファイルフォーマットで利用される。

#### セキュリティの考察:

本文書の 9 節を見よ。

#### 公開仕様:

本文書と 15 節で示された文書を参照せよ。

#### 付加情報:

なし。

ファイル拡張子:

なし。

マッキントッシュファイルタイプコード:

なし。

オブジェクト識別子またはOID:

なし

詳しい情報を得るための個人やemailアドレス:

stewe@stewe.org

所期の使用法:

COMMON

作者:

stewe@stewe.org

変更管理人:

IESG から委託された IETF ABT ワーキンググループ

## 8.2 SDP パラメータ

### 8.2.1 MIME パラメータの SDP へのマッピング

MIME メディアタイプの video/H.264 文字列はセッション記述プロトコル (SDP) [5]のフィールドで以下のようにマッピングされる。

SDP の「m=」行のメディア名は video でなければならない。

SDP の「a=rtpmap」行の符号化名は H264 (MIME サブタイプ) でなければならない。

「a=rtpmap」行のクロックレートは 90000 でなければならない。

オプションのパラメータである、「profile-level-id」、「max-mbps」、「max-fs」、「max-cpd」、「max-dpb」、「max-br」、「redundant-pic-cap」、「sprop-parameter-sets」、「parameter-add」、「packetization-mode」、「sprop-interleaving-depth」、「deint-buf-cap」、「sprop-deint-buf-req」、「sprop-init-buf-time」、「sprop-max-don-diff」、「max-rcmd-nalu-size」が存在するときは、SDP の「a=fmtp」行に含むべきである。これらのパラメータは、MIME メディアタイプ文字列として、parameter=value 対のセミコロンで区切られたリストの形で表現される。



SDP のメディア表現のサンプルが以下に示されている。(ベースラインプロファイル、レベル 3. 0、メインプロファイルのいくつかの制約は従わなくともよい。)

```
m=video 49170 RTP/AVP 98
a=rtpmap:98 H264/90000
a=fmtp:98 profile-level-id=42A01E;
    sprop-parameter-sets=Z0IACpZTBYmI,aMljiA==
```

## 8.2.2 SDP オファー/アンサーモデルの使用

ユニキャストのネゴシエーションに、SDP のオファー/アンサーモデル[7]を使用して RTP 上で H.264 を提供しようとしたとき、以下の制限とルールが適用される。

H.264 のメディアフォーマットコンフィグレーションを特定するパラメータは「profile-level-id」、  
「packetization-mode」、  
「packetization-mode」で必要ならば「sprop-deint-buf-req」である。これらの 3 つのパラメータは対称的に使われる。すなわちアンサー側は、もし、1 つ以上のパラメータ値がサポートされていないときには、すべてのコンフィグレーションパラメータを維持するか、メディアフォーマット (ペイロードタイプ) を完全に取り除くかのどちらかである。

注：対称性パラメータ使用の要求は、上記の 3 つのパラメータのみで、ほかのストリームプロパティや能力パラメータには適用しない。

簡単にこれらのコンフィグレーションを処理し、マッピングするために、[7]で明記されているように、オファー側で使用されたと同じ RTP ペイロードタイプ番号は、アンサー側でも使われるべきである。アンサー側はコンフィグレーション (「profile-level-id」、  
「packetization-mode」、もし存在すれば「sprop-deint-buf-req」) がオファー側と同じでない限り、オファー側で利用していたペイロードタイプ番号を含めてはいけない。

注：オファー側がアンサーを受信したときは、議論されているコンフィグレーションが新しいものか、すでに提供されたコンフィグレーションと同じものかどうかを決定するために、メディアタイプ (すなわち video/h264) と上記の 3 つのパラメータを基礎としたオファー側で宣言していないペイロードタイプと、すでに宣言されたペイロードタイプとを比較しなければならない。

「sprop-parameters-sets」、  
「sprop-deint-buf-req」、  
「sprop-interleaving-depth」、  
「sprop-max-don-diff」、  
「sprop-init-buf-time」のパラメータは、オファー側またはアンサー側が、このメディアフォーマットコンフィグレーションで送信する NAL ユニットストリームのプロパティを記述する。これの、オファー/アンサーパラメータの通常の使用との違いは、通常、ストリームのプロパティを宣言するようなパラメータが、オファー側やアンサー側は受信可能であることを示しているということである。H.264 の場合には、オファー側は、アンサー側が提供されたコンフィグレーションを用いて符号化されたメディアを受信することができると想定されている。

注：上記のパラメータは、同じコンフィグレーションをもつと宣言するエンティティによって送信されるい

ずれのストリームにも適用される。すなわちストリーム元に依存する。ペイロードタイプで縛られると言うよりはむしろ、それらの値が送られてきた場合、別のペイロードタイプに適用せざるを得ない。何故なら、コンフィグレーションを示しているからである。

能力パラメータ（「max-mbps」、「max-fs」、「max-cpb」、「max-dpb」、「max-br」、「redundant-pic-cap」、「max-rcmd-nalu-size」）は、拡張能力を宣言するのに利用してもよい。それらの解釈は、方向属性による。方向属性が `sendonly` の時は、パラメータは送信器が生成しうる RTP パケットと NAL ユニットストリームの限界を記述している。方向属性が `sendrecv` や `recvonly` の時には、パラメータは受信器が受け付ける限界を記述している。

上記で定めるように、オファー側は、インタリーブ H.264 ストリームのオファーに、逆インタリーブバッファのサイズを含めなければならない。オファー側とアンサー側が逆インタリーブバッファリングの能力をお互いに通知できるようにするために、両方に「`deint-buf-cap`」を含めることを勧める。「`sprop-deint-buf-cap`」が 2 度目のオファーとアンサーで選択されるときに、この情報が利用される。インタリーブのストリームに対しては、受信器の能力が不明の時、異なるバッファリングの要求を含めた複数のペイロードタイプの提案を考慮することを同様に勧める。

「`sprop-parameter-sets`」パラメータは上記に記述されたように使用される。さらに、アンサー側は、オファーで受け取ったすべてのパラメータセットをアンサーに含めなければならない。「`parameter-add`」パラメータの値によって、異なったルールを適用する。もし「`parameter-add`」が `false` (0) ならばアンサーは付加パラメータセットを加えてはいけぬ。もし「`parameter-add`」が `true` (1) ならば、アンサー側のアンサーは、「`sprop-parameter-sets`」パラメータを付加パラメータとして付加してもよい。アンサー側は「`parameter-add`」の値と無関係に、アンサーで宣言された「`sprop-parameter-sets`」を利用したビデオストリームを受信することも同様に受け入れなければならない。

注：パラメータセットを加えるとき、パラメータセット識別子の衝突によって、すでに伝送されていたパラメータセットの上書きを行わないように、注意しなければならない。

さらに、マルチキャストでストリームを送信した場合、以下のルールが適用される。

ストリームプロパティパラメータ（「`sprop-parameter-sets`」、「`sprop-deint-buf-req`」、「`sprop-interlaving-depth`」、「`sprop-max-don-diff`」、「`sprop-init-buf-time`」）は、アンサー側で変更してはいけぬ。このようにペイロードタイプは変更なしで受け入れられるか、削除かのどちらかである。

受信器能力パラメータ「`max-mbps`」、「`max-fs`」、「`max-cpb`」、「`max-dpb`」、「`max-br`」、「`max-rcmd-nalu-size`」は、アンサー側で `sendrecv` または `recvonly` として宣言されるすべてのストリームをサポートしなければならない、さもなければ以下いずれかの動作をしなければならない。

- メディアフォーマットの削除
- セッション拒絶

受信機能力パラメータ「`redundant-pic-cap`」は、アンサー側で以下のような `sendrecv` または `recvonly` として

宣言されたすべてのストリームに対してサポートすべきである。アンサー側は、もしオファー側が「redundant-pic-cap」が 0 を示すなら、送信されたストリームに冗長なコード化されたピクチャを含めてはいけない。さもなくば（「redundant-pic-cap」が 1 の時）、どのようにしてアンサー側が冗長化コードピクチャを利用すべきかの推奨は、この文書の範囲外である。

以下は、オファーやアンサーと方向属性の異なった組み合わせにおいて、異なったパラメータをどのように解釈しなければならないかという完全なリストである。

- 「a=sendrecv」を使用するか、もしくは方向属性を使用しないオファーとアンサーにおいて、または「a=recvonly」を使用するオファーとアンサーにおいて、次のパラメータの解釈を使用しなければならない。

受信機の実際のコンフィギュレーションやプロパティを宣言するもの：

- profile-level-id
- packetization-mode

送信ストリームの実際のプロパティを宣言するもの（「a=sendrecv」を使用するか方向属性を使用しない時のみ適用可能）：

- sprop-deint-buf-req
- sprop-interleaving-depth
- sprop-parameter-sets
- sprop-max-don-diff
- sprop-init-buf-time

受信機の実装能力を宣言するもの：

- max-mbps
- max-fs
- max-cpb
- max-dpb
- max-br
- redundant-pic-cap
- deint-buf-cap
- max-rcmd-nalu-size

