

卒業論文 2019 年度（令和元年度）

業務用オーディオシステムのための
AES67 アプリケーションの実装と検証

慶應義塾大学 環境情報学部

城 一統

村井・楠本・中村・高汐・バンミーター・植原・三次・中澤・武田
合同研究プロジェクト

2020 年 1 月

卒業論文 2019 年度（令和元年度）

業務用オーディオシステムのための AES67 アプリケーションの実装と検証

論文要旨

従来のオーディオ伝送の諸問題を解決するために、IP オーディオ伝送が登場した。のちに乱立するさまざまな IP オーディオ伝送規格を標準化する AES67 規格が策定された。ブラックボックスとなっている商用規格を使わずとも、自由に機器やアプリケーションを開発することができるようになった。

そこで、AES67 に準拠したアプリケーションを実装し、実際の業務用オーディオ現場における利用可能性を検証することにした。

業務用オーディオシステムに必要な要件は、遅延の少なさ、高音質、伝送の安定性が挙げられる。

本論文では、AES67 のソフトウェアオーディオ伝送は、業務用オーディオにおける要件を満たすと考えられる。

キーワード

AES67、Audio over IP

慶應義塾大学 環境情報学部

城 一統

Abstract Of Bachelor's Thesis Academic Year 2019

Implementation and Evaluation of AES67 sender and receiver

Summary

IP audio transmission has emerged to solve the problems of conventional audio transmission. The AES 67 standard was established to standardize the various IP audio transmission standards that were established later. It is now possible to freely develop devices and applications without using the commercial standard which is a black box.

Therefore, we decided to implement an application conforming to AES 67 and verify the availability of the application in the actual audio field for business use.

Requirements for professional audio systems include low latency, high sound quality, and transmission stability.

In this paper, it is considered that the software audio transmission of AES 67 satisfies the requirements for professional audio.

Keywords

AES67, Audio over IP

Bachelor of Arts in Environment and Information Studies

Keio University

Kazunori Jo

目 次

第 1 章	序論	1
1.1	背景	1
1.2	本論文の目的	2
1.3	本論文の構成	2
第 2 章	業務用オーディオの利用現場と関連技術	4
2.1	構成	4
2.2	オーディオ機器	4
2.2.1	マイク	4
2.2.2	ミキサー	5
2.2.3	パワーアンプ	5
2.2.4	スピーカー	5
2.2.5	映像機器	5
2.3	インターフェイス	5
2.3.1	アナログ入出力	5
2.3.2	AES/EBU	6
2.3.3	S/PDIF	7
2.3.4	SDI	8
2.3.5	CobraNet	8
2.4	ケーブル	8
2.4.1	マイクケーブル	8
2.4.2	スピーカーケーブル	8
2.5	帯域	8
2.6	IP ベースのデジタル伝送規格	8
2.6.1	Internet Protocol	9
2.6.2	Ethernet AVB	9
2.6.3	Dante	9
2.6.4	RAVENNA	9
2.6.5	AES67	10
2.7	業務用システムに必要な要件	10
第 3 章	設計	11

3.1	アプリケーションの構成	11
3.2	AES67 オーディオストリーム送信機能	11
3.2.1	IP マルチキャスト送信	11
3.2.2	IP ユニキャスト送信	12
3.2.3	SAP パケットの送信	12
3.2.4	PCM オーディオの送信	12
3.3	AES67 オーディオストリーム受信機能	12
3.3.1	RTP パケットの受信	13
3.3.2	PCM オーディオの保存	13
3.4	まとめ	13
第 4 章	実装	14
4.1	対応する環境とプラットフォーム	14
4.1.1	実装環境	14
4.1.2	対応 OS	14
4.1.3	対応アーキテクチャ	14
4.2	SAP プロトコルのアナウンス	15
4.3	RTP 通信	16
4.3.1	IP パケットにおける DSCP 値の付加	16
4.3.2	RTP パケットのペイロード	17
4.3.3	RTP パケットの送信間隔	17
第 5 章	評価	18
5.1	評価手法	18
5.2	評価環境	18
5.3	AES67 伝送の結果	18
5.3.1	UDP パケット転送遅延	18
5.3.2	パケットドロップ率	19
5.4	遅延の発生	19
第 6 章	結論	20
6.1	本研究のまとめ	20
6.2	本研究の結論	20
6.3	今後の課題	20
6.3.1	RTP パケット送信間隔の適正值	20

6.3.2	伝送経路の冗長化	20
6.3.3	ハードウェア実装	21
	謝辞	22
	参考文献	24

图 目 次

表 目 次

2.1	AES/EBU と S/PDIF の比較 [1]	7
5.1	コンピュータ評価環境	18
5.2	ネットワーク評価環境	18

第1章 序論

1.1 背景

今日の音楽のリスニング環境は、イヤホンやヘッドホンで聴くことはごく当たり前のものになっている。スマートフォンやコンピュータといった再生機器に、3.5mm ステレオミニプラグ¹を接続するか、あるいは Bluetooth といった無線伝送によってオーディオは伝送されている。

一般的な民生用オーディオ伝送では、インターフェイスや接続が手軽な利便性をとり、ある程度の音質劣化や遅延は許容されている。一方で、業務用オーディオの世界では妥協のない高音質の追求や、絶対に伝送される安定性が優先される。

プロのアーティストによるライブ、テレビ局やインターネットの動画配信などで用いられる業務用オーディオでは、民生向けで使われるオーディオ伝送に比べ、次にあげるような条件が要求される。

- 高音質
- 低遅延
- 伝送の安定性

デジタル伝送技術は、アナログ伝送に比べ伝送経路による劣化が起こりにくくなっている。ただし、オーディオ入力と出力の最初と最後は図??のように、必ずアナログとデジタルの変換（以下、A/D 変換、D/A 変換）が必要となる。そのため、すべての伝送経路をアナログで伝送するのに比べ、少なからず伝送するにあたり A/D 変換と D/A 変換にともなう遅延が発生する。しかしながら、技術の向上によって遅延量は少なくなっている。

