

Digitale Audiosignale in Netzwerken

Bachelorarbeit

im Studiengang

Audiovisuelle Medien

vorgelegt von

David Benjamin Lock

Matr.-Nr.: 30416

am 28. Mai 2019

an der Hochschule der Medien Stuttgart

Erstprüfer: Prof. Oliver Curdt

Zweitprüfer: Andreas Bohr

Ehrenwörtliche Erklärung

Hiermit versichere ich, David Lock, ehrenwörtlich, dass ich die vorliegende Bachelorarbeit mit dem Titel: „Digitale Audiosignale in Netzwerken“ selbstständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden.

Ich habe die Bedeutung der ehrenwörtlichen Versicherung und die prüfungsrechtlichen Folgen (§ 26 Abs. 2 Bachelor-SPO (6 Semester), § 24 Abs. 2 Bachelor-SPO (7 Semester), § 23 Abs. 2 Master-SPO (3 Semester) bzw. § 19 Abs. 2 Master-SPO (4 Semester und berufsbegleitend) der HdM) einer unrichtigen oder unvollständigen ehrenwörtlichen Versicherung zur Kenntnis genommen.

Stuttgart, den 27.05.2019

(David Lock)

Kurzfassung

Diese Arbeit beschäftigt sich mit aktuell eingesetzten Audioprotokollen, die die digitale Übertragung von Audiosignalen auf Netzwerkbasis ermöglichen. Hierbei sind Protokolle gewählt, die stark am Markt vertreten sind und durch Standards eine herstellerübergreifende Interoperabilität versprechen. Um ein grundlegendes Verständnis für die Interoperabilität zwischen den Protokollen zu schaffen, werden die dafür zuständigen Standards aufgeführt. Protokolle der Hersteller Audinate und ALC NetworX oder der Milan-Initiative werden auf funktionaler Ebene erläutert und im Anschluss anhand von technischen und praktischen Merkmalen verglichen. Besondere Beachtung findet dabei die Echtzeitfähigkeit und die Latenz. Aus diesen werden Vor- sowie Nachteile der Technologien gezogen. Ziel dabei ist es, einen Überblick über den Stand vorhandener Technologien sowie deren Interoperabilität zu erhalten und einen möglichen Trend in der weiteren Entwicklung am Markt festzustellen. Die Schwierigkeit besteht hierbei darin, einen Vergleich trotz unterschiedlicher Ansätze der Übertragungswege anzustellen. Dies wird durch eine Einteilung der Einsatzmöglichkeiten und Einordnung der Protokolle berücksichtigt.

Inhaltsverzeichnis

Ehrenwörtliche Erklärung	II
Kurzfassung	III
Inhaltsverzeichnis	IV
Abbildungsverzeichnis	VI
Tabellenverzeichnis	VII
Abkürzungsverzeichnis	VIII
1 Einleitung	1
2 Grundlagen der digitalen Audioübertragung	4
2.1 Digitalisierung eines Audiosignals	4
2.2 PCM-Signal	5
2.3 AES-3	6
3 Grundlagen der Netzwerktechnik	8
3.1 OSI-Modell	10
3.1.1 Layer-1: Bitübertragungsschicht	11
3.1.2 Layer-2: Sicherungsschicht	11
3.1.3 Layer-3: Vermittlungsschicht	12
3.1.4 Layer-4: Transportschicht.....	12
3.2 Ethernet nach IEEE802.3	13
3.2.1 CSMA/CA	15
3.2.2 Ethernet Frame.....	16
3.3 Protokolle	16
3.3.1 Das Internet Protocol	17
3.3.2 UDP	19
3.3.3 RTP	20
3.4 Komponenten	22
3.4.1 Switch	23
3.4.2 Router	25
3.5 Unicast, Broadcast, Multicast	27
3.6 Quality of Service (QoS)	28
3.6.1 Netzwerksicherheit	30
3.6.2 Latenz	31
4 Übertragung von Audiosignalen in Netzwerkstandards	33
4.1 Interoperabilität in Netzwerken	33

4.2	AES67	34
4.2.1	Synchronisation und Media Clock	34
4.2.2	Transport & QoS	36
4.2.3	Codierung & Stream	37
4.2.4	Sitzung	38
4.2.5	Discovery (Geräteerkennung) & Connection Management (Verbindungsmanagement)	38
4.3	SMPTE ST 2110	39
4.3.1	Synchronisation und Media Clock	40
4.3.2	Transport & QoS	40
4.3.3	Codierung & Stream	41
4.3.4	Sitzung	42
4.3.5	Discovery (Geräteerkennung) & Connection Management (Verbindungsmanagement)	42
4.4	Audio Video Bridging – AVB/TSN	42
4.4.1	Synchronisation und Media Clock	44
4.4.2	Transport & QoS	45
4.4.3	Codierung & Stream	46
4.4.4	Sitzung	47
4.4.5	AV Discovery, Enumeration, Connection Management and Control	47
4.5	NMOS IS-04/IS-05 by AMWA/ AES70	48
5	Betrachtung der Audionetzwerke	51
5.1	RAVENNA AES67	51
5.1.1	Technische Merkmale.....	52
5.1.2	Netzwerkeigenschaften.....	52
5.1.3	Betriebs-/Erweiterungsmöglichkeiten & Kompatibilität.....	53
5.2	Dante(-AV)	55
5.2.1	Technische Merkmale.....	56
5.2.2	Netzwerkeigenschaften.....	56
5.2.3	Betriebs-/Erweiterungsmöglichkeiten	57
5.3	AVB/Milan	59
5.3.1	Technische Merkmale.....	59
5.3.2	Netzwerkeigenschaften.....	60
5.3.3	Betriebs-/Erweiterungsmöglichkeiten	61
6	Allgemeiner Vergleich/Bewertung	62
7	Zusammenfassung und Ausblick	72
8	Glossar	76
9	Quellenverzeichnis	77
10	Anhang A: E-Mail Verkehr	83
11	Anhang B: Bewertungstabelle	85

Abbildungsverzeichnis

Abbildung 2-1 Aufbau eines 32-Bit AES-3 Pakets	7
Abbildung 3-1: Schematische Darstellung einer Netzwerkarchitektur	13
Abbildung 3-2: Frame Format für Ethernet.....	16
Abbildung 3-3: Datagram-Format für das IP	19
Abbildung 3-4: UDP Segment-Format.....	20
Abbildung 3-5: Aufbau eines RTP-Headers.....	21
Abbildung 3-6: Beschreibung verwendeter Netzwerkgeräte durch Schichtzuordnung	23
Abbildung 3-7: Paketversandt Unicast – Multicast.....	27
Abbildung 3-8: Toleranzbereich von verzögerten Video- und Audiodaten	32
Abbildung 4-1: Synchronisation im Precision Time Protocol.....	35
Abbildung 4-2: AVB-Protokollstack.....	44
Abbildung 5-1: Bedienoberfläche der ANEMAN-Oberfläche.....	54
Abbildung 5-2: GUI des Dante Controllers.....	58
Abbildung 7-1: Roadmap der JT-NM	74

Tabellenverzeichnis

Tabelle 1: Das OSI-Schichtenmodell	11
Tabelle 2: Sammlung aus gängigen Ethernet-Standards	15
Tabelle 3: Priorisierung der Datenströme.....	30
Tabelle 4: Priorisierung der Streams in AES67.....	37

Abkürzungsverzeichnis

SMPTE	Society of Motion Picture and Television Engineers
AAF	Advanced Authoring Format
AES	Audio Engineer Society
AMWA	Advanced Media Workflow Association
AoE	Audio Over Ethernet
AoIP	Audio Over IP
AVB	Audio Video Bridging
BMC	Best-Master-Clock
CSMA/CD	Carrier Sense Multiple Access /with Collision Detection
DDM	Dante Domain manager
DiffServ	Differentiated Services
DSCP	Differential Service Code Point
EBU	European Broadcast Society
HDM	Hochschule der Medien
IP	Internet Protocoll
IEEE	Institute of Electronics and Engineers
kHz	Kilohertz (physikalische Einheit)
LAN	Local Area Network
MAC	Multi Access Control
MAN	Metropolitan Area Network
NMOS	Networked Media Open Specifications
PDU	Protocol Data Units
PTP	Precision Time Protocol

QoS	Quality of Service
PCM	Pulse Code Modulation
RFC	Request for Comments
RTP	Real-Time Transport Protocol
SDP	Session Description Protocol
SDVoE	Software Defintion over Ethernet
SMPTE	Society of Motion Picture and Television Enigneers
ToS	Type of Service
TSN	Time-Sensitive Network
VLAN	Virtuelles LAN
WAN	Wide Area Network

1 Einleitung

Seit dem Vorstoß des Internets in den Zivilbereich und in die Wirtschaft am Anfang des 21. Jahrhunderts, ist dieses im Alltag angekommen. Audio- und Videostreams finden täglich über mobile Endgeräte wie das Smartphone und Personal Computer statt, unsere Kommunikation ist größtenteils digitalisiert und mit dem Internet of Things findet eine Vernetzung im Wirtschaftsbereich statt.

Diese Entwicklung findet ebenso im Bereich der Medienproduktion statt. Bis Ende des 20. Jahrhunderts fand die Übertragung von Audiodaten mit einfachen Punkt-zu-Punkt Verbindungen über elektrische Leiter statt. Hierfür wurden einfache unidirektionale XLR-Kabel verwendet, die auf große Patchbays liefen und über Kreuzschienen eine Zuweisung von Quelle und Ziel ermöglichten. Mit der Digitalisierung der Audiotechnik, welche durch den Marktstart der Audio-CD in den 80er Jahren vorangetrieben wurde, entstanden neue Workflows und die Möglichkeit, Audiodaten auf digitalen Trägern zu sichern. Bis zum Anfang des 21. Jh. wurde eine Kombination aus analogen Verbindungen und der digitalen Speicherung verwendet. Mit der Einführung des AES10 Standards MADI entstand die Möglichkeit, mehrere Kanäle digitalisierter Audiosignale über ein einziges Kabel zu übertragen. Dies bildete eine Alternative zu unhandlichen und teuren Multicore Kabeln. In den Folgejahren entstanden die ersten „Audionetzwerke“ mit Technologien wie AES50 oder Aviom Pro64. Diese bildeten mit dem Transport von Audio über das physische Ethernet Layer die ersten OSI-Layer-1 Netzwerke und benutzen Cat5-Ethernetkabel oder Lichtwellenleiter. Sie nutzten Ende-zu-Ende Verbindungen oder wurden in Reihe geschaltet und erforderten ein Routing. Diese Lösungen waren mehr Protokolle zur Organisation von Datenströmen als wirkliche Audionetzwerke.

Mit Sicht auf das immer bedeutendere Internet entstand das Bedürfnis nach einer Transportform, die Signale selbstständig in einem Netzwerk transportiert und über große Strecken transportiert. Mit der vorhandenen Ethernet-Technologie wurden Lösungen zum Transport über Bridges und Switche entwickelt. Daraus entstanden Lösungen zum Layer-2 Transport in kleinen lokalen Netzwerken und später der Transport über das Internet Protocol, das Daten nun über weitere Strecken transportiert.

Knapp 15 Jahre nach Entwicklung der ersten Ethernet-fähigen Audionetzwerke gab es von vielen Herstellern unterschiedliche proprietäre Lösungen zur Übertragung von Au-

diodateien. Dabei wurden diese von jedem Hersteller auf eigene Weise umgesetzt und es gab viele unabhängige Systeme. Um nicht auf eine einzige Systemlösung angewiesen zu sein, wurden Standards zur gemeinsamen Interoperabilität entwickelt. Die Audio Engineering Society (AES) entwickelte 2010 einen gemeinsamen Standard zur Interoperabilität beim Audiotransport, AES67. Auf diesem aufbauend, wurde von der SMPTE der Standard ST2110 entwickelt. Dieser Standardisiert den Transport von Video und Audio. Zeitgleich wurde AVB von der IEE 802.1 Group zum Transport von zeitkritischen Audio- und Videodaten aufgebaut und bildet die Grundlage für ein Netzwerk.

In Zukunft wird die Übertragung von Audiodateien in digitalen Netzwerken stattfinden. So stellte eine statistische Erhebung von RH Consulting einen Anstieg der netzwerkfähigen Produkte im Jahr 2013 von 398 auf 1583 Produkte im Jahr 2017 fest¹. Es ist noch unklar, ob weiterhin proprietäre Lösungen verwendet werden oder diese sich unter einem Standard vereinen. Ebenfalls ist die Netzwerktechnik nicht auf den Transport von zeitkritischen Daten ausgelegt. Aus diesem Grund wird eine Übersicht über die in den Standards verwendeten Protokolle und deren Funktionsweise gegeben. Um eine gerechte Bewertung der Interoperabilitäts-Standards zu geben, wird deren Aufbau und Funktion erklärt und nach einem Vergleich wird eine Bewertung dieser durchgeführt.

Diese Arbeit analysiert die auf den Standards beruhenden Audioprotokolle und bewertet diese nach einem technischen Vergleich. Hierbei sollen die Anwendungsgebiete sowie die Potenz der Protokolle nach dem Vergleich herausgearbeitet werden. Der Vergleich wird nach technischen und praktischen Merkmalen durchgeführt. Bestandteile dessen sind übertragbare Formate, die mögliche Größe der Netzwerke, Mechanismen, um Stabilität und sichere Übertragung zu gewährleisten und der gebotene Komfort. Ebenfalls wird die Echtzeitfähigkeit unter dem Aspekt der Übertragungsdauer geprüft und individuell auf die im Audionetzwerk vorhandene Latenz eingegangen.

Hierfür werden die Lösungen Dante (Audinate), Ravenna (ALC NetworX) und Milan (auf AVB basierend) betrachtet. Diese Netzwerklösungen stellten 2018 mit Cobranet die meist verfügbaren Geräte am Markt. Während die Produktverfügbarkeit von Ravenna und Dante im Zeitraum 2014 bis 2018 stetig anstieg, war bei Cobranet eine Stagnati-

¹ <http://rhconsulting.eu/blog/files/NetworkedAudio.html>

on zu beobachten. Dies ist ebenso bei AVB-fähigen Geräten zu beobachten.² Die Technologie wird dennoch in dieser Arbeit mit einbezogen, da mit d&b Audiotechnik, Meyer Sound und L-Acoustics eine markttreibende Kraft hinter dem 2018 veröffentlichten Milan System steht.

Am Ende der Arbeit wird neben der Einsatzmöglichkeit der Audionetzwerke auch auf deren Fähigkeit der Videoübertragung eingegangen. Ebenso wird eine Aussicht auf kommende Entwicklungen in diesem Bereich erläutert.

Ziel der Arbeit ist es, eine Übersicht über die vorhandenen Standards und daraus entwickelten Netzwerkprotokolle zu geben und einen Ausblick auf einen möglichen Favoriten zu geben, welcher für eine gemeinsame Interoperabilität zwischen den Netzen sorgt.

² <http://rhconsulting.eu/blog/files/NetworkedAudio.html>

2 Grundlagen der digitalen Audioübertragung

Das folgende Kapitel behandelt die für diese Arbeit relevanten Grundlagen der Audioübertragung, welche für die Netzwerktechnik wichtig sind. Es behandelt die Wandlung des Signals und die Formate, welche in Netzwerken übertragen werden. Dazu zählt das unkomprimierte digitale Audiosignal sowie das AES-3 (Audio Engineering Society) Format. Beide werden in der paketbasierten Übertragung in Netzwerken benutzt.

2.1 Digitalisierung eines Audiosignals

Am Anfang eines digitalen Tonsignals steht immer ein analoges Signal. Dieses ist eine wert- und zeitkontinuierliche elektrische Spannung. Da Computer mit diskreten Binärsignalen arbeiten, muss das kontinuierliche Signal (fortlaufend, ohne Abstufungen) in ein solches, diskretes, gewandelt werden.

Eine kontinuierliche Größe ist eine physikalische Größe, die in ihrem physikalisch-technischem System jede beliebige Größe annehmen kann, deren Amplitude durch die sogenannte Vollaussteuerung begrenzt wird und welche durch eine Funktion bestimmt werden kann.³

Eine diskrete Größe lässt sich in eine Zahlenfolge einordnen und wird durch diese messbar. So ist ein diskretes Signal eine Folge, welche durch regelmäßige Abtastung in Zahlenwerte umgewandelt wird. Jeder Abtastwert des Signals wird „Sample“ genannt. Die Größe des Samples wird durch die Abtasttiefe bestimmt.⁴

Wie oben aufgeführt, ist das Audiosignal zeit- und wertkontinuierlich. So wird bei der A/D-Wandlung das Signal durch Abtastung zeitdiskret gemacht. Dies geschieht durch die Unterteilung in kleinere Zeitblöcke, welche werden von der Abtastfrequenz bestimmt. Diese wiederum entsteht durch das abzutastende Frequenzspektrum und das Shannonsche-Abtasttheorem. Dieses besagt, dass die Abtastrate f_s größer als das Doppelte der höchsten wiederzufindenden Frequenz f_{\max} sein muss, somit folgt: $f_s > 2 f_{\max}$.

³ Vgl. Dickreiter, Michael et al. (2014): Handbuch der Tonstudioteknik. Band 2, Berlin/Boston, Walter de Gruyter GmbH, S. 659 f.

⁴ Vgl. Dickreiter (2014), S. 663

Das menschliche Gehör nimmt Schall im Bereich zwischen 20Hz und 20.000 Hz, also 20kHz wahr. Somit gilt: $f_s > 2 * 20.000\text{Hz} \rightarrow f_s > 40.000 \text{ Hz}$.⁵

Im Laufe der Zeit haben sich die Samplerates (Abtastfrequenzen) von 44,1kHz und 48kHz, sowie deren Vielfachen von 96kHz und 192kHz in der professionellen Audio-technik durchgesetzt.

Nach dieser Wandlung ist ein zeitdiskretes sowie ein wertekontinuierliches Signal vorhanden. Die vorhandenen analogen Augenblickswerte müssen nun in diskrete Bitwerte umgewandelt werden. Diesen Vorgang nennt man Quantisierung.

Die Quantisierung erfolgt durch den Vergleich der analogen Abtastwerte mit einer gestuften Skala. Diese Skala entspricht der Abtasttiefe, die die Genauigkeit des abgebildeten Signals bestimmt und nimmt einen Bitwert n an, welcher 2^n -Stufen erzeugt.⁶ Typische Quantisierungswerte in der heutigen Zeit sind 16, 24 oder 32Bit, wobei in der Aufnahme eines Signals meist mit 24 oder 32 Bit gearbeitet wird. Dies ermöglicht durch eine höhere Auflösung eine präzisere Bearbeitung des Signals.

2.2 PCM-Signal

Das Pulse Code Modulation-Signal (PCM) bildet „[...] binäre Signalfolgen mit konstantem Zeitintervall zwischen den Elementarsignalen (Datenbits oder -worte), die das analoge Tonsignal linear abbilden.“⁷

Als bekanntestes Format bildet es die Basis für Digitalaudioanwendungen, etablierte sich jedoch durch den Standard als CD-Format Anfang der 80er.

Der Prozess der Pulse Code Modulation wird durch ein Begrenzen des Eingangssignals mit einem Tiefpassfilter und dem oben beschriebenen Abtasten und Quantisieren beschrieben. Danach werden die ermittelten digitalen Tastwerte über eine Codierung in ein einheitliches Digitalsignal umgewandelt. Bei der Codierung wird der Binärcode gewählt, welcher eine digitale Übertragung von 1en und 0en ermöglicht. Das PCM-Verfahren ist ein Quellencodierungsverfahren, es liefert also das „rohe“ Quellsignal und dessen Eigenschaften ohne Informationsverlust über dieses.

⁵ Vgl. Zölzer, Udo (2005): Digitale Audiosignalverarbeitung, Wiesbaden, Teubner Verlag/ GWV Fachverlage GmbH, S. 67 f.

⁶ Vgl. Dickreiter (2014), S. 667 f.

⁷ Dickreiter (2014), S. 832 f.

Am häufigsten ist das PCM-Signal bei Audiodateien im WAV-Containerformat (.wav) vorzufinden.⁸

Datenrate eines Stereosignals:

- 16 Bit, 44,1 kHz: $16 \text{ Bit} * 44.100 \text{ Samples/s} * 2 \text{ Kanäle} = 1,411 \text{ Mbit/s}$ (CD-Qualität)
- 16 Bit, 48 kHz: 1,536 Mbit/s
- 24 Bit, 48 kHz: 2,304 Mbit/s (Studio-Qualität)
- 24 Bit, 96 kHz: 4,608 Mbit/s

2.3 AES-3

Das digitale AES-3 Tonformat ist von der AES und EBU (European Broadcast Society) genormt, umfasst zwei Tonsignale und arbeitet ohne Datenreduktion. Es wird zur Übertragung digitaler Signale zwischen Audiogeräten über die AES/EBU-Schnittstelle benutzt. Die Übertragung findet mit symmetrischen Leitungen und XLR-Steckern statt, ist jedoch auch mit Koaxialkabel oder einem Netzkabel möglich. Die Datenrate ist mit 48 kHz und 24 Bit genormt, jedoch sind auch Abtastraten von 32, 44,1, 96 und 192 kHz, sowie Auflösungen von 16 Bit und 20 Bit möglich. Mit einer Erhöhung der Abtastrate muss eine Senkung der Auflösung verbunden sein, so dass die 2,304 Mbit/s bei einem Stereosignal nicht überschritten werden.

Die kleinste Übertragungseinheit des AES-3 Signals ist eine 32 Bit große Gruppe, diese ist schematisch in Abbildung 2-1 abgebildet. Davon bilden die ersten vier Bits eine Präambel. Diese dient zu Synchronisierung. Dem angeschlossen folgt ein 24 Bit langes Audiosample (PCM codiert). Diesem sind 4 weitere Bits angehängt. Das Validity-bit (V), das signalisiert, ob die übertragenen Audiodaten ein gültiges Signal darstellen. Das User Data Bit (U), welches für einen unabhängigen Bitstrom verwendet wird. Das Channel-Status-Bit (C), über welches der Kanalstatus vermittelt wird. Das letzte Bit ist das Paritäts-Bit, welches zur Überprüfung der korrekten Übertragung von Bit 4-31 be-

⁸ Vgl. Dickreiter (2014), S. 674, S. 685

nutzt wird.

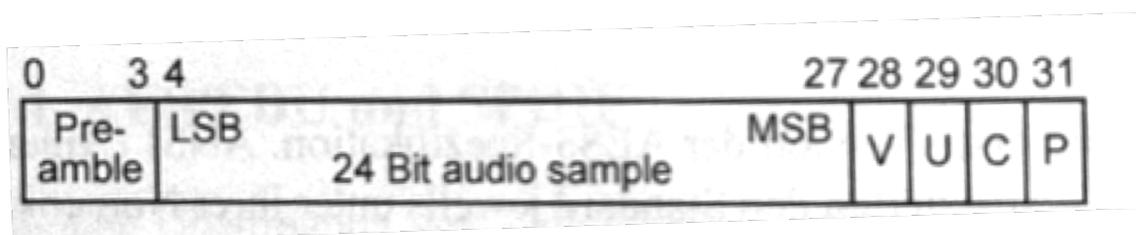


Abbildung 2-1 Aufbau eines 32-Bit AES-3 Pakets⁹

Zwei dieser Gruppen bilden einen Rahmen. Die beiden Gruppen sind i. d. R. der rechte und linke Kanal des Stereosignals. 192 Rahmen werden in einem Block zusammengefasst. Aus den einzelnen Channel-Status-Bits des Blocks wird die 24 Byte große Kanalstatusinformation gebildet.¹⁰

Durch die 8 Zusatzbits ergibt sich aus der Datenrate von 2,304 Mbit/s bei einem 24 Bit /48 kHz PCM-Signal ein AES-3 Signal mit einer Datenrate von 3,072 Mbit/s.¹¹

⁹ Dickreiter (2014), S. 690

¹⁰ Vgl. Dickreiter (2014), S. 690 f.

¹¹ <https://www.bet.de/lexikon/aes3/>

3 Grundlagen der Netzwerktechnik

Um die Funktionsweise eines Audionetzwerkes betrachten zu können, müssen zunächst die Grundlagen des Datentransportes und die Verwendung der verschiedenen Standards und Protokolle, sowie der grundsätzliche Aufbau eines Netzwerkes bekannt sein. Auf diese Themen wird in diesem Kapitel eingegangen. Ziel hierbei ist es nicht die kompletten Grundlagen der Netzwerktechnik zu vermitteln. Es werden die grundlegenden Funktionen für die später ausgeführten Audionetzwerke aufgeführt und erläutert, so dass ein einfacherer Vergleich der Netzwerke miteinander möglich ist.

Nach Schreiner ist ein Netzwerk: „eine Infrastruktur, die Datenendgeräten die Kommunikation, den Datenaustausch und die Nutzung gemeinsamer Ressourcen transparent ermöglicht.“¹². Dies bedeutet konkret, dass der Benutzer sich nicht um die zu versendenden Daten und deren Transportwege kümmern muss und somit alles automatisiert abläuft¹³.

Die Größe eines Netzwerkes wird durch ihren Nutzen und ihren Einsatzbereich bestimmt. So gibt es das Local Area Network (LAN), eine Netzwerkinfrastruktur, welche sich über kleinere Bereiche erstreckt und Arbeitsgeräte miteinander verbindet und ist meist auf Gebäude beschränkt. Das Metropolitan Area Network (MAN) verbindet größere Bereiche, die sich über Gebäude und Straßen hinweg erstrecken und liegt mit seiner Größe zwischen dem LAN und dem Wide Area Network (WAN). Das WAN erstreckt sich über große geographische Bereiche und verbindet einzelne Geräte sowie lokale Netzwerke miteinander.¹⁴

Ein Audionetzwerk agiert als Medienübertragungsnetz in einem besonders kritisch zu betrachtendem Umfeld und stellt besondere Grundkriterien zur Übertragung der Daten. Diese spielen eine Rolle für die unten folgende Betrachtung der Netzwerkfunktionen. Die Übertragung in einem solchen digitalen Netzwerk benötigt eine sichere und zuverlässige Infrastruktur, über welche die Datenströme fließen. Es ist es wichtig, dass die

¹² Schreiner, Rüdiger (2016): Computernetzwerke. Von den Grundlagen zur Funktion und Anwendung, München, Carl Hanser Verlag, S.3

¹³ Vgl. Schreiner (2016), S.3

¹⁴ Vgl. Scherff, Jürgen (2010): Grundkurs Computernetzwerke. Eine kompakte Einführung in Netzwerk- und Internet-Technologien, Wiesbaden, Vieweg + Teubner Verlag, S. 29

Audiodaten jederzeit an jedes der im Netzwerk angeschlossenen Geräte geschickt werden können, so dass diese wiedergegeben und/oder gleichzeitig verarbeitet werden können. Durch klassische Punkt zu Punkt Verbindungen ist dies nicht mehr gegeben. Um nicht eigene Router und Techniken zu entwickeln, wird bei den meisten Audionetzwerken auf die vorhandene Infrastruktur des Ethernets zurückgegriffen. Mit dieser werden Signale durch Netzwerkschicht und Router intelligent und deterministisch weitergeleitet (Siehe Kap. 3.4). Die Verarbeitung der Signale sollte ebenfalls im lokalen Netzwerk geschehen, da sonst die Gefahr entsteht größere Verzögerungen in der Datenübertragung zu erzeugen. Da die Übertragung der Signale meist nicht nur der Wiedergabe, sondern auch der Echtzeitkommunikation dient und eine Anforderung der Live-Produktion an Echtzeitnetzwerken besteht, muss diese zeitkritisch geschehen.