どのようにしてオファー／アンサー能力交換を行わなければならないかを宣言するもの：

- parameter-add

- メディアストリームについて方向属性「a=sendonly」を含むオファーまたはアンサーにおいて、次のパラメータの解釈を使用しなければならない。

送信しようとするストリームの実際のコンフィギュレーションやプロパティを宣言するもの：

- profile-level-id
- packetization-mode
- sprop-deint-buf-req
- sprop-max-don-diff
- sprop-init-buf-time
- sprop-parameter-sets
- sprop-interleaving-depth

ストリームを受信する時、送信機の能力を宣言するもの：

- max-mbps
- max-fs
- max-cpb
- max-dpb
- max-br
- redundant-pic-cap
- deint-buf-cap
- max-rcmd-nalu-size

どのようにしてオファー／アンサー能力交換が行わなければならないかを宣言するもの：

- parameter-add

その上、次の考慮が必要である。

- 受信機能力の宣言のために使用されるパラメータは一般的に「ダウングレードダブル」である。すなわち、パラメータは送信機の可能な振る舞いの上限を表している。このように送信機は、これらのパラメータ値以下の値を使用して、エンコーダの設定を選択しても良い。パラメータセット内で伝達される値の制限が、陽には示されないが使用されるプロファイルやレベルから明らかであるなら、「sprop-parameter-sets」を送信機の能力の宣言で使用してはいけない。
- コンフィギュレーションを宣言するパラメータは、「profile-level-id」パラメータのレベルの部分を除いて、「ダウングレードダブル」ではない。これは、受信機が使用しようとし、送信機側においてそのまま使用しなければならない値を表現する。
- 送信機の能力を宣言し、「非ダウングレードダブル」パラメータをこの宣言で使用する時、これらのパラメータは受け付け可能であるコンフィギュレーションを表現する。高い互換性レベルを達成するために、複数の選択可能なコンフィギュレーションをオファーするのが得策である。例えば、パケット化モードについてである。シングルペイロードタイプにおいて複数のコンフィギュレーションをオファーするこ

とは不可能である。したがって、複数のコンフィギュレーションがオファーされる時、それぞれのオファーは、それと付随する RTP ペイロードタイプを要求する。

- たとえ受信機が、あるペイロードフォーマットが持つ機能のサブセットをサポートするだけでも、すべての MIME パラメータを理解すべきである。これが保証するのは、メディアを受信するオファーが、いつ、オファーの受信機によってサポートされる能力へダウングレードできるかを、受信機が理解可能であるということである。
- アンサー側は、追加のメディアフォーマットコンフィギュレーションでオファーを拡張しても良い。しかしながら、それらの使用を可能にするために、多くの場合において2番目のオファーが、メディア送信機が使用するであろうストリームプロパティパラメータを提供するために、オファー側から要求される。これはまた、オファー側がこのメディアフォーマットコンフィギュレーションを、単に送信するのではなく受信可能でなければならないという効果を有する。
- もしオファー側が送信と受信の間で非対称な能力を持つとするなら、オファー側は異なった RTP セッションを提供しなければならない。つまり、「recvonly」と「sendonly」として宣言された異なったメディアの行である。これはシステム上に更なる影響を及ぼす場合がある。

### 8.2.3. 宣言型セッション記述 (Declarative Session Descriptions) における使用方法

RTP 上の H.264 が RTSP [27]や SAP [28]のような宣言型スタイルにおいて SDP で提供される時、以下の考慮が必要である。

- NAL ユニットストリームと受信機の両方のプロパティを示すことのできるすべてのパラメータは、NAL ユニットストリームのプロパティを示すために使用される。例えばこの場合、パラメータ「profile-level-id」は、送信機の能力ではなくストリームで使用される値を宣言する。これは、次のパラメータの解釈を使用しなければならないという結果になる。

実際のコンフィギュレーションまたはプロパティを宣言するもの：

- profile-level-id
- sprop-parameter-sets
- packetization-mode
- sprop-interleaving-depth
- sprop-deint-buf-req
- sprop-max-don-diff
- sprop-init-buf-time

使用できないもの：

- max-mbps
- max-fs
- max-cpb
- max-dpb

- max-br
  - redundant-pic-cap
  - max-rcmd-nalu-size
  - parameter-add
  - deint-buf-cap
- SDP の受信者は、指定されたすべてのパラメータとパラメータ値をサポートしなければならない。そうでなければ、受信機はセッションを拒否しなければならないか (RTSP)、セッションに参加してはいけない (SAP)。受信アプリケーションによってサポートされようとする値の使用は、セッションの生成者の責任に帰する。

### 8.3. 例

両側が送信と受信双方を行う場合の SIP のオファー／アンサー能力交換は、次のようであろう。SDP のメディアコーデック定義のみが示されている。いくつかの行は文字列制限のために折り返っている。

オファー側→アンサー側の SDP メッセージ：

```
m=video 49170 RTP/AVP 100 99 98
a=rtpmap:98 H264/90000
a=fmtp:98 profile-level-id=42A01E; packetization-mode=0;
    sprop-parameter-sets=Z0IACpZTBmI,aMljiA==
a=rtpmap:99 H264/90000
a=fmtp:99 profile-level-id=42A01E; packetization-mode=1;
    sprop-parameter-sets=Z0IACpZTBmI,aMljiA==
a=rtpmap:100 H264/90000
a=fmtp:100 profile-level-id=42A01E; packetization-mode=2;
    sprop-parameter-sets=Z0IACpZTBmI,aMljiA==;
    sprop-interleaving-depth=45; sprop-deint-buf-req=64000;
    sprop-init-buf-time=102478; deint-buf-cap=128000
```

上記のオファーは、3 つの異なったパケット化フォーマットで同一のコーデックコンフィギュレーションを表す。PT 98 はシングル NALU モードを、PT 99 は非インタリーブモードを表し、PT 100 はインタリーブモードを示す。インタリーブモードの場合では、アンサーが PT 100 のサポートを示す場合にオファー側が使用するかもしれないインタリーブパラメータも含まれる。すべての 3 つの場合においてパラメータ「sprop-parameter-sets」が伝達するのは、このコンフィギュレーション (profile-level-id とパケット化モード) を受け付けたオファー側から、ストリームを受信したアンサー側が必要とする初期パラメータセットである。

「sprop-parameter-sets」の値は、上記例においては同一値だが、各ペイロードタイプで異なっている場合があることに注意すること。

アンサー側→オファー側の SDP メッセージ：

```
m=video 49170 RTP/AVP 100 99 97
a=rtpmap:97 H264/90000
```

```

a=fmtp:97 profile-level-id=42A01E; packetization-mode=0;
    sprop-parameter-sets=Z0IACpZTB YmI,aMljiA==,As0DEWlsIOp==,
    KyzFGleR
a=rtpmap:99 H264/90000
a=fmtp:99 profile-level-id=42A01E; packetization-mode=1;
    sprop-parameter-sets=Z0IACpZTB YmI,aMljiA==,As0DEWlsIOp==,
    KyzFGleR; max-rcmd-nalu-size=3980
a=rtpmap:100 H264/90000
a=fmtp:100 profile-level-id=42A01E; packetization-mode=2;
    sprop-parameter-sets=Z0IACpZTB YmI,aMljiA==,As0DEWlsIOp==,
    KyzFGleR; sprop-interleaving-depth=60;
    sprop-deint-buf-req=86000; sprop-init-buf-time=156320;
    deint-buf-cap=128000; max-rcmd-nalu-size=3980

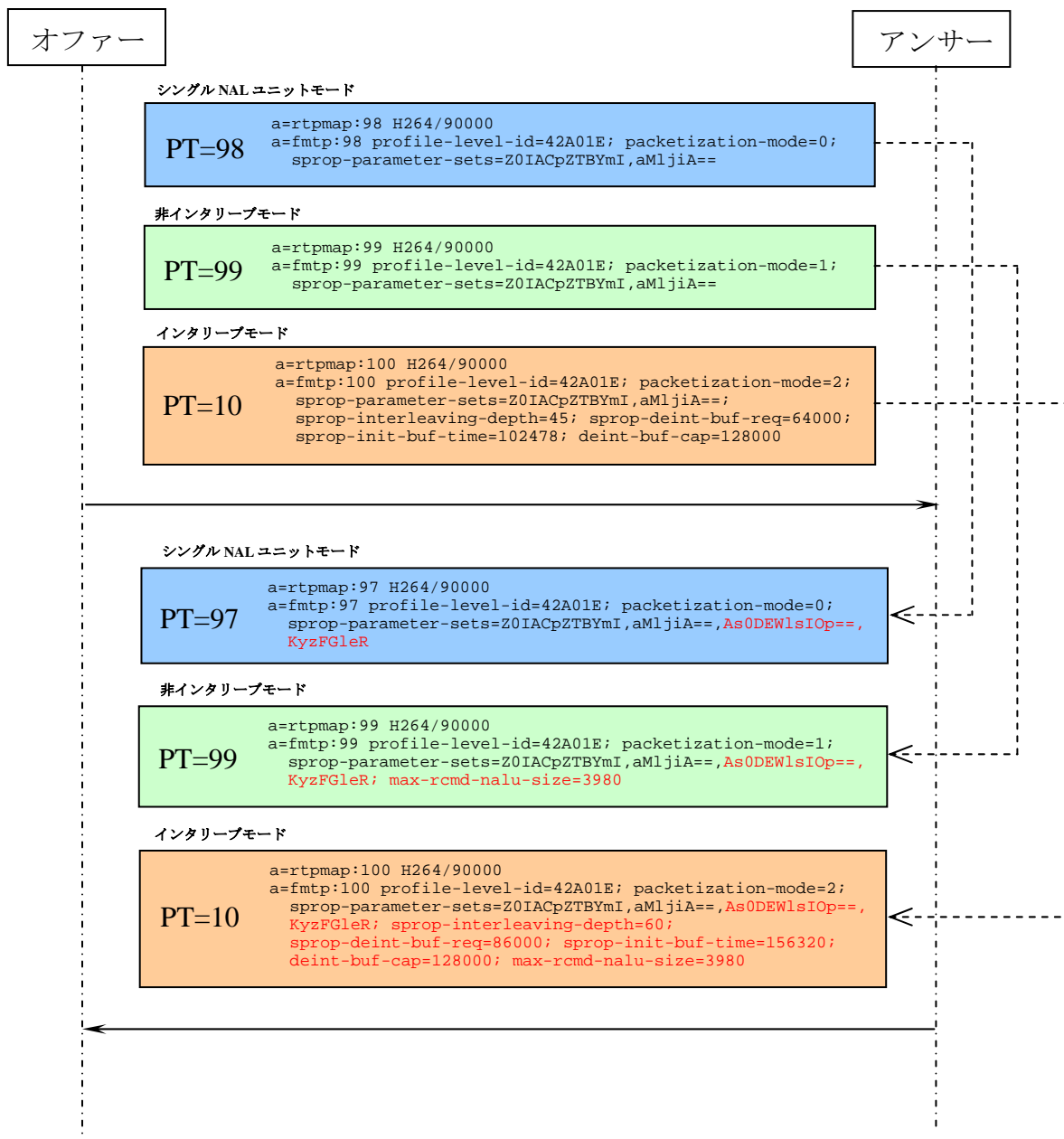
```

