

Bachelorarbeit im Studiengang Audiovisuelle Medien

**Systemdesign für mobiles Livestreaming mit IP-basiertem Workflow  
und binauralem Audio am Fallbeispiel *Stayin' Live Stuttgart***

vorgelegt von

Arne Morgner

Matrikelnummer: 34840

am 27.02.2021

an der Hochschule der Medien Stuttgart

zur Erlangung des akademischen Grades

Bachelor of Engineering

Erstprüfer: Prof. Dr. Frank Melchior

Zweitprüfer: Prof. Oliver Curdt

## Ehrenwörtliche Erklärung

„Hiermit versichere ich, Arne Morgner, ehrenwörtlich, dass ich die vorliegende Bachelorarbeit mit dem Titel: *Systemdesign für mobiles Livestreaming mit IP-basiertem Workflow und binauralem Audio am Fallbeispiel Stayin'Live Stuttgart* 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, 27.02.2021

---

Arne Morgner

## Kurzfassung

Im Zug der Covid-19-Pandemie sind Kulturschaffende, Medien- und Kreativbranche gezwungen, Produktionskonzepte umzudenken und neue Systeme und Systemabläufe zu entwickeln, um unter den Bedingungen der Pandemie überhaupt weiterarbeiten und produzieren zu können. Abläufe in Produktionen müssen grundlegend neu gedacht und umstrukturiert werden, um bestehenden Hygieneauflagen, personellen Beschränkungen und den daraus entstandenen neuen Produktionssituationen Genüge zu tun.

Diese Bachelorarbeit beschäftigt sich mit dem Thema *Livestream-Produktion* und beschreibt ein mögliches Konzept und Systemdesign, um unter den bestehenden Beschränkungen und den daraus resultierenden Anforderungen eine moderne Produktionsumgebung zu ermöglichen. Künstler\*innen und Produktionsbeteiligten soll es ermöglicht werden, trotzdem flexibel, agil und schnell zu arbeiten. Hierbei werden die Vorteile und Herausforderungen eines IP-basierten Workflows beleuchtet und an einer Planung und Durchführung eines Systemdesigns aus der Praxis betrachtet.

## Abstract

In the wake of the Covid-19 pandemic, cultural workers, the media and creative industries are forced to rethink production concepts and develop new systems and system processes in order to be able to continue working and producing under the conditions of the pandemic. Production processes have to be fundamentally rethought and restructured in order to meet existing hygiene requirements, personnel restrictions and the resulting new production situations.

This bachelor thesis deals with the subject of live stream production and describes a possible concept and system design to enable a modern production environment under the existing restrictions and the resulting requirements. Artists and those involved in production should still be able to work flexibly, agile and quickly. The advantages and challenges of an IP-based workflow are highlighted and considered in the planning and implementation of a system design from practice.

## Danksagung

Für die Umsetzung des praktischen Teils der Sendungen im Rahmen der Produktion *Stayin' Live Stuttgart* möchte ich mich recht herzlich bei meinem Team bedanken. Marco Selter und Christoph Hafner, die mir wunderbare Bilder an der Kamera lieferten. Christian Tobias, für die Ruhe und Besonnenheit am Sendeton im Umgang mit der Binauraltechnik. Sam Brauner für die ansprechende Gestaltung der Grafiken und einen zuverlässigen Sendebetrieb an der Livestream- und Sendetechnik. Ebenso bedanke ich mich bei Charlotte Schön für die erfrischende Moderation und natürlich gilt ein großes Dankeschön den Künstlerinnen und Künstlern Tabea Booz, Annique Göttler, Lisa Wilhelm, Daniel Pflumm, Peter-Philipp Röhm und Ramon Schmid.

Für die Unterstützung im technischen und gestalterischen Sponsoring bedanke ich mich bei dem Team der Dear Reality GmbH unter Leitung von Achim Fell bei der Integration ihres kostenlos zur Verfügung gestellten Plugins zur Live-Binauralisierung. Ebenso ein Dank an die Geschäftsführerin Bärbel Heck des Vitra Store Stuttgart, die uns kostenlos ihr Mobiliar als Requisite für das Studiodesign zur Verfügung gestellt hat. Außerdem möchte ich mich bei Walter Ercolino, Leiter des Pop-Büros Stuttgart, für den Support des Projektes und der Künstler\*innen bedanken.

An der Hochschule für Musik und Darstellende Kunst Stuttgart bedanke ich mich bei der Hochschulleitung und ganz besonders bei meinen Kollegen Prof. Piet Meyer und Diplom-Tonmeister Igor Stepanov für die Unterstützung und Bereitstellung von Material aus dem elektronischen Studio der Hochschule und den Rückhalt bei meiner studienbegleitenden Tätigkeit.

Größter Dank gilt meinen Professoren Oliver Curdt und Frank Melchior, die mit ihrer hohen Expertise in Gestaltung und Technik immer passende Impulse und Rat für die Umsetzungen meiner Projekte hatten, und sich nie zu schade waren auch persönlich dem Projekt beizuwohnen, manchmal auch bis spät in den Abend.

Herzlichen Dank für Ihre Unterstützung in meinem Studium bis hin zur Bachelorarbeit.

# Inhaltsverzeichnis

Ehrenwörtliche Erklärung .....	2
Kurzfassung .....	3
Abstract .....	3
Danksagung .....	4
Abkürzungsverzeichnis .....	7
1. Einleitung .....	8
2. Stayin'Live – Projektvorstellung .....	9
3. Entwicklung eines Sendungskonzepts – Rahmenbedingungen .....	11
4. Entwicklung des Systemdesigns für <i>Stayin'Live Stuttgart</i> .....	15
4.1 Personal und Besetzung der Positionen .....	16
4.2 Transportfähigkeit der Produktion .....	18
4.3 Flexibilität und Skalierbarkeit der Produktionskomponenten .....	19
4.4 Maximallast und Leistungsaufteilung der Produktion .....	20
4.5 Die zu übertragenden Quellen und benötigte Senken .....	22
4.5.1 Quellen und Senken der Bildtechnik .....	22
4.5.2 Quellen und Senken der Tontechnik .....	23
4.6 Wahl des Sendeformats - Stream Encoding und Bitraten .....	24
4.6.1 Die richtige Balance in der Formatwahl und der Bitrate .....	25
4.6.2 Streaming Workflow .....	27
4.6.3 Stream-Encoding per Software .....	28
4.7 Aufbau der Bildtechnik und ihre Komponenten .....	29
4.7.1 Der Bildmischer und dessen Anbindung im System .....	30
4.7.2 Anforderungen an die Kamera und Kameraumgebung .....	31
4.7.3 Tally – Aufnahmelicht-/Rotlichtsteuerung .....	34
4.7.4 Wandlung des SDI-Signals zur Softwareanbindung .....	34

4.7.5 Programmaufzeichnung - MAZ.....	36
4.8 Aufbau der Tontechnik mit binauralem Sendeton .....	38
4.8.1 Mikrofonierung der Künstler .....	39
4.8.2 Wahl des digitalen Audiosystems – Blackbox Betrachtung .....	41
4.8.3 Intercom und Abhörsituation der Produktion.....	45
4.9 IP-basierte Steuerung der Bild- und Tontechnik.....	48
5. Aufbau des Audiosystems mit binauralisierten Sendeton .....	50
5.1 Systemanforderungen der Live-Binauralisierung .....	50
5.2 Signalstruktur im Hybrid-Audiosystem .....	51
5.2.1 Signalfluss und Signalgestaltung des Sendetons.....	52
5.2.2 Signalfluss des In-Ear-Monitoring-Systems .....	54
5.2.3 Signalfluss des Kommunikationssystems .....	55
5.3 Aufbau und Struktur der „Digital Audio Workstation“ (DAW) .....	55
5.3.1 Anforderung und Wahl der passenden DAW .....	55
5.3.2 Anforderungen und Wahl des Plugins zur Binauralisierung.....	56
6. Postproduktion: Optimierung des Systems .....	58
6.1 Optimierung im Stream-Encoding per Hardware-Encoder.....	58
6.2 Wechsel des SDI-USB-Konverter .....	59
6.3 Einbindung einer individuellen HRTF .....	60
7. Fazit.....	61
Tabellenverzeichnis.....	65
Abbildungsverzeichnis .....	65
Literaturverzeichnis .....	65
Anhang - Abbildung.....	70