Dies bedeutet, dass die Übertragungsdauer zwischen Eingangssignal und Ausgangssignal nur wenige Millisekunden betragen darf. In einer Studie der AES wurde dieser Zusammenhang in Bezug auf die subjektive Störung einer Latenz bei Musiker*innen in Live-Umgebung untersucht. Die Studie kam auf einen Wert von 6,5ms. Ab diesem führte das verzögerte Signal bei den Musiker*innen zu einer spürbaren Verzögerung und subjektiven Störung des Spielgefühls.¹⁵

Der Datentransport in einem Netzwerk findet über Knotenpunkte statt. Diese verbinden die einzelnen Geräte und kümmern sich um die Weiterleitung der empfangenen Datenpakete vom Absender- zum Empfängergerät. Der Begriff Netzwerktopologie steht für die logische oder physische Verbindung der Knotenpunkte in einem Netzwerk. Diese können für den Datentransport in unterschiedlichen Anordnungen bestehen. Dabei werden die Bustopologie, die Ringtopologie, sowie die Sterntopologie am häufigsten verwendet. Bei der Bustopologie werden die Teilnehmer über sog. Transceiver, eine Verbindung aus Transmitter und Receiver, an eine gemeinsame Leitung angeschlossen. Fällt bei dieser Verbindungsart ein Teil der Leitung aus oder wird beschädigt, können die Teilnehmer beider Teile nicht mehr miteinander kommunizieren. Werden die an die Leitung angeschlossenen Teilnehmenden in eine Ringtopologie gesetzt, so besitzen sie eine gemeinsame Leitung, aber sind gegen Ausfälle des Netzes geschützt, da die Datenpakete ebenso in die andere Richtung übertragen werden können. Sind mehrere Teil-

¹⁵ Vgl. Lester, Michael und Boley (2007), Jon: The Effects of Latency on Live Sound Monitoring, Audio Engineering Society Convention 123, S.19

nehmenden an einem zentralen Verteiler (einem Switch) angeschlossen, können die Geräte gleichzeitig und parallel über den Knotenpunkt miteinander kommunizieren.¹⁶ Dadurch sind über weniger Knotenpunkte mehr Teilnehmer erreichbar, was als Nebeneffekt die Verzögerung der Datenpakete minimiert. Jeder überquerte Netzwerkknoten im Übertragungsweg wird „Hop“ genannt.¹⁷ Die Verzögerung entsteht bei der Paketverarbeitung im Knoten (siehe 3.4 Komponenten, Router) und ist bei echtzeitkritischen Anwendungen zu beachten.

Ein Netz muss beispielweise nicht aus einer reinen Ringtopologie bestehen, sondern kann aus einer Mischung der vorhandenen Topologien bestehen.

3.1 OSI-Modell

Von den Anfängen der Computertechnik bis zum heutigen Tage haben sich Programme, sowie deren Informationsverarbeitung weiterentwickelt und sind komplexer geworden. Dadurch haben sich Programmiersprachen verändert und Neue sind entstanden. Um es weiterhin zu ermöglichen, dass Systeme verschiedener Hersteller miteinander kommunizieren können, wurde 1984 das OSI-Referenzmodell (OSI für Open Systems Interconnection) entworfen. Diese Referenzarchitektur teilt die Kommunikation in sieben verschiedene Schichten auf. Dies ermöglicht, dass diese sich unabhängig voneinander weiterentwickeln können, ohne die Kommunikation der anderen Schichten zu beeinflussen.¹⁸

In Bezug auf diese Arbeit und Audionetze werden im Folgenden die Schichten eins bis vier genauer betrachtet, da diese als transportorientierte Schichten die Datenübertragung regeln, während Schichten fünf bis sieben anwendungsorientiert sind. In der folgenden Tabelle sind die Schichten eins bis sieben sowie deren Kommunikationsverlauf abgebildet. Eine allgemeine Eigenschaft des Transports über eine Netzwerkschicht ist hierbei zu beachten: Je höher die Transportschicht ist, desto mehr Protokolle sind zum Aufbau einer Verbindung nötig. Diese Protokolle müssen im Signalpfad abgehandelt werden und benötigen dadurch Verarbeitungszeit. Dies führt beim Versenden über höhere Schichten zu einer längeren Latenz im Netzwerk.

¹⁶ Vgl. Scherff (2010), S.2

¹⁷ Vgl. Schreiner (2016), S.79

¹⁸ Vgl. Schreiner (2016): S.3

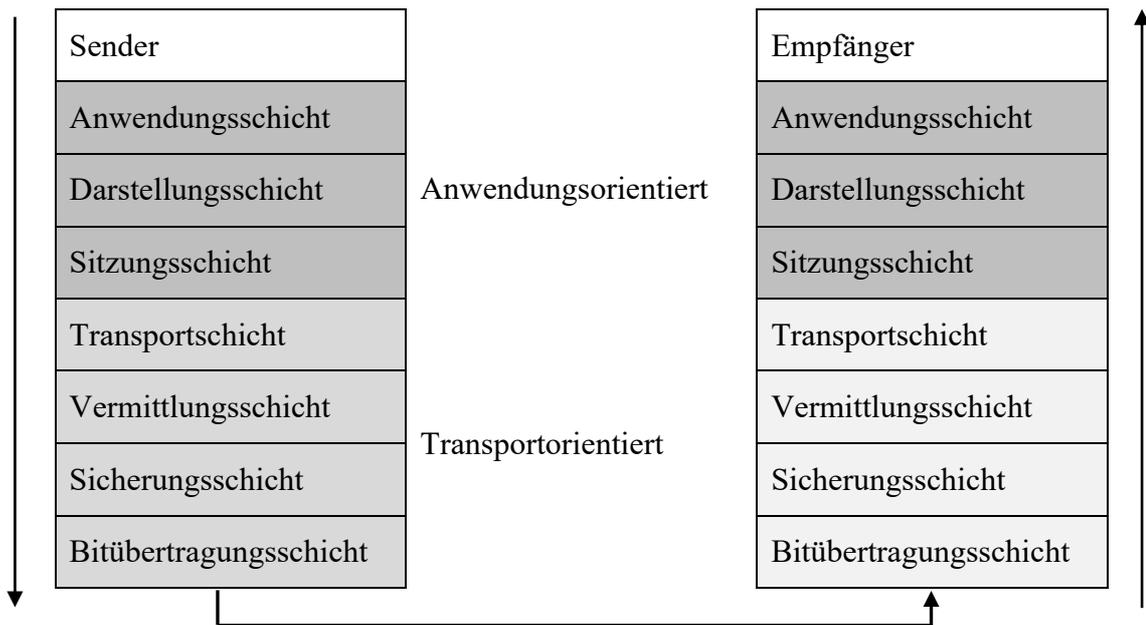


Tabelle 1: Das OSI-Schichtenmodell

Für folgende Arbeit ist es wichtig im Hinterkopf zu behalten, dass die Bezeichnung „Layer-X/Y-Netzwerk/Switch/Router“ beschreibt über welche Schicht die Kommunikationsprozesse in dem Netzwerk ablaufen, also ob z.B. die Adresse des Zielgerätes aus einer Liste mit den **MAC-Adressen** (Media Access Control Adresse) oder einer Liste mit den **IP-Adressen** gesucht wird.

3.1.1 Layer-1: Bitübertragungsschicht

Als unterste Schicht bildet die Bitübertragungsschicht die physikalische Schicht, über welche die Datenströme geleitet werden. Computer arbeiten auf der untersten Ebene mit Bits (Nullen und Einsen), diese werden durch physikalische Größen wie elektrische Spannungen oder Ströme, Lichtwellen oder magnetische Wellen weitergeleitet. Passive Netzwerkkomponenten wie Stecker oder Kabeltypen sowie aktive Netzwerkkomponenten wie Netzwerkadapter, Repeater, Switches oder Router gehören zu den Übertragungselementen der Bitübertragungsschicht. Die unterste Schicht wird in die drei oben genannten Bereiche aufgeteilt: LAN, MAN und WAN.

3.1.2 Layer-2: Sicherungsschicht

Die Sicherungsschicht (oder auch Datenverbindungsschicht) ist für die sichere Datenübertragung zuständig. Sie realisiert zuverlässige Verbindungen und sorgt für eine sichere Bitübertragung auf dem Übertragungsmedium, ebenso ist sie für die physikalische

Adressierung der Systeme in einem Netzsegment zuständig. Dies geschieht über die Adressierungen der MAC-Adressen (Media Access Control Adresse) einzelner Geräte. Die zu versendenden Daten werden in einzelne Pakete (Frames) gepackt und mit Header und Trailer (der Anfang und das Ende des Datensatzes), welche Felder mit Steuer-, Adress- und Fehlererkennungsdaten beinhalten.

3.1.3 Layer-3: Vermittlungsschicht

Layer-3 wird durch die Sicherungsschicht gebildet. Diese realisiert die logische Adressierung von Geräten. Die Sicherungsschicht vermittelt zwischen Systemen auf einem Übertragungsabschnitt, während die Vermittlungsschicht die Datenübertragung zwischen zwei Endsystemen über mehrere Übertragungsabschnitte definiert. Die Adressierung der Geräte, sowie die Routingprotokolle arbeiten mit IP-Adressen und sorgen dafür, dass beim Ausfall einer Verbindung im Netzwerk redundante Wege benutzt werden können.

3.1.4 Layer-4: Transportschicht

Die vierte Schicht stellt den Transportdienst für die darüber liegende Sitzungsschicht bereit und ist somit die Verbindungsschicht zwischen den transport- und anwendungsorientierten Schichten. Sie übernimmt die Bereitstellung eines universellen Transportdienstes durch die Regelung der Flusskontrolle des Datenstroms. So können mehrere Anwendungen gleichzeitig Daten über eine Verbindung senden.

Sie ist verantwortlich für das Senden der Daten in korrekter Reihenfolge und das unterteilen des Datenstroms in Segmente, welche an die unteren Schichten weitergegeben und einer Anwendung zugeordnet werden. Im empfangenen Computer stellt sie die Daten in der empfangenen Reihenfolge zur Verfügung. Implementiert sind verbindungsorientierte Dienste wie verbindungslose Dienste, von welchen Letztere für zeitkritische Medienübertragung von großer Wichtigkeit sind.

OSI-Layer 5-7: Diese Layer sind für die Kommunikation und Organisation innerhalb der Endgeräte zuständig. Ebenso werden die Daten für den Austausch über Layer 1-4 durch Kompression und Konvertierung optimiert.

Folgende Abbildung zeigt den Datenverlauf zwischen zwei Endgeräten über einen Knotenpunkt und die dabei zu durchlaufenden Schichten des OSI-Modells.

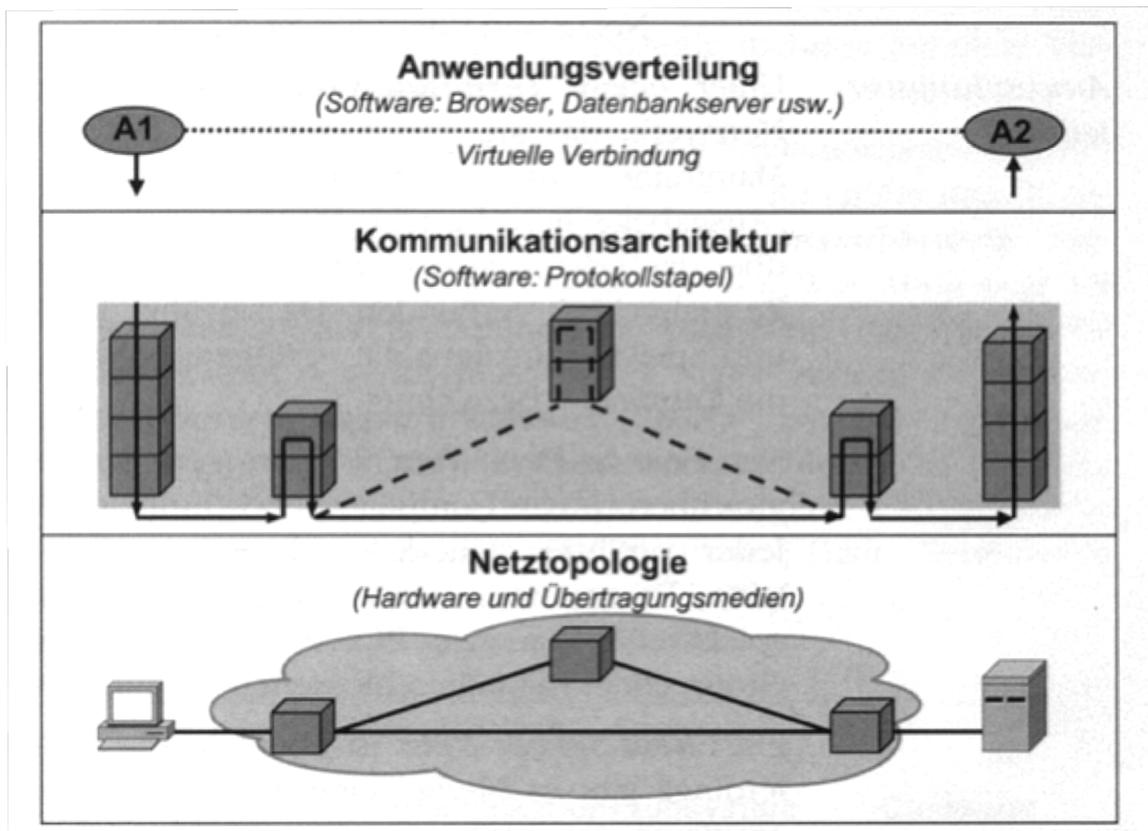


Abbildung 3-1: Schematische Darstellung einer Netzwerkarchitektur¹⁹

3.2 Ethernet nach IEEE802.3

Das Ethernet entstand in den 1970er Jahren. Als Übertragungsmedium für das Koaxialkabel und Rechner in einer Bus Topologie sollte es eine Alternative für das damals entwickelte ALOHA Funknetz sein, welches die Luft zur Übertragung benutzte. Dieses wurde weiterentwickelt und zu einem offenen Standard. Das Übertragungsmedium des Koaxialkabels wurde auf die Benutzung des Twisted-Pair Kabels sowie des Glasfaserkabels ausgebaut. Auf diesem beruht der in den 80er Jahren vom IEE (Institute of Electronics and Engineers) gebildete 802.3-Standard, welcher heute noch in Verwendung ist. Da ein Ethernet Netzwerk kostengünstig aufbaubar ist, setzte dieses sich gegen Token Ring und andere Konkurrenzprodukte durch und ist heute weltweit verbreitet. Das Ethernet arbeitet auf der OSI-Schicht 2, der Sicherungsschicht, über die der Austausch von Datenpaketen aus den ober liegenden Schichten erfolgt und wird gerne mit

¹⁹ Scherff (2010), S. 11

dem Protokollpaar TCP/IP verwendet. Teilnehmer in einem Netzwerk werden über die 48-Bit lange Mac-Adresse identifiziert.²⁰

Während die Geschwindigkeit des Ethernet-Standards anfangs eine Übertragungsrate von 10Mbit/s betrug, so entwickelte diese sich mit dem Standard und den Kabeltypen über die Jahre hinweg weiter. Übertragungsgeschwindigkeiten bis zu 100Gbit/s sind erreichbar und seit dem 21. März 2019 bietet DE-CIX als erster Internetknoten 400Gbit/s-Ethernet an.²¹

Für die verschiedenen IEEE 802.3-Spezifikationen wurde ein Klassifikationsschema eingeführt, welches eine Bezeichnung für den Ethernet Typ, dessen Kabel sowie Geschwindigkeit beschreibt.

Bezeichnung	Datenrate	Kabeltyp	Max. Leitungslänge
10Base-T	10 MBit/s	Twisted-Pair (mind. Cat3)	100 m
100Base-TX	100 MBit/s	Twisted-Pair (mind. Cat5)	100 m
100Base-FX/SX	100 MBit/s	Glasfaser	300 m
1000Base-T	1 GBit/s	Twisted-Pair (mind. Cat5e)	100 m
1000Base-SX/LX	1 GBit/s	Glasfaser	550 m/5 km
10GBase-T	10 GBit/s	Twisted-Pair (Cat6A/ Cat7)	100 m
10GBase-SR	10 GBit/s	Glasfaser	300 m
10GBase-LR	10 GBit/s	Glasfaser	10 km
40GBase-LR4	4 Wellenlängen a 10GBit/s	Glasfaser	10 km
100GBase-SR10/ER4	4 Wellenlängen a 25GBit/s	Glasfaser	10 km/40 km

²⁰ <http://www.elektronik-kompodium.de/sites/net/0603201.htm>

²¹Vgl. <https://www.de-cix.net/de/about-de-cix/media-center/press-releases/de-cix-first-internet-exchange-worldwide-to-offer-400-gigabit-ethernet-access-technology>

Tabelle 2: Sammlung aus gängigen Ethernet-Standards^{22, 23, 24}

Bei näherer Betrachtung der Tabelle fällt auf, dass ab einer gewissen Geschwindigkeit (>10MBit/s) die Übertragungsrate eines Koaxialkabels nicht mehr ausreicht und auf Twisted-Pair oder Glasfaserkabel umgestiegen werden muss. Während sich die Spezifikation für das Ethernet über die Jahre änderte, hat sich das grundlegende CSMA/CD-Verfahren (Carrier Sense Multiple Access /with Collision Detection) über die Jahre nicht verändert. Dieses kümmert sich um das Versenden der Datenpakete im Netzwerk und handelt in Fällen der Paketkollision.

3.2.1 CSMA/CA

Ein beim Versenden von Paketen häufig auftretendes Problem ist die Kollision derer. So wollen zwei benachbarte Stationen ein Paket über dieselbe Leitung versenden, die Pakete kollidieren und die enthaltene Information geht verloren. Beim CSMA/CD-Verfahren wird von beiden benachbarten Stationen überprüft, ob der zu benutzende Kanal frei oder belegt ist (Carrier Sense). Wenn es zu einer Kollision kommt kann diese erkannt und darauf reagiert werden (Collision Detection).

Die Kollisionserkennung läuft auf analoger Ebene ab, d.h. Die Hardware der Station muss den Kanal während der Übertragung abhören. Ist das zurückkehrende Signal ein anderes, wie das abgeschickte Signal, so vermittelt dies der Station die Kollision des Signals. Die Station bricht die Übertragung abrupt ab (statt wie beim Beenden Zeit und Bandbreite zu in Anspruch zu nehmen). Nach erkannter Kollision wird von der Station ein Signal abgesetzt, welches das Senden weiterer Daten verhindert. Nach diesem Signal wird ein Timer aus den MAC-Adressen der Geräte berechnet und gesetzt. Ist der Timer beendet wird erneut ein Signal abgesetzt.²⁵

Da heute Netzwerke kaum noch mit Hubs arbeiten, sondern mit Switches, können sie Vollduplex, also in beide Richtungen senden und vermeiden Kollisionen. Dies macht

²² <http://www.elektronik-kompodium.de/sites/net/1107311.htm>

²³ <http://www.elektronik-kompodium.de/sites/net/1406171.htm>

²⁴ <http://www.elektronik-kompodium.de/sites/net/1406261.htm>

²⁵ Tannenbaum, Andrew S., Wetherall, David J. (2012): Computernetzwerke, 5. Auflage, München, Pearson Deutschland GmbH

den CSMA/CD-Algorithmus theoretisch überflüssig, jedoch wird er immer noch verwendet, um die Kompatibilität zu gewährleisten.

3.2.2 Ethernet Frame

Das auf dem Layer 2 arbeitende Ethernet vermittelt und packt die Daten in gleich große Pakete. Der Versandt dieser Pakete geschieht über die Adressierung der MAC-Adresse. Der Inhalt ist nicht relevant. Die Pakete, auch Frames genannt werden in ein bestimmtes Format verpackt.

Jedes Paket beginnt mit der Präambel, einem Bitmuster zur Synchronisation. Dieses ist 7 Byte lang. Dem folgt das 1Byte lange SFD (Start of Frame Delimiter) (in der Abbildung 3-2 auch SOF, Start of Frame), welches den Beginn des eigentlichen Frames vermittelt. Angehängen ist jeweils eine 6 Byte großen Ziel- und Quelladresse der Netzwerkkarten beider Endsysteme. Das darauffolgende 2 Byte große Feld kann dem Typ zur Identifikation des Layer 3 Protokolls dienen oder es bestimmt die Anzahl der Datenbytes im Frame. Nun folgt das Datenfeld, welches die Nutzdaten der höheren OSI-Schichten ineinander verschachtelt beinhaltet. Es enthält gegebenenfalls Füllbytes, um die vorgeschriebene Frame-Länge einzuhalten. Das Ethernet-Frame endet mit der FCS, der Frame Check Sequence. Diese ist eine Prüfsumme zur Erkennung von Fehlern im Frame. Die Gesamtgröße des Frames darf 1518 Bytes nicht überschreiten.

Größe in Byte	7	1	6	6	2	46...1500	4
	Präambel	S O F	Zieladresse	Quelladresse	Typ	Daten	FCS

Abbildung 3-2: Frame Format für Ethernet²⁶

3.3 Protokolle

Wie bei den Menschen auch eine gemeinsame Sprache für die Kommunikation nötig ist, so ist es in der digitalen Welt wichtig eine gemeinsame Kommunikation zwischen verschiedenen Systemen zu definieren. Kommunikationsprotokolle in Rechnernetzen, auch Netzprotokolle genannt, regeln die Kommunikation und inhaltliche Struktur des Daten-

²⁶ <http://www.easy-network.de/bilder/ethernet-frame.jpg>

austausches. Netzwerkprotokolle werden auf den verschiedenen Layern des OSI-Modells benutzt, stellen der höheren Schicht Dienste bereit und benötigen Dienste der darunter liegenden Schichten. Für die Internetprotokollfamilie hat sich die Gliederung der Netzzugangsschicht (OSI-Layer 1&2), Internetschicht (Layer 3), Transportschicht (Layer 4) und Anwendungsschicht (Layer 5-7) durchgesetzt.²⁷

In dieser Internetprotokollfamilie stehen viele verschiedene Protokolle zur Auswahl. So beispielweise auf der Transportschicht das UDP (User Datagram Protocol) und das TCP (Transmission Control Protocol). Beide beschreiben ihren eigenen Kommunikationsablauf. Das UDP stellt keine Ende-zu-Ende Kontrolle zur Verfügung. Dadurch können bei Paketen verloren werden, die Übertragung geschieht jedoch schneller und verringert die Systemlatenz. Das TCP hingegen stellt eine zuverlässige Verbindung her, bei welcher alle Datenpakete beim Empfänger ankommen und im Falle des Verlustes noch einmal angefordert werden. Aufgrund des zeitkritischen Faktors in Mediennezwerken ist hier das UDP aufgrund der Performance das wichtigere Streaming-Protokoll.

In den folgenden Unterkapiteln werden nur die Protokolle IP, UDP und RTP behandelt.

3.3.1 Das Internet Protocol

Das Internet Protocol ist ein auf dem OSI-Layer 3 agierendes Standardprotokoll, welche Daten vom Sender zum Ziel leiten soll. Der Transport wird über die IP-Adresse geregelt. Würde dies über die MAC-Adresse stattfinden, müsste jeder Switch der Welt jedes Zielgerät und dessen Adresse wissen. So wird das Subnetz adressiert, in welchem die benötigte MAC-Adresse registriert ist.

Die Datenpakete werden verbindungslos und ungesichert transportiert, wodurch es zu fehlerhaften, doppelten oder gar keinen Übertragungen kommen kann (oder die Pakete werden in falscher Reihenfolge empfangen). IP stellt somit einen ungesicherten verbindungslosen Dienst dar, wobei „Sicherheit“ sich bei Protokollen auf das definitive Ankommen aller Datenpakete bezieht. Die Sicherung der Datenübertragung wird von Protokollen höherer Schichten übernommen.²⁸

Wie in Kapitel 3.2.2 erwähnt ist die maximale Paketgröße des Ethernet-Frames auf 1500 Bytes begrenzt. Ein IP-Paket kann bis zu 65.535 Bytes lang werden. IP ist fähig bei

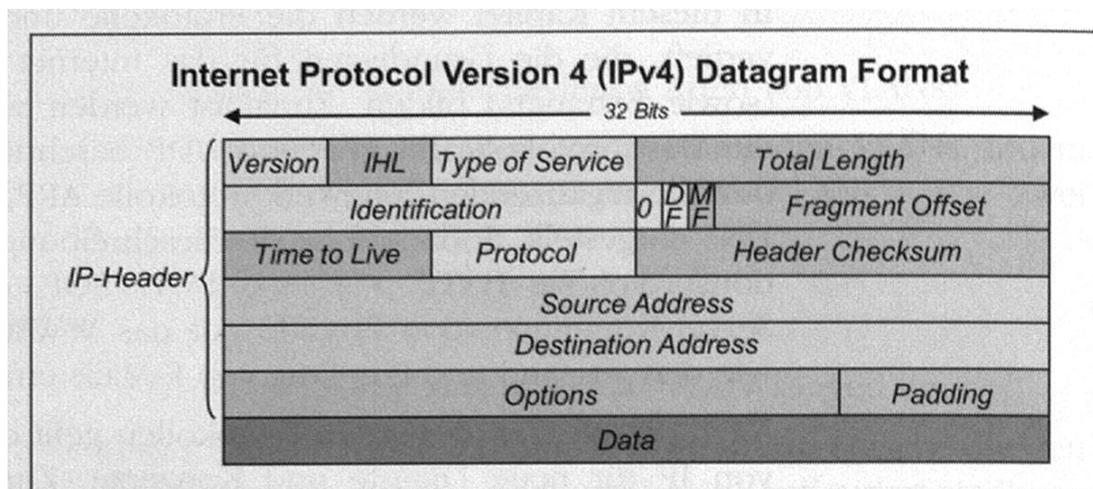
²⁷ Vgl. Dickreiter (2016), S. 1155

²⁸ Riggert, Wolfgang (2014): Rechnernetze, 5. Auflage, München, Carl Hanser Verlag, S. 116

Paketgrößen >1500 Byte diese zu fragmentieren, also zum Versand zu zerlegen und auf viele Pakete aufzuteilen. Die im Header zuständigen Felder „Flags“ und „Offset“ sind hierfür zuständig und geben an ob das Paket fragmentiert wurde (Flags) und welche Position das Fragment im Paket besitzt (Offset), um so am Ziel wieder zusammen gefügt werden zu können. Der Verlust eines Frames führt dazu, dass das ganze Paket erneut angefordert werden muss. Ebenso beeinflusst die Fragmentierung die Datendurchsatzrate. Einer Fragmentierung der Pakete bei Medienstreams ist abzuraten, da mit UDP (siehe unten) eine Neuanforderung der Fragmente nicht möglich ist und somit die Wahrscheinlichkeit, dass das ganze Paket unbrauchbar ist.²⁹

Das klassische IP-Paket wird durch einen mindestens 20 Byte langen Header angeführt, in welchem die für die Übertragung relevanten Felder enthalten sind. Dem Header folgend befinden sich die zu versendenden Daten. Im klassischen IPv4 beträgt der Adressraum der 32 Bits, also 4.294.967.296 Adressen. Dieser Adressraum konnte lange Zeit alle benötigten IP-Adressen abdecken. Mit der rasanten Verbreitung von Smartphones und dem Aufkommen des Internet of Things (IoT)³⁰ steigt die Anzahl der benötigten Internetadressen an. Dieser Knappheit soll durch die Adressraumerweiterung von IPv6 mit einer Größe von 128 Bits entgegengewirkt werden.

Die folgende Abbildung zeigt die Header der IPv4- und IPv6-Formate.



²⁹ Schreiner (2016), S. 85

³⁰ IoT ist ein Sammelbegriff für Technologien, die virtuelle und physikalische Dinge miteinander vernetzen. U.a. soll in der Industrie jede Maschine über eine Netzwerkadresse steuerbar sein. Dies erhöht die Anforderung an Netzwerkadressen.