もう一つ、デジタル伝送を行うことによる恩恵は高音質のまま伝送できることである。アナログ伝送のまま伝送を行えば、理論的には A/D、D/A 変換を挟まないため、劣化しないはずである。しかしながら、アナログ伝送は外部からのノイズに弱く、伝送経路中で劣化する可能性が否定できない。

そこで、デジタル伝送ではアナログ信号を 0 と 1 の信号にする。アナログ信号は、連続したデータ表現のため、デジタル信号にそのまま表現すると無限となってしまう。それでは現実的でないためデジタル信号では、量子化によってデータ量を減らすことができる [2]。したがって、従来アナログ伝送においてケーブル 1 本で 1 チャンネル伝送していたところをデジタル伝送により 2 チャンネル以上の多チャンネル伝送ができるようになる。AES/EBU というデジタル伝送規格では、接続に用いるケーブルにもよるが 2 ~ 16 チャンネル程度の伝送が行える。

¹いわゆるイヤホンジャック

インターネット、ネットワークの接続に利用する Ethernet（イーサネット）は、インターネットの高速化やコンピュータで扱うコンテンツのリッチ化にともない、通信速度の高速化が進んでいる。2019 年現在、市販の Ethernet ケーブル（図??）を用いた接続では、最大 10Gbps²の通信速度で転送することができる。

Ethernet を用いたデジタルオーディオ伝送では、高速な Ethernet の帯域を用いて従来のデジタルオーディオ伝送に比べ、より多くのチャンネルを伝送することができる。多チャンネル化が実現できるのみならず、IP ベースのエラー訂正が可能となる。

大規模な業務用オーディオ環境では、広帯域な Ethernet を用いた IP ベースの伝送への移行が進んでいる。主流となっているのは、Audinate 社が開発した Dante という規格だ。Audinate 社は Dante 機能のチップをオーディオ機器メーカーに卸す手法をとっている。オーディオ機器メーカーに対して Dante に関する詳細な仕様は公開されておらず、ブラックボックスとなっている現状がある。したがって、オープンな IP ベースのデジタルオーディオ伝送が必要である。

そこで、登場したのが AES67 だ。AES67 は、乱立する既存の IP ベースのデジタルオーディオ伝送技術を共通化し、相互運用性を高めるための規格である。今後、オーディオの IP 伝送が主流となっていくなかで、オープンな標準化技術によってさまざまなオーディオ機器を接続できることは重要である。

1.2 本論文の目的

本論文では、オーディオ機器の最初と最後の A/D、D/A 変換を行う前のアナログ伝送を除くすべての伝送経路を AES67 によって実現できるかの可能性を追求する。必要となる、AES67 オーディオストリームを送受信するアプリケーションを実装する。その上で、既存の商用規格から置き換えが可能なのか、業務用途で要求される厳しい条件に耐えられるかを検証する。

1.3 本論文の構成

本論文における以降の構成は次の通りである。

2 章では、本論文で扱う IP ベースのオーディオ伝送について理解ための前提技術について、解説する。業務用オーディオが使われている現場の構成、さらにはオーディオのデジタル伝送技術について触れる。

3 章では、AES67 を用いた IP ベースのオーディオ伝送の送受信を行うアプリケーションの設計を行い、その内容について述べる。

4 章では、3 章で設計したアプリケーションを実装し、その内容について述べる。

5 章では、4 章で実装したアプリケーションをもとに、実際の業務用オーディオ環境を模した実験を行い、評価した結果について述べる。

²Gbps: Gigabit per second、ギガビット毎秒

6 章では、本研究における結論と今後の展望、業務用オーディオの伝送技術の未来について述べる。

第2章 業務用オーディオの利用現場と関連技術

本章では、業務用オーディオが利用されている現場とそれらを取り巻く周辺環境、運用上の課題について解説する。

2.1 構成

業務用オーディオが利用されている現場は、町の小規模な祭りからイベント、ライブやコンサートなど多岐にわたる。本項では、同時に複数人がマイクを利用するトークイベントを例に、実際の現場で利用されている構成を紐解いていく。

オーディオの伝送経路にアナログ伝送を用いた、非常に簡素な構成を図??に示す。オーディオの入力から出力まで数多くの機器が存在するケースもあるが、接続図では省略する。

4本のマイクを1台のミキサーに入力し、複数のソースのオーディオをミックスしたオーディオが出力される。この構成では、すべての伝送経路をアナログ伝送によって行われている。アナログ伝送の場合、原則オーディオ1チャンネル¹につきケーブルが1本ずつ必要となる。

次に、IPベースのデジタル伝送を用いた場合オーディオ機器同士の接続は、図??のように変化する。

IPベースのデジタル伝送では、すべての伝送をL2スイッチを介して行う。コンピュータ同士を接続するネットワークに用いる汎用的な製品で対応できるため、同時にオーディオ以外のデータの送受信も可能だ。アナログ伝送では、オーディオ1チャンネルあたりにケーブル1本が必要となるため、ステレオの伝送に2本必要としていた。一方デジタル伝送では、多チャンネル伝送ができるためケーブルは1本で送信することができる。

2.2 オーディオ機器

本節では、図??と図??で示したオーディオ機器について、解説する。

2.2.1 マイク

マイクは、音声を電気信号に変換する装置である。伝送方式から、有線マイクと無線マイク（以下、ワイヤレスマイク）に分類できる。

¹オーディオ信号の入出力の数。マイクの場合、1本につき1チャンネル。スピーカーが一般的なステレオ構成であれば2チャンネルとなる。

ワイヤレスマイクの場合、送信機と受信機に分かれている。送信機は、マイクで変換した信号をアンテナを通して無線によって伝送する。受信機は、アンテナから受信した電波という構成だ。それらを総称してワイヤレスシステムと呼ばれることがある。