## Abkürzungsverzeichnis

AD-Wandler - Analog zu Digital Wandler

BMD - Blackmagic Design

CDN - Content Delivery Network

DA-Wandler - Digital zu Analog Wandler

DAW - Digital Audio Workstation

DHCP - Dynamic Host Configuration Protocol

EB - elektronische Berichterstattung

EBU - European Broadcast Union

HdM - Hochschule der Medien Stuttgart

HMDK - Hochschule für Musik und Darstellende Kunst Stuttgart

HRTF - Head-Related Transfer Function

IP - Internet Protocol

ITU - International Telecommunication Union

MAZ - Zuspelung und/oder Aufzeichnung von Video und Ton

OBS - Open Broadcaster Software Studio

RTMP - Real Time Messaging Protocol

Schuko – Schutzkontaktsteckverbindung

VRRP - Virtual Router Redundancy Protocol

## 1. Einleitung

Im Frühjahr 2020 ereilte eine Pandemie in kurzer Zeit die gesamte Bevölkerung und verändert bis heute das Zusammenleben, den Arbeitsalltag und das Privatleben auf unterschiedlichste Art und Weise. Zur Eindämmung des Covid-19-Virus und dessen unter Umständen lebensbedrohlichen Krankheitsverlauf traten massive Einschränkungen in Kraft. Im Bereich des öffentlichen Lebens wurden Großveranstaltungen abgesagt, um das Zusammentreffen von größeren Menschenansammlungen zu vermeiden. Im Verlauf des Jahres wurden die Einschränkungen immer schärfer und erstreckten sich auf den Berufsalltag bis hin zur Unterbindung der Ausübung beruflicher Tätigkeiten in vielen Branchen. Auch der private Alltag wurde zunehmend eingeschränkt und Deutschland erfuhr einen ersten Shutdown in der ersten Jahreshälfte 2020.

Von den Corona-Maßnahmen besonders betroffen waren von Anfang an die Medien- und Veranstaltungsbranche, sowie die Künstler\*innen im Live-Bereich durch die Unterbindung von Großveranstaltungen jeglicher Art, die ebenso eine technische Betreuung zur Übertragung von Bild und Ton intern, wie auch zur Außenübertragung für Rundfunk, Fernsehen und das Web benötigen. Allein die Kultur- und Kreativwirtschaft umfasst über eine Viertelmillion Unternehmen mit 1,8 Millionen Erwerbstätigen in Deutschland und erbrachte im Jahr 2018 einen Gesamtumsatz von 171,1 Milliarden Euro <sup>1</sup>.

Eine eigene Welt stand plötzlich still und musste sich neu orientieren, um weiter arbeiten zu können und wirtschaftlich zu überleben. Künstler\*innen mussten neue Inhalte und Konzepte schaffen, um sich den Gegebenheiten und den Beschränkungen anzupassen, um auch ein Minimum an Erwerb möglich zu machen. Ebenso musste technisch und wirtschaftlich eine Produktionswelt gestaltet und geplant werden, die andere und neue Inhalte überträgt und zugänglich macht. Online Konzerte, Livestreams und On-Demand-Plattformen gründeten sich rasch und entwickelten sich weiter. Die Qualität dieser Formate variierte sehr stark, da Übertragungsqualität, künstlerische Ausgestaltung und Content-Delivery-Strukturen sehr unterschiedlich waren.

---

<sup>1</sup> (Bundesministerium für Wirtschaft und Energie, 2020)

In meinem Studium war ich genauso betroffen und eingeschränkt, da durch die Einschränkungen und Hygienemaßnahmen der Forschungs- und Lehrbetrieb und damit die Produktion an der Hochschule im gewohnten Rahmen nicht mehr möglich war. Der Shutdown ereilte meinen gewohnten Studienverlauf daher auf gleiche Weise wie auch die Kultur- und Kreativbranche. Ein Ende dieser Situation war im März und April 2020 nicht absehbar und das Sommersemester hatte soeben begonnen. Wie sollte also eine nötige Produktion unter diesen Umständen möglich sein?

Mein Wunsch war es, ein Projekt zu initiieren, das eine Produktion im Rahmen der Pandemie ermöglicht und Künstler\*innen und Kreativschaffenden eine Plattform für ihre Kunst und ihre Situation gibt. Hieraus entstand die Forschungsfrage, ob es möglich sein würde, eine in Technik und Personal kleine Produktion zu planen und umzusetzen, wobei die Beschränkungen – die maximal zulässige Personenzahl und die geltenden Hygieneregeln - eingehalten werden können, ohne dass die Produktionsqualität oder Livestream-Standards in Bild und Ton leiden. Aufgabe war es nun, ein System zu konzeptionieren, zu planen und mit einem realen Programmkonzept zu testen. Hierzu sollte das fertig designte System mit einer echten Produktion in unterschiedlichsten Produktionssituationen wie verschiedene Locations, künstlerischen Genres und verschiedener Kunst getestet werden. Die Planung und Optimierung dieses Systemdesigns für modernes Livestreaming wird in dieser Bachelorarbeit zentral betrachtet, ausgewertet und in der Postproduktion optimiert.

## 2. Stayin'Live – Projektvorstellung

Stayin'Live ist ein deutschlandweites Projekt, das Künstler\*innen gewidmet ist, die nicht nur einen hohen Einkommensverlust erleiden, sondern auch ihre Bühne und damit ihr Sprachrohr durch die Covid-19-Pandemie nahezu vollständig verloren haben. Im Frühjahr 2020 hat mein ehemaliger Kommilitone und guter Freund Jonas Urvat die Plattform mit drei Kollegen in Berlin gegründet. Es wurden kleinere Livestreams mit minimalem Aufwand unter Eigenregie mit einzelnen Künstlern organisiert und umgesetzt: zwei Kameras, eine einfache Mikrofonierung mit einem kleinen analogen Mischpult und Livestream-Technik, betreut durch eine Person; ein Setup, das allein handhabbar ist und auch durch Laien bzw. auch durch die Künstler\*innen selbst betreut und bedient werden kann. Zu diesem Zeitpunkt war noch

unklar, wie sich die Pandemielage entwickeln würde. Das Setup sollte selbst bei stärkster Beschränkung – die die Bundesregierung zur Eindämmung der Infektionen ggf. vornehmen würde – auch als „One-Man-Show“ durch die Künstler\*innen zu bedienen sein. So wurde das erste Setup als Streaming-Lösung in einem Koffer weitergegeben, von Künstler\*in zu Künstler\*in, ohne zusätzlichen Kontakt über Dritte.

Das Projekt stieß schnell auf großes Interesse, da es ehrenamtlich aufgebaut war und für Künstler\*innen und Musiker\*innen eine kostenfreie Plattform für ihre Arbeit bot. Alle Spenden im Livestream gingen an die aufführenden Künstler\*innen im Stream. Bedingt durch den Lockdown im Frühjahr 2020, die weiteren Beschränkungen im Bewegungsradius und die Einschränkungen im sozialen Umfeld, bot es sich an, kleinere „Inseln“ aufzubauen, anstatt auf das Wachstum zu großen Streams in Berlin mit größerem Team zu setzen. So entstand ein weiteres Team in Hamburg und eines in Stuttgart. Hierzu hat mich Jonas Urvat im April 2020 angesprochen und mich auf das Projekt aufmerksam gemacht, woraufhin ich die Projektleitung übernahm.

Für die Unterstützung in der Projektleitung habe ich Charlotte Schön, Mediensprecherin und Studentin im „Master für Rhetorik“ an der Hochschule für Musik und Darstellende Kunst Stuttgart (HMDK), begeistern können. Ich kenne sie über Aufnahmen, die ich für sie an der HMDK im Rahmen ihres Studiums getätigt habe. Sie hat Aufgaben im Bereich der Organisation und Kommunikation mit den Künstler\*innen sowie die redaktionelle Arbeit übernommen. Im Livestream selbst war sie außerdem für die Moderation der Show zuständig. Somit bestand das Team der Projektleitung aus Jonas Urvat, der die Erfahrung als Gründungsmitglied einbrachte, Charlotte Schön, die die redaktionelle Arbeit übernahm, und mir selbst für die technische Leitung/Produktionsleitung. Dabei habe ich Technik und Team zusammengestellt sowie das System konzeptioniert und aufgebaut. Außerdem zählten die Disposition und das Projektmanagement zu meinen Aufgaben.

Für die Positionen im Livestream hatten wir das Team wie folgt eingesetzt:

Kamera: Marco Selter, Christoph Hafner      Ton: Christian Tobias

Grafik: Sam Brauner      Bildmischer/Regie: Arne Morgner

Umgesetzt wurde das Projekt im Juni 2020, wobei an den vier Sonntagen jeweils eine Sendung von etwa 60 Minuten Länge gestreamt wurde. Dem ging eine Planungs- und Erprobungsphase mit dem gesamten Team von etwa vier Wochen im Mai voraus. Von der ersten Idee und Konzeption mit Charlotte Schön und Jonas Urvat bis zur letzten Sendung vergingen 2,5 Monate. Grundsätzlich wurden alle im Team mit allen Positionen vertraut gemacht, um im Notfall jede Position ersetzen zu können und um für alle das Maximale an Erfahrungswert und Wissen aus dem Projekt zu generieren.