Abbildung 3-3: Datagram-Format für das IP³¹

Die hierbei wichtigsten Bestandteile des IPv4-Headers sind:

- **Version** des Internetprotokolls
- **Type of Service (ToS)**: Priorisierung/Wichtigkeit des Pakets im Stream
- **Total Length**: Gesamtlänge des Pakets
- **Flag & Fragment Offset**: (siehe oben).
- **Time to Live**: Zeit nach der Paket gelöscht wird wenn es beim Ziel nicht ankommt.
- **Protocol**: Definiert welches Protokoll von nächsthöheren Schicht benutzt wird.
- **Source-/ Destination Adress**: Sender-/ Empfängeradresse

Für den IPv6-Header gilt:

- **Version**: wie bei IPv4
- **Traffic Class**: entspricht "Type Of Service"
- **Flow Label**: spezifiziert Pakete als solche eines gleichen Datenstroms
- **Payload Length**: Länge der auf den Header folgende Daten
- **Next Header**: zeigt auf Protokoll des nächst höheren OSI-Layers
- **Hop Limit**: entspricht Time to Live
- **Source-/ Destination Adress**: Adressierung des Empfängers/ Ziel

Das „Type of Service“ ist vor allem für die Übertragung echtzeitkritischer Audio-/Mediendaten von wichtiger Bedeutung, da diese in normalen Datenflüssen priorisiert werden können (Mehr in Kapitel 3.6).³²

3.3.2 UDP

Das User Datagram Protocol (UDP/ über RFC768³³ spezifiziert) ist ein Kommunikationsprotokoll des Layer 4. Ähnlich IP bietet es eine verbindungslose, unzuverlässige Datenübertragung zwischen Diensten an. Es stellt somit keine sichere Verbindung her oder überprüft das Ankommen von versandten Paketen, wodurch diese verloren oder

³¹ Scherff (2010), S. 318

³² Scherff (2010), S. 317 ff.

³³ RFC = Request for Comments ist eine Reihe aus technischen und organisatorischen Dokumenten zum Internet

dupliziert werden können. Ohne die Empfangsbestätigung werden bei Paketverlust diese auch nicht nachgesendet. Die UDP-Anwendung muss mit diesem Problem umgehen können. UDP ist zwar unsicher, kann jedoch gegenüber dem sicheren TCP die dreifache Übertragungsgeschwindigkeit erreichen und führt zu keinen Verzögerungen.³⁴ Geht ein Audiopakete bei der Übertragung verloren kommt es mit dem UDP zum sog. „Drop-Out“, d.h. in der Wiedergabe ist kurz nichts zu hören, aber die Übertragung läuft mit dem nächst empfangenen Paket weiter.

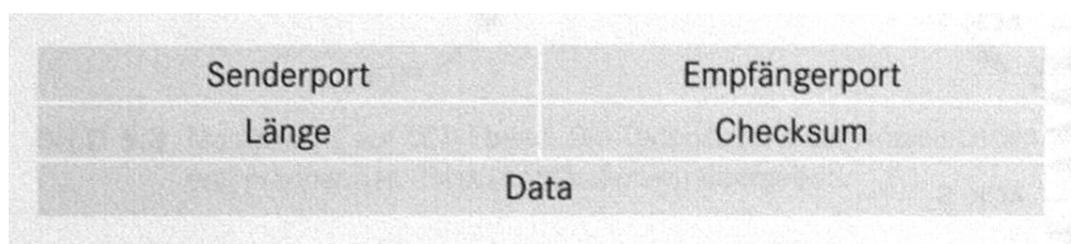


Abbildung 3-4: UDP Segment-Format³⁵

Der Header des Formats ist 64 Bit groß (16 Bit je Abschnitt) und vergleichsweise klein. Ähnlich den vorherigen Protokollen werden die Absender- und Empfänger-Ports beschrieben.

Ports dienen 2 Zwecken; sie können auf weitere Netzwerkprotokolle und deren Dienste verweisen oder werden zur Unterscheidung mehrerer Verbindungen zwischen demselben Paar Endpunkten benutzt. Vergleicht man die Zustellung des (Daten-)Pakets mit dem eines Briefes entspricht die IP-Adresse dem Familiennamen. Der Service (Port in Zusammenhang mit Protokoll) ist mit dem Vornamen vergleichbar, welcher bestimmt für wen der Brief im Haus/Wohnung bestimmt ist.

Nach den 16 Bit langen Sender-/ Empfänger-Ports wird mit „Länge“, die Gesamtlänge und „Checksum“ die Prüfsumme für das Segment gesendet.

3.3.3 RTP

Das Real-Time Transport Protocol (RTP/ im RFC 3550 spezifiziert) ist ein zur Echtzeitkommunikation entwickeltes Transportprotokoll und wird über UDP betrieben. Bei der Echtzeitübertragung von Audio- und Videosignalen werden diese als zusammen-

³⁴ Riggert (2014), S. 161

³⁵ Schreiner (2016), S. 114

hängende Folge von Segmenten verschickt. Es ermöglicht den Versand verschiedener Audio- oder Video-Codecs (z.B. Jpeg, PCM, oder MPEG-1/2-Audio³⁶), sowie die richtige Reihenfolge der Segmente und die Isochronität der Übertragung durch Zeitstempel im Header. Durch fortlaufende Nummerierung der Pakete wird es dem Empfänger ermöglicht Paketverluste oder die Zustellung in falscher Reihenfolge zu erkennen.³⁷

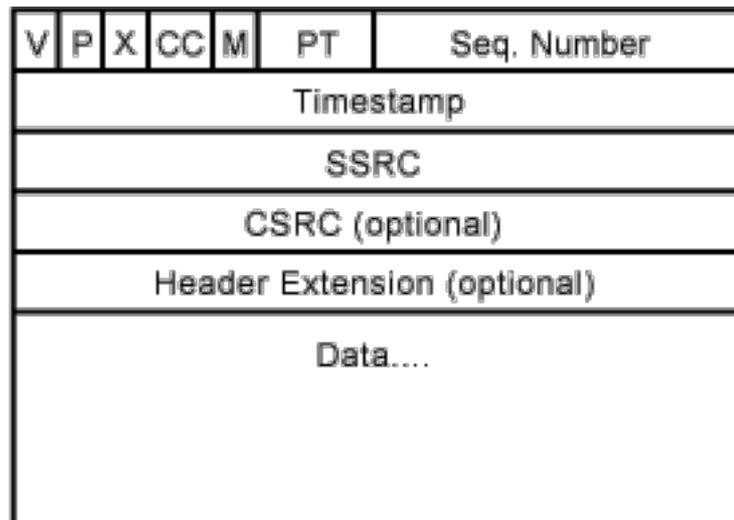


Abbildung 3-5: Aufbau eines RTP-Headers³⁸

Die Abkürzungen haben folgende Bedeutung: Version (V), Padding (P), Extension (X), CSRC Count (CC), Marker (M) und Payload Type (PT). Das CSRC-Feld dient zur Identifizierung der Quellen (bis zu 15 sind registrierbar) und das SSRC-Feld zur Identifikationsquelle.

Das RTP wird durch das ebenso im RFC 3550 spezifizierte RTCP (Real-Time Transport Control Protocol) ergänzt. Dieses tauscht zwischen den Endsystemen die Steuer- und Statusinformationen aus und überwacht den Kommunikationsablauf.

³⁶ <https://tools.ietf.org/html/rfc3551>

³⁷ Scherff (2010), S. 348

³⁸ <http://www.elektronik-kompodium.de/sites/net/bilder/11060711.gif>

3.4 Komponenten

Die Datenverteilung benötigt eine sichere und zuverlässige Infrastruktur. Statt nun eigene Geräte zur Verteilung von Daten zu entwickeln, greift man auf bereits vorhandene Geräte zurück. Aufgrund der oben beschriebenen Standards wurden über die Jahre Geräte zur Verteilung aber auch zur Kopplung von Netzwerken entwickelt und arbeiten auf den verschiedenen Ebenen im OSI-Modell.

Damit eine Kommunikation der Endgeräte funktioniert werden deren Daten über Netzwerkgeräte übertragen. Diese nehmen dabei unterschiedliche Funktionen ein. Switches, Repeater und Hubs übernehmen in einem Netzwerk die Rolle der Netzknoten und kümmern sich um die Weiterleitung der Pakete im Netzwerk. Router und Bridges dienen als Koppellemente zwischen den Netzwerken, indem sie die Daten sicher in ein anderes übertragen oder den besten Weg zum benötigten Netzwerk suchen. Repeater arbeiten auf OSI-Layer 1 und sorgen für die simple Verstärkung eines Signals, um dessen Reichweite zu erhöhen. Der Switch tut dies ebenso, besitzt jedoch mehrere Anschlüsse und steuert die Signale auf OSI-Ebene 2 über die MAC-Adresse. Auf der Ebene 3 sucht der Router die Verbindungen über IP-Adressen heraus und verbindet mehrere Netze zu einem Gesamtnetz. Auf OSI-Ebene 7 arbeitet das Gateway und verbindet über verschieden benutzte Protokolle hinweg Netze.³⁹

³⁹ Scherff (2010), S. 287 ff.

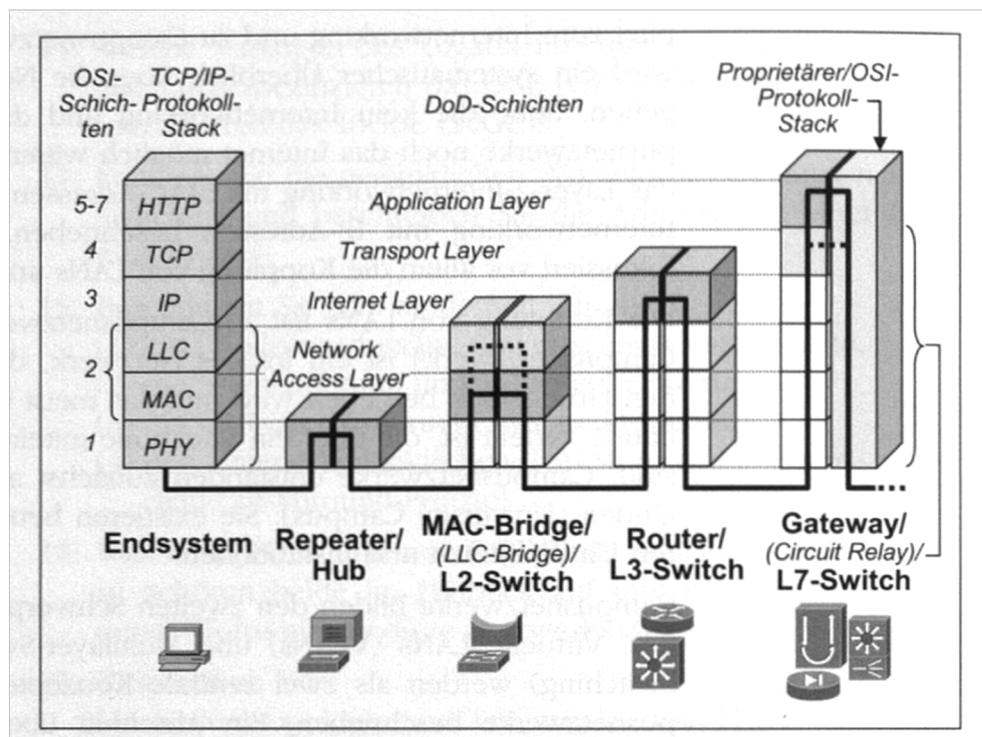


Abbildung 3-6: Beschreibung verwendeter Netzwerkgeräte durch Schichtzuordnung⁴⁰

3.4.1 Switch

Der Switch arbeitet auf dem Layer 2 des OSI-Modells und löst den Hub sowie den Repeater ab. Der Switch ist intern mit einer hochleistungsfähigen Bus-Matrix ausgestattet und leitet Datenpakete an die den Ports angeschlossenen Geräte weiter. Der Switch erkennt automatisch Unicast, Multicast und Broadcastpakete (siehe Kapitel 3.5). Die Ports sowie die angeschlossenen Geräte bilden eigene Netzwerksegmente, dem Switch sind nur angeschlossene Knoten und nicht das ganze Netzwerk bekannt. Jedem Port ist eine Bridge vorgeschaltet. Dies hat zur Folge, dass jedes am Switch angeschlossene Gerät in eigener CSMA/CD Umgebung kommuniziert und dadurch direkt angeschlossene Geräte auch im Duplexmodus arbeiten können. Im Duplexmodus muss sich um keine Kollision gekümmert werden, wodurch alle Leitungsdrähte benutzbar und somit in beide Richtungen gesendet werden kann.⁴¹

Wie in 3.2.1 ausgeführt, findet sich die Ziel- und Quelladresse im Ethernet Header wieder und wird durch die MAC-Adresse beschrieben. Der Switch leitet anhand dieser Ad-

⁴⁰ Scherff (2010), S. 288

⁴¹ Schreiner (2016), S. 51 ff.

resse die empfangenen Pakete zielgerichtet weiter. Dies findet in drei Schritten statt; der Netzknoten entscheidet ob das Paket vollständig ist und die angegebene Adresse eine Weiterleitung ermöglicht. Im zweiten Schritt wird die Quelladresse auf Vorhandensein in der Adressliste geprüft und zuletzt die Zieladresse ebenso auf einen Eintrag überprüft und versendet. Ist die Quelladresse nicht vorhanden wird ein Adresslernprozess ausgeführt, bei welchem er diese in eine Liste mit dem zugehörigen Port einträgt. Ist die Zieladresse nicht vermerkt, wird das Paket an jeden Port geschickt. Dieser Prozess wird „Flooding“ genannt.⁴²

Wird der Switch in eine redundante Umgebung wie einem Ring implementiert kann es zu Problemen kommen. Broadcastanfragen oder Flooding führen dazu, dass empfangenen Pakete immer wieder an die benachbarten Netzknoten geschickt werden. Durch die Topologie werden die Pakete in einer Endlosschleife durch das Netzwerk geschickt und ein Packet Storm entsteht, durch den das Netzwerk zusammenbricht. Durch das Implementieren den Spanning-Tree Algorithmus wird dies verhindert. Dieser sorgt dafür, dass es keine unerwünschten kreisenden Pakete gibt, indem die Topologie erkannt und redundante Wege in eine Baumtopologie überführt werden. Dazu werden die redundanten Ports deaktiviert und auf Bereitschaft gestellt, entfällt eine Verbindung im System, wird dies erkannt und blockierte Ports aktiviert.⁴³

Neben Layer 2 Switchen gibt es Switche die auf dem Layer 3, 4, und 7 arbeiten. Layer 3 Switche vereinen die Funktionen von Switchen und Routern und verwalten Pakete auf MAC-, sowie IP-Adressebene. Der Switch unterscheidet sich zum Router (siehe Kap. 3.4.2) im Arbeitsweisen, bedingt durch seine schnellere Latenzzeit und höhere Durchsatzrate. Jedoch steht er in einigen Funktionen dem Router nach und wird eher in LAN-Umgebungen eingesetzt, während der Router für WAN-Umgebungen benutzt wird.⁴⁴ Layer 4 Switche können über den TCP-/UDP-Header die enthaltenen Portnummern herauslesen. Anhand dieser, sowie der IP-Adressen ist es dem Switch möglich die Datenströme zu identifizieren und mit der gewünschten Dienstgüte, der sogenannten „Quality of Service“ (siehe Kap. 3.6), zu versehen.⁴⁵ Layer 7 Switche verarbeiten den Dateninhalt des Pakets bevor sie die Verbindung schalten. Dies bietet die Möglichkeit intelli-

⁴² Riggert (2014), S. 190

⁴³ Riggert (2014), S. 198 f.

⁴⁴ <https://www.itwissen.info/Layer-3-Switch-layer-3-switch.html>

⁴⁵ <https://www.itwissen.info/Layer-4-Switch-layer-4-switch.html>

gente Dienste zu unterstützen und Anfragen flexibler zu verteilen.⁴⁶

Je höher das OSI-Layer ist auf welchem der Switch arbeitet, desto mehr Daten müssen für die Weiterleitung verarbeitet werden und um so größer wird die Latenz, die im Switch entsteht.⁴⁷

Switche sind ebenso fähig ein Virtuelles LAN, ein sogenanntes VLAN, zu erstellen. In größeren LAN-Umgebungen ist das komplette Netz eine Broadcastdomäne, d.h. alle Endgeräte empfangen ein Broadcastframe, wenn dieser durch einen Flooding betreibenden Switch gesendet wird. Die vielen Broadcastanfragen der Geräte führen schnell zu einer Überlastung des Netzes. Durch eine Aufteilung des großen Netzes in kleinere VLANs lässt sich dieses Problem eindämmen, da die Broadcastbereiche schrumpfen und sich auf logische Bereiche beschränken lassen. Verbleibt die Kommunikation der Endgeräte innerhalb eines VLANs, wird geschwicht. Die Kommunikation zwischen verschiedenen VLANs wird geroutet. Für ein VLAN wird der Ethernet-Header auf Layer 2 mit dem Standard 802.1q erweitert. Um zwischen VLANs zu kommunizieren, muss ein Layer 3 Switch vorhanden sein.⁴⁸

3.4.2 Router

Router sind in der Lage Datenpakete bis auf das Layer 3 des OSI-Schichtmodells zu interpretieren. Sie verbinden verschiedene LANs und LAN-Technologien. Ebenfalls trennt der Router Broadcastdomänen und ist für Broadcastanfragen nicht durchlässig. Die Adressierung findet über die IP-Adresse statt, zur Weiterleitung ermittelt der Router den optimalen Weg. Um diesen zu ermitteln führt der Router sogenannte Routing-Tabellen, in welchen nachgeschlagen wird ob genannte IP-Adressen bekannt sind. Über die Routing-Tabelle weist der Router neben den angeschlossenen (Sub-)Netze und dem nächst liegendem Router, den bestmöglichen Weg für ein Paket durch das Netzwerk. Um die Tabellen nutzen zu können, müssen diese befüllt und konfiguriert werden. In kleinen Netzwerken ist dies manuell möglich, dies wird statisches Routing genannt. Hierbei ist der Weg durchs Netzwerk festgelegt, ein Topologiewechsel muss händisch eingetragen werden, dient aber der Sicherheit. Für größere Netzwerke benutzt man ein

⁴⁶ <https://www.itwissen.info/Layer-7-Switch-layer-7-switch.html>

⁴⁷ Schreiner (2016), S. 78

⁴⁸ Scherff (2010), S. 303 ff.

dynamisches Routing, bei welchem die Router durch Protokolle die Verbindungen zu anderen Netzwerken und -geräten erlernen.⁴⁹

Um Routing-Schleifen und damit eine damit verbundenen Überlastung des Netzes zu verhindern, wird der TTL-Block im IP-Frame benutzt. Die ihm zugewiesene Zahl bestimmt die Lebensdauer eines Paketes, also nach wie vielen Hops das Paket verworfen wird. Bei jedem Hop wird der Zahlenwert inkrementell um den Wert 1 verkleinert bis das Paket bei dem Wert 0 verworfen wird.

Zwei häufig beim dynamischen Routing verwendete Algorithmen sind das Distance-Vector-Routing und State-Link-Routing. Das *Distance-Vector-Routing* basiert auf einem System des Austauschs, in welchem die Router ihre kompletten Routingtabellen mit ihren direkten Nachbarn austauschen. Diese Routingtabellen enthalten den Distanzvektor zu einem Ziel-Netz. Die von anderen Routern erhaltenen Distanzvektoren werden erhöht (da man weiter von dessen Ziel-Netz liegt) oder es werden bei der Aktualisierung die Einträge mit kleineren Vektoren übernommen.

Flexibler und agiler gegenüber Änderungen, jedoch rechenaufwändiger ist das *State-Link-Routing*. Bei diesem senden die Router in einer Domäne die Tabelleneinträge zu ihren Nachbarn, per Broadcast an die anderen. Aus diesen Informationen bezieht jeder Router eine Topologie-Datenbank und einen Netztopologie-Baum, welcher aus allen Pfaden zu den Zielnetzen und ihrer Länge besteht. Mit Parametern wie z.B. Hops, Geschwindigkeit, Kosten und Bandbreite wird aus den empfangenen Informationen die Routing-Tabelle mit den kürzesten Pfaden zu den Ziel-Netzen gebildet. Mit diesem Algorithmus ist dem Router die Position jedes anderen Routers im Netz bekannt.⁵⁰

Aufgrund des Rechenaufwandes bei Routing-Algorithmen, besitzt dieser größere Latenzzeiten wie ein Switch. Aus diesem Grund ist es ratsam, statt eines Routers, Layer-3-Switche in einem echtzeitkritischen Netzwerk zu verwenden und mit dem Router in andere Netzwerke weiterzuleiten.

⁴⁹ Riggert (2014), S. 210 ff.

⁵⁰ Scherff (2010), S. 171 f.

3.5 Unicast, Broadcast, Multicast

Unicast, Broadcast und Multicast sind Verkehrsarten, die durch die Anzahl der Kommunikationsteilnehmer bestimmt werden. Bei Unicast handelt es sich um eine einfache Punkt-zu-Punkt Übertragung, bei welcher ein Paket von einem Sender an einen Empfänger gesendet wird. Im Gegensatz dazu steht Broadcast, eine Sendung, bei welcher das Pakete an alle Endgeräte im Netz zugestellt wird. Diese ungerichtete Übertragung stößt jedoch schnell an ihre Grenzen, da diese in größeren Netzwerken zur Überlastung des Netzwerks führen kann.

Ist eine Gruppe von Geräten das Ziel der Übertragung, benutzt man Multicast. Eine Unicast-Sendung wäre denkbar, hierbei würde das Datenpaket am Sender dupliziert und mehrfach parallel zu den Empfängern gesendet werden. Dies führt zu einer höheren Netzbelastung.

Beim IP-Multicasting wird das Paket einzeln versandt und möglichst spät vom Router vor dem Empfänger dupliziert, um die Übertragungskapazität des Netzwerkes gering zu belasten.⁵¹

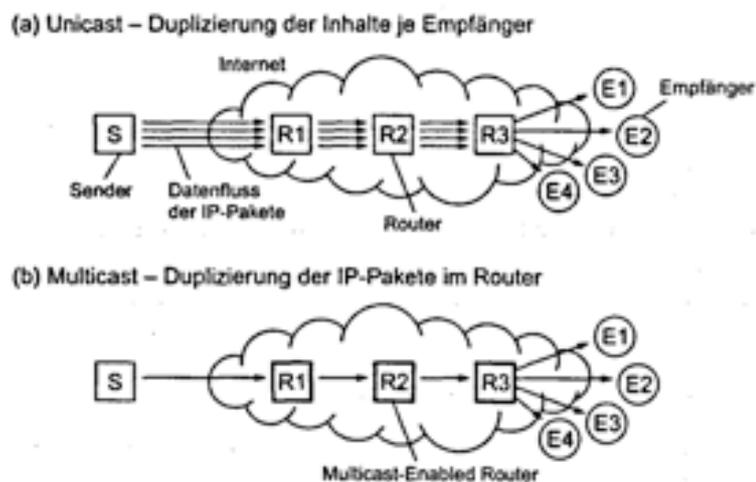


Abbildung 3-7: Paketversand Unicast – Multicast⁵²

Unicast wird hauptsächlich bei Informationsaustausch zwischen zwei Kommunikationspartnern verwendet. Multicast eignet sich zum Transport von Multimediaströmen

⁵¹ Dickreiter (2014), S. 1153 f.

⁵² Dickreiter (2014), S. 1154

und wird zur Steuerung von Netzwerkgeräten verwendet. Broadcastverkehr findet hauptsächlich im technischen Bereich statt, z.B. bei der Suche von Zielgeräten im Netzwerk.⁵³

3.6 Quality of Service (QoS)

Quality of Service (QoS) beschreibt die qualitative Güte eines Netzes aus der Sicht des Anwenders. Der Begriff unterliegt keiner einheitlichen Definition, sondern bildet einen Oberbegriff zu gestellten Anforderungen, sowie den Maßnahmen zu einer Verbesserung der Kommunikationsdienstleistung.

Diese Qualität lässt sich messtechnisch erfassen, wird aber subjektiv beurteilt. Erfassen lässt sich beispielweise die Dienstgüte als eine Mindestübertragungsqualität. Diese lässt sich über die Zuverlässigkeit des Netzwerkdienstes, der durchschnittlichen Verlust- und Fehlerrate der gesendeten Pakete und der Isochronität, der maximalen Verzögerung (Delay) und Verzögerungsschwankung (Jitter) angeben. Erweitern lässt sich dies um die gewünschten Eigenschaften, die das Netzwerk besitzen soll; die Qualitätsanforderungen. Beschrieben werden diese durch die Netzwerkverfügbarkeit, durch den Ausfall des Netzes, die Netzwerksicherheit, dem Schutz vor schädlichen Angriffen und Einwirkungen beim Datentransport. Ebenso dazugehören die Benutzermobilität, also die Möglichkeit sich als Nutzer überall im Netz einzuloggen und die Netzwerkskalierbarkeit, beschrieben durch die Fähigkeit des Netzwerkes sich an Erweiterungen und Veränderungen anzupassen.⁵⁴

Ein gutes QoS kann durch eine Kombination mehrerer Maßnahmen erreicht werden. Als grundlegendste ist die richtige Dimensionierung des Netzes zu nennen. Ist ein ausreichender Datendurchsatz vorhanden, läuft das Netzwerk nicht in Gefahr der Überlastung. Durch die Reservierung von Bandbreite für bestimmte Anwendungen oder Streams wird eine garantierte Übertragungsrate sichergestellt, welche auch im Falle der Überlast gilt. So können z.B. in einem 1Gbit/s Multimedienetzwerk mindestens 700Mbit/s für Medienstreams reserviert werden, während auf den anderen 300Mbit/s sonstiger Daten transportiert werden. Dies wird Integrated Service genannt⁵⁵

⁵³ Scherff (2010), S. 20

⁵⁴ Scherff (2010), S. 57 ff.

⁵⁵ Scherff (2010), S.182

Darauf aufgebaut kann eine Erhöhung der Qualität durch einen priorisierten Transport von Datenpaketen erreicht werden, dem Differentiated Service (DiffServ). Zur Priorisierung ist eine Einteilung der Datenpakete in Verkehrsklassen nötig. Die höheren Verkehrsklassen werden bevorzugt im Netzwerk versandt. Die Priorisierung funktioniert nur in Netzen, in denen die Verkehrsklassen gelten.⁵⁶

DiffServ findet über IP, also über das OSI-Layer 3 statt und ist im RFC1349⁵⁷ beschrieben. Es wird durch eine Markierung des Differential Service Code Point (DSCP) im IP-Header kenntlich gemacht. Dieser liegt im ToS-Feld, welches im Jahr 2000 umbenannt wurde und nun im RFC2474 beschrieben ist.⁵⁸ Das DSCP-Feld ist mit 6 Bit beschrieben, bietet also theoretisch 64 Klassen zur Weiterleitung der Datenpakete. Die Priorisierung findet beim Weiterleiten über den Router oder Layer-3-Switch statt.⁵⁹

Für eine Priorisierung von Datenströmen in Audionetzwerken wird empfohlen nach einer festen Ordnung zu priorisieren. Die DSCP Werte für die Priorisierung der Datenströme unterscheiden sich von Hersteller zu Hersteller und müssen händisch beim Einrichten des Netzwerkes eingetragen werden. Werden zwei Netzwerke mit unterschiedlich definierten Werten vernetzt, sollte „DSCP remarking“ verwendet werden. Dabei analysiert der Switch die DSCP-Werte der miteinander verbundenen QoS-Domänen und wandelt diese in für die Domäne entsprechende neuen Werte um. Folgende Tabelle zeigt die allgemein empfohlene Priorisierung der Datenströme.⁶⁰

Priorität	Datenstrom
Hoch	Synchronität der Uhr über PTP
Mittel	Audiodaten
Gering	Steuerdaten
Sehr Gering (Best-Effort-Prinzip ⁶¹)	Für Medienstrom nicht relevante Daten,

⁵⁶ <https://www.elektronik-kompodium.de/sites/net/0905131.htm>

⁵⁷ <https://tools.ietf.org/html/rfc1349>

⁵⁸ <https://tools.ietf.org/html/rfc2474>

⁵⁹ <https://www.itwissen.info/DiffServ-differentiated-services.html>

⁶⁰ Allen & Heath (2019), AES67 – One Standard to Unite Them All, S.6

⁶¹ Beim **Best-Effort-Prinzip** steht die im Netzwerk restliche Übertragungskapazität dem Datenverkehr dieser Stufe zu. Bei Überlast werden diese Pakete zuerst verworfen und müssen erneut gesendet werden.

	die über Netzwerk gesendet werden.
--	------------------------------------

Tabelle 3: Priorisierung der Datenströme

3.6.1 Netzwerksicherheit

Die hier beschriebene Sicherheit ist vom Rest der Arbeit abzugrenzen. Dort wird „Sicherheit“ im Kontext der Übertragungssicherheit oder Redundanz verwendet. Dieser Abschnitt beschäftigt sich mit der Absicherung gegenüber unerlaubten Zugriffen auf die im Netzwerk befindlichen Daten und Geräte. Dabei liegt der Schwerpunkt auf zeitkritischen Mediennetzwerken und dem Transport von Audiodaten. Um eine grundlegende Sicherheit in seinem Netzwerk zu erzielen, gibt es mehrere Methoden.