オファー／アンサー能力交換が送信ストリームと受信ストリームの両方を取り扱うので、オファーはオファー側が受信を望む正確なパラメータを示し、アンサーはアンサー側と同じストリームのパラメータを受け付けることを示す。この場合、オファー側はペイロードタイプ 98 を受信しても構わないと宣言した。アンサー側は同等のペイロードタイプ 97 を宣言することでこれを受け付ける。すなわちアンサーは、3 種類のパラメータ「profile-level-id」「packetization-mode」「sprop-deint-buf-req」について同じ値を有する。これは、プロパティを宣言するパラメータに関連して、オファー側とアンサー側の両方について、次の意味を持つ。オファー側は最初に、PT=98 のペイロード定義において、「sprop-parameter-sets」のある値を宣言した。しかしながら、アンサー側は PT=97 としてこれを受け付けたので、オファー側が PT=97 を送信する時に、PT=98 における「sprop-parameter-sets」の値を代わりに使用しなければならない。同様に、アンサー側がオファー側に PT=98 を送信する時、アンサー側は PT=97 において宣言されたプロパティパラメータを使用しなければならない。

アンサー側はまた、ペイロードタイプ 99 と 100 が表した 2 つのコンフィギュレーションの受信を受け付ける。アンサー側は、アンサー側からオファー側の方向に、そしてペイロードタイプのストリームを送信するために使用するバッファ関連パラメータに、初期パラメータセットを提供する。アンサー側はまた、「deint-buf-cap」パラメータを提供することによって、非インタリーブ操作のメモリ制限をオファー側に提供する。これは、オファー側が 2 番目のオファーの作成を決定する場合にだけ、新しい値を考慮に入れることのできる場所で有効である。「max-rcmd-nalu-size」は、アンサー側が NALU を 3980 バイトまで効率的に処理可能であることを示す。しかしながら、ネットワークがこのサイズをサポートするという保証は無い。

上記例におけるパラメータセットは、H.264 コーデックの正規の動作点を表現していないことに注意のこと。base64 文字列は例として使用されただけである。

訳注：下記の図は、訳者が追加したもので、原文には含まれない。



※赤文字は、アンサーで追加・変更されたパラメータ



#### 8.4. パラメータセットの考慮点

H.264 パラメータセットはビデオコーデックの基本部分で、そのオペレーションに不可欠である(1.2 節参照)。それらの特性および復号プロセスの重要性により、消失、或いは誤送信されたパラメータセットは、受信側でローカルに隠蔽するのは困難である。破壊されたパラメータセットの参照は、通常、復号プロセスへの誤った結果につながる。破壊は、例えば、誤送信或いはパラメータセットデータ構造の損失により生じ、あるパラメータセットの時宜を失した更新によっても生じる。したがって下記の提案は RTP 送信側の実装者へのガイドラインとして提供される。

パラメータセット NALU は、以下 3 つの異なった原理を用いて転送することができる。

- A. 実際の RTP セッションに先立ってセッション制御プロトコル(アウトバンド)を使用する。
- B. 進行中の RTP セッション中にセッション制御プロトコル(アウトバンド)を使用する。
- C. RTP セッション進行中のペイロード(インバンド)内の RTP ストリームに含まれる。

セッション制御プロトコル内では、原理 A および B を実装することが必要である。SIP と SDP は、SDP のオファ―/アンサーモデルで、および本文書の前節で記述されるように用いられる。この節では、原理 A と B がどのようにセッション制御プロトコル内で実装されなければならないかについてのガイドラインを示す。それは使用される特別のプロトコルに依存しない。原理 C は本仕様で定義された RTP ペイロードフォーマットによってサポートされる。

ピクチャおよびシーケンスパラメータセット NALU は、信頼できる転送が提供されない限り、RTP ペイロード中で送信されるべきでない。なぜなら、どちらかのタイプのパラメータセットが損失した場合、おそらく、対応する RTP ストリームの相当な部分の復号が出来なくなるであろうから。したがって、信頼できるセッション制御プロトコル(即ち上述の原理 A あるいは B の使用)を用いたパラメータセットの送信が推奨される。

本節の残りでは、アウトバンドシグナリングがパラメータセット NALU の信頼できる伝送を供給し、インバンドでは伝送しないことが仮定される。パラメータセットのインバンドシグナリングが使用される場合、送信側はエラー特性を考慮に入れ、高確率で正確にパラメータセットを提供するようなメカニズムを用いるべきである。正確な受信の確率を増加させるメカニズムにはパケット反復、FEC および再送信が含まれる。信頼性の低いアウトバンド制御プロトコルの使用にはインバンドシグナリング(損失の可能性がある)と同様の欠点があり、更に同期が困難になるかもしれない(下記参照)。したがって、それは推奨されない。

パラメータセットは、原理 B および C を用いたセッションの有効期間中に追加または更新してもよい。パラメータセットがそれらを参照する NAL ユニットより先にデコーダ側に存在することが要求される。パラメータセットの更新あるいは追加は更なる問題を引き起こす可能性がある。したがって、以下の推奨が考慮されるべきである。

- パラメータセットが追加あるいは更新される場合、原理 C は上述されるように送信エラーに弱い。したがって、原理 B が推奨される。

- パラメータセットが追加あるいは更新される場合、任意のパラメータセットがその使用に先立って伝えられることに注意すべきである。同期はアウトバンドシグナリングとインバンドトラフィック間で存在しないことは通常のことである。アウトバンドシグナリングが使用される場合、送信側は、シグナリングプロトコルからの配送の確認に先立ち更新されたパラメータセットを必要とするNALUの送信を始めないことが推奨される。
- パラメータセットが更新された場合、以下の同期問題を考慮に入れるべきである。受信側でパラメータセットに上書きする場合、送信側は、問題のパラメータセットが網内あるいは受信側のバッファ内に存在する任意のNALUによって要求されないことを保証しなければならない。そうではなくとも、誤ったパラメータセットでの復号が起こるかもしれない。この問題を減少させるために、十分に長い時間(関連するNALUがすべて消費されたことを保証するだけ)使用されなかったパラメータセットだけを上書きする、あるいは代わりに(ビデオ符号化の効率上、逆効果となるかもしれないが)新しいパラメータセットを加える、のどちらか一方が推奨される。
- 新しいパラメータセットが追加されたとき、以前に未使用のパラメータセット識別子が使用される。このことは、前節で提示された問題を回避する。しかしながら、マルチパーティセッションにおいて、同期制御プロトコルが使用されないならば、多数のエンティティが、同一の識別子のための異なるパラメータセットを追加しようとする危険がある。そして、このことは回避されなければならない。
- 同一のRTPセッション内で原理BおよびC双方の使用によってパラメータセットを追加あるいは修正することは、その制御とRTPチャンネル間で同期がなされていないことから、パラメータセットの矛盾につながるかもしれない。したがって、原理BおよびCは、十分な同期が提供可能でない限りは、同一のセッションでは共に使用してはならない。