### 3. Entwicklung eines Sendungskonzepts – Rahmenbedingungen

Die grundlegende Idee von *Stayin' Live Stuttgart* war es, lokalen, selbstständigen Künstler\*innen, die aufgrund der Covid-19-Pandemie einen hohen Einkommensverlust erleiden, trotzdem eine Bühne zu geben. Es sollten Künstler\*innen ohne laufende Verträge mit einem Label oder Agenturen sein, die selbstständig in der Stuttgarter Kulturszene agieren und damit am meisten Unterstützung in dieser Zeit benötigen. Auf der anderen Seite wollten wir Stuttgarter Lokalitäten eine Möglichkeit der Werbung geben und somit eine Win-Win-Situation für alle Beteiligten ermöglichen. Durch die Schließung von Lokalen, Geschäften, Clubs und Bars waren hier ebenso viele Unternehmer, Kleinunternehmer und Selbstständige betroffen, die in der Stuttgarter Kulturszene im Hintergrund arbeiten und weniger sichtbar sind: Die Barista, die Veranstaltungstechniker\*innen, die Caterer, Köche, Bedienungen in Restaurants und Bars, Clubbetreiber\*innen und die Angestellten, die ebenso einen hohen Einkommensverlust aufgrund des Lockdowns im Frühjahr 2020 hinnehmen mussten und am Existenzminimum um das finanzielle Überleben ringen.

Charlotte Schön, Jonas Urvat und mir war es hier von Anfang an ein großes Anliegen, lokale „Kulturhelden“ zusammenzubringen und gemeinsam ein gemeinnütziges Projekt auf die Beine zu stellen, das alle Seiten ehrenamtlich unterstützt. Damit war ein wichtiger Punkt in der Planung gesetzt: Das Projekt wurde unentgeltlich umgesetzt. Die Produktion erhielt keine Entlohnung und es gab kein Budget für Equipment und Location. Lediglich die Künstler haben per Spendenaufruf im Livestream über einen PayPal.me-Link direkt Spenden erhalten können, die Produktion jedoch nicht, da die Spenden direkt auf die Spendenkonten der Künstler\*innen liefen. Ebenso wurde der Live-Stream nicht monetarisiert. Da kein Budget in

der gesamten Produktion vorhanden war, war von Anfang an klar, dass der Livestream kostenlos gestreamt werden würde, ohne Tickets zu verkaufen oder andere Einnahmemöglichkeiten zu nutzen. Dies hatte für das Systemdesign zur Folge, dass das zur Verfügung stehende Equipment stark eingeschränkt war. Leihgaben von Firmen und Verleihern waren aus Hygienegründen nicht möglich oder konnten aus finanziellen Gründen durch die Firmen nicht gestellt werden, da die Verleiher bzw. Hersteller durch die Pandemie selbst in finanziellen Schwierigkeiten steckten und somit ein Sponsoring für die Technik nahezu ausgeschlossen war. Das eingesetzte Equipment wurde somit aus privater Hand, von der HMDK und der Hochschule der Medien Stuttgart (HdM) gestellt.

Um eine gewisse Vielfalt der Kunstszenen in und um Stuttgart abzubilden, war schnell klar, dass wir verschiedene Genres abdecken möchten. Außerdem wollten wir ein möglichst breites Publikum ansprechen und somit auch das Interesse von Zuschauern zwischen den Genres wecken, um Austausch zu schaffen und um die Probleme und Auswirkungen des Lockdowns auf die Kunst- und Kulturlandschaft in einer besonderen Breite darzustellen. Deshalb haben wir uns Künstler\*innen aus den Genres Klassik, Jazz und Pop gesucht. Um außerdem die darstellende Kunst einzubinden, haben wir den Bereich Performance Sprechen zusätzlich gewählt. Somit haben sich vier Sendungen mit entsprechendem Schwerpunkt entwickelt:

1. Sendung am 07.06.2020 um 19 Uhr mit Annique Göttler, Konzertpianistin – Klassik
2. Sendung am 14.06.2020 um 19 Uhr mit Tabea Booz, Singer/Songwriter – Pop
3. Sendung am 21.06.2020 um 19 Uhr mit Lisa Wilhelm (Drums), Peter Philipp Röhm (Piano) – Jazz
4. Sendung am 28.06.2020 um 19 Uhr mit Charlotte Schön, Ramon Schmid – Performance Sprechen

Der strukturelle Aufbau und Ablauf jeder Sendung orientierte sich an einem wiederkehrenden Gerüst:

- kurze Anmoderation zur Sendung durch Charlotte Schön
- musikalische Vorstellung/Eingangsperformance
- Vorstellung der Künstler im Gespräch mit Charlotte Schön
- zweiter künstlerischer/musikalischer Beitrag
- Gespräch: persönliche Lage/Schicksal durch die Pandemie

- dritter künstlerischer/musikalischer Beitrag
- Danksagung an Sponsoren, Verabschiedung und Werbung für die nächste Sendung

Ziel war es, das Programm auf etwa 60 Minuten zu beschränken und eine gute Balance zwischen Beitrag und Gespräch zu finden, aufgeteilt in etwa 30 Minuten Musik/Beitrag und 30 Minuten Gespräch.

Eine besondere Herausforderung und damit eine der wichtigsten Anforderungen war für uns das Hygienekonzept bzw. die zum damaligen Zeitpunkt geltenden Corona: Es durften sich maximal fünf Personen in einem Raum aufhalten. Außerdem mussten die Abstände von mindestens 1,5 Metern zu anderen Personen jederzeit eingehalten werden. Material und Bühnenbild wurden entsprechend angepasst und die Arbeitsstruktur, die räumliche Aufteilung und die Anordnung der Regie durch diese Vorgaben maßgeblich bestimmt, so dass die geltende Verordnung eingehalten werden konnte. Sowohl während des Auf- und Abbaus als auch in der Produktion herrschte für alle Maskenpflicht. Lediglich die Künstler\*innen durften in der Probe und während der Sendung mit entsprechenden Abständen die Masken abnehmen<sup>2</sup>.

Auf der Suche nach Veranstaltungsorten haben wir versucht, für jede\*n Künstler\*in passende Räumlichkeiten zu finden. Bei unserem ersten Livestream mit Annique Göttler haben wir in erster Linie nach einem akustisch geeigneten Raum gesucht, der über einen großen Konzertflügel verfügt. Da es in Stuttgart nicht viele Räumlichkeiten mit entsprechender Größe und Instrumentarium gibt, die wir kostenlos nutzen dürfen, lag nahe, hierfür den Konzertsaal der HMDK zu nutzen. Die mögliche Alternative – der Mozartsaal in der Liederhalle Stuttgart – stand in unseren präferierten Zeitfenstern nicht zur Verfügung.

Im Konzertsaal der Musikhochschule haben wir einen Raum mit drei verschiedenen Flügeln vorgefunden (zwei Steinway & Sons D-Flügel und ein Steinway & Sons B-Flügel). Ebenso kam Annique Göttler die Akustik des Konzertsaals sehr entgegen, da gerade in der klassischen Musik die akustische Umgebung für die Interpretation und Wahrnehmung der Werke maßgeblich ist, auch aus Sicht der Künstlerin im eigenen Monitoring. Der Saal fasst 418 Sitzplätze für das Publikum und hat eine Deckenhöhe von ca. zwölf Metern. Hier haben wir die Regie in einer der Seitenbühnen, gegenüber der eigentlichen Inspizientenposition,

---

<sup>2</sup> (BW, 2020)

aufgebaut. Strom war in ausreichender Dimension vorhanden, so stand uns in der Regie eine 16A-Schutzkontaktsteckdose, die mit einem C16-Sicherungsautomat abgesichert war, zur Verfügung. Auf der Bühne hatten wir sechs getrennte Stromkreise, ebenfalls 16A-Schutzkontaktsteckdosen, jeweils mit C16-Sicherungsautomaten abgesichert. Darüber hinaus ist der Konzertsaal mit einer professionellen Beleuchtung ausgestattet. Hier hingen zwölf Profilscheinwerfer mit jeweils 2kW als Auflicht und zehn Fresnel-Stufenlinsen Scheinwerfer mit 1kW als Spitzenlicht über der Bühne. Profitieren konnten wir zudem von einem Kabelschacht zwischen Bühne und Regieraum. Im Bühnenbereich war eine direkte Netzwerkverbindung per EtherCON-Buchse vorhanden.