Angriffe auf ein Netzwerk können verschiedene Formen haben. So ist es möglich gegen Switche vorzugehen, indem die eigene MAC-Adresse in die eines registrierten Gerätes geändert wird und geheime Daten über dieses angefordert werden. Um dem vorzubeugen muss ein Router vor die im Netz eingebundenen Switche gesetzt werden, da dieser mit einer IP Adresse arbeitet und somit nicht von dieser Form des Angriffs betroffen ist. Ein anderer Weg ist die Manipulation der ARP-Tabellen, bei denen mit einer korrekten IP- und gefälschten MAC-Adresse falsche Daten eingeschleust werden. Dies eröffnet weitere Angriffsmöglichkeiten auf das Netzwerk.

Eine Absicherung des Switches wird durch die PortSecurity möglich gemacht. Hierbei wird der entsprechende Port am Switch gesperrt und nur für Datenverkehr gegenüber dem zu erwartendem Gerät geöffnet.

Angriffe auf den Router können durch Access-Control-Lists (ACLs) abgewehrt werden. Diese beschreiben welchen IP-Adressen und Protokollen es erlaubt ist über den Router zu kommunizieren. Der Nachteil hiervon ist die Entstehung von Latenz.⁶² Dies ist in einem zeitkritischen Netzwerk unerwünscht. Aus diesem Grund sollten solche Router als Randgeräte im Netz, jedoch nicht als Knoten, bzw. Verteiler der Mediendaten dienen.

Ebenfalls kann die PortSecurity eines Switches durch direktes anschließen von Geräten an den Port umgangen werden und somit dessen Sicherungsmechanismen umgangen werden. Um das Abhören von Datenströmen zu verhindern, können diese verschlüsselt

⁶² Riggert (2014), S. 245 f.

werden. Dafür müssen Sender und Empfänger einen gemeinsamen Schlüssel kennen, mit dem Datenströme übersetzt werden. Der Einsatz von Verschlüsselungsmethoden führt jedoch ebenfalls zur Entstehung von Latenz, welche als kritisch zu betrachten ist.

Dies lässt auf Möglichkeiten schließen, die die Sicherheit in einem Netzwerk erhöhen. Um Attacken von außen zu verhindern sollten die Mediennetzwerke parallel zu bestehenden Netzen aufgebaut und nach außen durch Firewalls geschützt werden. Der Netzwerkverkehr sollte intern stattfinden. Durch Kontrollsoftware kann der Zugriff von außen beschränkt werden, so dass eine Konfiguration der angeschlossenen Geräte nur durch den Netzwerkadministrator stattfinden kann.

Um zu verhindern, dass externe Geräte an Knotenpunkte gelangen und dort über die physikalische Anbindung Angriffe durchführen können, sollten die Geräte nicht frei zugänglich sein. Nur autorisierte Personen sollen Zugänge zu den Räumen besitzen, in welchen die Netzwerkstruktur vorhanden ist. Der Zugang sollte durch Schutzmechanismen geschützt.

3.6.2 Latenz

Latenz spielt bei der Übertragung von Audiosignalen in professionellem Umfeld eine kritische Rolle. Kritisch ist diese, da Latenz in Echtzeitübertragungen subjektiv die Qualität beeinflusst. So haben Michael Lester und Jon Boley in einer von der Audio Engineering Society (AES) herausgegebenen Studie befunden, dass menschliche Empfindlichkeit gegenüber Verzögerungen im Millisekundenbereich liegt.⁶³ Hierfür spielten Musiker ihre Instrumente und bekamen ihr eigenes Signal mit einer Verzögerung zurückgespielt. Diese wurde anhand des subjektiven Spielgefühls bewertet und ausgewertet.

Die Latenz in einem Netzwerk muss demnach unter einem gewissen Wert liegen, um ein natürliches Gefühl beim Zusammenspielen oder in der Sprachübertragung zu erzeugen. Ein solcher Wert wird in Kapitel 6 unter dem Bewertungskriterium „Latenz“ angeführt.

⁶³ Lester, Michael und Boley, Jon (2007): The Effects of Latency on Live Sound Monitoring, Audio Engineering Society Convention 123

Latenz kann auf verschiedenen Strecken gemessen werden. Es kann die Bearbeitungsdauer eines Gerätes oder die Zeit gemessen werden, in welcher ein Signal von einem Ende des Netzwerks bis zum anderen benötigt. Die Minimallatenz im Netzwerk wird durch die Paketlänge kumuliert mit der Netzlatenz beschrieben. Dabei beträgt die Minimallatenz in Ethernet-Netzen bis zu 0,5ms.⁶⁴

Dies bedeutet, dass die Übertragung kleinerer Datenpakete mit einer geringeren Latenz verbunden ist und die Übertragung von großen Datenpaketen mit einer höheren Latenz.

Indirekt mit der Latenz zusammenhängend ist die Lippensynchronisation. Für diese gibt es einen Toleranzbereich, in dem ein asynchrones Video- und Bildsignal als zusammenhängend wahrgenommen wird. Dieser Toleranzbereich wird durch die Rec. ITU-R BT.1359 beschrieben und ist in Abbildung 3-8 beschrieben.

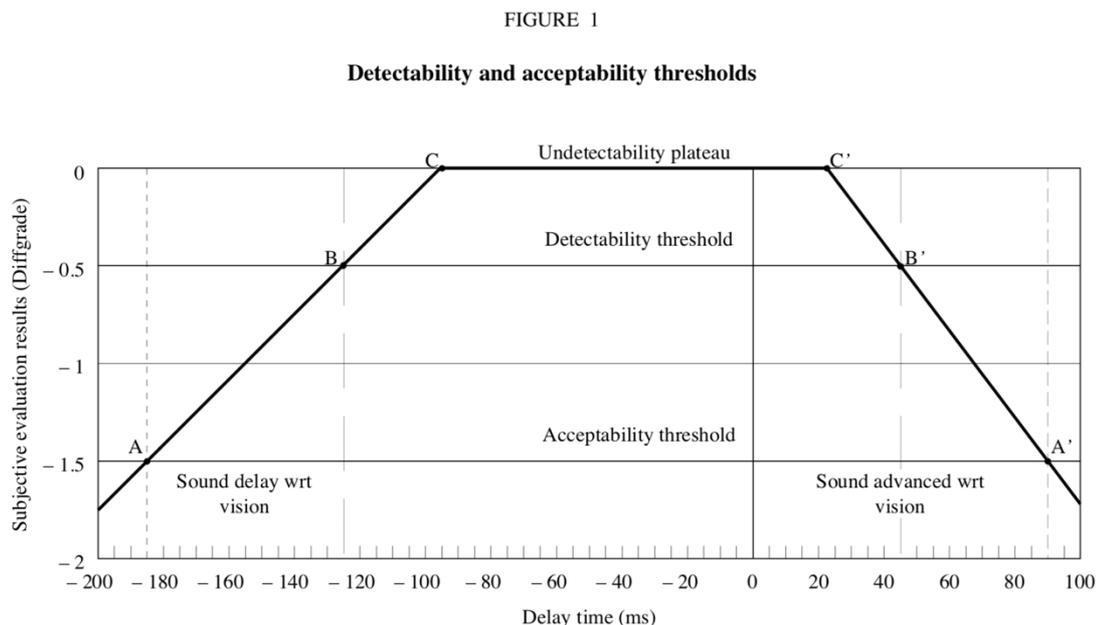


Abbildung 3-8: Toleranzbereich von verzögerten Video- und Audiodaten⁶⁵

Zu sehen ist eine größere Akzeptanz, wenn das Bild vor dem Ton wahrgenommen wird. Bei einer Bildwiederholrate von 50Hz sind dies 2-3 Frames Toleranz. Diese Synchronität gilt es in einem System durch die Synchronisierung von Audio- und Videomaterial zu erreichen.

⁶⁴ Vgl. Gross, K. (2013): AES67 standard for audio applications of networks, *Audio Engineering Society Convention 67(7)*, S. 18 ff.

⁶⁵ https://www.atsc.org/wp-content/uploads/pdf/audio_seminar/12%20-%20JONES%20-%20Audio%20and%20Video%20synchronization-Status.pdf

4 Übertragung von Audiosignalen in Netzwerkstandards

4.1 Interoperabilität in Netzwerken

1996 wurde mit Peak Audio's „CobraNet“-Protokoll eines der ersten erfolgreichen Protokolle zum Transport von Audiodaten über digitale Netzwerke veröffentlicht. CobraNet arbeitet mit Audio over Ethernet (AoE), auf dem Layer 2 des OSI-Modells. Nach dessen Erfolg begannen auch andere Firmen Ethernet-Protokolle zu entwickeln und am Anfang des 21. Jh. wurde auf den IT-Standards neue Protokolle auf dem IP basierend entwickelt. Diese auf dem OSI-Layer 3 arbeitenden Protokolle werden auch Audio over IP (AoIP) genannt. Die Hersteller verwendeten hauptsächlich proprietäre Lösungen, die grundsätzlich ähnlich arbeiteten, aber ein Austausch untereinander nicht ermöglichten. Aus diesem Grund bildeten sich Arbeitsgruppen, welche eine standardisierte Datenübertragung zwischen den Systemen ermöglichen sollten.

So wie die Datenübertragung im Netzwerk eine Standardisierung und somit festgelegte Strukturen besitzt, wird dies auch bei den Medienstreams angewandt. AES67, SMPTE ST 2110-30 und AVB (Audio Video Bridging) bezeichnen eine Reihe von Standards, welche einen Rahmen zur netzwerkbasierten Audioübertragung bilden. Diese Standards ermöglichen die Interoperabilität verschiedener Netzwerk-Protokolle durch eine gemeinsame Definition von IP-Protokollen.

Während SMPTE ST 2110 und AVB eine Sammlung an Standards für synchronisiertes Streaming von Audio- und Videodaten ist, ist AES67 ein Standard zur Audioübertragung in digitalen Netzwerken. Die Anforderungen von AES67 bilden die Basis für den Austausch von Audiodaten und die Synchronisation unterschiedlicher Audionetzwerk-Protokolle (z.B. einem Dante- und Ravenna-Protokoll). Zudem ist AES67 ein offener Standard, wodurch die Hersteller keine Kosten für die Implementierung übernehmen müssen, eine höhere Akzeptanz dieses Standards vorhanden ist und dieser sich so einfacher verbreiten kann.

Im folgenden Kapitel wird auf die technischen Spezifikationen des AS67-Standards, sowie dessen Zusammenhang mit SMPTE ST 2110 und Audio Video Bridging eingegangen. Ebenso werden diese beiden Technologien mit ihren Standards und Spezifikati-

onen näher betrachtet, da sie ebenso die Grundlage für (Video- und) Audioübertragungssysteme bilden.

4.2 AES67

Wie in Kapitel 4.1 beschrieben, gab es kaum die Möglichkeit des Datentransfers zwischen Netzwerken. Aus diesem Grund bemühte sich die Audio Engineering Society mit Firmen, Herstellern und Anwendern eine gemeinsame Lösung zu finden. Dies begann 2010 mit der Projektgruppe X192, die Gemeinsamkeiten der verschiedenen Protokolle suchte und die Anforderung hatte, Audiodateien mit einer Mindestauflösung von 44,1kHz und 16 Bit auf eine Latenz von <10ms zu versenden. Aus diesen Bemühungen heraus wurde 2013 das einheitliche Protokoll AES67 veröffentlicht und in den Folgejahren aktualisiert, erweitert und korrigiert. Im April 2018 wurde die aktuelle Spezifikation des Protokolls veröffentlicht und erhält, neben neuen Informationen, Empfehlungen für künftige Erweiterungen. Das AES67-Protokoll ist eine Sammlung aus Standards auf Zeit-, Übertragungs-, Formats- und Inhaltsebene. Um eine hohe Akzeptanz der Hersteller gegenüber AES67 zu erreichen, wurde die Reihe der Standardisierungen klein gehalten. Hauptsächlich wird es eingesetzt, um Layer 3 Netze zu verbinden. Es wurde so spezifiziert, um eine Interoperabilität mit Audio Video Bridging (AVB) Class A und B auf OSI-Layer 2 zu erlauben.

Der dritten und aktuellen Auflage des AES67-Protokolls wurden die Protocol Implementation Conformance Criteria (PICS) hinzugefügt. Mit diesem können einzelne Anforderungen und Spezifikationen für AES67 herausgesucht und abgeglichen werden.⁶⁶

283 Produkte sind laut der AIMS-Allianz, welche die Adoption von AES67 vorantreibt, mit AES67 verifiziert⁶⁷

4.2.1 Synchronisation und Media Clock

Um eine gemeinsame Synchronisation aller Geräte im Netzwerk zu gewährleisten, benutzt AES67 statt der herkömmlichen Wordclock das nach IEEE 1588-2008 definierte Precision Time Protocol (2002 als PTP und 2008 als PTPv2, im Text als PTP abge-

⁶⁶ Gross, K. (2013): AES67 standard for audio applications of networks, *Audio Engineering Society Convention 67(7)*, S. 42;

⁶⁷ <https://aimsalliance.org/aes67-resources/>

kürzt). Bei PTP werden im paketbasierten Netzwerk räumlich verteilte Uhren miteinander synchronisiert. Dies geschieht durch die Bestimmung der exaktesten Uhr, diese Referenzuhr wird „Grandmaster Clock“ genannt. Diese sendet die für das ganze Netzwerk geltende Uhrzeit an die anderen Uhren. Bei dem Paketaustausch von „Master“ (der sendenden Uhr) zu „Slave“ (die empfangende Uhr) werden die dabei entstehenden Verzögerungen mit Zeitstempeln ermittelt.

Der Master sendet eine Zeit-gekennzeichnete Nachricht an den Slave und sendet nach einer kurzen Zeit erneut seine Zeit $t1'$. Der Slave empfängt die erste Nachricht und merkt sich die Empfangszeit $t2$, sowie $t2'$ des Folgepakets. Durch diese beiden Zeiten kann der Offset, die Zeitabweichung zum Master, bestimmt werden. Der Slave sendet nun seinen Zeitstempel $t3$ zum Master und dieser die Empfangszeit $t4$ an den Slave, welcher aus $t3$ und $t4$ die Latenz der Übertragungsstrecke berechnet. Über den Offset-Wert und die Latenz synchronisiert sich der Slave mit dem Grandmaster.

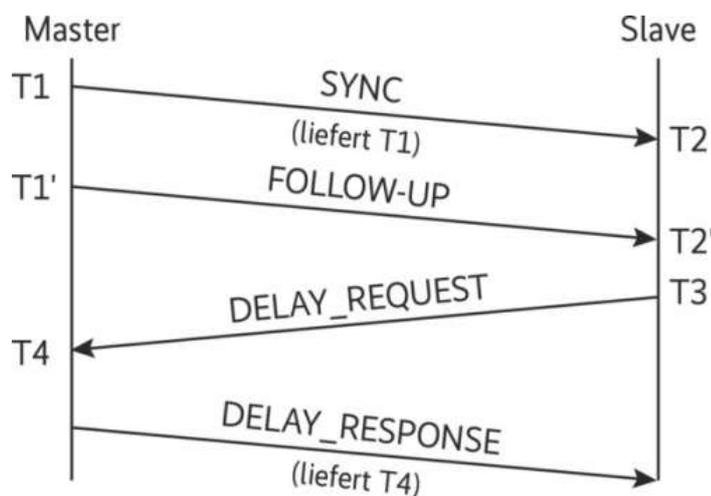


Abbildung 4-1: Synchronisation im Precision Time Protocol⁶⁸

Durch dieses Verfahren werden Synchronisationen vom Mikrosekundenbereich bis in den Nanosekundenbereich erzielt.⁶⁹

Das Ethernet Netzwerk AVB benutzt IEEE 802.1AS zur Synchronisation. Dieser Standard beschreibt ein IEEE 1588-2008-Profil. Aus diesem Grund darf das AVB Netzwerk das native IEEE 802.1AS-Profil zur Synchronisation benutzen.⁷⁰

⁶⁸ https://www.heise.de/scale/geometry/500x500/q60/select/ix/2016/6/1465011889222141/tn_sun.ke.ntp-tp.Z1.jpg

⁶⁹ <https://www.perlesystems.de/supportfiles/precision-time-protocol.shtml>

Die Media Clock (Medien-Uhr) bestimmt die Samplingrate des von ihr gesendeten Streams. Diese muss von dem Empfänger sowie dem Sender (diese Streams) unterstützt werden. Ihre Systemzeit bezieht sie von der Netzwerkzeit, gegeben durch den Grandmaster. Das zu transportierende Audiomaterial wird in der Samplingrate der Media Clock abgetastet. Dem in AES67 definierten Standard nach, muss die Abtastrate 48kHz sein und soll die Raten 44.1 und 96 kHz unterstützen.⁷¹

4.2.2 Transport & QoS

AES67 definiert hier den Transport der Datenpakete über die Netzwerkschicht (Layer-3/Vermittlungsschicht), sowie die Transportschicht (Layer-4). Es wird von einem Transport über IP ausgegangen, jedoch wird der Standard zur Übertragung von IP-Paketen über Ethernet (RFC 894) für Layer-2-Transport ausgeführt.

Die Layer-3-Übertragung geschieht über IPv4-Paketen (nach RFC 791, siehe Kapitel 3.3.1). Neben Unicast muss Multicast über IP-Multicast, nach RFC 1112, ermöglicht sein. Zum Verwalten des Multicast muss IGMPv2 (nach RFC 2236) unterstützt werden und IGMPv3 (nach RFC 3170) darf verwendet werden. Da diese Standards Rückwärtskompatibel sind, kann ein neuerer Standard in einem, auf dem älteren basierenden, Netzwerk arbeiten.

Für die Transportschicht wird in AES67 das User-Datagram-Protocol (nach RFC 769) definiert, welches das Real-Time Transport-Protocol (nach RFC 3550) nutzt. Das vom RTP verwendete Profil ist das nach RFC 3551 definierte „RTP Profile for Audio and Video Conferences with minimal Control“. Ebenso sollte RTCP verwendet werden. RTCP ist ein Protokoll zum Austausch von Steuer- und Verbindungsdaten. Wie in Kapitel 3.3.1 erwähnt, ist auch in AES67 eine Fragmentierung der Pakete unerwünscht. Dadurch muss die maximale Paketgröße bei 1440 Bytes liegen, so dass eine Kompatibilität zu IPv6 in Zukunft möglich ist.

⁷⁰ Gross, K. (2013): AES67 standard for audio applications of networks, *Audio Engineering Society Convention 67(7)*, S.13

⁷¹ Gross, K. (2013): AES67 standard for audio applications of networks, *Audio Engineering Society Convention 67(7)*, S. 13 f.

Verschlüsseltes Streaming, wie TLS⁷², wird in PTP unterstützt, jedoch nicht im AES67-Standard.

Neben den Transportprotokollen wird in AES67 Quality of Service durch priorisierten Transport von Datenströmen definiert. Dieser wird durch die DiffServ-Methode (siehe 3.6), nach RFC 2474, implementiert. Es wird empfohlen, die Priorisierung mit der unten stehenden Tabelle zu erzielen, durch Netz-Administratoren oder Steuerungsschnittstellen dürfen die Pakete auch mit anderen DSCP Werten versehen werden.

Klasse	Transport	DSCP (DiffServ) Werte
Clock	PTP-Synchronisation	EF (46)
Media	RTP/RTCP-Medienstreams	AF41 (34)
Best-Effort	PTP Steuer- und Verbindungsdaten	DF (0)

Tabelle 4: Priorisierung der Streams in AES67⁷³

4.2.3 Codierung & Stream

Der vorgeschriebene Codec zur Übertragung von Audiodaten in AES67 ist L16 und L24. Diese Codecs beschreiben 16 Bit und 24 Bit linear codierte PCM-Signale. Eine Abtastrate von 48 kHz muss unterstützt werden, 44.1 und 96 kHz sind empfohlen. Der Betrieb auf 48 kHz muss mit L16 und L24 unterstützt werden, eine Samplingrate von 44.1 kHz mit L16 (CD-Qualität) und 96 kHz mit L24 unterstützt.

Für das Streaming, die Latenz und Kanalanzahl im Stream ist die Paketzeit wichtig, welche durch die Echtzeitlänge des Audioteils im Datenpaket bestimmt wird. Kleinere Zeiten sorgen für einen höheren Transportaufwand und eine höhere Gerätelast, erzielen jedoch schnellere Latenz. Größere Zeiten sorgen für größere Latenz und Buffering. Hierbei zu beachten ist, dass die „Paketzeit“ in Zusammenhang mit Samples im Paket steht, d.h. je größer die Zeit, desto mehr Samples sind im Paket. Im Rückschluss bedeutet dies, dass mit sehr geringen Zeiten die benötigte Bandbreite des Streams steigt, da

⁷² TLS= Transport Layer Security, auch unter der Vorgängerbezeichnung SSL (Secure Sockets Layer) bekannt ist ein Verschlüsselungsprotokoll zur sicheren Datenübertragung im Internet.

⁷³ Vgl. Gross, K. (2013): AES67 standard for audio applications of networks, *Audio Engineering Society Convention 67(7)*, S. 15 f.

weniger Samples pro Frame übertragen werden, der Frame-Header jedoch immer die gleiche Größe besitzt.

Das Protokoll muss eine Paketzeit von 1ms mit 8 Kanälen pro Stream zur Verfügung stellen. Empfohlen werden ebenfalls Paketzeiten von 125 μ s und 250 μ s, um AVB Class A & B kompatibel zu sein, 333 μ s für niedrig-Latenz-Übertragung und 4ms zum Transport über WANs oder QoS-Eingeschränkte Netzwerke. Mit dem Halbieren oder Vierteln der Paketlänge wird die Kanalanzahl verdoppelt, bzw. vervierfacht.

Die Gesamtlatenz in einem Netzwerk wird durch die Verzögerung zwischen Eingang des Signals und Ausgang des Signals beschrieben.

4.2.4 Sitzung

Zur Beschreibung der Sitzung und der Eigenschaften von Multimediasströmen sieht AES67 das Session-Description-Protocol (SDP, nach RFC 4566) vor. SDP-Datensätze stellen Informationen der Streams über Netzwerkadressierung, Clock-Informationen, Abtastrate, Codierung, Kanalzahl und Dateiformat bereit. Die Datensätze werden für das Connection Management und Discovery verwendet und mit deren Protokollen übertragen.

4.2.5 Discovery (Geräteerkennung) & Connection Management (Verbindungsmanagement)

Um eine Verbindung zwischen zwei Geräten aufzubauen, verwendet AES67 das Session Initiation Protocol (SIP), um eine Unicast-Verbindung zwischen 2 Endpunkten aufzubauen.

Die Vereinheitlichung von Discovery und Connection Management hätte zu einer Inflexibilität des AES67-Protokolls und der Inkompatibilität anderer Protokolle geführt. Aus diesem Grund werden mögliche Standards aufgezählt.

Zur Geräteerkennung können folgende Protokolle verwendet werden: Bonjour, ein offener Standard mit Identifizierung über mDNS und DNS-SD, das Session Announce-

ment Protokoll (SAP) oder AMWA NMOS Discovery and Registration Specification (IS-04).

Wenn Connection Management stattfindet, müssen die Geräte über einen SIP URI⁷⁴ Eintrag aufgezeichnet werden. SIP kann für Unicast benutzt werden. Für Multicast kann SDP über z.B. SAP versendet werden.

Höher liegende Mechanismen, wie Kontrollmechanismen oder Gerätemanagement, sind nicht spezifiziert, so dass die Hersteller ihre eigenen Lösungen verwenden können.⁷⁵

4.3 SMPTE ST 2110

Die Society of Motion Picture and Television Engineers (SMPTE) ist ein 1916 gegründeter Verband zur Entwicklung von Standards und Videoformaten. Die SMPTE ist von internationalen Organisationen wie der International Standards Organisation (ISO) oder dem American National Standard Institute (ANSI) anerkannt.

SMPTE ST 2110 (ST für Standard; im Folgenden mit 2110 abgekürzt) ist eine Sammlung aus Standards, die der Übertragung unkomprimierter Medien-Streams dient. Derzeit sind über 60 Hersteller an der Integration des Standards beteiligt. Wie Andreas Hildebrand mir in einer Mail mitteilt, finden ähnlich des AES67 „Plugfest“ für ST2110 „Dirty Hands“ Workshops statt. In diesen testen AMWA, SMPTE, VSF, AES, JT-NM und EBU Verbindungen auf Basis von ST2110 sowie erweiterten Funktionalitäten wie NMOS (siehe Kapitel 4.5).⁷⁶

Beschrieben wird der Transport und die Synchronisation elementarer Datenströme, wie Audio, Video und Zusatzdaten⁷⁷. Diese werden im Gegensatz zum Vorgänger SMPTE ST 2022 in separaten Streams übertragen und vereinfachen den digitalen Workflow in IP-Umgebung. In dem 2018 aktualisierten Standards wird die Übertragung von Audio (2110-30 und 2110-31), Video (2110-20/21 und 2110-22), der Systemzeit (2110-10)

⁷⁴ SIP-URI dient der Adressierung von SIP-Teilnehmern, indem die Identifizierung über Nutzererkennung und Domain stattfindet. Vergleichbar ist dieser Eintrag mit dem einer Telefonnummer.

⁷⁵ Vgl. Gross, K. (2013): AES67 standard for audio applications of networks, *Audio Engineering Society Convention 67(7)*, S. 24 f.

⁷⁶ Email mit Hildebrand, Andreas, siehe Anhang A

⁷⁷ Engl. „Ancillary data“

und Zusatzdaten (2110-40) beschrieben. Ebenso ist der AES67-Standard zur Übertragung von PCM-Audiodaten integriert. In Entwicklung befinden sich Standards zum Transport von Bit-Raten codierter Videosignale (2110-22), dem Versand einzelner Videosignale über mehrere Streams (2110-23) und dem Versand dynamischer Metadaten (2110-41). 2110-21 beschäftigt sich dabei mit der „Burst-artigen“ Natur von Videosignalen und wie mit diesen umzugehen ist. Von 2110 wird dadurch eine Stream-Management-Funktion übernommen, welche eine Glättung der Netzbelastung mit sich führt und ähnelt damit dem Traffic-Shaping-Prozess des CBS (-> siehe AVB 4.4.2 Transport & QoS).

Für diese Arbeit ist dabei die Betrachtung des Transports von Systemzeit und Audiodaten relevant, ob sich dieser von AES67 unterscheidet und wie sich der 2110-Standard davon unterscheidet.