2.2.2 ミキサー

ミキサーは、複数チャンネルのオーディオ入力をミックスし、1チャンネル以上の出力をする装置である。アナログミキサーとデジタルミキサーの2つに分類できる。ミックス処理をアナログ回路で行うか、DSP²によって行うかの違いがある。

2.2.3 パワーアンプ

ミキサーからの出力では、スピーカーを駆動させるのに必要な電気信号のレベルが不足している。電気信号を増幅させるのがパワーアンプと呼ばれる装置だ。ただし、スピーカーにパワーアンプが内蔵されているものもあり、その場合パワーアンプは不要となる。

2.2.4 スピーカー

スピーカーは、電気信号を音に変換する装置だ。人の耳に音を届けるために必要である。

2.2.5 映像機器

オーディオ信号が、映像機器によって入力されることもある。たとえば、ビデオカメラの映像出力やPCの画面出力にオーディオ信号が入ることがある。ビデオスイッチャーによって映像・音声信号の分離が行われ、音声信号がミキサーに入力される構成が一般的だ。

2.3 インターフェイス

2.3.1 アナログ入出力

アナログ伝送は、音を電気信号に変換して伝送する。マルチケーブルを用いて1本で複数のチャンネルを束ねることもある。

²Digital Signal Processor

2.3.2 AES/EBU

AES/EBU は、オーディオ専門家の国際組織である米国オーディオ技術者協会 (AES)³と、欧州放送連合 (EBU)⁴の両者によって策定されたデジタルオーディオ伝送規格である。それぞれ AES3-1985[3] と EBU Tech.3250-E[4] という名称で、1985 年に策定された。なお、両者の名前をとって一般的に AES/EBU の名称で呼ばれている。

(1) IEC による標準化

また、AES3 は国際電気標準会議 (IEC⁵) によって標準化もされている。IEC による標準化では 4 つのパートで構成されており、IEC 60958-3 では AES3 から派生した民生用途向け (いわゆる S/PDIF) と、IEC 60958-4 では AES3 に基づいた業務用途向けのものが定められている。

接続するインターフェイスは IEC 規格によって策定されている。アナログオーディオ伝送にも用いられる XLR コネクタによるバランス (平衡) 接続や、BNC コネクタを用いた同軸ケーブルによるアンバランス (不平衡) 接続、さらには D-sub コネクタを用いたオーディオ機器メーカーによる独自規格が存在する。本稿執筆時点で最新の AES3 規格では、光ファイバーによる伝送も検討されている⁶。

(2) AES-3id-1995

BNC コネクタを用いたアンバランス接続は、AES3-id-1995 という名称で策定されている。AES/EBU では、インピーダンスが 110 Ω のケーブルを用いる。バランス接続は、ノイズの影響を受けにくく安定した伝送が期待できるものの、アンバランス接続の同軸ケーブルに比べ高価である。BNC コネクタが付いたインピーダンスが 75 Ω の同軸ケーブルは、業務用ビデオ伝送の分野で主流だ。

大規模なホールのような会場では予めさまざまな場所に配線しておき、パッチ盤で接続するという手法がある。パッチ盤のコネクタが BNC であれば何を伝送してもよいため、ビデオ伝送用に配線されている既存の同軸ケーブルを利用できる。AES/EBU で同軸ケーブルを用いて伝送するというアイデアは、ビデオ技術の団体米国映画テレビ技術者協会 (SMPTE)⁷から AES に対して要望が出ていたという経緯がある [5]。同軸ケーブルを用いた伝送は、信号を補償 (ブースト) することで長距離伝送が可能となる。同規格は、SMPTE から SMPTE 276M-1995 という規格にもなっている。

³Audio Engineering Society, Inc. <http://www.aes.org/>

⁴European Broadcasting Union <https://www.ebu.ch/home>

⁵International Electrotechnical Commission

⁶AES3-2009 (r2019): AES standard for digital audio engineering - Serial transmission format for two-channel linearly represented digital audio data <http://www.aes.org/publications/standards/search.cfm?docID=13>

⁷Society of Motion Picture and Television Engineers <https://www.smpete.org/>

2.3.3 S/PDIF

前述したとおり、業務用の AES/EBU から派生した民生用向けの規格は、IEC 60958-3 によって標準化されている。この規格は、一般的に S/PDIF (Sony/Philips Digital Interface) と呼ばれる。日本のソニーとオランダのフィリップスが開発したためそのように呼ばれている。

AES/EBU では、民生用では馴染みの薄い XLR コネクタや BNC コネクタが用いられてきたが、S/PDIF では光デジタル音声端子 (Optical) と同軸デジタル音声端子 (Coaxial) の 2 種類が存在する。光デジタル端子は、東芝が提唱した TOSLINK と呼ばれる角型コネクタ、Mini-TOSLINK と呼ばれる 3.5mm ステレオミニプラグの形状をしたコネクタがある。同軸端子は、RCA コネクタが使われている。

Mini-TOSLINK は、以前 Apple のパソコンである iMac や Mac mini、MacBook Pro にも搭載されていた⁸。

(1) AES/EBU と S/PDIF の違い

AES/EBU は業務用で S/PDIF は民生用だが、それぞれの違いを表 (1) に記す。業務用オーディオは、可能な限り伝送距離を長くし、XLR コネクタで接続した場合バランス接続による安定性をとっている。一方、民生用オーディオは出力レベルを低くし、伝送距離も 10m 程度である。ただし、著作権保護信号を伝送することができ、受信機が信号を認識するとその機器で録音ができなくなるようだ。

表 2.1: AES/EBU と S/PDIF の比較 [1]

	AES3-1992 (r1997)	AES-3id-1995 ^a	S/PDIF ^b
ケーブル インターフェイス コネクタ	110 STP バランス 3 ピン XLR	75 同軸 アンバランス BNC	75 同軸/光ファイバー アンバランス RCA / TOSLINK
信号レベル	2–7V peak to peak	1.0–1.2V peak to peak	0.5–0.6V peak to peak
最小信号レベル	0.2V	0.32V	0.2V
最大伝送距離	100m	1000m	10m
コピー制御	不可	不可	可能
ビット深度	24bit	24bit	20bit (24bit に拡張可能) ^c