Im Livestream mit Tabea Booz haben wir uns für das „Mókuska Caffè“ entschieden und eine Zusage für die kostenlose Nutzung erhalten. Die Rösterei und Kaffeebar befindet sich in Stuttgart-West und ist ein kleines, modernes Café, das wunderbar zum Stil alternativen Singer-Songwriterin Tabea Booz passt. Zufällig hat sich nach der Anfrage für den Veranstaltungsort herausgestellt, dass das „Mókuska Caffè“ das Lieblingscafé der Künstlerin ist. Die örtlichen Gegebenheiten unterschieden sich im Vergleich zum Konzertsaal der HMDK jedoch in hohem Maße. Als Regierraum stand nur ein kleiner Nebenraum zur Verfügung, zudem ein schlecht beleuchteter Kundenraum mit einem Stromkreis, der mit einem 16A B16-Sicherungsautomat abgesichert war. Hier mussten wir für eine entsprechende Ausleuchtung selbst sorgen und auch das Streulicht, das über die großen Schaufenster in das Lokal einfällt, von außen selbst in den Griff bekommen. Aufgrund paralleler Wände und des modernen Stils des Cafés mit Backstein- und Betonwänden, die zum Teil unverputzt waren und eine schallharte Oberfläche bildeten, kam es zu starken Reflexionen und ausgeprägten akustischen Moden (stehende Wellen). Hier war schnell klar, dass wir die Künstler mit In-Ear-Monitoring ausstatten müssen und eine Abnahme der Instrumente mit sehr dichter Mikrofonierung am Instrument erfolgen muss. Auch hier haben wir bei einer ersten Begehung die Infrastruktur überprüft. Die Netzwerkverbindung war durch einen Consumer-Router von Fritz.BOX mit einem RJ45-Anschluss gegeben.

Für den Livestream im Jazz-Genre mit Lisa Wilhelm an den Drums und Peter-Philipp Röhm wollten wir ursprünglich das Café Galao am Marienplatz Stuttgart nutzen. Hier hätten wir eine kleine Bühne mit Restaurant/Bar vorgefunden, ein Klavier und einfache Bühnenbeleuchtung. Unglücklicherweise war hier ein Coronafall bei den Betreibern aufgetreten, so dass wir eine

Woche vor dem geplanten Livestream auf unsere Backup-Lösung zurückgreifen mussten: der Konzertsaal der HMDK.

Für unseren vierten und letzten Livestream im Genre *Performance – Sprechkunst* konnten wir die *Kulturinsel Stuttgart* in Bad Cannstatt als Partner und Location gewinnen. Sie ist eine Plattform für Kunst und Kultur in der Region und bietet verschiedene Räumlichkeiten für eine große Vielfalt an Events aller Art. Neben der Vermietung dieser Räumlichkeit tritt die Kulturinsel auch selbst als Veranstalter auf und bezeichnet sich selbst als „multikulturelle Oase im Neckarpark“<sup>3</sup>. Neben einem Club, einem großen Atelier und einem Café mit Lesungsraum bietet die Kulturinsel auch ein Loft mit Bar und Tagungsraum mit Blick über die Kulturinsel und den Neckarpark. Als vierten Drehort haben wir uns für das Loft entschieden und konnten hier auf einen hellen Raum mit Bar-Charakter im Industriestil zurückgreifen. Eine Besonderheit war hier, dass sowohl Regie als auch Studiobereich in der ersten Etage lagen und nicht per Aufzug erreichbar waren. Das Equipment musste über ein enges Treppenhaus getragen werden. Strom war mit drei getrennten Stromkreisen ausreichend vorhanden, die jeweils mit 16A C16-Sicherungsautomaten abgesichert waren. Die Beleuchtungssituation war für uns nicht ausreichend, sodass zusätzliche Ausleuchtung notwendig war. Aufgrund glatter, mit Rauputz verputzter Wände, die parallel zueinander angeordnet waren, gab es starke Reflexionen und akustische Moden. Der vorhandene DSL-Business-Anschluss der Telekom mit entsprechend schnellem Internetzugriff kam uns hingegen zugute. Den offenen Gigabitswitch des Hausnetzwerks im Nebenraum zur Regie durften wir für den Livestream ebenfalls nutzen.

#### 4. Entwicklung des Systemdesigns für *Stayin'Live Stuttgart*

Bei der Entwicklung eines Systemdesigns gibt es unterschiedliche Herangehensweisen, die ein System aus verschiedenen Perspektiven betrachtet, es in verschiedene Kategorien einteilt, Gewerke bestimmt und die Verbindung zwischen den zu übertragenden Quellen zu entsprechenden Senken schafft. Berücksichtigt werden sollten wichtige Faktoren wie zum Beispiel Wirtschaftlichkeit (Kosten), Flexibilität, personeller Aufwand zur Betreuung des

---

<sup>3</sup> (Petzold, 2020)

Systems, Skalierbarkeit, Geschwindigkeit (Auf- und Abbauzeit), Komplexität, einfache Handhabung durch Nutzer, Transportfähigkeit und Energieeffizienz.

Unter diesen Faktoren möchte ich im Folgenden die Quellen und Anforderungen an das Livestreaming herausstellen, die sich durch das vorangegangene Sendungskonzept ergeben haben.

Somit ist zunächst die Betrachtung mit weitem Blick auf die Situation ein sinnvolles Herangehen an die Systemplanung und Strukturierung. Es steht ein noch offenes System, das durch die umstehenden Faktoren grundlegend in seiner Struktur bestimmt wird. Bei diesen Faktoren handelt es sich zum Beispiel um die besondere personelle Situation aufgrund der Pandemie, die technischen Möglichkeiten vor Ort und das zur Verfügung stehende Material. Dazu gehören außerdem künstlerische Anforderungen an die Sendung und den Inhalt sowie das Produktionsziel und die damit verbundenen Zielformate und Parameter am Ausgang des Systems<sup>4</sup>.

All diese Faktoren werden zunächst analysiert und betrachtet, bevor der Blick ins Detail erfolgt.

#### 4.1 Personal und Besetzung der Positionen

Durch die besondere Lage der Covid-19-Pandemie ging die erste Überlegung dahin, wie eine Bild- und Tonregie so klein gefasst werden kann, dass mit max. fünf Personen in einem Raum und insgesamt nicht mehr als fünf Personen im Team ein Livestream technisch realisierbar und handhabbar wird. Das war für das Projekt eine Kernanforderung, die durch die äußeren Umstände als gegeben hingenommen werden musste.

Die Bildregie des HdM-eigenen TV-Studios benötigt in der Regel sechs Personen für einen flüssigen Ablauf und mindestens vier Personen, um alle nötigen Geräte zu bedienen: einen Bildtechniker, einen Bildmischer, einen Operator für die MAZ, einen Grafiker und einen Regisseur. Zusätzlich werden mindestens ein Toningenieur für den Sendeton sowie zwei bis drei Kameraleute benötigt. Für einen Livestream muss zusätzlich noch ein Sendetechniker/Livestream-Techniker eingeplant werden. Welche der zehn Positionen

---

<sup>4</sup> (Prof. Dr. Reinhard Haberfellner, 2018, S. 56-57)

konnten bei *Stayin'Live Stuttgart* zusammengeführt werden, um einerseits Personal zu sparen, aber andererseits den Qualitätsverlust so gering wie möglich zu halten?

Um schnell und mobil agieren zu können, nutzten wir kleine Kameras (Sony PXW-Z150) aus dem Bereich der elektronischen Berichterstattung (abgekürzt mit EB). Diese bringen alle wichtigen Parameter für unseren Livestream mit, kommen aber ohne Camera Control Unit aus (Sony Corporation, 2016), so dass wir auf die Position der Bildtechnik verzichten konnten. Unsere Kameraleute übernahmen den Bereich der Bildtechnik in Absprache mit dem Bildmischer und stellten Parameter wie Blende, Belichtung und Gain selbst ein. Da unser Licht statisch war und wir für eine gleichmäßige Ausleuchtung des Studios mit Kunstlicht gesorgt hatten, musste hier während des Livestreams/Sendebetriebs auch keine Veränderung vorgenommen werden.