いくつかのシナリオにおいて(例えば、H.241に対応するこのペイロードフォーマット仕様のサブセットのみが用いられる場合)、アウトバンド・パラメータセット転送を採用することはできない。この場合、パラメータセットはインバンドで転送されなければならない。ビットストリーム内の非パラメータセットデータとの同期は暗黙的である。しかし、損失の可能性を考慮に入れなければならない。損失の可能性は上記で議論されたメカニズムを使用して少なくされるべきである。

- 原理AおよびBを使用してパラメータセットが最初に提供され、そしてその後インバンドで追加あるいは更新された(原理C)場合、すでにアウトバンドで提供されたパラメータセットが更新されることに伴う危険がある。受信側が(例えば損失または遅れた調整のために)いくつかのインバンド更新を逃す場合、それらの受信側は、時宜を失ったパラメータを用いてビットストリームの復号を試みる。パラメータセットIDをアウトバンドとインバンドのパラメータセットに分けることが推奨される。

H.264 エンコーダからの最大の柔軟性および最良のパフォーマンスを実現するために、もし可能ならば、任意の送信側にセッション内でそれ自身のパラメータセットの追加を許すことが推奨される。「parameter-add」パラメータを「偽」にセットすることは、そのセッショントポロジが参加者による自身のパラメータセット追加を許さない場合に限るべきである。

## 9. セキュリティの考察

本仕様で定義されたペイロードフォーマットを用いる RTP パケットは、RTP 仕様[4]において、および任意の適切な RTP プロフィール(例えば[16])において議論されたセキュリティの考察に従う。このことは、例え

ば、SRTP[26]の適用を通じて、メディアストリームの機密性が暗号化によって達成されることを示す。このペイロードフォーマットと共に使用されるデータ圧縮がエンドーエンドで適用されるために、全ての暗号化が圧縮後に実行される必要がある。

受信端の均一でない計算負荷を持つ圧縮技術を用いるデータ符号化には、潜在的な DoS の脅威が存在する。攻撃者は、復号するのに複雑でかつ受信器に過負荷を生ずるようなストリームに感染能力のあるデータグラムを注入することができる。H.264 は、多くの将来の NAL ユニットの復号プロセスに影響する NAL ユニットを含むデータグラムを生成することが非常に容易なので、そのような攻撃に特に弱い。したがって、少なくとも RTP パケットのデータ発生元認証およびデータの完全性保護の使用（例えば、SRTP[26]によって）が推奨される。

RTP パケットおよびそれらのペイロードの機密性および完全性を保証する適切なメカニズムはアプリケーションおよびその転送/シグナリングプロトコルに非常に依存することに注意のこと。したがって、SRTP は上記の例として与えられるが、他の可能な選択は存在する。

特にそれらが能動素子を含んでいる場合、デコーダは、扱いに関して注意しなければならない。また、デコーダは、ストリームを含む内容提示に対してユーザデータ SEI メッセージの適用可能な領域を制限しなければならない。

認証、完全性あるいは機密性保護を備えたエンドーエンドのセキュリティは、MANE に対し、完全なパケット廃棄以外のメディアを意識した操作を許さない。そして機密性保護の場合、メディアを意識した方法でパケットを廃棄することすら許さない。任意の MANE にその操作を許すためには、セキュリティコンテキストの確立において信頼されたエンティティであることが要求される。

## 10. 輻輳制御

RTP のための輻輳制御は RFC 3550[4]および任意の適用可能な RTP プロファイル(例えば RFC 3551[16])に従って使用されるべきである。

ベストエフォートサービスが利用されているならば、追加要求は次のとおり。

このペイロード・フォーマットのユーザは、パケット損失率を許容可能なパラメータ内に納めるためにパケット損失をモニターしなければならない。許容可能なパケット損失とは、仮に TCP フローが同一のネットワーク条件で同一のネットワーク経路を経由して流れたとして、現に RTP フローがなすより小さくない平均スループットを得るような損失である。この条件は伝送速度(あるいは階層化されたマルチキャストセッションが参加した複数の階層)を調節するような輻輳制御機構を実行することによって、および損失率が許容できないほど大きいなら、受信側にそのセッションを解放するよう調整することによって満足させることができる。

輻輳制御原理を動作させるために必要なビットレート調整は、実時間符号化が用いられる時、容易に実行できる。しかし、符号化済みのコンテンツが伝送される時、帯域調整は、異なったビットレートでの同一のコンテンツの2個以上の符号化表現、あるいは非参照ピクチャまたはサブシーケンスの存在を要求する。異なった表現間のスイッチングは通常同一の RTP セッションにおいて実行される。即ち、拡張プロファイルの SI/SP スライスとして知られている概念の採用によって、あるいは IDR ピクチャ境界におけるストリームのスイッチングによって実行される。ダウングレード不可能なパラメータ(プロファイル/レベル ID のプロフ

ファイル部分のような)が変更されるよう要求された時のみ、メディアストリームを終了させ再起動することが必要になる。このことは異なった RTP ペイロードタイプを用いることによって実現しても良い。

MANE は 7.3 節で概説した提案に従って、そのストリームが直前のパケットロスによって破壊された時、パケットストリームから該当の使用できないパケットを取り除いてもよい。これにより、ある特別な場合には、ネットワーク負荷を軽減できる。

## 11. IANA の考慮点

IANA は 1 個の新しい MIME タイプを登録した。8.1 節参照。

## 12. 付録:アプリケーション例

このペイロード仕様は、H.264 のために見越された広範なアプリケーション領域をカバーするために、使用上とても柔軟性がある。しかしながら、この大きな柔軟性で実装者が合理的なパケット化スキームを決めることを困難にすることにもなる。実世界シナリオにこの仕様をいかに適用するかという観点上のいくつかの情報が、近い将来、学術論文やテストモデルソフトウェアおよび記述の形で発表されるだろう。しかしながら、いくつかの当面の使用方法のシナリオは本文書で同様に述べられている。

### 12.1. ITU-T 勧告 H.241 付属資料 A に従ったテレビ電話

オプションで映像圧縮方式として H.264 を使用する H.323 に基づいたテレビ電話システムは、パケット化方式として H.241 Annex A [15]をサポートするよう要求される。この付属資料で定義されるパケット化メカニズムは、この仕様の 1 個の小サブセットと技術的に同一である。

あるシステムが、H.241 Annex A に従い動作する時、パラメータセット NAL ユニットのインバンドで送信される。シングル NAL ユニット・パケットだけが使用される。多数のそのようなシステムは IDR ピクチャを規則的には送らない。しかし、ユーザからの指示によって、あるいは制御プロトコル手段によって要求される時(例えば、マルチポイント制御ユニットにおけるスイッチング、あるいはフィードバックによって要求された誤り訂正のためのビデオチャンネル間のスイッチング時)のみ送られる。

### 12.2. テレビ電話、スライス・データパーティションなし、NAL ユニットアグリゲーションなし

このスキームの RTP 部分は実装されテストされている。(但し、制御プロトコル部分を除く。下記参照)

ほとんどの実世界テレビ電話アプリケーションにおいて、ピクチャサイズやオプションのモードのようなピクチャパラメータはコネクションの継続期間中には決して変更されない。したがってすべての必要なパラメータセット (通常 1 個だけ) は、能力交換/通知プロセスの副次作用として (例えば、この文書の 8.2 節に規定された SDP シンタックスに従って) 送られる。

すべての必要なパラメータセット情報が RTP セッション開始前に確立されるので全くパラメータセット NAL ユニットの送信する必要はない。スライスデータパーティションも使用されない。したがって RTP パケットストリームは、基本的に単一の符号化スライスを運ぶ NAL ユニットの構成される。

エンコーダは、それらがベストパフォーマンスを提供するように、符号化スライス NAL ユニットのサイズを選ぶ。しばしば、このことは IP ネットワークの MTU サイズへ符号化スライスのサイズを適合させることにより行われる。小さなピクチャサイズでは、結果として、1 パケット当り 1 ピクチャの構成をとることと

なる。イントラ・リフレッシュ・アルゴリズムはパケット損失およびその結果生じるドリフトに関連した歪を解消する。

### 12.3. テレビ電話：NALユニット集合を用いたインタリーブパケット化

本スキームは、よりよいエラー隠蔽を可能にするもので、RFC 2429のパケット化[10]を用いたITU-T H.263ベースの設計に用いられている。これは既に実装されたものであり、良い結果が報告されている[12]。

VCL エンコーダは、1マクロブロック(MB)ラインの全てのMBが一つのスライスに割り当てられるように、原画像を符号化する。偶数MB行アドレスを持つ全てのスライスは一方のSTAPに結合され、奇数MB行アドレスを持つ全てのスライスはもう一方のSTAPに結合される。それらのSTAPが、RTPパケットとして伝送される。パラメータセットの確立は、上記のように行なわれる。