^aSMPTE 276M-1995

^bIEC 60958-3

^c4bit 分予備領域として確保されており、20+4bit を合わせて送信できる。

⁸MacBook Pro は 2015 年発売の機種まで、3.5mm ステレオミニジャックにアナログ入力とコンボジャックで搭載していた。 https://support.apple.com/kb/SP719?locale=ja_JP&viewlocale=ja_JP

2.3.4 SDI

SDI⁹は、ビデオ信号を伝送する規格である。非圧縮のビデオとオーディオを伝送することが可能だ。

本来はビデオ伝送用の規格だが、AES3 に準拠した 48kHz/24bit のオーディオを 16 チャンネル伝送できる。

2.3.5 CobraNet

CobraNet は、1996 年に米 Peak Audio が開発した Ethernet を用いたオーディオ伝送規格である。Ethernet を使うが、IP ベースでないのが特徴的だ [6]。

2.4 ケーブル

2.4.1 マイクケーブル

さまざまなコネクタがあり、バランス接続では XLR、TRS (モノラル)、アンバランス接続では TRS (ステレオ)、RCA などが存在する。コネクタを用いずに銅線そのままに接続する方法もある。

2.4.2 スピーカーケーブル

XLR、フォン、スピコンなどのインターフェイスがある。なお、スピーカーケーブルはパッシブスピーカーで使うが、パワードスピーカーはマイクケーブルで接続する。

2.5 帯域

民生用オーディオは圧縮されている場合がほとんど。MP3 や AAC オーディオは 256 ~ 320kbps が一般的だ。一方、業務用オーディオは非圧縮のものが非常に多い。不可逆圧縮オーディオと比べ、情報量に大きな差がある。

2.6 IP ベースのデジタル伝送規格

本項では、本研究と関連するオーディオのデジタル伝送を行う関連技術について、IP ベースの伝送について示す。

⁹Serial Digital Interface

2.6.1 Internet Protocol

2.6.2 Ethernet AVB

Ethernet AVB は、IEEE¹⁰により策定された Ethernet でオーディオとビデオを伝送する規格である。正式には IEEE 802.1 Audio/Video Bridging という。

2.6.3 Dante

Dante は、米 Audinate¹¹が開発した IP ベースのオーディオ伝送規格である。最大ビット深度は 32bit、最大サンプリング周波数は 192kHz で伝送できる [6]。

2019 年現在、ミキサーやパワーアンプなど、総合音響機器メーカーのヤマハや¹²、パナソニック¹³、マイクメーカーの米 Shure¹⁴、独 Sennheiser¹⁵、オーディオテクニカ¹⁶等のデジタルワイヤレスシステム、業務用レコーダー分野ではティアック (TASCAM ブランド)¹⁷など、幅広いメーカーと製品が対応している Audio over IP のデファクトスタンダードともいえる規格だ。

2.6.4 RAVENNA

RAVENNA¹⁸は、独 ALC NetworX が開発した IP ベースのオーディオ伝送規格である。

他の IP ベースの商用伝送規格と異なり、仕様がオープンになっており RAVENNA Partner Network として各国のオーディオメーカーが参加している。

以下は、RAVENNA の Web サイト¹⁹に掲載されている RAVENNA Partner Network に関する記述だ。

RAVENNA was developed by ALC NetworX who continues to be leading developer of the technology. However, given that RAVENNA is an open technology based on existing standards, this means that anyone is free to develop their own solutions. Indeed, this is actively encouraged in order to give customers as much choice as possible and not have to rely on ALC NetworX as the exclusive solution provider.

¹⁰Institute of Electrical and Electronics Engineers

¹¹<https://www.audinate.com/?lang=ja>

¹²<https://jp.yamaha.com/products/proaudio/index.html>

¹³<https://sol.panasonic.biz/sound/index.html>

¹⁴<https://www.shure.com/ja-JP>

¹⁵<https://ja-jp.sennheiser.com/>

¹⁶<https://www.audio-technica.co.jp/proaudio/>

¹⁷<https://tascam.jp/jp/>

¹⁸<https://www.ravenna-network.com/>

¹⁹<https://www.ravenna-network.com/about-ravenna/overview>

RAVENNA は、無償でバーチャルオーディオデバイスを公開している²⁰。Windows、macOS、Linux に対応しており、インストールすることで RAVENNA と AES67 機器とオーディオの送受信ができるようになる。

2.6.5 AES67

AES67 は、AES が策定した IP ベースのオーディオ伝送規格である。本研究において使用する規格でもある。2.3.2 節の AES/EBU 同様、策定当初から仕様が公開されている。

Dante には AES67 Mode オプションがあり、オプションを有効にすることで AES67 対応機器とオーディオの送受信を行うことができる。RAVENNA においても、AES67 に準拠して動作することが可能だ。AES67 は、乱立する IP ベースのオーディオ伝送規格を統一する目的で策定された。

AES67 では、IP 伝送においてマルチキャストとユニキャストをサポートしている。AES67 で使用する IP バージョンは IPv4 のみであり、IPv6 はサポートされていない。しかしながら、将来 IPv6 をサポートする可能性を見据え、要所に対応のための準備が行われていることが仕様に書かれている。

2.7 業務用システムに必要な要件

業務用オーディオシステムは民生用オーディオシステムに比べ、厳しい要求が存在する。

はじめに、遅延の少なさが求められる。業務用の既存の IP ベースのオーディオ伝送規格で、Dante においてはデフォルトのレイテンシを 1 ミリ秒としている。1 ミリ秒程度の遅延では人間が知覚することはないが、業務用オーディオの運用が想定されるライブステージなどにおいて、マイクから入った音が最終的にスピーカーに出るまでの間にエフェクタのような信号処理が加わることがある。したがって、伝送経路では遅延量を最小にすることが理想だ。