Des Weiteren konnten die Positionen von Zuspiegelung, Grafik und Livestream-Technik zu einer zusammengefasst werden. Meine Wahl fiel auf die Open Broadcaster Software Studio (OBS), da sie das Encoding für den Livestream sowie das Senden an einen beliebigen Server bzw. direkt an eine Social-Media-Plattform per Real Time Messaging Protocol (abgekürzt RTMP) ermöglicht. Außerdem verfügt sie über zusätzliche Funktionen wie Grafik-Overlays, das Abspielen verschiedener Ton- und Bildaufzeichnungen (abgekürzt MAZ) sowie das Live-Schalten von Bild- und Tonquellen. Einige dieser Abläufe konnten wir so automatisieren, dass ein Cue der Regie mit „Go“ in der Sendung ausreichte.

Den Arbeitsbereich des Bildmischers fusionierten wir mit dem Bildregisseur, da aufgrund unseres Sendekonzeptes – das mit wenigen Protagonisten und einer Moderation auskommt – nur vier Kameras benötigt wurden. Zudem konnten die Grafiken und Einspielungen von unserem Livestream-Techniker übernommen werden. Die Bildregie schaltete die Kameras und gab die Schuss-Kommandos direkt an die Kameraleute weiter, so dass die oftmals hohe Kommunikation in der Bildregie zwischen Mischer und Regisseur vermieden werden konnte. Dies wirkte sich überdies positiv auf die Konzentration und das Arbeitsklima in der Regie aus.

Im Bereich Studio/Bühne war es problemlos möglich, mit zwei Kameraleuten zu arbeiten: Zwei statische (für den Bereich der Bühne/Musik und die Talkrunde) und zwei handgeführte Kameras (für Detailschüsse und Schuss/Gegenschuss Situation) reichten aus, um einen flüssigen Bildschnitt zu gewährleisten. Als Besonderheit ist jedoch zu erwähnen, dass jede

Kamera zwei Positionen abdecken musste, ohne dabei ihren Standort zu wechseln. Daher mussten die handgeführten Kameras auf Stativen so positioniert werden, dass jederzeit sowohl ein guter Schuss auf die Bühne als auch einer auf die Talkrunde möglich war.

Im Bereich des Tons erwies sich trotz zusätzlichem Aufwand ein Toningenieur als Operator als ausreichend. Zu nennen ist an dieser Stelle die Live-Binauralisierung des Audios für Kopfhörer. Hierzu haben wir alle Signale einzeln in der digitalen Audio Workstation (abgekürzt DAW) *MAGIX Sequoia 15* mittels *dearVR Pro* binauralisiert. Diese Kanäle wurden anschließend im Mischpult summiert, das Stereosignal sendefähig gemacht und an den Bildmischer zum Einbetten in das Programm übergeben.

Als entbehrlich zeigte sich darüber hinaus auch ein Tontechniker/Backliner im Studiobereich, da das im Remotenetzwerk integrierte Funksystem aus der Regie durch den Toningenieur überwacht werden konnte. Auch konnten mit Blick auf das Infektionsrisiko nicht zulässige Mikrofonwechsel umgangen werden, indem alle Künstler\*innen und die Moderatorin ihr eigenes Mikrofon über die gesamte Produktion einer Sendung behielten.

Insgesamt konnte die Personalstärke daher von zehn (vier Kameraleute, jeweils ein Grafiker, MAZ Operator, Regisseur, Bildmischer, Toningenieur und Tontechniker) auf fünf Personen (zwei Kameraleute, jeweils ein Livestream-Techniker, Toningenieur und Bildregisseur) dem Hygienekonzept entsprechend reduziert werden.

## 4.2 Transportfähigkeit der Produktion

Aufgrund der unterschiedlichen Produktionsräumlichkeiten musste das gesamte Equipment sehr flexibel und einfach zu transportieren sein. Da in der *Kulturinsel Stuttgart* kein Aufzug vorhanden ist, musste das Equipment tragbar sein, weshalb auf große Mischpulte verzichtet werden musste. Die Bundesanstalt für Arbeitsschutz und Arbeitsmedizin gibt an, dass eine maximale Belastung durch Heben, Halten und Tragen von Lasten von 40 kg bei Männern und 25 kg bei Frauen nicht überschritten werden sollte <sup>5</sup>. Ebenso mussten bei allen Cases und Touring-Cases die Maße der Türen vor Ort beachtet werden. Das bedeutete im Fall der Produktion im *Mókuska Caffè*, dass die Racks und Cases eine Breite von 600 mm nicht

---

<sup>5</sup> (Mantej, 2019)

überschreiten durften. Als Lösung hierfür bot sich ein Stack-Rack-System an: 19“-fähige Racks, die sich ohne Werkzeug stapeln und zu einer Einheit verbinden lassen, die mit einem Rollbrett versehen werden können und die zusätzlich die Möglichkeit bieten, die Geräte vorverkabelt zu transportieren. So konnte überdies vor Ort viel Zeit eingespart werden. Außerdem setzten wir Quick-Slide-Door-Stack-Racks von *ProCase* ein, welche zusätzliches Leergut durch Rackblenden vermieden, was Gewicht, Zeit zum Verstauen und unnötigen Platz in der Produktion verbraucht hätte. Da Bildmischer, Tonmischpult, Streaming-Technik, Stageboxen und Funksystem in Quick-Slide-Door-Stack-Racks verbaut wurden, konnten sie in einem Stapel transportiert und gerollt werden. Am Produktionsort wurden die Racks voneinander getrennt und an die jeweiligen Positionen geschoben. Das schwerste Rack wog 16 kg und konnte von zwei Personen problemlos getragen werden. Dieses QSD-Stack-Rack 600 mit 5 HE Einbauhöhe – unser größtes zu transportierendes Element des Systems – wies nach dem Trennen von allen anderen Racks ein Maß von 600 x 368 x 600 mm (BxHxT) auf.

### 4.3 Flexibilität und Skalierbarkeit der Produktionskomponenten

Durch die verschiedenen räumlichen Gegebenheiten und unter den besonderen Gesichtspunkten der Corona-Verordnung mussten unsere Arbeitsplätze in der Regie frei positionierbar wie auch skalierbar sein. Auf Basis eines Gigabit-Routers mit integriertem WLAN-Access-Point auf 2,5 GHz und 5 GHz wurden alle Positionen der Regie per IP-Netzwerk gesteuert. Ebenso waren alle Geräte per Web-Oberfläche oder Anwendung über einen Mac oder Windows-PC steuerbar und Touchscreen-fähig. Dies hatte zwei Vorteile: Zum einen konnten wir die Arbeitsplätze unabhängig der Position der Geräte in der Regie frei wählen und kurzfristig neu anordnen. Zum anderen waren auf allen Rechnern alle Anwendungen installiert, sodass sie im Havariefall Zugriff auf alle anderen Komponenten und Gewerke gehabt hätten, um schnell eingreifen zu können. Beispiel: Im Fall des Ausfalls des Rechners mit der Tonsteuerung, wäre der Ton dennoch weitergelaufen, da nur der Remote-Zugriff unterbrochen gewesen wäre. Folglich hätte ein anderer Rechner bzw. Arbeitsplatz der Regie diese Aufgabe umgehend übernehmen können. Somit war neben einer hohen Flexibilität auch eine Redundanz im Steuerungssystem gegeben. Sowohl in den Räumlichkeiten der *Kulturinsel* als auch im Konzertsaal der Musikhochschule stand uns wiederum mehr Platz zur Verfügung; auch konnten wir neben den Rechnern Hardwaresteuerungen nutzen. Außerdem

war ein Midi-Over-Ethernet-Controller für das Tonmischpult vorhanden, der sich direkt mit dem DSP-Mischer per Netzwerk verbunden hat.

Bei entsprechender Platzkapazität konnte der Bildmischer eine Hardwaresteueroberfläche nutzen, die sich per IP-Netzwerk mit ihm verband. Diese Hardwareoberflächen waren technisch zur Steuerung nicht notwendig und boten keinen zusätzlichen Funktionsumfang. Die Hardware kam dennoch so oft wie möglich zum Einsatz, da Steuerungs-Komfort sowie Genauigkeit und Schnelligkeit im Arbeiten für alle Beteiligten einen Mehrwert darstellte. So musste beispielsweise, um eine Verbindung aufzubauen lediglich ein Kaltgerätestecker zur Stromversorgung verlegt werden und ein Netzkabel mit RJ45-Buchse vom Gerät zum nächstgelegenen Switch gezogen werden. Die Hardwarecontroller und Mischer waren trotz aktivem Dynamic Host Configuration Protocol (abgekürzt DHCP) mit statischen IP-Adressen vorkonfiguriert, wodurch sich die Geräte sofort erkannten und wiederfanden.