ここでは、STAPを使用することが重要な点である。スライス数(1枚のCIFピクチャなら18個)が大きくなってくることが、IP/UDP/RTPヘッダのオーバーヘッドが許容できなくなるほど大きくなることにつながるためである。(ただし、ソースコーディングツールFMOを使用しない場合についてである。これは本シナリオでは想定しない。) さらに、H.324Mや、3GPPで規定されるIPベースのテレビ電話等の無線映像伝送システムでは、比較的小さい伝送パケットサイズが用いられる傾向にある。例えば、H.223 AL3 SDUの一般的なMTUサイズは、およそ100byteである[17]。本パケット化スキームに従って個々のスライスを符号化することにより、有線-無線ネットワーク間通信においてさらなる利点が生じる。それは、個々のスライスが、無線システムで推奨される最大パケットサイズよりも小さくなりやすいからである。従って、ゲートウェイは、有線ネットワークで使用される(複数個の)STAPを、たった一つのNALユニットを持つ複数個のRTPパケットに変換することができる。それは、無線ネットワークでは望ましいことであり、その逆(無線→有線への変換の場合)も同様である。

### 12.4. データパーティションを行うテレビ電話

本スキームは既に実装されたものであり、特に比較的高いパケットロス発生率の状況下では、良好なパフォーマンスを提供することが示されている[12]。

データパーティションは、何らかの形式の不均一なエラー保護が利用できる場合のみに有効なものとして知られている。通常、単一セッションのRTP環境では、均一なエラー品質が仮定されている。すなわち、そのセッションの全パケットでのパケットロスの発生確率が統計的に同じであるということである。しかし、一つのRTPセッションで個々のパケットのパケットロスの発生確率を下げる手段が存在する。例えば、RFC 2733に準拠したFECパケットでは、どのメディアパケットがそのFECパケットと関連しているかが明記される。

全てのケースにおいて、被るオーバーヘッドは相当なものではあるが、それはFECを用いない場合にイントラ情報に費やすであろうビット数と同じ大きさのオーダーである。

しかし、本メカニズムは、システムにはいかなる遅延も加えるものではない。

本スキームも同様に、完全なパラメータセットの確立は、制御プロトコルの手段を用いて実行される。

### 12.5. FUと前方誤り訂正によるテレビ電話またはストリーミング

本スキームは既に実装されたものであり、特に比較的高いパケットロス発生率において、良いパフォーマンスを提供することが示されている[19]。

再送が適用されないシナリオにおいて、パケットロスに対抗するための最も効率的な手段は、前方誤り訂正(FEC)である。アプリケーションレイヤにおいてエンド・トゥ・エンドでFECを利用することは、個別リッ

ク(特に伝送パスにおいて各リンクが異なる特性を示すような場合)での FEC ベースの保護を行うよりも時に非効率ではあるものの、いくつかのシナリオでは、アプリケーションレイヤのエンド・トゥ・エンドの FEC は不可避である。RFC 2733[18]は、パケットロス環境における一般的なアプリケーションレイヤのエンド・トゥ・エンドでの FEC を利用する手段を提供する。前方誤り訂正のバイナリコードは、異なるパケット間での同一ビット位置に対して XOR 演算を適用することにより生成される。

バイナリコードは、パラメータ(n,k)により明示される。ここで、k はそのコネクションに用いられている情報パケットの数、そして n は k 個の情報パケットに対して生成されたパケット数を示している。すなわち、k 個の情報パケットに対して、n-k 個のパリティパケットが生成されることになる。

RFC 2733 フレームワーク内で、パラメータ(n,k)のコードが用いられる場合について、以下の性質がよく知られている：

- a) 1 個の RTP パケットに対して適用する場合、RFC 2733 はパケットの複写のみを提供する。
- b) XOR で連結されたパケット同士が同じ長さを持つ場合に、RFC 2733 が最もビットレート効率が良くなる。
- c) 同一のパケットロス確率 p で、かつ固定された k に対しては、n の値が大きくなればなるほど、訂正後のエラー率は小さくなってゆく。例えば、パケットロス確率 10%、k=1、n=2 の場合、訂正後のエラー率は約 1%になるが、n=3 になると、訂正後のエラー率は約 0.1%となる。
- d) 同一のパケットロス確率 p で、符号率 k/n を固定にした場合、n の値が大きくなればなるほど、訂正後のエラー率は小さくなってゆく。例えば、パケットロス確率 p=10%、k=1、n=2 の場合は訂正後のエラー率は約 1%になるが、k=12、n=24 による拡大 Golay 符号の場合は、訂正後のエラー率は 0.01%である。

FU を用いずに、H.264 ベースライン符号化映像と組み合わせて RFC 2733 を適用する場合、いくつかのオプションが考慮される可能性がある：

- 1) ビデオエンコーダが、各映像フレームが単一スライスで符号化されている NAL ユニットの生成する場合。FEC を適用しようとする、単純な符号、すなわち例えば(n=2, k=1)といった符号しか使えなくなってしまう。すなわち、各 NAL ユニットの単に複写されているだけとなってしまう。上記 d)に従い、この欠点により明らかに符号パフォーマンスが悪くなり、かつ柔軟性も低くなる。なぜなら、(n, k=1)の符号しか使用できないためである。
- 2) ビデオエンコーダが、各映像フレームが 1 個以上の連続したスライスで符号化されている NAL ユニットの生成する場合。FEC を適用しようとする、NAL ユニットのシーケンスに渡って、例えば(n=24, k=12)のようなより良い符号を使うことができる。しかし、フレーム当たりの RTP パケットの数によっては、1 個のパケットロスが重大な遅延につながるかもしれない。この遅延は、フレーム当たりの RTP パケット数が多くなれば、軽減される。全く異なる長さのパケットも接続されるかもしれない、そのため、上記 b)に従ってビットレート効率が下がってしまう。しかし、注意深く処理を行えば、1KB 以上のスライスに対しては、同じような長さ(差異が 100~200byte 程度)のパケットを生成する可能性があり、その場合は壊滅的にビット効率を下げてしまうようなことは起こらないであろう。
- 3) ビデオエンコーダが、あるフレームにおいて、おそらくほぼ等しい長さの k 個のスライスを含むような、NAL ユニットの生成する場合。FEC を適用すると、各フレームの NAL ユニットのシーケンスに全体に渡って、より良い符号、例えば(n=24, k=12)等を用いることができる。上記 2)の場合と比較した場合の遅延はおそらく減少するであろう。しかし、いくつかの欠点が明らかになっている。まず、符号化された映像の符号化効率がかなり低くなってしまふ。なぜなら、スライス構造化された符号化により、フレーム内予測が減少し、スライスのオーバーヘッドが追加で必要になるためである。次に、あらかじめ符号化されたコンテンツや、あるいはゲートウェイを介した運用の場合は、その映像は通常は、

FEC が適用できるように  $k$  個のスライスに適切に符号化されることはない。最後に、同じ長さの  $k$  個のスライスを生成する映像符号化は直接には実行できず、複数回のエンコーディングパスが必要となってしまう。

これまでに述べた欠点の多くは、FEC と組み合わせて FU を適用することで回避することができる。各 NAL ユニットの、原則として同じ長さの任意数の FU に分割することができる。よって、エンコーダが特段同じ大きさのスライスを生成する処理を行わずとも、適切な  $k$  と  $n$  を適用することができる。例えば、1 個のフレーム全体を含むような 1 個のスライス NAL ユニットの、 $k$  個の FU に分割でき、パリティチェックコード( $n=k+1,k$ )が適用できる。しかし、生成された全てのフラグメントが復元されなければ、スライス全体が失われるという欠点も存在する。このように、フレームがいくつかのスライスに分割された場合と比べ、より大きなセクションが失われてしまう。

提示したテクニックにより、(定期的に発生するイントラフレームのような)追加の情報源符号化レイヤでの冗長性が存在しない場合であっても、良好なエラー耐性を得ることが可能である。従って、エラーフリーの伝送とエラーの発生しやすいネットワーク上での伝送の両方に対して、同じ符号化映像シーケンスを最大限の圧縮効率と品質を得る目的で用いることが可能である。

さらに、本テクニックにより、前符号化(蓄積メディア用にあらかじめ符号化)されたシーケンスに対して、遅延を加えることなしに、FEC を適用することが可能である。この場合、エラーの発生しやすいネットワーク向けに符号化されたものでないような前符号化されたシーケンスについても、さらなる遅延を追加することなく、ほぼ信頼性を保ったまま伝送することが可能である。さらに、FU を同じ長さにすることが、RFC 2733 での効率的なビットレート利用につながっている。

エラー発生率が伝送パケットの長さに依存する場合(例：モバイル伝送の場合[14])、FU に対して FEC を適用することによる利点は、なお一層明らかである。基本的に、FU サイズの柔軟性により、各 NAL ユニットの適切な FEC を適用し、複数 NAL ユニットの不均一なエラー保護を行うことが可能である。