また、伝送するオーディオは劣化を避けるため、前述した遅延の問題から非圧縮伝送で行うことが求められる。

²⁰<https://www.ravenna-network.com/resources/>

第3章 設計

本章では、AES67 を用いた IP ベースのオーディオ伝送をするアプリケーションの設計について述べる。

3.1 アプリケーションの構成

本研究で実験を行うために必要な、アプリケーションの機能は大きく分けて2つに分類できる。AES67 オーディオストリームの送信と、受信を行うものだ。

3.2 AES67 オーディオストリーム送信機能

AES67 オーディオストリームの送信機能について述べる。AES67 オーディオストリームは、Real-time Transport Protocol (RTP)[7][8] を用いて、UDP で伝送される。

3.2.1 IP マルチキャスト送信

はじめに、AES67 オーディオストリーム送信機能のうち、IP マルチキャストによって送信する機能について述べる。マルチキャストを用いることで、同一のオーディオストリームを複数のクライアント（オーディオ機器）で受信することが可能になる。

マルチキャスト送信の利点は、オーディオストリームを受け取る機器が複数台におよぶ場合、一つのオーディオストリーム（パケット）を複数の機器で受け取れるようになる。その一方で、オーディオストリームを受け取る機器が1台のみの場合は当該するオーディオストリームを受け取る必要のないオーディオ機器にもパケットが送信されてしまい、マルチキャスト送信は欠点となる。

IPv4 マルチキャストアドレスは RFC5771[9] において、224.0.0.0/4¹のアドレス空間を用いると定められている。Dante の AES67 モードで動作させるには 239.69.0.0/16²のアドレス空間を用いると定められているため、本実装では 239.69.0.0/16 のアドレス範囲を使うものとする。

¹224.0.0.0 から 239.255.255.255 までのアドレス。クラス D アドレスともよばれる。

²239.69.0.0 から 239.69.255.255 までのアドレス。

3.2.2 IP ユニキャスト送信

次に、AES67 オーディオストリーム送信機能のうち、IP ユニキャストによって送信する機能について述べる。ユニキャストでは、オーディオストリームを一つのクライアントに向けて伝送する。マルチキャストが1対多の伝送であるのに対し、ユニキャストは1対1の伝送となる。

ユニキャスト送信の利点は、オーディオストリームを受け取る機器が1台のみの場合、他のオーディオ機器やそれ以外のネットワーク機器にパケットが流れず、効率よく伝送を行うことができる。

本稿執筆現在、Dante の AES67 モードはユニキャスト伝送をサポートしていないため、AES67 規格をフルサポートしている機器同士の機能である。

3.2.3 SAP パケットの送信

AES67 規格では、5 つのネットワーク内において対応機器を検出する方法が定められている [10]。

- Bonjour
- Session Announcement Protocol
- Axia Discovery Protocol
- Wheatstone WheatnetIP Discovery Protocol
- AMWA NMOS Discovery and Registration Specification

そのうち、Session Announcement Protocol(SAP) を用いた機器検出方法は、Dante の AES67 モードで唯一サポートされている方法であり、今回の実験においても SAP を使用する。

3.2.4 PCM オーディオの送信

オーディオストリームの内容は、AES67 規格に準拠した PCM 24bit/48kHz の 1 チャンネル(モノラル)オーディオを伝送する。あらかじめ FFmpeg³に代表される動画・音声のエンコーダを用いて、WAVE 形式などから PCM フォーマットのオーディオを生成したものをアプリケーションに入力する。

3.3 AES67 オーディオストリーム受信機能

前述した AES67 オーディオストリーム送信機能で、伝送されたオーディオストリームを受信するための受信機能について述べる。

³<https://www.ffmpeg.org/>

3.3.1 RTP パケットの受信

UDP マルチキャストアドレスやユニキャストアドレスにおいて AES67 オーディオストリーム伝送に用いるデフォルトのポート番号 5004 番で待ち受け、RTP パケットを受信する。

3.3.2 PCM オーディオの保存

RTP パケットに含まれる PCM オーディオを取り出し、保存する。

3.4 まとめ

本章では、実験で用いる PCM 形式のオーディオを AES67 オーディオストリームで送受信するアプリケーションについての設計を述べた。次章では、本章で設計したアプリケーションの実装を行う。

第4章 実装

本章では、3章で示した設計を実装した際の詳細について述べる。

AES67の送受信アプリケーションは、Go言語で実装する。Go言語を採用した理由は、大きく3つの理由がある。

- 現代的な静的型付け言語である
- 容易にシステムコールを呼び出すことが可能
- 異なるOSやアーキテクチャの上で動くバイナリを生成できる

なお、実装コードはGitHubの `itochan/aes67-transmitter` リポジトリ (<https://github.com/itochan/aes67-transmitter>) で公開している。

4.1 対応する環境とプラットフォーム

実装するにあたり、Go言語の特色であるクロスコンパイル可能な特徴を生かし、複数のプラットフォームで動作可能とした。Raspberry PiのようなARMベースのプロセッサを搭載したコンピュータでも動作させるため、ARMv7アーキテクチャもサポートしている。

4.1.1 実装環境

- go version go1.13.6

4.1.2 対応OS

- Windows (`$GOOS=windows`)
- macOS (`$GOOS=darwin`)
- Linux (`$GOOS=linux`)

4.1.3 対応アーキテクチャ

- amd64 (`$GOARCH=amd64`)
- armv7 (`$GOARCH=arm $GOARM=7`)

4.2 SAP プロトコルのアナウンス

SAP では、RFC 2974[11] で定められているマルチキャストアドレスとポートに機器の対応プロトコルや IP アドレス、ポート番号をアナウンスする。その内容には RFC 4566[12] で策定された Session Description Protocol (SDP) を用いる。