4.3.1 Synchronisation und Media Clock

ST 2110 benutzt für die Synchronisation und die Zeitinformationen das PTP 1588-2008. In ST 2110 wird die Unterstützung des SMTPE ST 2059-2 PTP Profils festgelegt. Durch Beschränkungen des Profils wird eine Kompatibilität mit dem PTP Media Profile aus AES67 hergestellt. Diese Kompatibilität wird durch AES-R16-2016, einem technischen Report über das „Plug Fest“, sichergestellt.

Die Medienstreams werden getrennt über das Netzwerk gesendet und zur Synchronisation mit RTP Zeitstempeln versehen. Über diese werden die Datenpakete am Signalempfänger synchron zusammengefügt.⁷⁸

Der Takt der Mediuhr wird, wie in AES67, durch die Samplingrate des Audiomaterials bestimmt. Die zu unterstützenden Abtastraten entsprechen dem AES67 Standard.

4.3.2 Transport & QoS

Wie in AES67 wird in 2110 der Transport der Daten über IP/UDP und RTP definiert. Der Audio-Transport wird, wie bei AES67 beschrieben, abgehandelt. 2110 unterscheidet sich im Transport der RTP-Clock durch einen veränderten Offset von AES67. Der

⁷⁸ SMPTE Standard (2017): ST 2110-10:2017 - Professional Media Over Managed IP Networks: System Timing and Definitions, S. 9 f.

Unterschied muss vom Geräteersteller beachtet werden.⁷⁹

QoS ist wie in AES67 für Audio implementiert. Ein Standard zum managen des Multimedia-Streams ist nicht gegeben, kann jedoch durch das Implementieren von AMWA NMOS IS-06 erreicht werden.⁸⁰

4.3.3 Codierung & Stream

Neben dem in AES67 vorgeschriebenen Versand eines PCM-Streams ist es in ST 2110 möglich, Audiodaten über AES3 (siehe Kap. 2.3) zu versenden. In dem Paket sind PCM Daten oder nach ST 337/338 definierte Daten (z.B. Dolby E) enthalten. Das 24 Bit große Paket enthält keine L24 Daten, sondern benutzt 3 Byte für PCM24-Daten und 1 Byte für die AES3 Metadaten. Diese Packmethode basiert auf dem RAVENNA-AM824-Standard.

In 2110-30 orientiert der Standard sich auch bei der Codierung, Auflösung und Samplingrate an AES67. Hier wird in Konformitätsgraden (Conformance Level)⁸¹ unterschieden. Ein Stream muss zwischen 1-8 Kanälen mit 48kHz, L24 und 1ms Paketzeit übertragen können (Level A), Level AX, B, BX, C, CX sollten implementiert werden. Level B ergänzt dieses um die Paketzeit von 125µs bei 1-8 Kanälen und Level C erweitert den Transport bei einer Paketzeit von 125µs auf 1-64 Kanälen. Eine erweiterte Übertragung mit der Abtaste von 96kHz ist ebenfalls möglich, dabei wird die Kanalanzahl halbiert und das Level mit einem „X“ ergänzt (Level AX, BX, CX).

In ST2110-31 muss ebenfalls Level A implementiert werden, jedoch mit 1-6 Kanälen bei 48Khz und bei 96kHz wird die Kanalanzahl halbiert und von einer ungeraden Kanalanzahl auf die nächst tiefere gerundet wird.

Die -X Level werden für die 44.1kHz Abtaste erweitert und es wird das D-Level (und DX-Level) mit kürzeren Paketzeiten angefügt.⁸²

⁷⁹ Im AES67-RTP-Protokoll wird die Stream-Clock, also die Uhrzeit des Datenstroms per Definition mit einem zufälligen Offset belegt. In 2110 wird dieser auf den festen Wert „0“ gesetzt. Ein AES67 Sender muss diesen Wert festlegen können, ansonsten ist keine Kompatibilität gegeben.

⁸⁰ AMWA/NMOS = Advanced Media Workflow Association / Networked Media Open Specification entwickelt Standards, die Arbeiten im Netzwerk beschreiben, z.B. Registration & Discovery, Communication Control, Network Control

⁸¹ Zur Vereinfachung und einfacheren Lesefluss wird der Konformitätsgrad mit „Level“ (von Conformity Level) bezeichnet

⁸² SMPTE Standard (2018): ST 2110-31:2018 - Professional Media Over Managed IP Networks: AES3 Transport, S. 10 f.

4.3.4 Sitzung

ST 2110-30 nutzt wie AES67 SDP-Pakete zum Übertragen von Sitzungsinformationen. In 2110 wird die Möglichkeit gegeben, im SDP-Paket Informationen über die Formatierung der Audiokanäle zu übertragen. So kann die Kanalanzahl und Anordnung der Audiogruppe mitübertragen werden, um das Format, wie Stereo oder Surround zu bestimmen.⁸³

4.3.5 Discovery (Geräteerkennung) & Connection Management (Verbindungsmanagement)

Zu Standards von Discovery und Connection Management sind keine Angaben im ST 2110 Standard gemacht. Für zukünftig zu verwendende Methoden zur Erkennung von Geräten und Streams werden RTSP, SAP und SIP/IGMP vorgeschlagen.

Für die Zukunft ist eine Zusammenarbeit mit der AMWA geplant. Mögliche APIs⁸⁴ könnten die AMWA IS-04 (Geräte und Stream-Erkennung beim Anschluss an das Netzwerk), IS-05 (Connection Management zur Konfiguration der Verbindung zwischen Sender und Empfänger) und IS-06 (Netzwerkkontrolle, Visualisierung der Topologie, QoS und Netzwerkd Diagnose) sein.⁸⁵

4.4 Audio Video Bridging – AVB/TSN

AVB (Audio Video Bridging) ist 2011 von der Arbeitsgruppe IEEE 802.1BA entwickelt worden und bildet den Standard für Echtzeitstreaming zeitkritischer Audio- und Videodaten. Der Standard ist offen und lizenzfrei. Hinter AVB steht die AVNU-Allianz, welche diesen Standard verbreitet, weiterentwickelt und Geräte zertifiziert. 2012 entstand durch das Erweitern des Arbeitsgebiets der AVB Task Group die Time-Sensitive Network Task Group (TSN). Diese arbeitet an der Erweiterung der AVB-Mechanismen und ist ebenso von der AVNU-Allianz getrieben. AVB ist in TSN spezi-

⁸³ SMPTE Standard (2017): ST 2110-30:2017 - Professional Media Over Managed IP Networks: PCM Digital Audio, S. 5 F.

⁸⁴ API = Application Programming Interface; Die Programmierschnittstelle; Ein Programmteil zur Anbindung an das System

⁸⁵https://www.smppte.org/sites/default/files/section-files/BBTB2018-W04-Henry%20Goodman-AES67%20in%20SVIP_0.pdf

fiziert und ist für Echtzeitkommunikation ausgelegt, TSN strebt weiter in den industriell-/technischen Arbeitsbereich, in dem noch kritischere Ansprüche gestellt werden.⁸⁶

AVB streamt über den Ethernet-Standard und benutzt das OSI-Layer 2. Neben dem Einsatz im Live-Bereich wird es aufgrund der geringen Latenz auch im Automotive-Bereich eingesetzt. Im AVB-Netz wird eine andere Wortwahl für die Teilnehmer verwendet. So wird statt Sender und Empfänger, die Bezeichnung der Talker und Listener verwendet. Ein Switch (für Layer-2 Übertragung) wird AVB-Bridge (bzw. Bridge) genannt. Der Begriff Stream, für Datenstrom aus Audio-/Videokanälen, und Controller, für ein Gerät, das Talker und Listener verbindet und konfiguriert, bleiben gleich. Das Netzwerk aus Bridges, in welchem Talker und Listener kommunizieren wird AVB-Cloud (bzw. Cloud) genannt.

Jedes sich in der Cloud befindende Gerät ist mit den in der unteren Abbildung angezeigten Protokollen versehen und kann als Talker, sowie als Speaker agieren.

AVB stützt sich auf 4 Haupt-Standards: IEEE 802.1AS/ 802.1Qat/ 802.1Qav und 802.1BA. Auf diesen aufbauend wurden zum professionellen Transport von Video- und Audiodaten IEE 1722(.1) (Layer 2-Transport) und 1733 (Layer 3-Transport) spezifiziert. Diese Standards wurden für das Protokoll Milan entwickelt, um Interoperabilität zwischen AVB nutzenden Geräten zu gewährleisten. Um diese Protokolle nutzen zu können, müssen alle Bridges (bzw. Switche) in der Cloud mit diesen ausgestattet sein. Die Zertifizierung für AVB-fähige Produkte wird von der AVNU ausgestellt und ist online auf deren Website abrufbar.⁸⁷

⁸⁶ <https://www.itwissen.info/TSN-time-sensitive-networking.html>

⁸⁷ <https://avnu.org/certified-products/>

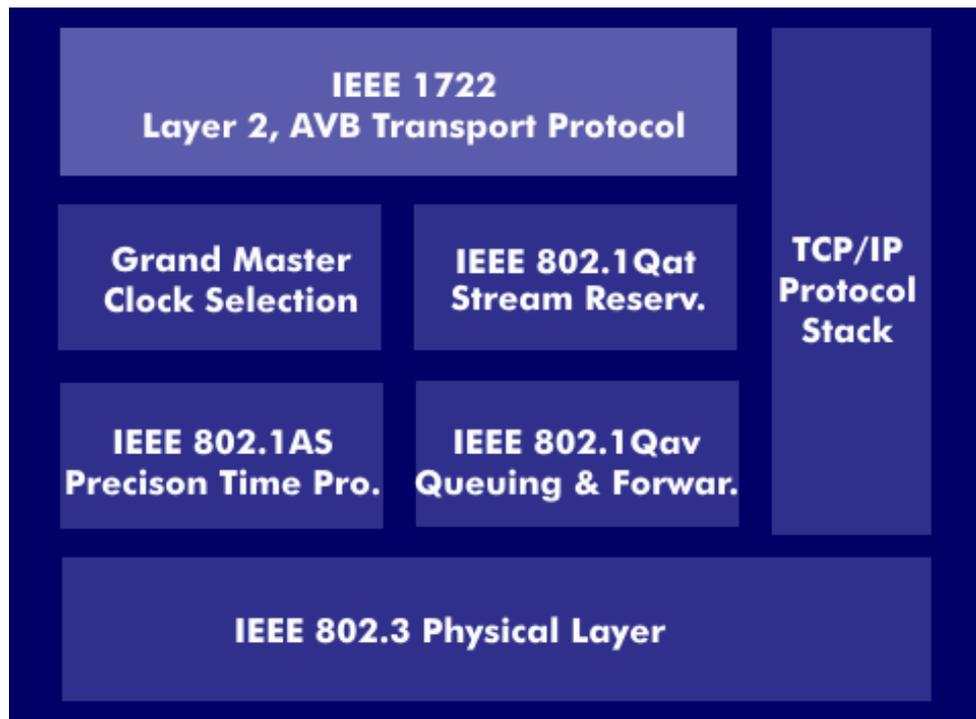


Abbildung 4-2: AVB-Protokollstack⁸⁸

4.4.1 Synchronisation und Media Clock

Ebenso wie bei AES67 und ST 2110 wird PTP bzw. dem IEEE 802.1AS-Standard verwendet. Konkret ist hier das „general Precision Time Protocol“- gPTP in Verwendung. Dieses nutzt die „Grandmaster Clock“, welche nach dem Best Master Clock Algorithmus ausgewählt wird. Ebenso wie bei den anderen Standards wird von dieser die gültige Zeit und der gültige Takt vorgegeben.

Zur Synchronisation des Streams gehörend ist der IEEE1722a Standard AVPT (Audio Video Transport Protocol). Dieser kümmert sich um die Verpackung und Sendung der Daten. Ebenfalls wird hier die Synchronisation der Streams beschrieben. Dies geschieht durch Timestamps in den Paketen. Diese werden durch die in 802.1AS beschriebene MasterClock synchron miteinander gewandelt und abgespielt. Zwar kann die Latenz normalisiert werden, jedoch können keine Codec-bedingten (oder AD/DA Wandlung, Delay durch räumliches Delay) Wandlungszeiten mit einberechnet werden. Die Normalisierung kann beispielweise durch die „Default Presentation Time“ von 2ms gesetzt werden. In dieser Zeit kann ein Paket durch die Cloud übertragen werden (das 1722-

⁸⁸ <https://www.itwissen.info/AVB-audio-video-bridging.html>

Protokoll definiert das maximale Netzwerk-Delay mit 2ms über 7 Hops) und müsste auch abgespielt werden.

4.4.2 Transport & QoS

Der Transport findet, wie erwähnt, auf Ethernetbasis statt und wird über das IEEE 1722-2011 AVB Transport Protocol beschrieben. Die zu versendenden Daten werden auf den anliegenden Takt synchronisiert in ein AVBTP-Paket gepackt, das 1542 Bytes groß ist und mit einem Zeitstempel versehen. Vor dem Verpacken wird überprüft, ob Talker und Listener den gleichen Takt besitzen und beim Controller eine Übertragungsanfrage gestellt. Ebenfalls wird über das Stream Reservation Protocol die nötige Bandbreite angefordert. Der Stream erhält eine eigene ID und das Paket wird an die MAC-Adresse des Listener über alle Bridges (bzw. Switches) geschickt. Die Adresse wird hierbei zufällig aus einem reservierten Adressraum bestimmt, auf ihre Einzigartigkeit geprüft und geschützt. Der Transport findet über einen Spanning Tree mit einer maximalen Hop-Zahl von 7 statt. AVTP stützt sich auf gPTP, SRP und FQTSS.⁸⁹

Der Stream läuft über AVTP (1722-2016), jedoch ist auch das Layer 3 Protokoll 1733-2011 im Standard enthalten. Dieses soll es ermöglichen, über einen RTCP-Format RTP-Zeitstempel mit der 802.1AS-Zeit zu korrelieren und ermöglicht eine Übertragung im eigenen Netz. Dies soll eine Interoperabilität mit AES67-RTP ermöglichen.⁹⁰ IEEE 1733 sei jedoch „eine unvollständige Definition und wird meines Wissens nach auch nirgendwo real verwendet“, erklärt Andreas Hildebrand auf die Frage, ob durch dieses Protokoll auch eine Interoperabilität mit ST2110 möglich sei.⁹¹

AVB bietet ebenso Sicherung von Bandbreite, Priorisierung von Datenströmen und somit QoS-Parameter. Dies geschieht über 802.1Qav FQTSS (Forwarding and Queuing of Time-Sensitive Streams) und 802.1Qat SRP (Stream Reservation Protocol).

Mit FQTSS wird eine Priorisierung der zeitkritischen Daten beschrieben. Diese werden gegenüber zusätzlichen Daten über ein Klassensystem individuell bevorzugt. Es werden in AVB die zeitkritischen SR-Klassen (SR für Stream Reservation) A und B definiert.

⁸⁹ https://avnu.org/wp-content/uploads/2014/05/AVnu-AABAC_IEEE-1722-Media-on-AVB-Networks_Rob-Silfvast.pdf

⁹⁰ <http://www.ieee802.org/1/files/public/docs2009/avb-stanton-rtp-over-avb-aka-ieee1733-0309.pdf>

⁹¹ Email mit Hildebrand, Andreas, siehe Anhang A

Nicht zeit-kritische Daten werden nach dem Best-Effort Prinzip übertragen, jedoch können diese Ströme nach VLAN-tagging priorisiert werden.

Die QoS-Eigenschaften bei Klasse A setzen sich wie folgt zusammen:

125µs Paketlänge, eine Maximallatenz von 2ms bei 7 Hops, Pufferzeit von 2ms.

Für Klasse B gilt:

250µs Paketlänge, eine Maximallatenz von 50ms bei 7 Hops, Pufferzeit von 2ms.

Um Bursting und Bunching⁹² entgegenzuwirken wird in AVB ein CBS (Credit-based Traffic Shaper) verwendet. Dieser sorgt (durch das Verschachteln von Streams und Auffüllen der Lücken mit Best-Effort-Paketen) für ein gleichmäßiges Versenden der Daten und damit einer konstanteren Datenrate.⁹³

Die Bandbreitenreservierung findet über SRP statt. Hier werden 75% für Class A&B Verkehr zugesichert, 25% für andere Daten. Die 75% sind per Default eingestellt, können jedoch manuell verändert werden. Die Bandbreite wird vom Talker und Listener am AVB-Controller angefragt. Dieser reserviert diese und hält sie bis zum Ende des Streams.⁹⁴

4.4.3 Codierung & Stream

Die Codierung sowie der Stream finden über IEEE 1722a AVTP statt. AVTP transportiert Audio-, Videoformate, die Mediuhr, Netzwerk-Diagnosedaten und Kontrolldaten. Über das Multicast Adress Allocation Protocol (MAAP) werden die Streams einer MAC-Adresse zugeordnet. Im Videostream können SDI-codierte Bitstreams nach SMPTE Standards wie ST 259 (270Mbit/s), 292 (1,5Gbit/s), 424 (3Gbit/s)⁹⁵, etc. übertragen werden. Komprimierte Formate sind MJPEG, JPEG 2000 und H.264.

Für den Transport von Audiodaten sind im Stream eine hohe Zahl an Kanälen zulässig, die Paketzeit beträgt nach Klasse 125µs oder 250µs. Samplerates von 44.1–192kHz sind möglich, sowie eine Datentiefe von bis zu 32 Bit. Übertragene Codecs sind das unkom-

⁹² Bursting und Busting beschreibt den Vorgang des Anstauens von Datenpaketen und dem folgenden unmittelbaren Versand über einen Netzwerkknoten

⁹³ <https://www.elektroniknet.de/elektronik-automotive/bordnetz-vernetzung/schnelle-datenuebertragung-121066-Seite-2.html>

⁹⁴ https://avnu.org/wp-content/uploads/2014/05/AVnu-AABAC_Traffic-Shaping-Infrastructure-Planning_Andre-Fredette.pdf

⁹⁵ https://avnu.org/wp-content/uploads/2014/05/AVnu-AABAC_IEEE-1722-Media-on-AVB-Networks_Rob-Silfvast.pdf

primierte PCM-Format und das Advanced Authoring Format (AAF). AAF enthält Mediendaten sowie zugehörige Metadaten und speichert beispielweise die Schnitte in einer Mediendatei. Eine konkrete Angabe zum Verpacken der Daten in den Stream ist nicht gegeben, jedoch ist die Packweise von Paketlänge, Samples im Paket, Auflösung des Signals, Kanalanzahl im Stream und Datenübertragungsrate abhängig und wird vom Hersteller definiert.⁹⁶

4.4.4 Sitzung

Die Beschreibung der Medienströme, sowie deren Metadaten werden über AVTP übertragen. Der Aufbau der Streams, deren Handling und Shaping wird vom Stream Reservation Protocol (SRP) übernommen.

4.4.5 AV Discovery, Enumeration, Connection Management and Control

SRP ist, aufbauend auf dem Multiple Reservation Protocol (MRP), ein Protokoll zum Netzwerkmanagement und ermöglicht dem Talker und Listener ein Finden des Partners, des VLANs und der Stream-Priorisierung.

In AVB sind die Mechanismen der Geräteerkennung und des Verbindungsmanagement über den Standard 1722.1 (IEEE 1722.1-2013) gegeben. Neben dem Verfahren, Geräte zu erkennen (Discovery) und untereinander zu verbinden (Connection Management), ist es ebenfalls beschrieben, deren Eigenschaften und Fähigkeiten zu ermitteln (Enumeration) und diese zu konfigurieren und zu steuern (Control). Diese Mechanismen werden unter AVDECC – AV Discovery Enumeration, Connection Management and Control zusammengefasst. Geräte arbeiten mit 1722.1/AVDECC auf der Basis von 1722/AVTP.

Im Standard wird jede an der in der Cloud teilnehmende Einheit als „Entity“ bezeichnet. Diese kann vier Rollen einnehmen; Listener & Talker (Die wie oben beschrieben dem SRP nach handeln), Controller oder Responder (Entity, welche weder Listener, Talker oder Controller ist, aber AVDECC verwendet).

AVDECC arbeitet mit dem Entity Model (AEM), Discovery Protocoll (ADP), Enumeration and Control Protocol (AECP) und dem Connection and Management Protocol (ACMP).

⁹⁶ https://avnu.org/wp-content/uploads/2014/05/AVnu-AAA2C_Audio-Video-Transport-Protocol-AVTP_Dave-Olsen.pdf

AEM überführt die Eigenschaften der Geräte in die Baumstruktur und beschreibt die Konfigurationen und technischen Eigenschaften dieser.

ADP bestimmt die Entity Findung über den Controller. Dieser sendet über die Cloud die „Entity Discover“ Aufforderung, dass alle Entitys sich über die „Entity Available“ Nachricht als vorhanden melden. Diese Meldung wird periodisch zum Controller gesendet, da dieser die Entity sonst abmeldet. Ebenfalls kann sich die Entity durch „Entity Departing“ abmelden.

Darauf aufbauend kann der Controller das AECP nutzen. Mit diesem lassen sich Datenströme befehligen, Informationen anfordern, Konfigurationen durchführen und Kontrollelemente bedienen.⁹⁷

ACMP definiert Verfahren und Kommandos des Controllers, um eine Verbindung zwischen Talker und Listener aufzubauen oder diese Stream-Verbindungen aufzulösen. Hierbei findet wie bei AECP eine Nachrichtensicherung durch Rückmeldung des Erfolgs oder eines Fehlers statt.

Neben der Möglichkeit, die Controller über ein TCP/IP Netzwerk anzusteuern, ermöglicht AVDECC es ebenfalls über eine XML-Dokument Offline-Konfigurationen der Geräte zu tätigen, bevor diese in einem Netzwerk angeschlossen werden. Die Remote-Steuerung findet über HTML1.1 statt und wird über SSL/TLS-Verfahren verschlüsselt.⁹⁸

4.5 NMOS IS-04/IS-05 by AMWA/ AES70

Wie bei den bisherigen Standards zu sehen ist, gibt es bisher einen geringen Nenner zur Interoperabilität beim Transport von Audiodaten. Höhere Protokolle zur Gerätesteuerung, dem Versenden von Steuerdaten, der Geräteerkennung, etc. sind (zum Teil abgesehen von AVB) noch nicht standardisiert.

Derzeit gibt es zwei verfolgbare Ansätze, in welche Richtung es weiter gehen könnte. Die Open Communication Standard-Allianz (OCA) unterstützt das entwickelte AES70. Dieses dient, aufbauend auf einem Transportprotokoll (wie beispielweise AES67) für das Verbindungsmanagement, der Gerätesteuerung und der Überwachung der Daten-

⁹⁷ <https://avb.statusbar.com/article/1722.1-features/>

⁹⁸ <https://avb.statusbar.com/files/presentation/avdecc-aes-ny-2013.pdf>

ströme.

D&b-Audiotechnik setzt auf den OCA/AES70 Standard, mit welchem sich d&b-Produkte in Steuerungssysteme von QSC, Crestron, AMX und Beckhoff einbinden lassen.⁹⁹ Die OCA verspricht dadurch eine Kontroll-Interoperabilität zwischen verschiedenen Geräteherstellern, so dass deren Geräte über ein Netzwerk Steuer- und Kontrolldaten schicken können. AES70 kann in Netzwerken mit implementierten Transportprotokollen verwendet werden, d.h. in AES67, SMPTE 2110 und auch in AVB-Umgebung. Der Standard beschränkt sich auf die Kontrolle von Audiogeräten.¹⁰⁰

Gegenüber einer Erweiterung mit AES70 steht die AMWA. Die Advanced Media Workflow Association ist eine Vereinigung, die die Steigerung der Effizienz in der Medienproduktion im Broadcastbereich vorantreibt, mit Organisationen wie der EBU, SMPTE oder VSF zusammenarbeitet und das AAF-Format und das MXF-Format entwickelt. Die AMWA ist eine Non-Profit Organisation deren Mitglieder aus Bereichen wie der Netzwerktechnik (Cisco, Dell, ...) und Medienproduktionshäusern (BBC, NBC, ProSiebenSat1, Walt Disney, ...) kommen. Mit Networked Open Media Specifications (NMOS) strebt die AMWAS eine Entwicklung von offenen interoperablen Spezifikationen im IP-basierten Umfeld an.

Im SMPTE ST2110 werden die NMOS Standards IS-04/05/06 in Aussicht gestellt. IS-04 übernimmt hierbei die Erkennung und Registrierungsmechanismen der Geräte, IS-05 das Connection Management mit Auf- und Abbau von Uni- und Multicast Streams. Durch Dezentralisierung und Aufteilung auch in einem WAN funktionierend.¹⁰¹ IS-06 übernimmt die Netzwerkkontrolle, wie es auch schon bei AVB ähnlich 802.1Qat mit dem SRP. IS-06 ermöglicht es dabei dem Broadcast Controller die Netzwerkstruktur zu analysieren, zu überwachen, Streams zu erzeugen, zu verändern, Bandbreite zu reservieren und mit Stream-Priorisierung weitere QoS-Parameter auszubauen.

NMOS entwickelt weitere Spezifikationen zum Ausbau der Standards. Von weiterem Interesse könnte auch IS-07, zeitgesteuerte Übertragung von Steuerbefehlen oder IS-08, eine Spezifikation zum zeitgesteuerten patchen von Audiokanälen, sein. NMOS arbeitet

⁹⁹ <https://www.dbaudio.com/global/de/loesungen/signalverteilung-und-audio-netzwerke/ocaaes70/>

¹⁰⁰ <https://www.oaalliance.com/oca-vision/>

¹⁰¹ Ähnlich wie bei PTP werden Bereiche aufgestellt in welchem zentrale Registrierung stattfindet und diese von einem übergeordneten Controller überwacht wird.

offen an den Spezifikationen, deren Dokumentationen und Stände online über Github einsehbar sind.¹⁰²

¹⁰² <https://amwa-tv.github.io/nmos/>

5 Betrachtung der Audionetzwerke

Die oben genannten Standards bieten eine Grundlage für eine gemeinsame Kommunikation. Diese Grundlage reicht jedoch nicht aus, dass zwei Geräte miteinander verbunden werden und anfangen zu kommunizieren. Dazu benötigt es weitere Protokolle und Software, welche dies ermöglicht. Im Folgenden werden die größten Audioprotokolle, die auf die obigen Standards setzen, erläutert. Dabei werden die technischen Eigenschaften und Spezifikationen der Audionetzwerke, sowie die für Funktion sorgenden Protokolle mit ihren Aufgaben in Zusammenhang gestellt. Neben dem strukturellen Aufbau des Protokolls und dem Aufbau der darin versendeten Daten werden Sicherungsmechanismen für Stabilität und Zusammenspiel der verwendeten Netzwerkteilnehmer aufgezeigt.