FU と FEC を用いた場合、発生するオーバーヘッドは相当なものとなるが、それは、もし FEC を全く適用しない場合にイントラ符号化されたマクロブロックに対して費やすべきビット数と同じオーダーの規模である。[19]では、同じエラー率かつ同じ全体ビットレートでの場合に、FEC ベースのアプローチにより、オーバーヘッドを含めて全体的なパフォーマンスを向上させたことが示された。

## 12.6. 低ビットレートのストリーミング

本スキームは、H.263 および非標準の RTP パケット化により実装されたもので、良い結果が得られている [20]。同様の良い結果が H.264 でも得られない技術的理由は存在しない。

今日のインターネットストリーミングは、端末がダイヤルアップモデムでコンテンツにアクセス可能にするために、提供されているビットレートのいくつかは比較的低いものである。有線 IP ネットワークでは、ネットワークの混雑度を下げるため、およそ 500~1500 バイトといった比較的大きめのパケットの方が、より小さいパケットにして送出頻度を上げるよりも望ましい。さらに、大きいパケットを使うことで、RTP/UDP/IP ヘッダのオーバーヘッド量が減少する。低レートの映像では、大きいパケットを使うことは、時に 2、3 枚程度までのピクチャをまとめて 1 個のパケットにカプセル化すべきであることを意味する。

しかし、多数の符号化ピクチャを含むパケットが損失することは、視覚上での品質において重大な結末をもたらすことになりかねない。なぜなら、実質的に、直前のピクチャを繰り返し表示する以外に、ピクチャ全体のロスを隠蔽する方法がないためである。比較的大きなパケットを構成し、かつロスの隠蔽を成功させる可能性を維持するための方法の一つは、何枚かのピクチャよりインタリーブを行ったスライスを含む MTAP を構成することである。個々の MTAP は、同一ピクチャに属する空間的に隣り合うスライスや、任意のピクチャでの同一空間にあるスライスを含むべきではない。もし一個のパケットが損失した場合、スライスの

損失が、同一ピクチャ上での空間的に隣り合うスライスや、時間軸上で直前や直後のピクチャにある同一位置のスライスに波及しがちになってしまうからである。結果的に、スライス損失に対する隠蔽が比較的連続したものとなりがちである。

### 12.7. ビデオストリーミングにおける強固なパケットスケジューリング

強固なパケットスケジューリングは、MPEG-4 Part 2 で実装されており、無線ストリーミング環境でシミュレーションが行われている。同様の良い結果が H.264 でも得られない技術的理由は存在しない。

ストリーミングクライアントは一般に、比較的大量のデータを蓄積することが可能な受信バッファを持っている。最初に、ストリーミングセッションが確立される段階で、クライアントは即座にストリームを再生することはなく、一般的には入力されたデータを数秒間バッファリングする。このバッファリングは、途切れない再生の維持に役立つ。なぜなら、一時的に伝送遅延が増大したりネットワークのスループットが落ちたりしている場合でも、クライアントはバッファリングしたデータを復号・再生することが可能であるからである。そうではなく初期バッファリングがない状態だと、クライアントは表示をフリーズさせ、復号を停止し、入力データを待つ必要が出てくる。またバッファリングは、任意のプロトコルレベルでの自動的または選択的な再送信においても必要である。一枚のピクチャのどの部分が失われたとしても、失われたデータを再送信するために再送メカニズムが用いられる場合がある。もし再送されたデータがスケジューリングされた復号あるいは表示時刻よりも前に受信されれば、損失は完全に修復される。

符号化されたピクチャは、復号されたシーケンスの主観的品質における重要さに基づき、ランク付けすることが可能である。例えば、通常の B ピクチャなどの非参照ピクチャは、主観的には最も重要性が低い。なぜなら、そのピクチャが存在しないことで他のいかなるピクチャの復号にも影響を及ぼさないからである。非参照ピクチャに加え、ITU-T H.264 | ISO/IEC 14496-10 標準には、サブシーケンスと呼ばれる時間軸上のスケラビリティ手法が含まれている[22]。符号化されたスライスデータパーティションまたはスライスグループに基づく主観的ランク付けもまた可能である。主観的に最も重要な符号化スライスや符号化スライスデータパーティションは、その復号順序より以前に送信することが可能である一方、主観的に最も重要性の低い符号化スライスと符号化スライスデータパーティションは、通常示される復号順序よりも後に送信することが可能である。従って、最も重要なスライスと符号化スライスデータパーティションのうちの任意の再送部分については、最も重要度の低いスライスやスライスデータパーティションに比べ、スケジュールされた復号・再生時刻より以前に受信される可能性が高いことになる。

## 13. 復号順序番号に対する論理的根拠

### 13.1. 序論

復号順序番号(DON)の概念は、主に、効率的な複数ピクチャ間のスライスインタリーブ(12.6 章参照)と、強固なパケットスケジューリング(12.7 章参照)を可能にするために導入された。これらのアプリケーションの両方においては、NAL ユニットの復号順序と関係なく伝送される。DON は、NAL ユニットの復号順序を示したものであり、受信機で復号順序を復元するために用いられるべきである。効率的な複数ピクチャ間のスライスインタリーブと強固なパケットスケジューリングに対する利用例は、それぞれ 13.2 章、13.3 章で示されている。13.4 章は、冗長符号化ピクチャによって得られるエラー回復性における DON の概念の利点を示したものである。13.5 章は、考えられる DON の代替についてまとめたものであり、なぜ DON が本 RTP ペイロード仕様で採用されたかの根拠を示すものである。



### 13.2. 複数枚ピクチャのスライスインタリーブの例

複数枚ピクチャのスライスインタリーブの例を以下に示す。以下に符号化映像シーケンスのサブセットが、出力順序で示されている。R は参照ピクチャを表し、N は非参照ピクチャを表している。番号は相対出力時刻を示している。

... R1 N2 R3 N4 R5 ...

これらのピクチャの復号順序を左から右へと並べると次のようになる：

... R1 R3 N2 R5 N4 ...

ピクチャ R1、R3、N2、R5、N4 の NAL ユニットが、それぞれ 1、2、3、4、5 に等しい DON で番号づけられている。

各参照ピクチャは、以下のように散らばる 3 種類のスライスグループより構成される(数字は、1 枚の QCIF フレーム内の各マクロブロックのスライスグループ番号を示す)：

```
0 1 2 0 1 2 0 1 2 0 1
2 0 1 2 0 1 2 0 1 2 0
1 2 0 1 2 0 1 2 0 1 2
0 1 2 0 1 2 0 1 2 0 1
2 0 1 2 0 1 2 0 1 2 0
1 2 0 1 2 0 1 2 0 1 2
0 1 2 0 1 2 0 1 2 0 1
2 0 1 2 0 1 2 0 1 2 0
1 2 0 1 2 0 1 2 0 1 2
```

簡単化のため、一つのスライスグループの全てのマクロブロックが一つのスライスに含まれると仮定する。3 種類の MTAP が 3 枚の連続した参照ピクチャから構成されているものとする。従って、各 MTAP に 3 個の集合ユニットが含まれ、各集合ユニットは一つのスライスグループに属するの全てのマクロブロックを含む。最初の MTAP は、ピクチャ R1 のスライスグループ 0、ピクチャ R3 のスライスグループ 1、ピクチャ R5 のスライスグループ 2 を含む。2 番目の MTAP は、ピクチャ R1 のスライスグループ 1、ピクチャ R3 のスライスグループ 2、ピクチャ R5 のスライスグループ 0 を含む。3 番目の MTAP は、ピクチャ R1 のスライスグループ 2、ピクチャ R3 のスライスグループ 0、ピクチャ R5 のスライスグループ 1 を含む。各非参照ピクチャは、STAP-B にカプセル化されている。

従って、NAL ユニットの伝送順序は次のようになる：

```
R1, slice group 0, DON 1, carried in MTAP, RTP SN: N
R3, slice group 1, DON 2, carried in MTAP, RTP SN: N
R5, slice group 2, DON 4, carried in MTAP, RTP SN: N
R1, slice group 1, DON 1, carried in MTAP, RTP SN: N+1
R3, slice group 2, DON 2, carried in MTAP, RTP SN: N+1
```

R5, slice group 0, DON 4, carried in MTAP, RTP SN: N+1  
 R1, slice group 2, DON 1, carried in MTAP, RTP SN: N+2  
 R3, slice group 1, DON 2, carried in MTAP, RTP SN: N+2  
 R5, slice group 0, DON 4, carried in MTAP, RTP SN: N+2  
 N2, DON 3, carried in STAP-B, RTP SN: N+3  
 N4, DON 5, carried in STAP-B, RTP SN: N+4

NAL ユニット			RTP 伝送	
ピクチャ	スライス グループ	DON	パケット Type	シーケンス 番号(SN)
R1	0	1	MTAP	N
R3	1	2		
R5	2	4		
R1	1	1	MTAP	N+1
R3	2	2		
R5	0	4		
R1	2	1	MTAP	N+2
R3	1	2		
R5	0	4		
N2		3	STAP-B	N+3
N4		5	STAP-B	N+4