Eine weitere Möglichkeit der Skalierbarkeit bestand im Bereich des Tonmischpultes, wobei wir uns für einen DSP-Mischer der Firma MIDAS entschieden hatten. Mit für unsere Produktion ausreichend Leistung, aber aufgrund des Verzichts auf physikalische Anschlüsse der Inputs und Outputs von AD- und DA-Wandlern und mit einer Höheneinheit im 19“-Rackmaß ein sehr platzsparender Mischer. Je nach Produktionsgröße, Anzahl und benötigten Positionen der Mikrofonvorverstärker mit AD-Wandlern und DA-Wandlern konnten wir per AES50-Netzwerk Stageboxen im 16in/8out-Format (MIDAS DL16) oder im 48in/16out-Format (MIDAS DL251) mitnehmen. So konnten wir je nach Größe des Ensembles und der benötigten Quellen und Senken unser System anpassen und mussten kein überflüssiges Material transportieren <sup>6</sup>.

#### 4.4 Maximallast und Leistungsaufteilung der Produktion

Die maximale Last in der Produktion war durch die Situation im *Mókuska Caffè* begrenzt. Hier waren im Studiobereich eine 16A-Schutzkontaktsteckdose (abgekürzt Schuko) mit B16-Sicherungsautomaten und eine weitere im Bürobereich des Cafés – in unserer Regie – eingebaut. Der höchste Stromverbrauch ergab sich durch die Beleuchtung. Hierfür nutzten

---

<sup>6</sup> (L MUSIC Group IP, 2016)

wir zwei 1 kW- und zwei 650 W-Stufenlinsenscheinwerfer, um unser Auflicht mit 3200 Kelvin Warmweiß zu generieren. Um in der Beleuchtung flexibel zu bleiben, betrieben wir alle weiteren Leuchten – darunter drei LED-Panels als Spitzenlicht und acht RGBWA-LED-StudioPAR-Scheinwerfer für Ambientebeleuchtung – mit Akkus, wodurch Arbeitszeit im Bereich der Verkabelung eingespart wurde. Mit einer Betriebszeit von fünf Stunden bei maximaler Last der LED-Lampen waren die Akkus bei einer gesamten Produktionszeit von acht Stunden mit Auf- und Abbau pro Tag absolut ausreichend für einen zuverlässigen Betrieb. Abgesehen davon lief das wichtigste Licht, das Auflicht, auf Netzspannung.

Leistungsbedarfsrechnung in Studio-/Bühnenbereich und der Regie:

Stückzahl	Gerät	Verbrauch	Verbrauch gesamt
	<b>16A Kreis Bühne (Betriebsspannung Beleuchtung)</b>		
02	DTS Scena 1000 MK2 Fresnel mit Botex UP-2 Dimmer	1kW	2kW
02	DTS Scena 650 MK2 Fresnel mit Botex UP-2 Dimmer	650W	1,3kW
	<b>Beleuchtungstechnik gesamt</b>	<b>Summe:</b>	<b>3,3kW</b>
	<b>16A Kreis Regie</b>		
01	BMD ATEM Studio HD	40W	40W
01	BMD Hyperdeck Mini (Annahme)	40W	40W
02	BMD SmartView 4K Monitor	24W	48W
04	BMD SDI – HDMI Converter Micro	1,6W	6,4W
01	MIDAS M32C (von M32 übernommen)	120W	120W
01	MIDAS DL16	45W	45W
01	Behringer X-Touch Faderwing	30W	30W
04	Laptops/Tablets div.	100W	400W
01	Nowsonic Stage Router Pro	25W	25W
02	Netgear GS108E Switch	6W	12W
	<b>16A Kreis Regie und zur Stagebox an der Bühne</b>		
04	Sennheiser ew500g4 (über ASA 214 betrieben)	-	-
01	Sennheiser ASA 214	28W	28W
01	MIDAS DL251 (redundantes Netzteil)	2x 150W	300W
04	Sony AC-L100	15W	60W
01	Netgear GS108E Switch	6W	6W
	<b>Ton- und Videotechnik gesamt</b>	<b>Summe:</b>	<b>1160,4W</b>

Tabelle 1 - Leistungsbedarfsrechnung der Produktion

Die Leistungsaufnahme mit dem gesamten Equipment musste sich auf zwei getrennte Stromkreise verteilen. Abgesehen von der Leistung war eine Trennung der Beleuchtungstechnik von der Ton- und Videotechnik durch die verwendeten Dimmer an den

Stufenlinsenscheinwerfern sinnvoll. Die verwendeten Dimmer *Botex UP-2* arbeiten nach dem Regelprinzip des Phasenanschnitts und können dadurch ein Netzbrummen verursachen, das auf dem Ton unter Umständen hörbar sein kann <sup>7</sup>. Bei hoher Belastung der Phase und des Stromkreises mit entsprechendem Dimmer Wert kann die veränderte Phase zu hörbaren Obertönen führen, wobei das Lastkabel des Dimmers ein so starkes elektromagnetisches Feld erzeugen kann, dass umliegende Kabel und Geräte durch Induktion zum Schwingen angeregt werden. In der Folge kommt es zu Störungen im Audiosignal, das als Brummen wahrgenommen wird <sup>8</sup>. Aus diesem Grund haben wir neben den Cat5e-Leitungen von der Regie zur Bühne zusätzlich ein Schukokabel mit 2,5mm<sup>2</sup> Leitungsquerschnitt gezogen, um die Stageboxen, die Backline der Musiker und das Funksystem mit der Betriebsspannung aus der Regie zu versorgen. So war auch gewährleistet, dass sich die Audiogeräte auf derselben Phase mit gleichem Nullleiter befinden, um weitere Störsignale zu vermeiden. Damit hatten wir auch für die Backline der Musiker wie Verstärker, Instrumente und Laptops genügend Spielraum mit über 2,4 kW Headroom.

## 4.5 Die zu übertragenden Quellen und benötigte Senken

Eine Herangehensweise im Systemdesign ist die Betrachtung der zu übertragenden Quellen auf die benötigten Senken, um passende Arbeitsabläufe mit dem gewählten Equipment zu entwickeln und daran ein passendes System zu gestalten. Aufgrund der Eingänge und Ausgänge am System mit den benötigten Anforderungen, wie zum Beispiel das Zielformat, orientiert sich die Wahl der Komponenten und deren Verbindung untereinander. Hier habe ich die Quellen und Senken in Bild- und Tonsignale getrennt betrachtet und anhand des Sendungskonzeptes definiert.

### 4.5.1 Quellen und Senken der Bildtechnik

Nachdem die großen Eckpunkte der Produktion durch Location, das Programm und äußere Gegebenheiten vorgegeben sind, folgte der nächste Schritt: Es wurden die möglichen Quellen

---

<sup>7</sup> (Thomann GmbH, 2020)

<sup>8</sup> (Bernstädt, 2019)

definiert und Eingangssignale bestimmt, die die Dimensionen an die jeweiligen Audio- und Videosystem vorgaben. Im Folgenden liste ich alle relevanten Quellen in der Bildtechnik auf, die durch das geplante Sendungskonzept und Programminhalte erforderlich wurden:

Als Bühnenkonzept wurde für alle geplanten Sendungen die Bühne in zwei Bereiche unterteilt: einer für die Musik und darstellende Kunst, einer für die Talkrunde/das Gespräch mit den Künstler\*innen. Für das Bild ergab sich hieraus der Wunsch, zwei statische Kameras aufzustellen. Eine der beiden bildete die gesamte Bühne der künstlerischen Darbietung ab, die andere die gesamte Gesprächsrunde. Zwei handgeführte Kameras lieferten die Nahaufnahmen. Somit lagen vier Kamerasignale per 3G-SDI am Bildmischer an. Erforderlich war außerdem die Möglichkeit für Zuspelungen und Grafik-Overlays.

Für die Senken musste der Bildmischer uns ein Programm mit eingebettetem Programmton in gewünschtem Sendeformat und einen AUX-Weg per SDI, der mit einer beliebigen Quelle des Bildmischers beschaltet werden konnte, übergeben. Der AUX-Weg lief auf den zweiten Eingang des Videomonitors für die Bildregie, um die Kamerabilder in der Probe zu prüfen. Zusätzlich wurde der AUX-Weg auf den zweiten Monitor in der Regie geschickt, der dem Toningenieur zur Verfügung gestellt wurde. Diese Flexibilität war nötig, damit der Toningenieur immer die gesamte Bühne überblicken konnte, indem er wahlweise Totale oder Multiview auf den Monitor schalten konnte.