5.1 RAVENNA AES67

RAVENNA ist ein AoIP-Netzwerk und wurde von ALC NetworX entwickelt, einer zu LAWO gehörenden Firma. Die Technologie wurde Ende 2010 veröffentlicht, über die Jahre weiterentwickelt und wird im Bereich der professionellen Audio- und Broadcastübertragung eingesetzt. Es ist nativ auf dem AES67 Standard aufgebaut, erweitert diesen und unterstützt inzwischen mit ST2110-30 und ST2110-31 den Audioteil von SMPTE 2110.¹⁰³

RAVENNA ist eine offene und lizenzfreie Technologie, für die Hersteller eigene Lösungen entwickeln können. Die Technologie wird von Andreas Hildebrand, Senior Produkt Manager, verbreitet. Dieser setzt sich als Mitglied der AES Arbeitsgruppe für die Entwicklung, Verbreitung, Wartung und Interoperabilität verschiedener Geräte mit dem AES67-Standard ein und arbeitet an der SMPTE ST 2110-Standardisierung. Im Jahr 2018 waren 166 Ravenna unterstützende Produkte von 25 Herstellern verfügbar.¹⁰⁴

¹⁰³ <https://www.ravenna-network.com/adopting-ravenna/overview>

¹⁰⁴ <http://rhconsulting.eu/blog/files/NetworkedAudio.html>

5.1.1 Technische Merkmale

RAVENNA versendet die Daten nach AES67-Standard über einen RTP-Stream. Es werden PCM-Daten im L16 und L24-Format mit einer Samplingrate von 32 kHz bis zu 192 kHz und mit einer Paketlänge von 125µs, 250µs, 1/3ms, 1ms und 4ms versendet. Ebenfalls ist dies mit 32 Bit AES3 Dateien möglich. Im AES3 Format können neben PCM auch Dolby E komprimierte, DSD/DXD-Formate (hochauflösendes Audio mit bis zu 384kHz Samplingrate) und die gängigen AES3 Codecs versendet werden. Mehr als 512 Audiokanäle können in einen Stream gepackt werden. Die Kanalanzahl ist mit angebotener Software erweiterbar.¹⁰⁵

Der Netzwerkverkehr läuft über Layer-3 Switche und für die Mindestgeschwindigkeit werden 100Mbit/s (Fast Ethernet) empfohlen. Ravenna kann auch auf 1Gbit/s oder 10Gbit/s für große Datenströme skaliert werden. Ravenna verfügt über keine feste Topologie und kann in ein bestehendes Netzwerk integriert werden. Ebenfalls kann ein weiter entferntes Ravenna-Teilnetzwerk über das WAN mit einem anderen Ravenna-Netz verbunden werden, so dass beispielsweise eine Regie mit einem Veranstaltungsort in Verbindung stehen kann. Hierfür ist eine ausreichende Bandbreite zum Transport der Daten nötig¹⁰⁶

5.1.2 Netzwerkeigenschaften

Ähnlich wie bei anderen Layer 3 Netzwerken, wird in Ravenna RTP über UDP/IP transportiert und QoS findet mit DiffServ statt. Die Synchronisation aller Geräte im Netzwerk geschieht über PTPv2.

Die Ende-zu-Ende Latenz in Ravenna ist durch die Anzahl der Samples in einem Frame einstellbar. So ist, wie in Kapitel 4.2.3 beschrieben, bei einem kleinen Paket die Latenz sehr gering. Aus diesem Grund ist es mit Ravenna möglich, bis zu 1 Sample pro Paket zu übertragen, um eine so geringe Latenz wie möglich zu erzielen (deutlich <1ms). Aufgrund des mitgesendeten Frame-Header ist dies jedoch sehr ineffizient und erhöht die nötige Bandbreite. Um eine richtige Ausspielung der Daten zu gewährleisten, muss eine Buffer-Latenz im System eingestellt werden. In Ravenna kann diese für jeden Stream individuell eingestellt werden und der größte Wert bestimmt die Gesamtlatenz.

¹⁰⁵ Andreas Hildebrand (2014): RAVENNA&AES67_V1.0, München, ALC NetworX GmbH S. 8 ff.

¹⁰⁶ Andreas Hildebrand (2017): AES67 Practical Guide, München, ALC NetworX GmbH, S. 5

Dies kann über die Preset-Werte Klein, Mittel und Groß geschehen, oder individuell mit Zahlenwerten eingestellt werden. Des Weiteren muss Latenz berücksichtigt werden, die nicht im Netzwerk entsteht, sondern bei A/D-, D/A-Wandlung, in einer DSP oder beim Codieren.¹⁰⁷

Ein Quality of Service findet in Ravenna wie in AES67 über DiffServ und die damit verbundene Zuordnung von DSCP-Werten statt. Die Priorisierung lautet wie bei AES67 (Mit Werten): PTP Verkehr (46) > RTP Verkehr (34) > restlicher Verkehr (0).¹⁰⁸

5.1.3 Betriebs-/Erweiterungsmöglichkeiten & Kompatibilität

Zur Erkennung von Geräten und deren Einrichtung ins Netzwerk werden von Ravenna (neben dem manuellen Eintragen) zwei Methoden geboten. In größeren Netzwerken mit Servern in der Infrastruktur kann DHCP/DNS verwendet werden. In kleineren lokalen Netzen kann der Zeroconfig-Mechanismus/Bonjour (DNS-SD basierter Dienst) verwendet werden. Zur Interoperabilität mit Dante Netzen ist ein SAP-Konverter vorhanden.¹⁰⁹

Der Transport der Stream-Metadaten, sowie der Verbindungsaufbau des Streams findet über RTSP/SDP statt. RTSP ist für die Kontrolle zuständig, während über SDP Daten übertragen werden. Low-Level Gerätekonfiguration ist über das HTTP-Protokoll vorhanden und kann über den Webbrowser getätigt werden.¹¹⁰

Für Ravenna wird Control-Software angeboten. Darunter „DashBoard“ von ROSS Video, sowie der SAP-Converter von ALC NetworX, mit welchem in Ravenna AES67-fähige Geräte und Dante Geräte miteinander verbunden werden können.

MERGING hat die kostenlose kompetente Steuersoftware ANEMAN veröffentlicht. Mit dieser ist eine graphische Oberfläche zur Verbindungskontrolle gegeben, ähnlich der Dantes, sowie eine Remote-Funktion, Netzwerkkonfiguration und Netzkontrolle.

¹⁰⁷ <https://www.ravenna-network.com/using-ravenna/technology/>

¹⁰⁸ Andreas Hildebrand (2017): AES67 Practical Guide, München, ALC NetworX GmbH, S. 3

¹⁰⁹ <https://www.ravenna-network.com/aes67/rav2sap/>

¹¹⁰ Andreas Hildebrand (2017): AES67 Practical Guide, München, ALC NetworX GmbH, S. 8 f.

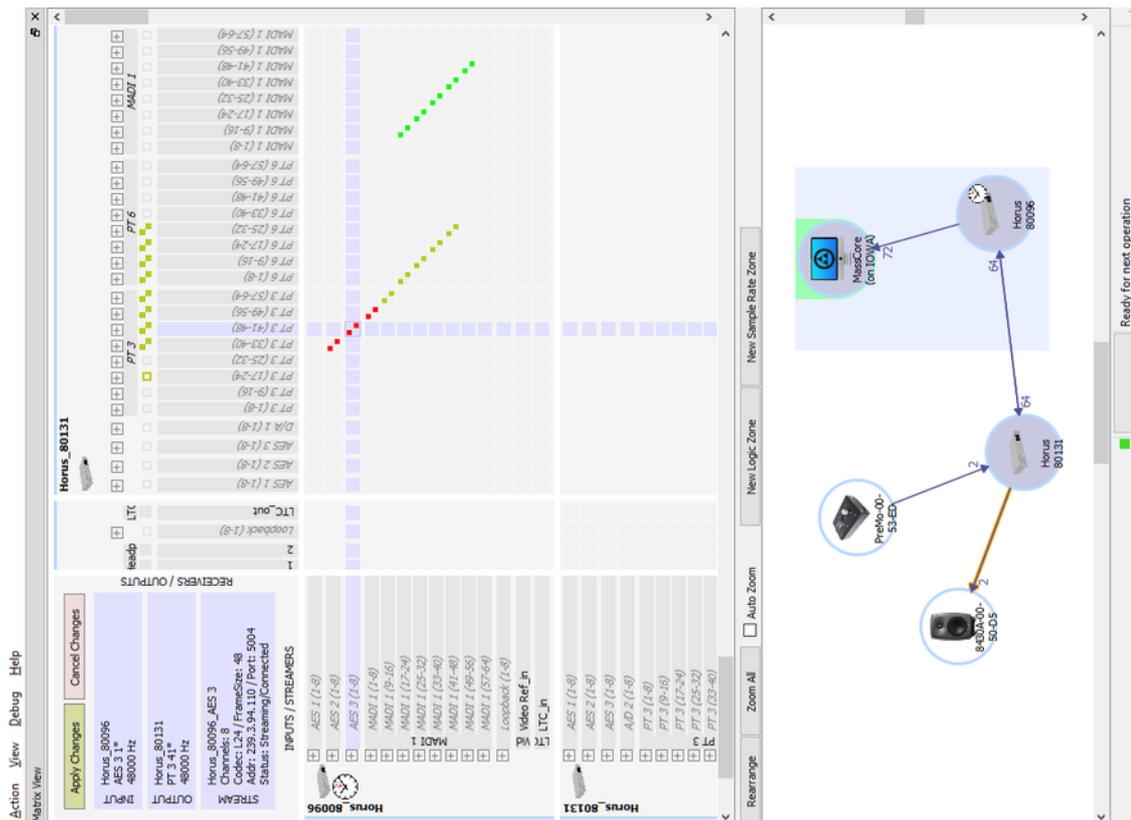


Abbildung 5-1: Bedienoberfläche der ANEMAN-Oberfläche¹¹¹

Die Implementierung der Technologie kann über Softwaretreiber erfolgen. So gibt es kostenfrei die RAVENNA Virtual Sound Card, welche 8x16 Kanäle mit 44.1/48kHz verwalten kann. Von LAWO und MERGING gibt es noch weitere Software (für Windows-Systeme), mit der man bis zu 384 Kanäle verwalten kann und von MERGING gibt es Software für OSX mit 64x64 Kanälen bei 48kHz. Hardware-Lösungen von ALC, ARCHWAVE, DIGIGRAM, ROSS und MERGING können bis zu 512x512 Kanäle auf allen Sample Rates verwalten.¹¹²

Hierbei ist bei vielen Herstellern schon NMOS IS-04/05 integriert.¹¹³

Über die Steuerungssoftware müssen die Übertragungen alle auf Sampling Rate, Paketzeit und Auflösung eingestellt werden. Neben dem Standardprofil gibt es Profile für hohe Performance, Ultra-low Latency, sowie AES67- und ST2110-Profile. Profile sind

¹¹¹ <https://www.ravenna-network.com/adopting-ravenna/aneman/>

¹¹² <https://www.ravenna-network.com/adopting-ravenna/oem-solutions/>

¹¹³ <https://www.rossvideo.com/oem-developer/bach-openmodule-st2110-aes67-audio-networking-module/>

individuell anpassbar und erstellbar. Um miteinander zu funktionieren, müssen diese sich gleichen.

Latenzzeiten müssen manuell über den Controller und mithilfe des Netzwerkanalyse-tools eingestellt werden.

Eine Kompatibilität mit AES67- und ST2110-fähigen Geräten ist gegeben und Video-versand ist bei Hardware mancher Hersteller implementiert. In den RAVENNA-Whitepapers wird SMPTE ST2110-30/31 zum Audiotransport erwähnt. Dadurch ist der Anschluss an ST2110-fähige Streams möglich. Ein einfaches Einbinden des Streams ist durch die Struktur des VSF TR-03 gegeben, da die Medienstreams getrennt über das Netzwerk laufen und RAVENNA diese direkt abgreifen kann.

Neben der Implementierung des IS-04 und IS-05 ist auch eine Nutzung des OCA-Kontrollprotokolls möglich.

5.2 Dante(-AV)

Dante ist wie Ravenna eine IP-basierte Lösung zum Transport echtzeitkritischer Mediendaten. Dante wurde im Jahr 2006 von der australischen Firma Audinate veröffentlicht und ist inzwischen das meist verbreitetste Audionetzwerk. Mit über 260 lizenzierten Herstellern und über 1600 Dante-fähigen Produkten besitzt Audinate derzeit den größten Marktanteil¹¹⁴. Die Technologie ist geschlossen und Lizenzgeschützt. Das Netzwerk arbeitete anfangs auf proprietären Protokollen, 2016 fand eine Integration von AES67 in die Produkte statt, im April 2018 wurde die Integration von SMPTE ST2110-10/30 angekündigt sowie von ST2059-2 (Standard zur Synchronisation von PTP-uhren).¹¹⁵

Auf der ISE-Messe (Integrated Systems Europe) im Januar 2019 stellte Audinate das Dante AV Modul vor. Damit stößt Dante in den AV over IP Bereich vor. Dante AV soll über ein 1Gbit-Netzwerk laufen, Codec unabhängig sein und HDCP Verschlüsselung-/Entschlüsselung-Support besitzen. Die geplante Veröffentlichung findet im Juni 2019 statt.¹¹⁶

¹¹⁴ <https://www.audinate.com/products/dante-enabled#catalog>

¹¹⁵ <https://www.prosoundnetwork.com/live/aes67-avb-tsn-dante-power-aoips-rise>

¹¹⁶ <https://www.professional-system.de/news/ise-2019-dante-kann-jetzt-auch-irgendwie-video/>

5.2.1 Technische Merkmale

Dante unterstützt den Transport von 16 Bit, 24 Bit und 32 Bit PCM codierten Audiodateien bei Sampling Raten von 44.1, 48, 88.2, 96, 176.4 und 192 kHz. Audiokanäle werden für den Versand in „Flows“ organisiert. In einen Unicast-Flow können bis zu 4 Kanäle gepackt werden, bei einem Multicast-Flow bis zu 8 Kanäle.¹¹⁷ Zu der Paketgröße im Stream werden keine Angaben gemacht, jedoch kann diese für das AES67 auf die geforderten 1ms gesetzt werden und für sehr geringe Latenz auf 4 Samples.

Für ein Netzwerk ist die Mindestanforderung von 100Mbit/s gegeben. Gigabit-Switches werden empfohlen, um eine bessere Latenz zu erreichen und die Bandbreite sollte entsprechend der Größe des Projektes angepasst werden.¹¹⁸

Dante kann in jeder Netzwerktopologie aufgebaut werden, eine bessere Leistung wird jedoch beim Bau eines sternförmigen Netzes erzielt, da hier die Anzahl der Hops für die Übertragungen geringgehalten werden können. Durch ein PRIMARY und SECONDARY Port an jedem Gerät ist eine zweite Leitung für Redundanz vorhanden. So kann im Falle eines Ausfalls direkt auf die andere Leitung umgeschaltet werden.

5.2.2 Netzwerkeigenschaften

Dante arbeitet auf dem OSI-Layer 3 und benutzt zur Übertragung UDP/IP. Im Gegenteil zu anderen Netzwerken wird hier nicht das RTP, sondern ein eigens entwickeltes proprietäres Protokoll verwendet.¹¹⁹ Die Synchronisation findet wie bei anderen Netzwerken über PTP statt, QoS über DiffServ und Streams werden mit SDP beschrieben.

Die Ende-zu-Ende Latenz in einem Dante Netzwerk ist von dessen Größe abhängig. Als typischer Latenzwert werden 150µs angegeben, die mögliche Minimallatenz in einem Netz ist 83.3µs (< 1/10ms). Für den Transport wird die Faustregel von 100µs pro Hop angegeben, garantiert wird eine Latenz von unter 1ms bei 10 Hops im Netzwerk. Hierbei empfiehlt Dante:

¹¹⁷ Audinate (2018): Dante Certification Program - Level 3 – Module 22: Dante Flows. Youtube, 05.05.2018, Web, 11.05.2019, in: <https://www.youtube.com/watch?v=Vi8B1KV0AJM&list=PL2UzvxHZNUyLbXLopZyF5IhOelyH8hFoP&index=22>, 01:55 Minuten

¹¹⁸ <https://www.audinate.com/sites/default/files/PDF/advanced-dante-networking-avnw-2015-audinate.pdf>

¹¹⁹ Arne Bönninghoff (2017): „IP Media Standards“, In: *Vdt-Magazin* 4/2017, S. 21

1 Switch Hop: 150µs

3 Switch Hops: 250µs

5 Switch Hops: 500µs

10 Switch Hops: 1ms

+10 Switch Hops: 5ms

Die eingestellte Latenz gilt immer für die Verbindung zwischen Empfänger und Sender. In einem System können so mehrere Latenzen bestehen. Zum Einstellen der richtigen Latenz bietet Dante eine Monitoring Funktion in der Controller Software.¹²⁰

Für eine stabile Latenz soll QoS in einem Dante Netzwerk aktiviert sein. Dies geschieht über die DiffServ-Funktion. Dabei wird in 4 Klassen folgend priorisiert (mit Dezimalwerten): zeitkritische PTP-Ereignisse (56) > Audiodaten, PTP (47) > Interne Daten (8) > Restliche Daten im Netzwerk (0).¹²¹

5.2.3 Betriebs-/Erweiterungsmöglichkeiten

Zur Anmeldung von Geräten in Dante gibt es wie bei Ravenna die Möglichkeit, Geräte über einen DHCP Server anzumelden und deren IP-Adresse zu konfigurieren. Ist ein DHCP-Server nicht vorhanden, findet die Zuweisung über Link-Local statt.¹²² Nach der IP-Anmeldung findet die eigentliche Discovery der Geräte statt. Diese findet in Dante über ein Proprietäres Protokoll (Dante Discovery) statt, welches eine automatische Erkennung, Anmeldung und Anzeige im Controller ermöglicht und wird als „Plug and Play“ Feature bezeichnet.¹²³

Die Kontrolle über Geräte- und Netzmanagement wird vom DANTE Controller übernommen und läuft ebenfalls über ein proprietäres Protokoll. Neben der Konfiguration der Geräte ist der Controller fähig, die Kanäle zu routen sowie die Latenz und Sample-rate einzustellen. Des Weiteren können mit der Überwachung des Netzwerkes Analyse-daten, wie Latenz und verwendete Bandbreite betrachtet werden. Es gibt die Funktion der Kanalbenennung, Gerätebenennung und es können Subgruppen gebildet werden. Diese Einstellungen sind ebenfalls in Presets speicherbar.

¹²⁰ https://dev.audinate.com/GA/dante-controller/userguide/webhelp/#latency_tab.htm

¹²¹ <https://www.audinate.com/networks-and-switches>

¹²² https://dev.audinate.com/GA/dante-controller/userguide/webhelp/#discovery_and_auto-configuration.htm#kanchor167

¹²³ <https://www.audinate.com/faq/what-dante-discovery>

Der Controller wird durch den Dante Domain Manager (DDM) erweitert. Dieser ist Desktop- oder Server basiert und bietet administrative Einstellungen, sowie verschlüsselten Zugriff bei Großprojekten.¹²⁴

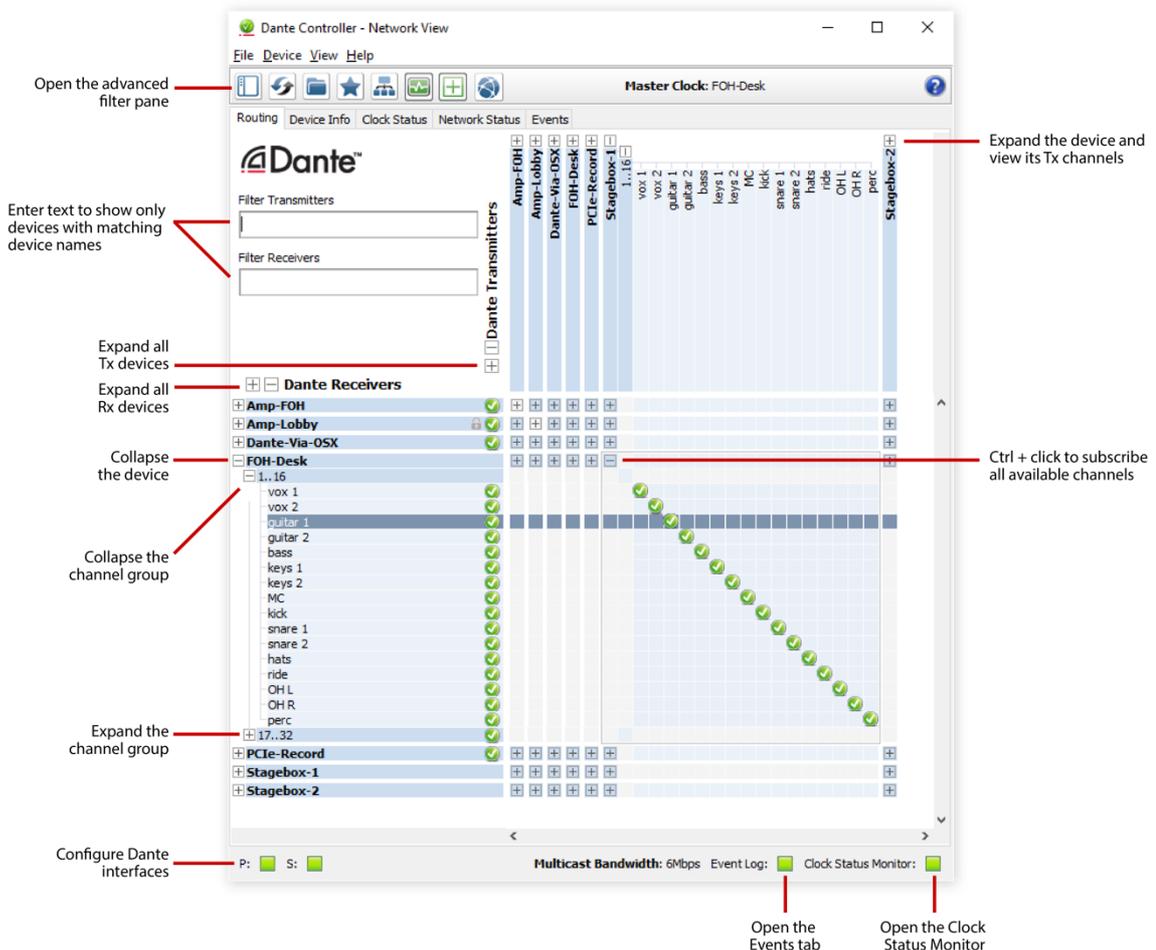


Abbildung 5-2: GUI des Dante Controllers¹²⁵

Mit der Virtual Soundcard und Dante Via sind zwei weitere Schnittstellen für das Abgreifen von Programmdatenströmen und die Einbindung eines Personal Computers gegeben.

Über ein Profil ist die AES67-Kompatibilität gegeben, jedoch muss ein SAP-Converter für die Geräteerkennung eingesetzt werden und es sind nur die Grundlagen vorhanden. Eine Kompatibilität mit AVB war geplant und wurde in vorausgehenden Abschlussarbeiten erwähnt. Nach Sven Schuhen sei diese jedoch um 2017 stillschweigend gestri-

¹²⁴ <https://dev.audinate.com/GA/dante-controller/userguide/pdf/latest/AUD-MAN-DanteController-4.1.x-v1.0.pdf>

¹²⁵ <https://dev.audinate.com/GA/dante-controller/userguide/pdf/latest/AUD-MAN-DanteController-4.1.x-v1.0.pdf> S. 43

chen worden.¹²⁶ Auf Audinates Website wird eine Interoperabilität mit AVB ausgeschlossen und nicht weiter erläutert.

5.3 AVB/Milan

Milan ist eine anwendungsbasierte Lösung zum Transport von AV-Mediendaten, die auf AVB basiert und Interoperabilität zwischen Geräteherstellern ermöglichen soll. Das Protokoll wurde 2018 auf der InfoComm von den Herstellern D&B-Audiotechnik, Avid, Meyer Sound, L-Acoustics und Biamp angekündigt. Hierbei muss erwähnt werden, dass Milan den AVB-Standard um IEEE 1722-2016 und 1722.1-2013 erweitert, um einen funktionalen interoperablen Rahmen zu bilden, da AVB zwar die Übertragung definiert, aber nicht die Ende-zu-Ende Interoperabilität oder ein gemeinsames Stream Format. AVB-Lösungen verschiedener Hersteller können standardmäßig nicht miteinander arbeiten, Milan-Geräte verschiedener Hersteller können gemeinsam in einem Netzwerk funktionieren. Das Milan-Protokoll ist offen und dessen Kontroll-Werkzeuge sind frei zugänglich. Wie bei AES67 sind Treffen geplant, auf denen die Kompatibilität aller Geräte geprüft wird.¹²⁷

Für eine gewährleisteteste Interoperabilität in einem Netzwerk müssen AVB-zertifizierte Switche vorhanden sein. Milan-fähige Geräte werden wie die Switche über die AVnu-Allianz zertifiziert. Derzeit sind 25 Geräte¹²⁸ von der Avnu zertifiziert.¹²⁹

5.3.1 Technische Merkmale

Das Milan-Protokoll schreibt den Transport von PCM codiertem Audio über AAF vor. Pakete werden nicht als solche beschrieben, sondern als PDUs (Protocol Data Units), also als Protokoll-Datenpakete. Diese beinhalten eine Anzahl an Samples sowie einen Zeitstempel. Im Protokoll wird ein Standardformat, ein 32Bit High Capacity-Format (HC32) und 24Bit HC-Format (HC24) beschrieben. Das Standardformat sichert die Interoperabilität mit anderen Milan-Geräten, HC32 ermöglicht den Transport vieler Kanäle bei geringer CPU-belastung und HC24 den Transport vieler Kanäle bei geringer

¹²⁶ Email mit Schuhen, Sven, Siehe Anhang

¹²⁷ <http://www.soundandcommunications.com/an-introduction-to-milan/>

¹²⁸ Hierbei handelt es sich um Modelle, bzw. deren Reihe. Nachfolgergeräte eines Modells werden nicht mitgezählt.

¹²⁹ <https://avnu.org/certified-products/>

Bandauslastung.

Das Standard Format definiert 48 kHz mit 6 Samples/PDU, 96 kHz mit 12 Samples/PDU und 192 kHz mit 24 Samples/PDU. Übertragen werden zwischen 1-8 Kanäle/Stream.

Das HC32-Format wird mit 48 kHz bei 6 Samples/PDU und 16-56 Kanälen/Stream (in 8er Schritten) beschrieben und bei 96 kHz 12 Samples/PDU mit 16 oder 24 Kanälen/Stream).

HC24 enthält bei 48 kHz 6 Samples/PDU und zwischen 1-64 Kanälen/Stream, bei 96 kHz 12 Samples/PDU mit 1-40 Kanäle/Stream und bei 192 kHz 24 Samples/Stream mit 1-16 Kanälen/Stream.

Die Unterstützung von 48 kHz-Signalen muss vorhanden sein, 96 kHz und 192 kHz sind optional. Wird 192 kHz unterstützt muss 96kHz unterstützt werden. Dies gilt für alle Übertragungsformate. Das Standardformat muss vorhanden sein, HC32 und HC24 können unterstützt werden. Wird das Format HC24 unterstützt, muss HC32 ebenfalls unterstützt werden.

Eine Redundanz des Netzwerkes ist durch den möglichen Aufbau eines primären, sowie sekundären Netzwerkes gegeben, welche parallel zueinander bestehen. Fällt in einem Netz die Verbindung aus, so kann das zweite Netz dieses wiederherstellen.¹³⁰ Wie bei AVB wird ein Milan Netzwerk über das Entity Modell (AEM) ebenfalls in der Baum-Topologie aufgebaut.

5.3.2 Netzwerkeigenschaften

Da Milan auf AVB aufbaut, wird AVTP zum getrennten Transport der Medienströme über OSI-Layer 2 verwendet und QoS über IEEE 802.1Q, mit Bandbreitenreservierung und Priorisierung, gegeben.