本実装では、次のセッション告知を行なっている。なお、IP アドレスとホスト名に関しては一部マスキング処理を施している。

Session Description

```
v=0
o=- 4 0 IN IP4 203.178.XXX.YYY
s=wifi-XXX-YYY.sfc.wide.ad.jp
c=IN IP4 239.69.XXX.YYY/15
t=0 0
m=audio 5004 RTP/AVP 97
c=IN IP4 239.69.XXX.YYY/15
a=rtpmap:97 L24/48000/2
a=sync-time:0
a=framecount:48
a=ptime:1
a=recvonly
```

本来 AES67 では、PTP (Precision Time Protocol) v2 経由でクロック同期する必要がある。今回の実装では PTPv2 によるクロック同期は行わないものとしているが、PTPv2 の同期が行う場合は、上記マニフェストに以下のマニフェストを追加する。

Session Description with PTPv2

```
a=mediaclock:direct=1810024580
a=ts-refclk:ptp=IEEE1588-2008:39-A7-94-FF-FE-07-CB-D:0
```

v SDP のプロトコルバージョン

o オリジン: [ユーザネーム] [セッション ID] [セッションバージョン] [ネットワークタイプ] [アドレスタイプ] [ユニキャストアドレス]

s セッション名

c コネクションデータ: [ネットワークタイプ] [アドレスタイプ] [コネクションアドレス]

t タイミング: [スタートタイム] [ストップタイプ]

m メディア説明: [メディア] [ポート] [プロトコル] [フォーマット]

a 属性: [属性名]:[値]

4.3 RTP 通信

AES67 のオーディオストリーム伝送には、RTP を用いる。RTP 通信を行うにあたり、GoRTP¹ ライブラリを利用した。後述する理由により、オリジナルの GoRTP から AES67 に必要な改変を行ったものを利用している²。

4.3.1 IP パケットにおける DSCP 値の付加

AES67 では、DiffServ ベースの QoS(Quality of Service) が必須条件に含まれている。QoS は、パケットに優先順位を付け、ネットワークスイッチ内でどのパケットの処理を優先するか決めるものだ。DiffServ は、IP ヘッダにおける DS フィールド中の DSCP 値を参照し、優先度のクラスによりキューイングするものである [13]。

GoRTP の実装では、UDP パケットのヘッダの値を変更することはできないため、AES67 で必須条件の DSCP 値を付加するための変更を行っている。

(1) QoS クラスと DiffServ アソシエーション

AES67 で使用されている QoS クラスと DiffServ アソシエーションを記す [10]。本実装では、以下のうち RTP パケットにおいて DSCP 値 34 を付加して送信している。

クラス名 Clock

トラフィックタイプ IEEE 1588-2008 *Announce, Sync, Follow_Up, Delay_Req, Delay_Resp, Pdelay_Req, Pdelay_RespandPdelay_Resp_Follow_Up* packets

デフォルト DiffServ クラス (DSCP10 進数値) EF(46)

クラス名 Media

トラフィックタイプ RTP and RTCP media stream data

デフォルト DiffServ クラス (DSCP10 進数値) AF41 (34)

¹wernerd/GoRTP <https://github.com/wernerd/GoRTP>

²itochan/GoRTP <https://github.com/itochan/GoRTP>

クラス名 Best effort

トラフィックタイプ IEEE 1588-2008 signaling and management messages. Discovery and connection management messages.

デフォルト DiffServ クラス (DSCP10 進数値) DF (0)

4.3.2 RTP パケットのペイロード

RTP パケットのペイロードは1 パケット中で 48 サンプル送る場合において、PCM 24bit/48kHz であれば 288 バイトとなる [14]。必要な RTP ペイロードの計算式を式 4.1 に示す。

$$P = bsc \quad (4.1)$$

- P : total payload bytes [byte]
- b : bit depth [byte] (24bit: 3[bytes])
- s : sampling rate [kHz] (48kHz: 48[kHz])
- c : audio channels [channel] (Stereo channel: 2[channel])

送信機能では、PCM で記録されたオーディオファイルを 288 バイトずつ取り出し、RTP ペイロードに載せたものを伝送する。

受信機能では、RTP ペイロードからオーディオを取り出し、保存する。

4.3.3 RTP パケットの送信間隔

RTP パケットを送信する間隔は、安定したオーディオストリームの送信において重要だ。本実験で実装したアプリケーション同士で送受信を行えば、パケットの間隔を気にせずとも受信できると考えられる。しかし、受信側が AES67 対応のオーディオ機器の場合において、同時に多くのパケットを受信するとバッファサイズを超え、処理能力の限界を超えることがある。一度に多くのオーディオストリームを受信し処理能力を超えると受信したオーディオストリームは切り捨てられ、音の断続的な途切れが発生する。

48kHz であれば 1 秒間に 48 サンプルを 1000 回送信することができれば、理想とする低遅延な伝送を行うことができる。実装したアプリケーションでは、1 ミリ秒ごとに送信する実装となっている。

第5章 評価

本章では、本研究における評価手法とその結果について述べる。

5.1 評価手法

4章で実装した AES67 送受信アプリケーションを使用する。2 台の PC でそれぞれ送信機能と受信機能を動かし、伝送にかかる遅延秒数を計測する。

5.2 評価環境

次の構成のコンピュータを 2 台、ネットワークスイッチを介して接続する。

今回使用するネットワークスイッチは、Dante 最適設定が搭載されている。Dante 最適設定は、有効にした状態で計測を行った。

表 5.1: コンピュータ評価環境

OS	Windows 10 Pro バージョン 1909
CPU	Intel Core i7-9700K 3.60GHz
RAM	32GB
SSD	Samsung SSD 970 EVO Plus 250GB
NIC	StarTech ST1000SPEXI (Intel I210-AT)