Eingänge:	Ausgänge:
4x Kamera per SDI	1x Programm
1x Zuspelung per SDI	1x Multiview
	1x AUX

*Tabelle 2 - Quellen und Senken in der Bildtechnik*

#### 4.5.2 Quellen und Senken der Tontechnik

Im Bereich Ton gab das größte Instrumentarium die Dimensionierung des Audiosystems vor. Das Duo aus Tabea Booz (Vocals und E-Piano) und ihrem Begleiter, Daniel Pflumm (Elektronik, Bass und Kontrabass), konnten wir mit acht Kanälen abbilden: Kontrabass, E-Bass, E-Piano (Stereo), Elektronik (Mono), Back Vocals, Main Vocals für die Musik sowie das Moderationsmikrofon von Charlotte Schön.

Zusätzlich realisierten wir die Kommunikation zwischen Regie, Kameraleuten und Künstlern über das Audiosystem und verzichteten aus Platzgründen auf ein eigenes Intercom. Da wir genügend Eingänge und Busse im Pult zur Verfügung hatten, konnten wir auf moderne digitale Audiosysteme zurückgreifen, die über genügend Leistung verfügen, um in unserem Setup diese Aufgabe zu übernehmen. Für Kameraleute, Regie und Künstler benötigten wir fünf weitere Eingänge und fünf Busse, um so eine Matrice zwischen den Kommunikationspartnern aufzubauen. Nicht mehr nötig war infolgedessen die zusätzliche Betreuung des Kommunikationssystems, ebenso wie der Auf- und Abbau einer Intercom-Basisstation und deren Einrichtung in der Produktionsplanung. Bestehen blieb lediglich der Aufwand für die Einrichtung der Hör-Sprech-Kombinationen als Kopfhörer mit Schwanenhalsmikrofon und die Verbindung mit dem Audiosystem.

Als Ausgänge übergaben wir zunächst den Programmton in Stereo an den Bildmischer, wohingegen der Monitorausgang des Mischpultes dem Toningenieur zum Abhören zur Verfügung stand. Für die Musiker hatten wir zwei getrennte Stereokopfhörmischungen vorbereitet. Die Bildregie erhielt ebenfalls eine eigene Stereomischung auf Kopfhörer, wobei die Rückwege der Kameraleute zur Kommunikation zusätzlich auflagen. Den Kameraleuten wurde jeweils eine Monokopfhörmischung mit dem Kommando aus der Bildregie zugewiesen, da sie eine Einohr-Hör-Sprech-Kombination mit nur einer Ohrmuschel trugen. So war es ihnen dennoch möglich, umliegendes Geschehen im Bühnenbereich über das freiliegende Ohr wahrzunehmen.

Eingänge	Ausgänge
8x Mikrophon/Line per XLR (+48V)	1x Stereo Programmton per XLR
4x Intercom Mikrophon per XLR (+48V)	2x Stereo Kopfhörmischung Künstler per XLR
	1x Stereo Monitorweg für Toningenieur per XLR
	1x Stereo Kopfhörmischung Bildregie per XLR
	2x Mono Kopfhörmischung Kamera per XLR

Tabelle 3 - Quellen und Senken in der Tontechnik

#### 4.6 Wahl des Sendeformats - Stream Encoding und Bitraten

Als weiteren Schritt bei der Planung des Systems stellte sich die Frage, welches Format gesendet werden sollte, welche Auflösung, Bitraten, Audioauflösung und welche Plattformen welche Formate vorgeben würden. Zielplattform war in unserem Fall *YouTube* sowie als

Aufmerksamkeit versprechende Nebenplattform *Facebook*. Zu beachten war auch die langsamste Upload-Geschwindigkeit in der jeweiligen Location, die das Maximum für die gesamte Produktion vorgab. Durch das Ausbleiben eines angefragten Sponsorings bei Vodafone und Telekom, war es uns nicht möglich einen eigenen Anschluss oder eine Dark Fiber für das Event in jede Location zu bekommen. Wir mussten mit dem leben, was wir vor Ort vorgefunden haben.

#### 4.6.1 Die richtige Balance in der Formatwahl und der Bitrate

Entscheidend für die richtige Balance in der Formatwahl und der Bitrate ist zunächst, welche Internetgeschwindigkeit in deutschen Haushalten vorhanden ist bzw. welche Downloadgeschwindigkeiten möglich sind. Daran orientiert sich das Zielformat im Encodieren des Videosignals mit passender Bitrate. Da in Deutschland der Ausbau von digitaler Infrastruktur nach wie vor Aufgabe der Länder und Kommunen ist, liegen sehr unterschiedliche Werten vor. Zwar gibt es ein bundesweites Förderprogramm – „Das Breitbandförderprogramm des Bundes“ – die Umsetzung für die letzten Meter in die Haushalte liegt allerdings in der Hand der Länder und Kommunen <sup>9</sup>. Die durchschnittlichen Werte der Downloadraten variieren zwischen ca. 14,45 Mbit/s in Niedersachsen und 27,22 Mbit/s in Bremen <sup>10</sup>. Ausgehend vom niedrigsten Wert von knapp 15 Mbit/s ist eine Übertragung von Full HD in 1920 p x 1080 p theoretisch möglich. *YouTube* gibt diesbezüglich mögliche 2560 p x 1440 p mit einer Bitrate von 6 bis 13 Mbit/s bei 30 fps an <sup>11</sup>. Die Spanne von 6 bis 13 Mbit/s indiziert einen gewissen Spielraum in der Kompression und dient zusätzlich als Puffer, bei auftretenden Schwankungen der Uploadrate auf den Server. Eine höhere Bitrate zu wählen, die durch *YouTube* wieder komprimiert werden würde, wäre unnötiger Datentransfer. *YouTube* kann hier lediglich als Orientierung dienen, sodass eine Balance zwischen Auflösung und Bitrate gefunden werden muss, um eine gute Auflösung für Detailschärfe zu gewährleisten. Die Bitrate sollte entsprechend hoch sein, damit Artefakte vermieden werden können und genügend Information für Farbtiefe und Kontraste in den

---

<sup>9</sup> (BMVI, 2021)

<sup>10</sup> (Heuzeroth, So schnell ist das Internet in Ihrem Bundesland, 2019)

<sup>11</sup> (Google LLC, 2021)

Pixeln vorhanden ist. Dennoch sollte die angegebene maximale Bitrate für die jeweilige Auflösung von *YouTube* nicht überschritten werden.

Zu klären ist ferner, wie oben bereits angedeutet, welche Geräte in den Zielhaushalten zum Empfangen des Livestreams neben mobilen Endgeräten vorhanden sind. In den meisten deutschen Haushalten findet man TV-Geräte in Full HD mit einer Auflösung von 1920 p x 1080 p. Dabei lag der Anteil im Jahr 2020 bei 83,9 %. Dem gegenüber stehen 24,3 % TV-Geräte in UHD. 44,8 % aller Geräte verfügen über Smart-TV-Funktionen, die Zugriff auf *YouTube* und andere Plattformen über das Internet erlauben <sup>12</sup>. Es lässt sich also auf eine optimale Auflösung von 1920 p x 1080 p bei bis zu 30 fps mit einer von *YouTube* empfohlenen Bitraten zwischen 6 und 13 Mbit/s schließen. Das wäre auch im Rahmen der niedrigsten durchschnittlichen Downloadrate von 14,45 Mbit/s in Niedersachsen auf dem deutschen Markt.

An den Veranstaltungsorten maßen wir bei einer technischen Begehung die Down- und Uploadraten und nahmen einen Test mit unserer Encoding- und Streaming-Software *OBS* vor. Dies lieferte folgende Ergebnisse:

1. HMDK Stuttgart (20.05.20, 12:52 Uhr)  
Download: 20.044kbit/s                      Upload: 18.968 kbit/s
2. Mókuska Caffè (19.05.20, 19:58 Uhr)  
Download: 46.650 kbit/s                      Upload: **7.810 kbit/s** (geringster Wert)
3. Kulturinsel Stuttgart (12.06.20, 14:21 Uhr)  
Download: 56.880 kbit/s                      Upload: 11.530 kbit/s

Hätten wir auf einen Headroom im Upload verzichtet, wäre die Auflösung von Full HD mit 1920 p x 1080 p nur mit großer Kompression oder der Gefahr eines Abbruchs im Stream möglich gewesen. Um die bestmögliche Auflösung im Audio zu verwenden und einen stabilen Livestream ohne Framedrops, stotternde Bilder und Schwarzbild-Sequenzen zu ermöglichen, entschieden wir uns für eine geringere Auflösung mit bestmöglicher Bitrate und einem entsprechenden Puffer zwischen benötigter und tatsächlich gemessener Uploadrate. Denn zur Überwachung des Uploads und des Streams auf den Social-Media-Plattformen sowie zur

---

<sup>12</sup> (Heuzeroth, 2020)

Steuerung des Streaming-Servers musste unser Stream-Techniker regelmäßig auf die Accounts der Plattformen zugreifen und die Streams abhören.