Die Bildung der Netzwerkehren findet über gPTP statt. Deren Synchronisation wird in AVB nicht erwähnt und wird in Milan über den AES11 Standard ergänzt.¹³¹

Bei der Ende-zu-Ende Latenz wird, wie in AVB, eine Latenz von 2ms auf 7 Hops ver-

¹³⁰ Cole Peterson (2018), Network redundancy interoperability specification, AVNU MILAN, S. 4

¹³¹ Marc Illouz (2018): Media clocking functional & interoperability specification, AVNU MILAN, S. 4

sprochen, bei einem Ethernet Gigabit-Switch wird die Latenz noch weiter gesenkt und dabei bei 25µs/Hop liegen.¹³²

5.3.3 Betriebs-/Erweiterungsmöglichkeiten

In Milan findet, wie in AVB, Discovery über ADP, Connection Management über ACMP und das Control Protokoll über AECP statt.¹³³ ACMP steuert zu Teilen das SRP; beide Protokolle wurden in Milan so definiert, dass eine „Plug-and-Play“-Fähigkeit innerhalb des Netzes vorhanden ist.¹³⁴

Eine gemeinsame Controllersoftware für Milan ist zum derzeitigen Zeitpunkt nicht veröffentlicht, Produkte werden von herstellereigener Software gesteuert. So ist für Produkte von L-Acoustics, wie der „P1 Processor and Measurement Platform“ und deren LA4X & LA12X Verstärker, die „Soundvision“ und „LA Network Manager“-Software verfügbar. D&b Audiotechnik wollen in Zukunft komplett auf Milan, als Standard zur Audioverteilung, setzen und bieten für die „DS20 Audio Network Bridge“ die R1 Fernsteuer-Software an. Diese erweitert Milan mit dem OCA (Open Control Architecture) /AES70 Steuerprotokoll. Mit diesem ist eine aus der Ferne steuerbare Audiotbearbeitung möglich.¹³⁵ Als Open-Source Lösung wurde von L-Acoustics eine AVDECC-Bibliothek zur Steuerung online gestellt, ebenfalls steht der Open-Source Controller „Hive“ von Kikisoft zum Download verfügbar.¹³⁶

Milan ist nicht AES67 oder ST2110 kompatibel.¹³⁷ Um in einem AVB/Milan-Netzwerk arbeiten zu können, müssen zertifizierte Switche verwendet werden.

¹³² Kreifeldt, Rick (2009): AVB for professional A/V Use, AVNU MILAN, S. 4

¹³³ Marc Illouz (2019): Discovery, connection and control specification for talkers and listeners, AVNU MILAN

¹³⁴ Marc Illouz (2019): Discovery, connection and control specification for talkers and listeners, AVNU MILAN, S. 57

¹³⁵ <https://www.dbaudio.com/global/de/produkte/processing-und-matrix/ds20/>

¹³⁶ <http://milanavcommunity.proboards.com/thread/9/available-milan-controllers>

¹³⁷ <http://rhconsulting.eu/blog/files/Milan.html>

6 Allgemeiner Vergleich/Bewertung

Im Folgenden werden die erläuterten Audioprotokolle unter technischen, sowie anwendungspraktischen Aspekten miteinander verglichen und eine Bewertung dieses Vergleichs durchgeführt. Eine tabellarische Darstellung dieser Bewertung befindet sich im Anhang-B. Um eine vereinfachte Bewertung darstellen zu können, wurden die Signalfarben Grün, Gelb, Orange und Rot verwendet. Rot steht dabei für ein ungenügendes oder negativ zu betrachtendes Ergebnis. Die Farbe Gelb wird für das Erreichen der gestellten Anforderung einer Kategorie verwendet und wird als „ausreichend“ bewertet. Bei der Farbe Orange gilt selbiges, jedoch müssen diskutierte Beschränkungen beachtet werden. Mit der Farbe Grün wird ein Erreichen der Anforderung, sowie ein Übertreffen dieser signalisiert und wird mit „gut“ bewertet.

Zum Vergleich wurde eine Tabelle angefertigt. Diese befindet sich in Anhang B. Der Vergleich findet in Kategorien statt, die Teile der Tabelle zusammenfassen. Die Bewertung erfolgt mit dem Vergleich. Der Vergleich und die Bewertung finden in folgenden Kategorien statt: Netzwerk (Größe, Redundanz und Topologie), Latenz, Flexibilität (Samplerate, Bittiefe und Dateiformat), Quality of Service, Kompatibilität, Zugänglichkeit (Lizenz, Standard und Steuerprotokoll), Komfort (Plug & Play und Controller) und der Videofähigkeit.

Netzwerk

AES67 und ST2110 definieren, über welche Protokolle eine Zusammenarbeit in einem Netzwerk stattfinden kann und sind keine Netzwerke. Die verwendeten Protokolle arbeiten auf Layer 3 des OSI-Modells. AVB beschreibt den Aufbau eines Netzwerkes, welches getaktet Daten über die Sicherungsschicht transportiert. In AVB wird eine Redundanz durch das STP erreicht, welches den Spanning-Tree in der Cloud bildet. Für das darauf aufbauende Milan Protokoll gilt selbiges. Als Layer-2 Technologien sind diese auf LAN-Netzwerke beschränkt. Die Layer-3 Protokolle Dante und Ravenna sind LAN-fähig. Ravenna kann, bei gegebener Infrastruktur, Daten in WANs transportieren. Dante kann, wenn mit dem Dante Domain Manager gesteuert wird, dies ebenfalls. Jedoch ist DDM als serverbasierte Methode mit deutlichen Mehrkosten verbunden.

Ravenna oder Dante haben keine festen Topologien, sollten aufgrund von Latenz jedoch bevorzugt als Stern aufgebaut werden. Durch den Einsatz von Managed Switches sind Routing und redundante Wege möglich. Redundanz erhalten die Geräte in Ravenna, Dante und Milan durch einen zweiten Netzwerkport, durch welchen ein paralleles und redundantes Netz aufgebaut wird. Bei Ausfall des primären Netzes wechselt der Datentransport auf das Sekundäre, die Controller aller drei Protokolle könne dies melden.

Latenz

Die Bewertung der Kategorie „Latenz“ ist von mehreren Faktoren abhängig. An erster Stelle steht die Echtzeitfähigkeit. Echtzeitsysteme sind nach Claus und Schwill als „Systeme zur unmittelbaren Steuerung und Abwicklung von Prozessen“, definiert und müssen quantitative gestellte Zeitanforderungen erfüllen.¹³⁸ Die hierfür verwendete Zeitanforderung wird über die subjektive Wahrnehmung von Musikern bestimmt. In einer Studie des AES¹³⁹ wird eine Verschlechterung des Zusammenspiels der Musiker bei zu hoher Latenz festgestellt. Lester und Boley, die Autoren der Studie, schlussfolgern dabei auf einen Latenzwert von 6,5ms.

Um Latenz, die außerhalb des Netzwerkpfades liegt, zu kompensieren, wird der für die Bewertung kritische Wert auf 5ms gesenkt. Ebenso ist hierbei zu erwähnen, dass bei Live-Beschallung weitere Komponenten (z.B. Plug-Ins für Tonbearbeitung) im Netzwerk hängen und Verzögerungen verursachen. Aus diesem Grund sollte Latenz für Live-Beschallung unter 1ms liegen. Im Broadcastbereich wird noch mehr Hardware zwischengeschaltet, wodurch die Latenz hier noch kritischer bewertet werden muss.

In AES67 wird nur vorgeschrieben, dass die Geräte in einem Bereich unter 10ms arbeiten müssen. Ebenfalls werden Paketgrößen mit 125µs Länge für kurze Latenz definiert (siehe 4.2). AES67 und ST2110 werden für diesen Vergleich jedoch nicht gewertet, da diese kein komplettes Netzwerk beschreiben.

Die Maximallatenz bei AVB sowie Milan wird durch die 2ms bei 7 Hops beschrieben,

¹³⁸ Claus, Volker/ Schwill, Andreas (2003): Informatik – Ein Fachlexikon für Studium und Praxis, Dudenverlag, Mannheim, S. 537

¹³⁹ Lester, Michael und Boley, Jon (2007): The Effects of Latency on Live Sound Monitoring, Audio Engineering Society Convention 123

minimal sind sogar $25\mu\text{s}$ pro Hop erreichbar. Dantes Maximallatenz liegt bei 5ms . Diese entsteht jedoch bei mehr als 10 Hops. Minimal wird in Dante $84\mu\text{s}$ Latenz bei 4 Samples im Paket erreicht. Ravenna gibt nur allgemein an, dass eine Latenz $<1\text{ms}$ erreicht wird. Verglichen mit Dante, welches ebenfalls über IP/UDP transportiert, kann die Aussage getroffen werden, dass Ravenna, im Ultra-Low-Latency Modus mit 1 Sample pro Paket, ähnliche Zeiten erreichen muss.

Alle zu vergleichenden Protokolle können die Anforderungen erfüllen. Obwohl in Dante 5ms Latenz entstehen können, geschieht dies nur in einem großen Netzwerk mit mehr als 10 Hops. Eine Live-Beschallung soll in einem kleinen Netzwerk stattfinden, in dem die Stage-Box, das FOH und Mischpult über einen Switch verbunden sind. Bei Dante sind $150\mu\text{s}$ für einen Hop angegeben, was für ein LAN-Netzwerk vollkommen ausreicht. Ebenso ist dies für Milan/AVB zu bewerten, da diese Netzwerke mit $25\mu\text{s}/\text{Hop}$ eine außerordentlich geringe Latenz besitzen. Ravenna als WAN-fähiges Broadcastnetzwerk kann ebenfalls mit dem Ultra-Low-Latency Modus deutlich die gestellten Anforderungen erfüllen. Als einzige Anmerkung zu den Layer-3 Lösungen ist hinzuzufügen, dass mit der Verkleinerung der Pakete die angeforderte Bandbreite ansteigt, da hier immer der Paket-Header mitgesendet wird. AVB/Milan sind durch den fehlenden IP-Header sparsamer und effizienter.

Flexibilität

Die für diesen Vergleich zu ermittelnde Flexibilität setzt sich aus den Kombinationsmöglichkeiten der Samplingrate und der Bittiefe im Audioprotokoll zusammen. Ein weiterer Faktor wird durch die übertragenen Datenformate gegeben. Eine zuverlässige Audioübertragung sollte mit 24 Bit und 48 kHz stattfinden, wie es auch im professionellen Audiobereich üblich ist. Für High Fidelity-Anwendungen und eine gute Bearbeitungsmöglichkeiten werden Aufnahmen in 96 kHz, 192 kHz oder 32 Bit Auflösung benötigt. Die Audiodatei soll verlustfrei PCM codiert sein, so dass uneingeschränkt bearbeitet wird. Weitere Transportformate sind nicht erforderlich, können jedoch Bandbreite einsparen oder erweiterte Datensätze transportieren.

AES67 ermöglicht den Transport mit 16 Bit Tiefe bei 44.1 kHz und 48 kHz und 24 Bit Tiefe bei 48 kHz und 92 kHz und setzt damit die gesetzte Anforderung um. ST2110 nimmt den AES67-Standard komplett auf und übertrifft die Anforderungen mit dem AES3-Format. Dante erweitert die mögliche Samplingrate auf 192 kHz und überträgt

im L24 und L32 Audioformat. Ravenna deckt von allen Protokollen den größten Frequenzbereich von 32-192 kHz ab und versendet (auf ST2110 aufbauen) neben PCM noch im AES3-Format und erweitert mit dem hochauflösenden SA-CD-Format DSD/DXD die transportierbare Samplingrate auf 384 kHz. AVB transportiert ebenfalls eine erweiterte Samplingrate von 44.1-192 kHz, jedoch nur auf 32 Bit. Die PCM Daten werden über AAF versandt, welches dieses Format um Metadaten erweitert. Milan transportiert Samplerates von 48-192 kHz mit 24 & 32 Bit PCM über AAF.

Alle Protokolle transportieren den erweiterten Frequenzbereich und die angeforderte Bittiefe. Dante, Milan und Ravenna erweitern die Auflösungstiefe auf 32 Bit, AVB spezifiziert den Transport auf 32 Bit. 24 Bit Daten werden in AVB durch anhängen von Nullen auf 32 Bit vergrößert, was zu einer erhöhten Bandbreite führt und als kritisch zu betrachten ist. Aus diesem Grund wird AVB in dieser Kategorie nur als ausreichend bewertet.

Neben dem ausreichenden Transport von unkomprimierten Audiosignalen zeichnet sich Ravenna mit dem Transport des AES3-Formats aus. Dieses kann neben Dolby-E weitere AES3-Codecs und PCM über AES3 in MADI überführen. Mit DSD/DXD wird ein Highfidelity-Format im Protokoll aufgenommen, welches im Heimbereich eingesetzt werden kann.

Milan und AVB senden PCM über AAF. AAF wird vor allem für die Medienpostproduktion eingesetzt, da es die Metadaten von Schnittprogrammen mitliefert. Das Format wird von Sony, Avid, BBC, CNN, etc. unterstützt. Somit kann Milan einen vereinfachten Workflow in einem Medienbearbeitungssystem mit sich bringen.

Quality of Service

Der hier betrachtete Quality of Service wird über implementierte Mechaniken in den Protokollen beschrieben, welche es ermöglichen, eine sichere und schnelle Übertragung des Medienverkehrs zu betätigen. Mögliche Mechaniken sind die Reservierung der Bandbreite oder die Priorisierung der Datenströme. Ist eine Mechanik vorhanden, gilt dies als den Anforderungen ausreichend zu bewerten. Sind weitere Sicherungsmechaniken vorhanden, wird dies als ein Übertreffen der Anforderungen bewertet.

In AES67 und somit ST2110 ist das QoS über die DiffServ Methode beschrieben. Ebenso bei Dante und Ravenna. Hier wird eine Stabilität durch Priorisierung der Datenströme erreicht. Die WorldClock steht dabei an erster Stelle, da es sonst zu Jitter kom-

men würde und die Geräte nicht synchron sind, danach kommen Audiodaten und folgend alle weiteren transportierten Daten. AVB und Milan stellen über IEEE 802.1Q einen Quality of Service zur Verfügung. Mit dem SRP wird bei Verbindungsaufbau eines Streams eine Bandbreitenreservierung durchgeführt. Ebenfalls werden Streams getagged und über ein Klassensystem priorisiert. Des Weiteren wird mit dem Credit Based Shaper eine Glättung des Streams erzielt und damit eine stabilere Auslastung der Bandbreite erzielt.

AES67, ST2110, Ravenna und Dante erfüllen mit dem DiffServ Prinzip die gestellten Forderungen und stellen eine isochrone Übertragung sicher. Dante ist laut Definition AES67-kompatibel, verwendet jedoch eine andere Wertezuordnung bei den DSCP-Werten, wodurch native AES67 Clock-Streams mit der Priorität Mittel, statt Hoch versendet werden. Hier muss eine Überführung der Werte von Dante in das andere AES67-System erfolgen.

AVB/Milan übertreffen die Anforderung. Es findet nicht nur die Priorisierung statt, sondern ein effektiveres Ausnutzen der Bandbreite durch den CBS. Mit der Bandbreitenreservierung für die Streams wird ein Integrated Service/garantierter Service zur Verfügung gestellt.

Kompatibilität

Die Kompatibilität beschreibt die Fähigkeit in einem gemeinsamen Netzwerk zu agieren und zu kommunizieren. Die Implementierung eines Standards, der zwischen Audioprotokollen vermittelt, ist hierfür ausreichend. Wenn über mehrere Standards kommuniziert wird, übertrifft dies die Anforderungen.

AES67 selbst ist eine Standardisierung von Protokollen, die Interoperabilität über ein Netzwerk ermöglicht. Obwohl es für Layer 3 konzipiert wurde, sind Klassen für den Audiotransport auf Layer 2 AVB spezifiziert. ST2110 baut darauf auf und integriert AES67 zum Audiotransport in seine Standards. Als Audioprotokoll integriert Ravenna AES67, sowie ST2110 vollständig. Dante ist ebenfalls als AES67 kompatibel zu bezeichnen, benötigt für den Datenaustausch einen SAP-Adapter. AVB beschreibt die Struktur eines Netzes und den Transport, verfügt über keine automatischen Erkennungsmechanismen anderer AVB-Geräte oder spezifiziert eine gemeinsame Zusammenarbeit. AVB kann AES67 Kompatibilität erreichen, wenn das Class A- und B-

Format bereitgestellt wird. Milan erweitert dies und ermöglicht die Zusammenarbeit von Milan-Geräten verschiedener Hersteller in einem Netzwerk.

AES67 wird als Grundlage für rudimentäre Übertragung nicht bewertet, da es selbst das verbindende Glied darstellt. ST2110 verwendet dieses für den Transport, so dass AES67-Geräte mit ST2110-Geräten Audiodaten austauschen können. Hierbei ist jedoch zu beachten, dass ST2110 trotz gestellter Beschränkungen mit AES67 kompatibel ist. Ravenna beachtet dies und ist gegenüber AES67- und ST2110-Geräten kompatibel.

Dante ist streng genommen AES67 kompatibel, verwendet jedoch nur die minimalen Anforderungen für den AES67 Transport. So ist es möglich nur mit einem 48 kHz auf 24 Bit aufgelösten Signal, Dante und AES67-Geräte in einem System zu verbinden. Hinzu kommt die Pflicht, einen AES67 zu SAP Adapter zu verwenden, da sonst keine Discovery stattfindet. Eine AVB-Kompatibilität wurde angekündigt, jedoch nie umgesetzt. Eine Begründung oder Absage dieser Kompatibilität fand nie statt, auf Audinates Website wird die Kompatibilität mit AVB ausgeschlossen.

AVB definiert das gemeinsame Netz. Kompatibilität ist erst über die Hersteller gegeben. Milan ermöglicht über die Standards völlige Interoperabilität über das Netz und erfüllt die Anforderung der Kompatibilität.

Zugänglichkeit

Um individuell auf Problemstellungen eingehen zu können und an Lösungen sowie Weiterentwicklungen eines Protokolls arbeiten zu können, ist die öffentliche Zugänglichkeit ein entscheidender Faktor. Offene Standards und freie Lizenzen ermöglichen eine freie Entwicklung des Produkts durch die Öffentlichkeit und erleichtern Interoperabilität. Ebenso liegt dadurch die Technologie nicht bei einem Akteur und ist von diesem abhängig, sondern kann bei Ausfall dessen weitergeführt werden.

AES67, ST2110, Ravenna, sowie AVB/Milan sind für die Entwicklung offen zugänglich und lizenzfrei. Dante verlangt Lizenzgebühren und definiert keinen offenen Standard.

Wie oben erwähnt, ermöglichen offene Standards einfachere Lösungsansätze und vereinfachtere transparentere Problembehandlung. Die geschlossene Systemstruktur von Dante führt zwar zu einem homogenen, in sich funktionierenden System, lässt jedoch kaum Ansätze zu, dieses aus Eigenbedarf weiter zu entwickeln und anzupassen. Die Verantwortung über das Produkt liegt bei Dante. Wird eine Änderung der Firmenpolitik

oder Preisstaffelung entschieden, sind die Gerätehersteller von der Dante-Technologie abhängig. Sichtbar ist dies beispielweise bei dem herausgegebenen AES67 Format in Dante, welches stark beschränkt wurde. Eine eigenständige Erweiterung dessen ist nicht möglich. Fällt bei der Milan-Group eine Firma aus oder geht Bankrott kann die Technologie von allen anderen weiterentwickelt werden. Ebenso ist dies bei AES67 der Fall, da dies offen zugänglich ist und alle Akteure damit entwickeln können. Aus diesem Grund werden die offenen, lizenzfreien Standards/Technologien als positiv bewertet, während ein geschlossenes, abhängiges System negativ bewertet wird.

Komfort

Der Komfort eines Systems lässt sich durch mehrere Faktoren beschreiben. Diese sind durch folgende Fragen gestellt: Liegt dem System eine einfach zu bedienende Oberfläche vor, die Übersicht bietet und in die Technologie integriert ist? Finden eine automatische Erkennung und Verbindung mit den im Netz angeschlossenen Geräten, auch Plug & Play genannt, statt? Benötigt die Bedienung ein ausgeprägtes Vorwissen? Hierbei werden AES67 und ST2110 als Standards nicht bewertet.

Für AVB liegt keine Bedienoberfläche vor. Dies wird erst mit der Integration von IEEE 1722.1 möglich. Ebenfalls ist in AVB kein Standard nativ vorhanden, welcher Plug & Play ermöglicht. Es ist auf Seite der Hersteller möglich, diese Optionen zu integrieren, wie es auch bei Milan spezifiziert wurde.

Dante verwendet den Dante Controller oder den Domain Manager zur Kontrolle von Netzwerken und Verbindungen. DDM bietet dabei mehr Möglichkeiten der Kontrolle eines großen Netzwerkes und den darin liegenden Domains. Da Dante proprietäre Protokolle wie Dante Discovery verwendet, wird die Steuerungssoftware ebenfalls von Dante gestellt. Dante Discovery ermöglicht eine automatische Erkennung, Anmeldung und eine Anzeige der Geräte in der Software und ermöglicht eine direkte Konfiguration. Dantes Steuerungssoftware bietet ebenso vorgefertigte Presets, ermöglicht Routing und Beobachten der Netzwerkuhr und Latenz im Netzwerk und stellt Informationen zu den angeschlossenen Geräten. Die Bedienungsoberfläche des Controller ist simpel und übersichtlich aufgebaut. Die Grundfunktionen sind über Video-Tutorials auf der Webseite von Audinate erklärt.

Ravenna bietet zwar Discovery der Geräte im Netz an, die Konfiguration muss jedoch über eine Controllersoftware geschehen. Derzeit gibt es drei verschiedene Lösungen,

die von verschiedenen Firmen angeboten werden. Der „RAV-2-SAP Converter“ ermöglicht die oben beschriebene AES67-Verbindung von Ravenna zu Dante. „ANEMAN“ ist ein Controller zum Patchen, ausführen von Remote Control der Geräte, sowie kontrollieren und konfigurieren des Netzwerks. „Dash Board“ ist eine offene Plattform zur Aufsicht und Kontrolle über ein Netzwerk (Hierbei für Audio, sowie Videogeräte). Die Steuersoftware für Ravenna bietet voreingestellte Presets und lässt viele Einstellungsmöglichkeiten zu, im Direktvergleich zum Dante Controller ist die Oberfläche weniger übersichtlich. Durch den Aufbau ist ebenfalls ein tieferes Wissen nötig, um die angezeigten Daten zu verstehen und interpretieren zu können.

Für Milan ist wie bei Ravenna kein zentrale einheitliche Steuersoftware vorhanden. Diese wird von den Geräteherstellern zur Verfügung gestellt. So stellen d&b sowie L-Acoustics die Softwarelösungen „LA Network Manager“ und die „R1 Remote Control“ für ihre Geräte zur Verfügung oder der Open Source Controller „HiveE von Kikisoft steht zur Verfügung. Ebenfalls wurde von L-Acoustics eine AVDECC-bibliothek zur eigenen Entwicklung und Weiterentwicklung von Lösungen online verfügbar gestellt. Im Milan Protokoll sind über das AVDECC Protokoll Plug & Play Mechaniken implementiert. Diese lassen sich über die Softwarelösungen nutzen. Neben dem Monitoring des Systems und Einstellmöglichkeiten der Geräte sowie Routing ist es über das OCA-Protokoll möglich, Signalbearbeitung, wie die Einstellung eines Equalizers oder die Veränderung des Abstrahlverhaltens eines Arrays, vorzunehmen. Der LA Network Manager beinhaltet interaktive Tutorials zum Erlernen der Bedienung.

Im Umfang der Steuerung-Softwares der verschiedenen Audioprotokolle ähneln sich alle und erfüllen die grundlegenden Funktionen. Jedoch ist anzumerken, dass Dante mit übersichtlichen gestalteten Oberflächen und dem Einsatz von Farbcodes die vorhandene Lage schnell vermitteln kann und einfachen Bedienkomfort bietet.

Die für Ravenna angebotenen Lösungen entsprechen den Anforderungen, zeichnen sich jedoch durch keine Besonderheiten aus.

Die für Milan angebotene Software erfüllt auch grundlegende Forderungen, zeichnet sich jedoch durch die Nutzungsmöglichkeit der AES70/OCA-Protokolle und den damit Verbundenen Steuerungsmöglichkeiten, die Plug & Play-Fähigkeit und die offene Steuerungsbibliothek aus.

Videofähigkeit

Die Videofähigkeit eines Systems wird durch die Fähigkeit beschrieben, Videostreams innerhalb des Netzwerkes zu transportieren.

ST2110 beschreibt in der Sammlung der Standards die Übertragung unkomprimierter Videostreams. Derzeit steht der Standard für die Übertragung komprimierter Formate aus, wird jedoch nachgereicht. In Dante ist es für Geräte mit der Dante AV Hardware möglich, Videos zu übertragen. In Milan/AVB ist dies im Standard beschrieben.

In einem Ravenna Netzwerk werden über ST2110 die Videostreams übertragen. Durch die definierte Netzwerkstruktur ist es Audiogeräten möglich die Audiodaten aus den TR-03 Streams direkt zu lesen, zu verarbeiten und bei Bedarf einzugliedern. Ein Vorteil gegenüber SDI-Streams in älteren SMPTE Standards ist es, dass kein De-Embedden getätigt werden muss. Die Latenz wird dadurch geringer gehalten und die Synchronisation steht am Ende der Übertragungskette.

Dante AV gibt nicht an, wie die Videostreams übertragen werden, jedoch ist es weiterhin möglich, mit Dante-geräten in einem solchen Netzwerk zu arbeiten.

AVB/Milan unterstützen Videoversand nativ und durch das Protokoll erhalten Geräte die nötigen Streams.

Mit AES67 ist ein solider Standard geboten, welcher die Interoperabilität zwischen Audioprotokollen vorantreibt und einen gemeinsamen Austausch von Audiodaten ermöglicht. Da es damit nur den kleinsten gemeinsamen Nenner definiert und keine Mechanismen für Erkennung von Geräten oder Netzwerkmanagement vorschreibt, liegt die Implementation bei den Geräteherstellern. Mit der Integration von AES67 gilt dies ebenfalls für ST2110. Jedoch wird hier die Kompetenz eines Protokolls durch die Erweiterung hinsichtlich des Videotransportes gesteigert und die Kompatibilität mit mehr Systemen gegeben. Die Interoperabilität von AES67 wird regelmäßig mit dem von der AES ausgeführten „Plugfest“ geprüft und sichergestellt. Diese Verifizierungsmethode ermöglicht den Herstellern eine Rückmeldung und Verbesserung ihrer Produkte.

Ebenso wird dies mit den „Dirty Hands“ Workshops für ST2110 getan, um Rückmeldung einzuholen und die Interoperabilität zu verifizieren.

AVB beschreibt eine gemeinsame Übertragung, aber keine Ende-zu-Ende Interoperabilität. Die mit AVB ausgestatteten Geräte sind standartmäßig nicht kompatibel miteinander und eine Verifizierung dessen gibt es nicht. Dies ist eben der Grund für die Grün-

dung der Milan-Gruppe, welche eine Interoperabilität aller Milan-Geräte in einem Netzwerk ermöglicht.

7 Zusammenfassung und Ausblick

Jedes der behandelten Protokolle bietet Vorteile sowie Nachteile, wodurch sich weder Ravenna, Dante oder Milan als Favorit für alles empfehlen lassen. Anhand der gegebenen Unterschiede können sich die Netzwerke auf Einsatzgebiete zusammenfassen lassen.

Milan, ein aus AVB entwickeltes Layer-2 Netzwerk, bietet durch seinen Aufbau und die niedrige Transportebene eine schnelle gesicherte Übertragung in kleineren Netzwerken. Ebenfalls können neben hochaufgelösten Audiodaten, Videodateien versendet werden. Eine volle Interoperabilität für Milan-fähige Geräte ist gegeben, jedoch funktioniert diese Technologie nur über zertifizierte Switche. Kleinere Netze, in denen Milan eingesetzt wird, können von Rauminstallationen bis zu Konzerthallen reichen. Durch den Transport von hochaufgelöstem Audio und dem Mitsenden von Video über das multimediale AAF-Format bietet sich Milan für die Arbeit mit Bild- und Videomischern, bzw. für eine Arbeitsumgebung mit Medienbearbeitungsprogrammen an. Der bisherige Einsatz und die vorhandenen Produkte sind auf den Einsatz in größeren Beschallungsanlagen ausgelegt.