受信機は、各 NAL ユニットに付随する DON の値を基にして、NAL ユニットの復号順序に再構成することができる。

もし MTAP のうちの 1 個が損失した場合は、空間的に隣り合いかつ同時刻に位置するマクロブロックが受信され、その損失を効率的に隠蔽するのに利用することが可能である。もし STAP のうちの 1 個が損失した場合は、その損失の影響は、時間軸上では伝搬しない。

### 13.3. 強固なパケットスケジューリングの例

強固なパケットスケジューリングの例は以下の通り。例において使われる通信システムは、ビデオが情報源から受信側までに処理される順で、次のコンポーネントから構成される:

- oカメラ、キャプチャリング
- o符号化前バッファ
- oエンコーダ
- o符号化ピクチャバッファ
- o送信機
- o伝送チャンネル
- o受信機

- o受信バッファ
- oデコーダ
- o復号ピクチャバッファ
- oディスプレイ

例において使われる画像通信システムは、次のとおりに動作する。ビデオストリームの処理が段階的に、そして、システムの全てのコンポーネントにおいて同時に起こることに注意すること。情報源ビデオシーケンスは、撮影され、符号化前バッファにキャプチャされる。例えば、符号化前バッファは、サンプリング順から、符号化順にピクチャ順序を並べたり、ビットレート制御の目的のために、複数の非圧縮フレームの解析を行う事に利用しても良い。ある場合においては、符号化前バッファが存在しない代わりに、直ちにサンプリングピクチャは符号化されるかもしれない。エンコーダは、符号化前バッファからのピクチャを符号化し、出力、すなわち、符号化ピクチャを符号化ピクチャバッファに格納する。送信機は、符号化ピクチャを、符号化ピクチャバッファから伝送 PACKET へとカプセル化し、それらを伝送チャンネルを通じて受信機に伝送する。受信機は、受信 PACKET を受信バッファに格納する。受信バッファ処理には、典型的には、伝送遅延ジッタのためのバッファリングが含まれている。受信バッファは、符号化データの正しい符号化順序を回復するためにも利用しても良い。デコーダは、受信バッファから符号化データを読み出し、出力として、復号ピクチャバッファに復号ピクチャを出力する。復号ピクチャバッファは、ピクチャの出力（または、表示）順序を回復するために利用する。最後に、ピクチャは、表示される。

次の図例において、**I** は、**IDR** ピクチャを示し、**R** は、参照ピクチャを示している。**N** は、非参照ピクチャを示し、**I,R,N** に続く数値は、復号順序において、前の **IDR** ピクチャと関係を示すサンプリング時間を示している。一連のピクチャの下の数値は、正規化されたシステムクロックタイムスタンプを示している。システムクロックは、この例においては、任意に初期化され、時間は、左から右に進む。それぞれの **I,R,N** ピクチャは、もし、符号化、伝送、復号に時間がかかっていないとするならば、直前の処理ステップと比較して、同じ時間軸上にマッピングされる。このように、同時に発生するイベントは、全ての図例中で同じカラムに位置している。

符号化ピクチャシーケンスのサブセットは、サンプリング順序で、以下に表現している。

```

...  N58 N59 100 N01 N02 R03 N04 N05 R06 ...  N58 N59 100 N01 ...
...  --|---|---|---|---|---|---|---|---|  ...  -|---|---|---|  ...
...  58  59  60  61  62  63  64  65  66 ... 128 129 130 131 ...

```

図 16. サンプリング順序でのピクチャシーケンス

サンプリングされたピクチャは、符号化順序に並べ替えられて符号化前バッファに格納される。この例では、非参照ピクチャが、出力順序における、前、及び、次の参照ピクチャの両方から予測されている、そして、**IDR** ピクチャに直後の非参照ピクチャを除くと、出力順序において、前の参照ピクチャからのみ予測されていると仮定する。このように、符号化前バッファは、少なくとも 2 つのピクチャを含んでいなくてはならず、そして、バッファリングは、2 つのピクチャ間隔の遅延の原因となる。

符号化前バッファリング処理の出力と、ピクチャの符号化（及び、復号）順序は、以下の通りである。

```

... N58 N59 I00 R03 N01 N02 R06 N04 N05 ...
... -|---|---|---|---|---|---|---|---| ...
... 60 61 62 63 64 65 66 67 68 ...

```

図 17. 符号化前バッファで再順序づけされたピクチャ

エンコーダ、または、送信器は、各ピクチャ DON 値を、復号順序における直前のピクチャの DON 値に 1 を加えて設定することができる。

簡単化のために、以下のことを仮定する。

- シーケンスのフレームレートは、一定である。
- 各ピクチャは、ただ 1 つのスライスから構成される。
- 各スライスは、ただ 1 つの NAL ユニットパッケージにカプセル化される。
- 伝送遅延がない。
- ピクチャは、一定間隔（すなわち、1/フレームレート）で伝送される。

ピクチャが復号順序で伝送される時、ピクチャは、次のとおりに受信される。

```

... N58 N59 I00 R03 N01 N02 R06 N04 N05 ...
... -|---|---|---|---|---|---|---|---| ...
... 60 61 62 63 64 65 66 67 68 ...

```

図 18. 復号順序での受信ピクチャ

伝送（または、受信）順序が復号順序と一致するので、オプション `sprop-interleaving-depth MIME` タイプパラメータは、0 に設定される。

デコーダは、以下に記述するように、ピクチャを復号順序から出力順序に配置し、復号ピクチャバッファを 1 つのピクチャ間隔でバッファする必要がある。

```

... N58 N59 I00 N01 N02 R03 N04 N05 R06 ...
... -|---|---|---|---|---|---|---|---| ...
... 61 62 63 64 65 66 67 68 69 ...

```

図 19. 出力順序

復号ピクチャバッファにおいて必要とされる初期バッファリング量は、バッファリング期間 SEI メッセージ

において、もしくは、H.264 ビデオユーザビリティ情報の `num_reorder_frames` シンタックスエレメントによって示すことができる。`num_reorder_frames` が示すのは、フレーム、相互補完するフィールドペア、ペアでないフィールドの最大数であり、それらは、復号順序におけるシーケンスのあらゆるフレーム、相互補完するフィールドペア、ペアでないフィールドに先行する最大数、そして、出力順序において、それに続く最大数を示している。簡単化のために、`num_reorder_frames` を復号ピクチャバッファにおいて必要とされる初期バッファ量を示すために利用すると仮定する。この例において、`num_reorder_frames` は、1 に等しい。

IDR ピクチャ I00 が伝送中に損失し、そして、システムクロック値が 62 であるとき、再送要求が行われるならば、再送された IDR ピクチャ I00 を受け取る時間は、1 ピクチャ間隔（システムクロックがタイムスタンプ 63 に達するまで）であることに注目すべきである。

IDR ピクチャがそれらの復号ポジションより 2 フレーム間隔で早く伝送されたと仮定する、すなわち、それらのピクチャが次のとおりに伝送される。

```

...  I00 N58 N59 R03 N01 N02 R06 N04 N05 ...
...  --|---|---|---|---|---|---|---|---|---| ...
...  62  63  64  65  66  67  68  69  70  ...

```

図 20. インターリービング：送信順序における早期 IDR ピクチャ

オプション `sprop-interleaving-depth` MIME タイプパラメータは、定義に従って 1 に等しく設定される。（この例における `sprop-interleaving-depth` 値は、以下のとおり、生成される。ピクチャ I00 は、伝送順序におけるピクチャ N58、または、N59 に先行し、復号順序において、それに引き続いている唯一のピクチャである。ピクチャ I00, N58, N59 を除いて、伝送順序は、ピクチャの復号順序と同じである。

符号化ピクチャがちょうど 1 つの NAL ユニットにカプセル化されている場合、`sprop-interleaving-depth` 値は、伝送順序におけるあるピクチャに先行するピクチャ、及び、復号順序における次に引き続くピクチャの最大数に等しい。）

受信バッファ処理は、`sprop-interleaving-depth` parameter 値に従って、同時に 2 つのピクチャを含んでおり、各ピクチャに付随する `DON` 値に基づいて、受信順序から正しい復号順序にピクチャを並び替える。受信バッファ処理の出力は、以下のとおりである。

```

...  N58 N59 I00 R03 N01 N02 R06 N04 N05 ...
...  -|---|---|---|---|---|---|---|---|---| ...
...  63  64  65  66  67  68  69  70  71  ...