表 5.2: ネットワーク評価環境

ネットワークスイッチ	YAMAHA SWR2311P-10G
------------	---------------------

5.3 AES67 伝送の結果

5.3.1 UDP パケット転送遅延

最小 0.2ms、最大 1.3ms の遅延が発生した。

5.3.2 パケットドロップ率

パケットのドロップ率は、0%だった。

5.4 遅延の発生

実験に使用したアプリケーションでは、1 ミリ秒間隔で1 チャンネルあたり 48 サンプルずつ送信している。しかし、実験の結果から頻繁に1 ミリ秒を超える間隔でパケットが送信され、受信したオーディオストリームを再生すると途中で音声途切れる結果となった。

第6章 結論

6.1 本研究のまとめ

本研究では AES67 のソフトウェア実装が業務用オーディオシステムにおいて適用可能か検証するため、AES67 オーディオストリームの送受信アプリケーションを実装し、評価を行った。

6.2 本研究の結論

5 章で行った評価結果から、AES67 のソフトウェアオーディオ伝送は、2.7 節で述べた業務用オーディオにおける要件を満たすと考えられる。

6.3 今後の課題

6.3.1 RTP パケット送信間隔の適正值

5.4 節で、1 ミリ秒間隔で 1 チャンネルあたり 48 サンプルずつ送信した結果遅延が発生したことを述べた。

1 ミリ秒のタイマーを設定し、1 ミリ秒間隔で送信できないのはコンピュータのタイマ性能に起因するものではないかと考えられる。したがって、5 ミリ秒間隔で 240 サンプルずつ、10 ミリ秒間隔で 480 サンプルずつ、といったように同時に送信するサンプル数を増やすことで安定した間隔で行えるのではないだろうか。

6.3.2 伝送経路の冗長化

業務用オーディオでは、音を高音質で伝送するとともに、確実に伝送することが必要である。SMPTE は、ST2110[15] というオーディオとビデオをまとめて送る規格を策定した。その中に、冗長化に関する仕様が定められている。冗長化を行うことで、より業務用途に適した IP オーディオ伝送が実現するだろう。

6.3.3 ハードウェア実装

AES67 の伝送を、ハードウェアで実装することでより高速に伝送できるのではないかと考えている。Field Programmable Gate Array (FPGA) で実装することを今後検討したい。

謝辞

本研究を進めるにあたり、ご指導いただいた慶應義塾大学環境情報学部教授 村井純博士、同学部教授 中村修博士、同学部教授 楠本博之博士、同学部教授 高汐一紀博士、同学部教授 Rodney D. Van Meter III 博士、同学部准教授 植原啓介博士、同学部教授 三次仁博士、同学部教授 中澤仁博士、同学部教授 武田圭史博士、同大学政策・メディア研究科特任准教授 佐藤雅明博士、同大学政策・メディア研究科特任教授 鈴木茂哉博士、同大学 SFC 研究所上席所員 斉藤賢爾博士に感謝いたします。Arch 研究グループでは、日頃より研究をご指導いただきました松谷健史博士、空閑洋平博士、大江将史博士に感謝いたします。特に、村井純博士は研究活動以外においても未熟な私に人生や恋愛とは何かということを教えていただきました。

本研究のきっかけとなったヤマハ株式会社 深沢豪氏、山根章生氏からは製品を開発する立場からのご意見をいただくことができ、大いに助けられました。重ねて感謝申し上げます。

プロフェッショナルオーディオへの興味を持つきっかけとなった、福利厚生団体音像工房での活動で知り合った金子浩幸氏は私に多くの音響や映像の仕事を紹介していただきました。また、大学のイベントで機材の特別貸出を実施していただいた、慶應義塾大学 湘南藤沢メディアセンター および同メディアセンター マルチメディアサービス担当 長坂功氏には深く感謝申し上げます。

村井・楠本・中村・高汐・バンミーター・植原・三次・中澤・武田合同研究プロジェクトに所属している学部生、大学院生、卒業生の皆さまに感謝いたします。長い研究室生活で多くの時間を過ごした、菅藤佑太氏、安井瑛男氏、石川達敬氏、阿部涼介氏、豊田安信氏、Korry Luke 氏、小西遼氏、矢内洋祐氏、用澤玄汰氏、真島大樹氏、山田真也氏、橋直雪氏、水野史暁氏、平野孝徳氏、さらには1年休学したため一緒に卒論を戦った栗原祐二氏、深川祐太氏、鈴木雄祐氏、島津翔太氏、井手田悠希氏、勝又海氏、西田亘氏、山田航太郎氏に感謝いたします。楽しいことも苦しいことも、どんな出来事でも包み隠さずに話したり相談できる大切な仲間です。また、研究室を日々きれいに使うとともに居心地をよくしてくれた学部生と大学院生の皆さまには大変感謝しています。本論文執筆に際して L^AT_EX テンプレートを作成し研究室内で知見を共有してくれた、研究室の同期と先輩方に感謝申し上げます。

研究室に所属している間はずっと Arch 研究グループにいましたが、ICAR 研究グループの皆さんと佐藤雅明博士とは合宿を通じて夜通し語り合ったり、ICAR に所属していないながらも仲良くしていただきました。

本研究は慶應義塾大学環境情報学部、つまり SFC (湘南藤沢キャンパス) に入学していなければ、おそらく行うことはなかったのではないかと思います。AO (アドミッションズ・オフィス) 入試を受けるにあたり、当時通っていた東京都立一橋高等学校で3年次担任の榎本津加子氏、高校から大学までの5年間アルバイトをしていたピクシブ株式会社の小芝敏明氏には私の推薦文を

書いていただくなど、多大なるご支援をいただきました。お二方がいなければ、今の SFC での生活と研究活動はありませんでした。深く感謝いたします。

中学卒業後、最初に入った高校は通信制の高校でした。1 年通って転学することになってしまいましたが、SFC にいる今の私がいるのはそのとき転学を決断していたのも関係しています。転学の前に通っていた科学技術学園高等学校で担任だった金子知史氏には、転学するにも関わらず相談やその後押しをしてくださり、感謝しています。