Das IP-Netzwerk Ravenna bietet durch seine Systemstruktur ebenfalls den Versand von hochaufgelöstem Audiomaterial über mehrere Datenformate. Durch AES67- und ST2110-Kompatibilität ist eine standardisierte Audio- und Videoübertragung, sowie die Integration unterschiedlicher Geräte in einem Netzwerk gegeben. Durch die native Fähigkeit über WAN-Netzwerke zu transportieren und im Interesse der Entwickler, ist das Einsatzgebiet von Ravenna nicht nur auf lokale Netzwerke beschränkt, sondern lässt einen Einsatz in der Broadcastindustrie zu. So ist diese Technologie auch bei Großveranstaltungen einsetzbar. Im Veranstaltungsort wird ein Netzwerk aufgebaut und die Daten werden über dieses an Ü-Wägen bzw. externe Regien und Mischorte weitergeleitet. Dennoch können auch Ravenna-Netze im Kleinen aufgebaut werden.

Dante, welches ebenfalls über das OSI-Layer-3 kommuniziert, ist das Netzwerk mit den meisten verfügbaren Geräte und dem stärksten Wachstum der letzten Jahre. Audinate führt mit seinem geschlossenen System den Markt an und baut dies weiter aus. Mit der Veröffentlichung des Dante AV-Systems im Frühjahr 2019 und dem Beitritt in die SDVoE-Allianz Mitte 2018 ermöglicht Audinate den Videotransport über ein Audio-Netzwerk. Zwar ist die Technologie nicht offen zugänglich oder einfach modifizierbar,

jedoch bietet Audinate mit dem Dante Controller, sowie der Plug&Play Funktion, einfache Bedienbarkeit des Systems und höhere Nutzerfreundlichkeit. Durch den Domain Manager wird der Transport über Router und größere Netzwerke gegeben.

Die bisher vorgestellten Audioprotokolle werden wahrscheinlich die nächsten Jahre noch parallel nebeneinander bestehen. Dies liegt zum einen an ihren unterschiedlichen Einsatzgebieten, zum anderen stehen hinter den vorgestellten Lösungen unterschiedliche treibende Kräfte, die diese längerfristig unterstützen werden. AVB bietet bisher nicht sehr viele zertifizierte Produkte an, jedoch wird es in der Industrie durch Implementierung in Automotive-Ethernet genutzt und von dieser weiterhin unterstützt.

Die Milan Gruppe ist noch recht jung und befindet sich gerade in einer Phase des Aufbaus und es sind noch nicht viele Produkte vorhanden. Big Players wie d&b Audiotechnik wollen dies jedoch vorantreiben und auf diese Technologie setzen. Mit der Unterstützung der Firmen und offenen Entwicklungsumgebung birgt Milan ein großes Potential. Jedoch ist es hier noch abzuwarten, wie weitere große Audiofirmen sich dazu bekennen. Ähnlich wie bei AVB wurde anfangs viel Hoffnung in die Technologie gesetzt, über Jahre gab es jedoch nicht die gewünschte Umsetzung und erst in den letzten Jahren wurden viele Produkte veröffentlicht. Werden im nächsten Jahr seitens der Allianz-Mitglieder umfangreiche Lösungen angeboten, ist ein aufspringen anderer Firmen wahrscheinlich.

Mit der Integration des SMPTE Standards hat Ravenna ebenfalls ein mächtiges System für die Zukunft übernommen, welchem die nächsten zwei Jahre weitere Mechanismen für eine standardisierte Netzwerkstruktur zur Verfügung gestellt wird. Die Joint Task Force on Networked Media (JT-NM) hat in diesem Zuge eine Roadmap zur Veröffentlichung aller weiter enthaltenen Standards veröffentlicht. Unter anderem die in Kapitel 4.5 beschriebenen AMWA Protokolle.

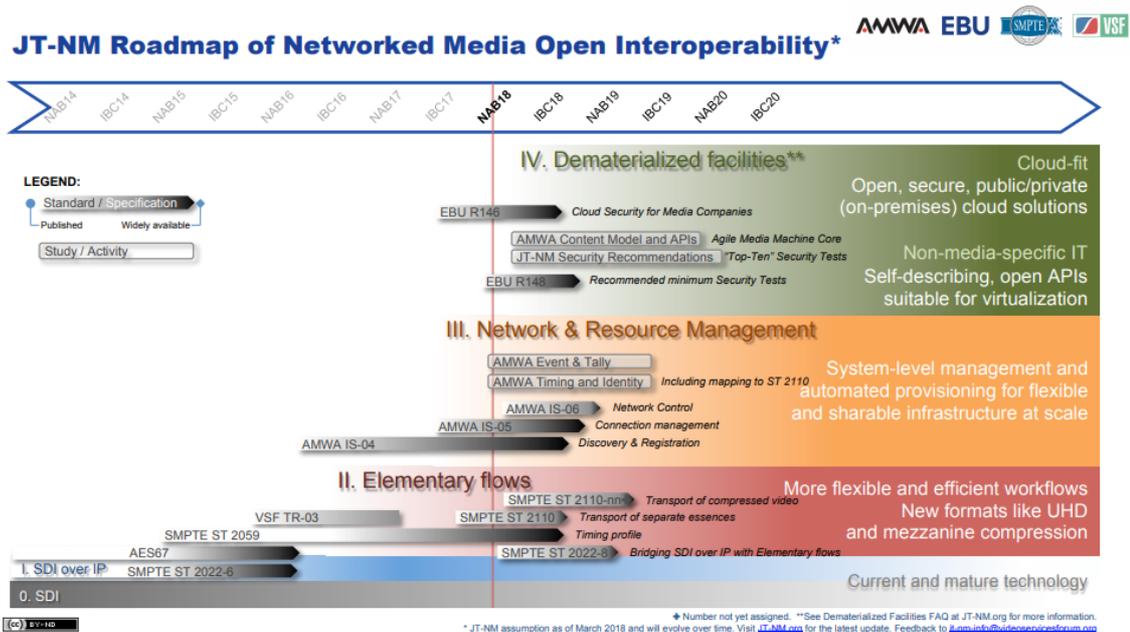


Abbildung 7-1: Roadmap der JT-NM¹⁴⁰

Neben der Standardisierung in der Audioübertragung ist eine Annäherung an die Videoübertragung und deren Integration in ein bestehendes Netzwerk zu sehen. Der Trend wird weiter in die Richtung AV-over-IP und somit der Verschmelzung von Audio und Video gehen.

Mit ST2110 ist ein solider Standard von Videosignalen in einem großen Netzwerk möglich. Zwar ist dieser bisher auf unkomprimierte Signale (und große Bandbreite) beschränkt, jedoch wird über den noch folgenden ST2110-22 auch auf einen Versand codierter Signale abgezielt, welcher die Bandbelastung verringern wird.

Wie oben angeführt ist der Videotransport in Milan möglich. Es bleibt herauszufinden, welche Produkte von Anbietern in naher Zukunft entwickelt werden.

Audinate entfernt sich mit Dante AV vom klassischen AoIP-Markt. Durch den Eintritt in die SDVoE-Allianz bleibt es zu vermuten, dass Dante seine Stellung im Pro-AV Bereich weiter ausbauen wird. Hierbei werden auf 10 GbE Systeme gesetzt, mit denen eine Videoausstellung erfolgt. Interessant wird hierbei sein, ob mit Dante AV weiterhin die AES67 Interoperabilität gegeben sein wird und ob sich die 2018 angekündigte Unter-

¹⁴⁰ <http://jt-nm.org/roadmap/>

stützung von SMPTE ST 2110 wirklich durchsetzen wird oder diese ebenfalls wie bei AVB fallen gelassen wird.

Die Weiterentwicklung der Netzwerklösungen ist somit mit großem Interesse zu betrachten, da diese einem stetigen Wandel unterliegt. Während der Recherchen und der Entstehung dieser Arbeit wurden neue Systeme angekündigt und neue Geräte vorgestellt. Das 2018 angekündigte Milan wurde ausgebaut und kompatible Produkte wurden auf dem Markt veröffentlicht. Audinate kündigte im Januar mit Dante AV ihr erstes videofähiges Modul an und wird Mitte Juni 2019 weitere Informationen veröffentlichen. Ravenna unterstützt ST2110, für welches in Zusammenarbeit mit der AMWA, EBU und VSF weitere Standards und Verbesserungen entwickelt werden.

8 Glossar

IP-Adresse: Logische Adresse für die Netzwerkerkennung, sowie die Hosterkennung. Die Adresse wird frei vergeben ist nicht definitiv. Sie wird zur Erkennung eines Gerätes im Internet benutzt. Das alte IPv4-Format bestand aus 32 Bit und konnte somit 4.294.967.296 Adressen weltweit vergeben, dieses wird durch das IPv6-Format abgelöst bei welchem eine 128-Bit-Adressierung benutzt wird und somit $3,4 \cdot 10^{38}$ also ca 340 Sextillionen Adressen zur Verfügung gestellt werden.

MAC-Adresse: Eindeutige, unverwechselbare 6-Byte (48 Bit) große Adresse eines Ethernet fähigen Gerätes welche zur Identifizierung eines Gerätes in einem Netzwerk dient. Sie besteht aus einer 3 Byte großen Hersteller- sowie 3 Byte großen Adaptererkennung. Dargestellt wird diese durch 12 hexadezimale Ziffern nach folgendem Schema:
FF-FF-FF-FF-FF-FF

9 Quellenverzeichnis

Allen&Heath: Allen & Heath (2019), AES67 – One Standard to Unite Them All (Whitepaper); unter: <https://www.allen-heath.com/media/AES67-One-Standard-to-Unite-Them-All.pdf>, abgerufen am 29.04.2019

AES67 verifizierte Geräte: <https://aimsalliance.org/aes67-resources/>, abgerufen am 04.05.2019

AES70: <https://www.ocaalliance.com/oca-vision/>, <https://www.ocaalliance.com/wp-content/uploads/2018/10/Whats-New-in-AES70-for-2018-.pdf>, abgerufen am 09.05.2019

AMWA: <https://amwa-tv.github.io/nmos/>, abgerufen am 09.05.2019

Andreas Hildebrand (2014): RAVENNA&AES67_V1.0, München, ALC NetworX GmbH. URL: <https://www.ravenna-network.com/app/download/13999773923/AES67%20and%20RAVENNA%20in%20a%20nutshell.pdf?t=1540389553>

Andreas Hildebrand (2017): AES67 Practical Guide, München, ALC NetworX GmbH. URL: <https://www.ravenna-network.com/2017/10/11/your-practical-guide-to-aes67-part-1/>, abgerufen am 10.05.2019

Arne Bönninghoff (2017): „IP Media Standards“, In: *Vdt-Magazin* 4/2017

AVB AACP: <https://avb.statusbar.com/article/1722.1-features/>, abgerufen am 09.05.2019

AVB Audio Video Transport: https://avnu.org/wp-content/uploads/2014/05/AVnu-AAA2C_Audio-Video-Transport-Protocol-AVTP_Dave-Olsen.pdf, abgerufen am 08.05.2019

AVB-Standard: <https://www.itwissen.info/TSN-time-sensitive-networking.html>, abgerufen am 07.05.2019

AVB-Protokollstack: <https://www.itwissen.info/AVB-audio-video-bridging.html>, abgerufen am 07.05.2019

- AVB: Traffic Shaping & Infrastructure planning: https://avnu.org/wp-content/uploads/2014/05/AVnu-AABAC_Traffic-Shaping-Infrastructure-Planning_Andre-Fredette.pdf, abgerufen am 08.05.2019
- AVB zertifizierung: <https://avnu.org/certified-products/>, abgerufen am 09.05.2019
- BET Onlinelexikon: <https://www.bet.de/lexikon/aes3/>, abgerufen am 09.04.2019
- Claus, Volker/ Schwill, Andreas (2003): Informatik – Ein Fachlexikon für Studium und Praxis, Dudenverlag, Mannheim
- Codecs RTP: <https://tools.ietf.org/html/rfc3551>, abgerufen am 16.04.2019
- Dante: <https://www.audinate.com/products/dante-enabled#catalog>, abgerufen am 11.03.2019
- <https://www.prosoundnetwork.com/live/aes67-avb-tsn-dante-power-aoips-rise>, abgerufen am 11.05.2019
- <https://www.professional-system.de/news/ise-2019-dante-kann-jetzt-auch-irgendwie-video/>, abgerufen am 11.03.2019
- <https://www.audinate.com/sites/default/files/PDF/advanced-dante-networking-avnw-2015-audinate.pdf>, abgerufen am 11.05.2019
- Audinate (2018): Dante Certification Program - Level 3 – Module 22: Dante Flows. Youtube, 05.05.2018, Web, 11.05.2019, in: <https://www.youtube.com/watch?v=Vi8B1KV0AJM&list=PL2UzvxHZNUyLbXLopZyF5IhOelyH8hFoP&index=22>
- https://dev.audinate.com/GA/dante-controller/userguide/webhelp/#latency_tab.htm, abgerufen am 11.05.2018
- <https://www.audinate.com/networks-and-switches>, abgerufen am 11.05.2019
- https://dev.audinate.com/GA/dante-controller/userguide/webhelp/#discovery_and_auto-configuration.htm#kanchor167, abgerufen am 11.05.2019
- <https://www.audinate.com/faq/what-dante-discovery>, abgerufen am 11.05.2019
- <https://dev.audinate.com/GA/dante-controller/userguide/pdf/latest/AUD-MAN-DanteController-4.1.x-v1.0.pdf>, abgerufen am 11.05.2019
- D&B Audiotechnik: <https://www.dbaudio.com/global/de/produkte/processing-und-matrix/ds20/>, abgerufen am 12.05.2019
- <https://www.dbaudio.com/global/de/loesungen/signalverteilung-und-audio-netzwerke/ocaaes70/>, abgerufen am 12.05.2019

-
- Dickreiter, Michael/ Dittel, Volker/ Hoeg, Wolfgang/ Wöhr, Martin (Hrsg.) (2014):
Handbuch der Tonstudioteknik. Band 2, 8. Auflage, Berlin/Boston, Walter de Gruyter GmbH
- DiffServ: <https://www.itwissen.info/DiffServ-differentiated-services.html>, abgerufen am 27.04.2019
- Ethernet-Technologie, 400Gbit/s: <https://www.de-cix.net/de/about-de-cix/media-center/press-releases/de-cix-first-internet-exchange-worldwide-to-offer-400-gigabit-ethernet-access-technology>, abgerufen am 11.04.2019
- Ethernet-Standards: <http://www.elektronik-kompodium.de/sites/net/1406261.htm>, abgerufen am 11.04.2019
<http://www.elektronik-kompodium.de/sites/net/0603201.htm>, abgerufen am 11.04.2019
<http://www.elektronik-kompodium.de/sites/net/1406171.htm>, abgerufen am 11.04.2019
- Frameformat Ethernet: <http://www.easy-network.de/bilder/ethernet-frame.jpg>, abgerufen am 12.04.201
- Gross, K. (2013): AES67 standard for audio applications of networks, *Audio Engineering Society Convention 67(7)*
- IEEE 802.3/Ethernet-Grundlagen: <http://www.elektronik-kompodium.de/sites/net/0603201.htm>, abgerufen am 11.04.2019
- IEEE 1722 AVTP: https://avnu.org/wp-content/uploads/2014/05/AVnu-AABAC_IEEE-1722-Media-on-AVB-Networks_Rob-Silfvast.pdf, abgerufen am 08.05.2019
- IEEE 1722.1-2013: <https://avb.statusbar.com/files/presentation/avdecc-aes-ny-2013.pdf>, abgerufen am 09.05.2019
- IEEE 1733, RTP over AVB: <http://www.ieee802.org/1/files/public/docs2009/avb-stanton-rtp-over-avb-aka-ieee1733-0309.pdf>, abgerufen am 08.05.2019
- JT-NM: <http://jt-nm.org/roadmap/>, abgerufen am 14.05.2019
- Lester, Michael und Boley, Jon (2007): The Effects of Latency on Live Sound Monitoring, *Audio Engineering Society Convention 123*
- Lippensynchronität: https://www.atsc.org/wp-content/uploads/pdf/audio_seminar/12%20-%20JONES%20-

[%20Audio%20and%20Video%20synchronization-Status.pdf](#), angerufen am 14.05.2019

MILAN: <http://www.soundandcommunications.com/an-introduction-to-milan/>, abgerufen am 12.05.2019

Marc Illouz (2018): Media clocking functional & interoperability specification, AVNU MILAN, <https://avnu.org/specifications/>

Marc Illouz (2019): Discovery, connection and control specification for talkers and listeners, AVNU MILAN, <https://avnu.org/specifications/>

Cole Perterson (2018), Network redundancy interoperability specification, AVNU MILAN, <https://avnu.org/specifications/>

Kreifeldt, Rick (2009): AVB for professional A/V Use, AVNU MILAN, https://avnu.org/wp-content/uploads/2014/05/AVnu-Pro__White-Paper.pdf, abgerufen am 12.05.2019

Fabian Braun (2018), Formats Interoperability Specification, AVNU MILAN, <https://avnu.org/specifications/>

<https://avnu.org/certified-products/>, abgerufen am 12.05.2019

<http://rhconsulting.eu/blog/files/Milan.html>, abgerufen am 12.05.2019

<http://milanavcommunity.proboards.com/thread/9/available-milan-controllers>, abgerufen am 12.05.2019

PTP: <https://www.perlesystems.de/supportfiles/precision-time-protocol.shtml>, abgerufen am 02.05.2019

Quality of Service: <https://www.elektronik-kompodium.de/sites/net/0905131.htm>, abgerufen am 27.04.2019

QoS in AVB: <https://www.elektroniknet.de/elektronik-automotive/bordnetz-ernetzung/schnelle-datenuebertragung-121066-Seite-2.html>, abgerufen am 08.05.2019

RAVENNA: <https://www.ravenna-network.com/adopting-ravenna/overview>, abgerufen am 10.05.2019

<https://www.ravenna-network.com/using-ravenna/technology/>, abgerufen am 10.05.2019

<https://www.ravenna-network.com/adopting-ravenna/aneman/>, abgerufen am 10.05.2019

<https://www.ravenna-network.com/adopting-ravenna/oem-solutions/>, abgerufen am

10.05.2019

RhConsulting: <http://rhconsulting.eu/blog/files/NetworkedAudio.html>, abgerufen am 14.05.2019

RH Consulting: <http://rhconsulting.eu/blog/files/NetworkedAudio.html>, abgerufen am 14.05.2019

RFC: <https://tools.ietf.org/html/rfc2474>, abgerufen am 27.04.2019
<https://tools.ietf.org/html/rfc1349>, abgerufen am 27.04.2019

Riggert, Wolfgang (2014): Rechnernetze, 5. Auflage, München, Carl Hanser Verlag

ROSSVIDEO: <https://www.rossvideo.com/oem-developer/bach-openmodule-st2110-aes67-audio-networking-module/>, abgerufen am 10.05.2019

Schreiner, Rüdiger (2016): Computernetzwerke. Von den Grundlagen zur Funktion und Anwendung, 6. Auflage, München, Carl Hanser Verlag

Scherff, Jürgen (2010): Grundkurs Computernetzwerke. Eine kompakte Einführung in Netzwerk- und Internet-Technologien, 2. Auflage, Wiesbaden, Vieweg + Teubner Verlag

Schuhe, Sven (2018): Avnu Alliance: „Das Milan-protokoll“, In: Production Partner 09/2018

SMPTE Standard (2018): ST 2110-10:2017 - Professional Media Over Managed IP Networks: System Timing and Definitions

SMPTE Standard (2017): ST 2110-30:2017 - Professional Media Over Managed IP Networks: PCM Digital Audio

SMPTE Standard (2018): ST 2110-31:2018 - Professional Media Over Managed IP Networks: AES3 Transport

SMPTE ST2110: https://www.smpte.org/sites/default/files/section-files/BBTB2018-W04-Henry%20Goodman-AES67%20in%20SVIP_0.pdf, abgerufen am 06.05.2019

Switch Layer 3, 4, 7: <https://www.itwissen.info/Layer-3-Switch-layer-3-switch.html>, abgerufen am 24.04.2019

<https://www.itwissen.info/Layer-4-Switch-layer-4-switch.html>, abgerufen am 24.04.2019

<https://www.itwissen.info/Layer-7-Switch-layer-7-switch.html>, abgerufen am 24.04.2019

Tanenbaum, Andrew S., Wetherall, David J. (2012): Computernetzwerke, 5. Auflage, München, Pearson Deutschland GmbH

Zölzer, Udo (2005): Digitale Audiosignalverarbeitung, 3. Auflage, Wiesbaden, Teubner Verlag/ GWV Fachverlage GmbH

10 Anhang A: E-Mail Verkehr

Mail 01:

Schuhen, Sven

Redakteur / Editor; Proffesional-System.de

Ich [...] hätte eine Frage zu einem Deiner Artikel "AVnu Allianz - Das Milan-Protokoll" (07.09.2018). Darin schreibst du, dass Dante AVB integrieren wirst. Dies habe ich auch in einer Bachelorthesis von 2014 gelesen, jedoch danach nirgendwo mehr und auf Audinate Seite steht, AVB sei nicht Kompatibel mit Dante. Gibt es Pläne, dass AVB integriert wird oder hat Audinate dies verworfen?

Leider hat Audinate das Thema AVB irgendwann um 2017 stillschweigend gestrichen. Und auch die Kompatibilität zu AES 67 ist schwierig. Ich habe mich dieses Jahr schon mit mehreren Audinate Mitarbeitern sowie mit Experten bei AES 67 unterhalten. Audinate sagt die Spezifikationen bei AES 67 wären so schwammig, sie hätten die einzig richtige Interpretation daraus gezogen und in Dante integriert. Die AES 67 sagen, es gäbe in der Branche schon lange eine einheitliche Linie, was die Integration des Standards angehe und Audinate würde sich nur die Rosinen rauspicken, bzw. gerade so das Thema ankratzen, um noch sagen zu können, sie wären kompatibel. Darüber hinaus kann AES 67 in Dante nur genutzt werden, wenn man sein Netzwerk über den kostenpflichtigen Dante Domain Manager verwaltet. Aber da habe ich auch schon Widersprüche gehört.

Mail 02:

Hildebrand, Andreas

Senior Product Manager ALC NetworX GmbH

Wird NMOS IS-04/05/06 in Ravenna integriert oder ist dies geplant?

RAVENNA beinhaltet bereits die Funktionalitäten Advertisement & Discovery sowie Connection Management, welche mit IS-04/05 als alternative Ansätze definiert werden. Da RAVENNA von Grund auf modular aufgebaut ist, können die verschiedenen Funktionalitäten durch beliebige alternative oder ergänzende Ansätze ergänzt werden. So haben wir z.B. bereits in einigen RAVENNA-Geräten parallel NMOS IS-04/05 oder das von Dante im AES67-Modus verwendete SAP-Protokoll eingebaut. Die Entscheidung zur Unterstützung bleibt jedoch dem jeweiligen Hersteller vorbehalten. IS-06 betrifft das Management von Netzwerk-Infrastrukturen, um im Netzwerk das Routing für bestimmte Streams festzulegen. Dies betrifft nur die bandbreiten-hungrigen Video-Streams; Audio-Streams werden praktisch nur mittels IGMP geroutet. Daher ist eine

Unterstützung von IS-06 für RAVENNA nicht notwendig. Eine verpflichtende Forderung nach Unterstützung für NMOS in RAVENNA-Implementierungen wird es vorerst nicht geben.

Ravenna stützt sich auf den AES67-Standard. Ist eine Unterstützung/Integration von AES70 geplant oder vorstellbar?

Hier gilt das zuvor Gesagte. AES70 ist ein Steuerungs- und Konfigurationsprotokoll, welches weit über das Verbindungsmanagement hinaus geht. Wenn ein Hersteller dieses Protokoll unterstützen möchte, kann es parallel zu RAVENNA implementiert werden (auch dies wurde bereits von einigen RAVENNA-Partnern gemacht).

Gibt es Pläne für die Erweiterung/den Ausbau von Ravenna?

Wir beobachten die Standardisierungsaktivitäten der diversen SDOs (bzw. gestalten diese aktiv mit). AES67, SMPTE ST2110 und auch NMOS sind in großen Teilen identisch zu RAVENNA-Definitionen oder basieren auf gleichen Funktionsprinzipien. Solange es keine hierzu widersprüchlichen Standards oder Funktionserweiterungen gibt, besteht kein konkreter Handlungsbedarf. Implementierungen werden selbstverständlich an die laufenden Entwicklungen angepaßt (dies obliegt aber – wie oben schon ausgeführt – den jeweiligen RAVENNA-Partnern).

Da es mir derzeit nicht möglich ist SMPTE ST338 zu lesen: Welche Formate kann St2110-31 neben AES3 und DSD/DXD noch transportieren?

Wie schon richtig ausgeführt, läßt sich mit AES3 neben dem eigentlichen AES/EBU Datenformat (L24 + meta data bits) auch non-PCM Audio gemäß SMPTE ST338 transportieren. Die Liste der möglichen Formate ist lang (u.a. auch das wohl bekannteste Format DOLBY-E). Für RAVENNA AM824 / SMPTE ST2110-31 ist der eigentliche Inhalt bzw. das tatsächliche Datenformat irrelevant, da es transparent transportiert wird. Was schlußendlich als Datenformat enthalten ist (i.e. AES/EBU oder Dolby-E oder sonstwas), wird auf übergeordneter Ebene festgelegt / ausgewertet. Das RAVENNA DXD/DSD Payload-Format ist übrigens nicht identisch mit AM824 / ST2110-31, sondern ein von Merging Technologies proprietär definiertes Format, welches aber offen gelegt ist.

Wird es wie beim AES-Plugfest 2017 auch Interoperabilitätstest für ST2110 fähige Geräte geben?

Hierzu hat es in der Vergangenheit bereits zahlreiche sog. „Dirty Hands“ Workshops seitens AMWA/SMPTE/VSF/AES/JT-NM/EBU gegeben, zuletzt kurz vor der NAB. Hier werden nicht nur Verbindungen auf Basis ST2110 sondern auch erweiterte Funktionalitäten wie NMOS etc. getestet.

Wäre über AVB IEEE 1733 eine Interoperabilität mit ST2110 möglich?

Nein. AVB ist eine L2-Technologie, ST2110 eine L3-Technologie. Darüber hinaus ist 1733 eine unvollständige Definition und wird meines Wissens nach auch nirgendwo real verwendet.

11 Anhang B: Bewertungstabelle

	AES67	SMPTE ST2110	AVB	Dante	Ravenna	Milan
Größe	-	-	Layer 2-Netzwerk	LAN, WAN-fähig mit DDM	LAN, WAN	Layer 2-Netzwerk
Redundant - Topologie	-	-	Ja / Spanning Tree	Ja, / Stern > Ring > Daisy Chain	Ja / Stern, Ring, Daisy Chain	Ja / Spanning Tree
Latenz	-	-	25µs- 2ms bei 7 Hops	84µs-5ms; 100µs/ Hops	<1ms	25µs- 2ms bei 7 Hops
Samplerate	44.1kHz - 96kHz	44.1kHz - 96kHz	44.1kHz - 192kHz	44.1kHz - 192kHz	32kHz - 192kHz (PCM), 384kHz DSD	48kHz - 192kHz
Bittiefe	16 Bit, 24 Bit	16 Bit, 24 Bit	32Bit	24 Bit, 32 Bit	16 Bit, 24 Bit, 32 Bit	24 Bit, 32Bit
Dateiformat	PCM	PCM, AES3	PCM über AAF	PCM	PCM, AES3, DSD	PCM über AAF
GoS	DiffServ	DiffServ	IEEE 802.1Q	DiffServ	DiffServ	IEEE 802.1Q
Kompatibilität	-	AES67	quasi AVB	AES67	AES67, ST2110	AVB
Lizenz - offener/ geschlossener Standard	Lizenzfrei - offen	Lizenzfrei - offen	Lizenzfrei - offen	Lizenzgebühren - geschlossen	Lizenzfrei - offen	Lizenzfrei - offen
Steuerprotokoll	-	-	offen	proprietär	offen	offen
Plug&Play	-	-	-	✓	-	✓
Controller	-	-	Extern, wenn 1722.1 vorhanden	Integriert	Extern	Extern, offen verfügbar
Videofähig	-	✓	✓	✓	✓	✓