```

図 21. インターリービング：受信バッファ

この場合も先と同様に、復号順序から出力順序に配置するために、以下に示すように、1 ピクチャ間隔の初期バッファリング遅延が必要とされる。

```
... N58 N59 100 N01 N02 R03 N04 N05 ...  
... -|---|---|---|---|---|---|---|---| ...  
... 64 65 66 67 68 69 70 71 ...
```

図 22. インターリーブング：再順序づけ後の受信バッファ

IDR ピクチャが伝送中に受け得る最大遅延は、アプリケーション、トランスポート、リンク層の再送が含まれる可能性があり、3 つのピクチャ間隔に等しくなる事に注意すること。従って、ピクチャが復号順序で伝送される場合と比較すると、IDR ピクチャの損失耐性は、再送をサポートするシステムでは、改善される。

#### 13.4. 冗長符号化スライスの強固な伝送スケジューリング

冗長符号化ピクチャは、ピクチャの全体の、あるいはその一部の符号化表現で、対応する原符号化ピクチャが正しく復号された場合には、復号には利用されない。復号された原ピクチャのあらゆる領域と、同一アクセスユニットにおけるあらゆる冗長ピクチャに H.264 復号処理を適用した結果の対応領域との間に顕著な違いが有ってはならない。冗長符号化スライス、冗長符号化ピクチャの一部として符号化されたスライスである。冗長符号化ピクチャは、誤りを含むビデオ伝送における不均等な誤り保護を提供するために利用することができる。ピクチャの原符号化表現が正しく復号されない場合、対応する冗長符号化ピクチャが復号される。冗長符号化ピクチャ機能を利用したアプリケーションや符号化技術の例として、ビデオ冗長符号化 [23] や、マルチキャストストリーミング [24] における「キーピクチャ」の保護が含まれる。

多くの誤りを含むビデオ通信システムの一つの特性として、伝送エラーはしばしばバースト的である。従って、伝送順序における 2 個以上の連続した伝送パケットに影響を及ぼすかも知れない。低ビットレートビデオ通信においては、符号化ピクチャの全体が一つの伝送パケットにカプセル化されることが比較的、よく見られる事である。従って、原符号化ピクチャと冗長符号化ピクチャは、伝送順序において、連続したパケットで伝送されるかも知れない。伝送の仕組みを、更にバースト伝送エラーに耐性を持たせるためには、原符号化ピクチャと冗長符号化ピクチャを 1 つ以上のパケット間隔において伝送する事が有益である。DON の概念は、これを可能とする。

#### 13.5. その他の設計可能性に関する備考

H.264 符号化規格のスライスヘッダシンタックス構造は、`frame_num` シンタックスエレメントを含んでおり、それは、符号化フレームの復号順序を示す事が可能である。しかしながら、次の理由のために、`frame_num` シンタックスエレメントの利用は、復号順序を回復するためには、実行可能ではなく、望ましくない。

- 受信器は、(デコーダに符号化データを渡す前に) 符号化ピクチャ当たり少なくとも 1 つのスライスヘッダを構文解析することが必要とされる。
- フレーム番号シンタックスエレメントは、各 IDR ピクチャにおいて、0 にリセットされるため、複数の符号化ビデオシーケンスからの符号化スライスは、インタリーブする事ができない。

- o frame\_num シンタックスエレメントの同じ値が相互補完フィールドペアの符号化フィールドで共有される。そのため、相互補完フィールドペアの符号化フィールドの復号順序は、frame\_num シンタックスエレメントや、H.264 符号化シンタックスの他のシンタックスエレメントに基づいて回復することができない。

MPEG-4 エレメンタリストリーム伝送のための RTP ペイロードフォーマット[25]は、同一 RTP パケット内において、アクセスユニットのインタリーブング、及び、複数アクセスユニットの伝送を可能とする。

H.264 符号化規格において、アクセスユニットは、[1]の 7.4.1.2 節に従って、原符号化ピクチャに関連する全ての NAL ユニットから成ること、として規定されている。従って、異なるピクチャのスライスは、インタリーブすることができず、改良されたエラー耐性のための複数ピクチャ・スライス・インタリーブング技術（12.6 節を参照）は、利用することができない。

## 14. 謝辞

著者は、注意深いレビューに対し、Roni Even, Dave Lindbergh, Philippe Gentric, Gonzalo Camarillo, Gary Sullivan, Joerg Ott, and Colin Perkins の各氏に感謝する。

## 15. References

### 15.1. Normative References

- [1] ITU-T Recommendation H.264, "Advanced video coding for generic audiovisual services", May 2003.
- [2] ISO/IEC International Standard 14496-10:2003.
- [3] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [4] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.
- [5] Handley, M. and V. Jacobson, "SDP: Session Description Protocol", RFC 2327, April 1998.
- [6] Josefsson, S., "The Base16, Base32, and Base64 Data Encodings", RFC 3548, July 2003.
- [7] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model with Session Description Protocol (SDP)", RFC 3264, June 2002.

### 15.2. Informative References

- [8] "Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC)", available from [http://ftp3.itu.int/av-arch/jvt-site/2003\\_03\\_Pattaya/JVT-G050r1.zip](http://ftp3.itu.int/av-arch/jvt-site/2003_03_Pattaya/JVT-G050r1.zip), May 2003.
- [9] Luthra, A., Sullivan, G.J., and T. Wiegand (eds.), Special Issue on H.264/AVC. IEEE Transactions on Circuits and Systems on Video Technology, July 2003.
- [10] Bormann, C., Cline, L., Deisher, G., Gardos, T., Maciocco, C., Newell, D., Ott, J., Sullivan, G., Wenger, S., and C. Zhu, "RTP Payload Format for the 1998 Version of ITU-T Rec. H.263 Video (H.263+)", RFC 2429, October 1998.
- [11] ISO/IEC IS 14496-2.
- [12] Wenger, S., "H.26L over IP", IEEE Transaction on Circuits and Systems for Video technology, Vol. 13, No. 7, July 2003.
- [13] Wenger, S., "H.26L over IP: The IP Network Adaptation Layer", Proceedings Packet Video Workshop 02, April 2002.
- [14] Stockhammer, T., Hannuksela, M.M., and S. Wenger, "H.26L/JVT Coding Network Abstraction Layer and IP-based Transport" in Proc. ICIP 2002, Rochester, NY, September 2002.
- [15] ITU-T Recommendation H.241, "Extended video procedures and control signals for H.300 series terminals", 2004.
- [16] Schulzrinne, H. and S. Casner, "RTP Profile for Audio and Video Conferences with Minimal Control", STD 65, RFC 3551, July 2003.
- [17] ITU-T Recommendation H.223, "Multiplexing protocol for low bit rate multimedia communication", July 2001.
- [18] Rosenberg, J. and H. Schulzrinne, "An RTP Payload Format for Generic Forward Error Correction", RFC 2733, December 1999.
- [19] Stockhammer, T., Wiegand, T., Oelbaum, T., and F. Obermeier, "Video Coding and Transport Layer Techniques for H.264/AVC-Based Transmission over Packet-Lossy Networks", IEEE International Conference on Image Processing (ICIP 2003), Barcelona, Spain,



September 2003.

- [20] Varsa, V. and M. Karczewicz, "Slice interleaving in compressed video packetization", Packet Video Workshop 2000.
- [21] Kang, S.H. and A. Zakhor, "Packet scheduling algorithm for wireless video streaming," International Packet Video Workshop 2002.
- [22] Hannuksela, M.M., "Enhanced concept of GOP", JVT-B042, available [http://ftp3.itu.int/av-arch/video-site/0201\\_Gen/JVT-B042.doc](http://ftp3.itu.int/av-arch/video-site/0201_Gen/JVT-B042.doc), January 2002.
- [23] Wenger, S., "Video Redundancy Coding in H.263+", 1997 International Workshop on Audio-Visual Services over Packet Networks, September 1997.
- [24] Wang, Y.-K., Hannuksela, M.M., and M. Gabbouj, "Error Resilient Video Coding Using Unequally Protected Key Pictures", in Proc. International Workshop VLBV03, September 2003.
- [25] van der Meer, J., Mackie, D., Swaminathan, V., Singer, D., and P. Gentric, "RTP Payload Format for Transport of MPEG-4 Elementary Streams", RFC 3640, November 2003.
- [26] Baugher, M., McGrew, D., Naslund, M., Carrara, E., and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)", RFC 3711, March 2004.
- [27] Schulzrinne, H., Rao, A., and R. Lanphier, "Real Time Streaming Protocol (RTSP)", RFC 2326, April 1998.
- [28] Handley, M., Perkins, C., and E. Whelan, "Session Announcement Protocol", RFC 2974, October 2000.
- [29] ISO/IEC 14496-15: "Information technology - Coding of audio-visual objects - Part 15: Advanced Video Coding (AVC) file format".
- [30] Castagno, R. and D. Singer, "MIME Type Registrations for 3rd Generation Partnership Project (3GPP) Multimedia files", RFC 3839, July 2004.

Authors' Addresses

Stephan Wenger  
TU Berlin / Teles AG  
Franklinstr. 28-29  
D-10587 Berlin  
Germany  
Phone: +49-172-300-0813  
EMail: stewe@stewe.org

Miska M. Hannuksela  
Nokia Corporation  
P.O. Box 100  
33721 Tampere  
Finland  
Phone: +358-7180-73151  
EMail: miska.hannuksela@nokia.com

Thomas Stockhammer  
Nomor Research  
D-83346 Bergen  
Germany  
Phone: +49-8662-419407  
EMail: stockhammer@nomor.de

Magnus Westerlund  
Multimedia Technologies  
Ericsson Research EAB/TVA/A  
Ericsson AB  
Torshamsgatan 23  
SE-164 80 Stockholm  
Sweden  
Phone: +46-8-7190000  
EMail: magnus.westerlund@ericsson.com

---