また、入学に関して欠かすことのできなかったものが奨学金です。奨学金なくして大学に進学することは不可能でした。高校で推薦をいただき、公益財団法人日本教育公務員弘済会 東京支部から 20 万円、公益信託中村奨学基金から 18 万円（いずれも給付奨学金）のご支援をいただきました。さらに、独立行政法人日本学生支援機構からは入学時特別増額貸与奨学金 50 万円と、第二種奨学金毎月 12 万円（機関保証）48 ヶ月の貸与奨学金を利用しています。奨学金受け取りに関して、推薦をしていただいた東京都立一橋高等学校教諭の皆さま、公益信託中村奨学基金の審査を行った三菱 UFJ 信託銀行 リテール受託業務部公益信託課の皆さま、日本学生支援機構 入学時特別増額貸与奨学金のつなぎ融資をしていただいた中央労働金庫 新宿支店の皆さまにはお礼申し上げます。

私が SFC に通い続けた、学部の 4 年間と 1 年間の休学で得られた体験は一生の宝物です。印象深い授業や、研究室やサークルを通して出会えた友人からは多くの刺激を受けるとともに深く考えさせられました。SFC で記憶に残る授業「データ・ドリブン社会の創発と戦略」「データ・ドリブン社会の創発と戦略（応用）」を通じて知り合うことになった、慶應義塾大学環境情報学部教授安宅和人博士からは、激動の時代を生き抜く知恵と勇気をいただくことができました。私が所属していたサークル・団体、福利厚生団体音像工房、SFC CLIP、アカペラシンガーズ K.O.E. の皆さまには大変お世話になりました。

1 年次の必修授業「情報基礎」で SA（スチューデント・アシスタント）をしていたのもいい経験でした。3 年間・6 学期と計 8 クラスも受け持つことになりました。授業の担当講師であった独立行政法人産業技術総合研究所 橋本尚久博士、慶應義塾大学理工学部 栗原聡博士、同大学政策・メディア研究科 阿部涼介氏は、SA として私を迎え入れてくれるのみならず、さまざまな相談にも乗っていただきました。一緒に SA/TA をしていた、尾崎周也氏、高塚大暉氏、深川祐太氏、山口航平氏、加藤太一氏、大塚崇夫氏、鈴木雄祐氏、栗原祐二氏、種谷望氏と教えた日々もいい思い出です。ここで出会えた履修者も、キャンパスで出会ったときに声をかけてくれるなど、教えることができてよかったと思う次第です。

そして、語らずにはいられない何よりも忘れられないことがあります。最後まで成し遂げられなかった SFC でしか、私にしか味わうことのできない恋愛の思い出です。これから先の人生でどんなことがあっても、私の心には残り続けることでしょう。ここに、好きな人の名前を書ければカッコよかったのかもしれませんが、しかし残念ながら、それが叶うことはありませんでした。

最後に、すべての人の名前をここで挙げることはできませんが、大学で出会うことができたすべての人に感謝します。SFC で出会えたみんなが大好きです。本当にありがとうございました。

参考文献

- [1] Dennis Born. Interfacing AES3 and S/PDIF. <https://www.rane.com/note149.html>, 2001.
- [2] 株式会社コンテック. アナログ入出力 - マルチファンクション — コンテック. <https://www.contec.com/jp/support/basic-knowledge/daq-control/analog-io/>.
- [3] Audio Engineering Society and American National Standards Institute. *A.E.S. Recommended Practice for Digital Audio Engineering: Serial Transmission Format for Linearly Represented Digital Audio Data*. American National Standards Institute standard. Audio Engineering Society, 1985.
- [4] European Broadcasting Union. Specification of the digital audio interface (The AES / EBU interface). pp. 1–20, 2004.
- [5] 株式会社トライテック. コラム : AES3ID-1995 / SMPTE-276M. <https://www.tritech.tv/column/aesebu.html>.
- [6] Nicolas Bouillot, Elizabeth Cohen, Jeremy R. Cooperstock, Andreas Floros, Nuno Fonseca, Richard Foss, Michael Goodman, John Grant, Kevin Gross, Steven Harris, Brent Harshbarger, Joffrey Heyraud, Lars Jonsson, John Narus, Michael Page, Tom Snook, Atau Tanaka, Justin Trieger, and Umberto Zanghieri. Best Practices in Network Audio. *J. Audio Eng. Soc.*, Vol. 57, No. 9, p. 735, 2009.
- [7] Henning Schulzrinne, Stephen L. Casner, Ron Frederick, and Van Jacobson. RTP: A Transport Protocol for Real-Time Applications. RFC 3550, July 2003.
- [8] Stephen L. Casner and Henning Schulzrinne. RTP Profile for Audio and Video Conferences with Minimal Control. RFC 3551, July 2003.
- [9] David Meyer, Michelle Cotton, and Leo Vegoda. IANA Guidelines for IPv4 Multicast Address Assignments. RFC 5771, March 2010.
- [10] Audio Engineering Society. *AES67-2018: AES standard for audio applications of networks - High-performance streaming audio-over-IP interoperability*.
- [11] Mark J. Handley, Colin Perkins, and Edmund Whelan. Session Announcement Protocol. RFC 2974, October 2000.

- [12] Colin Perkins, Mark J. Handley, and Van Jacobson. SDP: Session Description Protocol. RFC 4566, July 2006.
- [13] ヤマハ株式会社. DiffServ ベース QoS. http://www.rtpro.yamaha.co.jp/RT/docs/qos/diffserv_qos.html.
- [14] Thomas Edwards. Listen to Audio from AES67/ST 2110-30 PCAP Files — LinkedIn. <https://www.linkedin.com/pulse/listen-audio-from-aes67st-2110-30-pcap-files-thomas-edwards/>, 2018.
- [15] Society of Motion Picture and Television Engineers. *SMPTE 2110*.