Schließlich wählten wir nach einigen Tests die Auflösung 1280 p x 720 p bei 5 Mbit/s. Als Audio-Codec war AAC vorgegeben; wir streamten mit der bestmöglichen Qualität von 320 Kbit/s.

#### 4.6.2 Streaming Workflow

Um mehrere Plattformen flexibel ansprechen zu können (Multistreaming), hatten wir einen Streaming-Server angemietet. Dies brachte den Vorteil, dass wir die Uploadrate von *OBS* zum Server nur einmal benötigen und sich nicht die Uploadrate mal die Anzahl der Server der anzusprechenden Social-Media-Plattformen multiplizierte. Der Service *Restream.io* bietet einen schnellen Workflow als Content Delivery Network (abgekürzt CDN) und einen einfachen Anschluss an alle gängigen Streaming-Plattformen wie *YouTube*, *Twitch*, *Facebook*, *LinkedIn* u.a. per RTMP zur Übertragung des Livestreams in H.264<sup>13</sup>.

Da sich mit *Restream.io* ein Teil unserer Arbeit bzgl. des Livestreamings automatisieren ließ, bot es sich als adäquate Lösung für unseren mobilen Livestream an. Darüber hinaus war es möglich, auf geplante Events auf *YouTube* zu streamen, was normalerweise nicht im direkten RTMP-Stream auf *YouTube* möglich ist. Ein Livestream kann in diesem Fall zuvor erstellt und als Event mit passendem Link zum Livestream im Voraus beworben werden, was zur schnellen und einfachen Verbreitung entscheidend beiträgt. Ebenso ließen sich mit dem CDN-Server automatisiert Beiträge zur Werbung auf *YouTube* und *Facebook* veröffentlichen, was das Produktionsteam am Tag des Livestreams stark entlastete, so dass sich das kleine Produktionsteam mit Doppel-Funktionen ganz auf die Produktion vor Ort konzentrieren konnte. Sobald unser Server auf *Restream.io* einen RTMP-Stream mit passendem Schlüssel empfangen hatte, wurde sowohl auf *YouTube* als auch auf *Facebook* ein Event mit vordefinierten Angaben und weiterführenden Links zum Livestream erzeugt. Einem Zeitplan entsprechend wurden automatisch vorproduzierte Beiträge mit Spendenlink und Werbung der beteiligten Künstler veröffentlicht, um Zuschauer auf den Livestream aufmerksam zu

---

<sup>13</sup> (Restream Inc., 2021)

machen. Somit konnten wir am Produktionstag den Bereich der Werbung und des Sendens/Content Delivery ab dem RTMP Stream aus *OBS* voll automatisieren.

#### 4.6.3 Stream-Encoding per Software

Im Bereich des Encodings strebte ich eine Lösung an, die für das Projekt kostenfrei ist und zusätzliche Funktionalität bietet: das Einspielen von Grafiken für Bauchbinden und Titel, das Abspielen von Videos sowie das Synchronisieren von Bild und Ton vor dem Encodieren und das Senden per RTMP an den CDN-Server *Restream.io*. Aus technischer Sicht war vorgegeben, mit dem entsprechend kompatiblen Audiocodec AAC bis zu einer Auflösung von 1920 p x 1080 p mit 50 fps in h.264/x.264 zu encodieren, da dies eine effektive und zugleich gängige Art der Kompression ist <sup>14</sup>.

*OBS* ermöglicht es, die abzuspielenden Elemente, wie bewegte Titel, Abspann, statische Bauchbinden und Videozuspielungen/MAZ, in Szenen und Reihenfolgen anzulegen und damit den Arbeitsaufwand für den Livestream-Techniker und Grafiker überschaubar zu halten. Zu seinen Aufgaben gehörte es, in der Produktion den RTMP Stream zu überwachen, Grafiken zu schalten und die Zuspielung zu betreuen. Dies wiederum ließ sich so nur realisieren, da zuvor ein halbautomatisierter Ablauf über Macros und programmierte Szenenfolgen eingerichtet worden war. Zum Abspielen von Zuspielungen und Schalten von Grafiken musste somit nur ein Cue getätigt werden, sofern sich die Programmreihenfolge nicht spontan änderte.

Das Encodieren von Videos ist ein aufwendiger Rechenprozess, der CPU und GPU (falls kompatibel zum Encoder) stark beanspruchen kann. Da wir mobil und möglichst platzsparend agieren wollten, war eine wichtige Frage, ob *OBS* auf einem Consumer-Rechner würde arbeiten und unseren Anforderungen genügen können. Hierfür hatten wir vier verschiedene Laptops bzw. Tower-PCs des Teams getestet, die alle mit dem Betriebssystem *Microsoft Windows 10* ausgestattet sind:

- *Dell XPS 15"* (7590 mit i7-Prozessor, 16 GB RAM, GPU: Nvidia Geforce GTX 1650)
- *Lenovo ThinkPad T480s* (i7-Prozessor, 16 GB RAM, GPU: Intel GPU)
- *Surface Pro 7* (i7-Prozessor, 16 GB RAM, GPU: Intel GPU)

---

<sup>14</sup> (Fischer, 2016, S. 566)

- *Surface Book* (i7-Prozessor, 16 GB RAM, GPU: Intel GPU)
- *Tower-PC* (AMD Threadripper 1950 x 16-Kern-Prozessor, 64 GB RAM, GPU: NVIDIA Geforce 1080)

Dabei konnten keine offensichtlichen Unterschiede zwischen den Rechnern festgestellt werden; die Performance hat sich weder im Livestream noch bei der Nutzung negativ im Umgang mit der Software gezeigt. Letztlich wurde *OBS* auf dem beschriebenen *Dell XPS 15*“ ausgeführt und funktionierte stabil ohne Abstürze oder Ausfälle.

Ich möchte nicht unerwähnt lassen, dass die Möglichkeit bestanden hätte, einen Hardware-Encoder zu nutzen. Diese Hardwarelösung wäre grundsätzlich auch zu bevorzugen, da sie eine höhere Betriebssicherheit gewährleistet. Abgestimmte Hardware auf das Encoding und dedizierte Komponenten, welche für Video optimiert sind und nur eine bestimmte Aufgabe übernehmen. Im Falle verfügbaren Budgets hätten ich dies der Software Lösung vorgezogen. Die Nutzung von Consumer-Hardware mit Windows oder anderen Betriebssystemen ist jedoch eine kostengünstige Alternative, die in unserem Fall zuverlässig arbeitete. Allerdings fordern Wartung und Optimierung des *Microsoft* Betriebssystems für den professionellen Betrieb einen sehr hohen zeitlichen Aufwand in der Vorbereitung und Einrichtung.

#### 4.7 Aufbau der Bildtechnik und ihre Komponenten

Nachdem die äußeren Parameter der Ein- und Ausgabe die Wirkungsorientierte Betrachtung des Systems abdecken, folgt nun eine Strukturorientierte Betrachtung; die Wahl des Equipments, deren struktureller und lokaler Aufbau und die Steuerung der entsprechenden Systeme <sup>15</sup>. Eine mögliche Betrachtung mittels der Up-Down-Methode führt vom Groben ins Detail und ermöglicht eine nachvollziehbare Darstellung in komplexen Systemen <sup>16</sup>. Es geht eine Stufe ins Detail und ich Teile die Regie nun in passende Gewerke. Eines davon ist die Bildtechnik.

---

<sup>15</sup> (Prof. Dr. Reinhard Haberfellner, 2018, S. 36-37)

<sup>16</sup> (Prof. Dr. Reinhard Haberfellner, 2018, S. 56-57)

#### 4.7.1 Der Bildmischer und dessen Anbindung im System

Bei der Wahl unserer Komponenten der Bildtechnik bildete der Bildmischer ein zentrales Element, da bei ihm alle Bildsignale ankommen, geschnitten, bearbeitet und wieder ausgegeben werden. Nachdem das Sendeformat bestimmt war und die Anzahl der Quellen und Senken feststand, konnte nun das Gerät gewählt werden. Auf der Suche nach einem passenden Bildmischer für unser benötigtes System bei *Stayin' Live* kam in der Phase des aktuellen Livestream-Hypes aufgrund der Corona-Pandemie nur das Produkt der Firma *Blackmagic Design* infrage. Das lag in erster Linie an unseren Kameras aus dem Bereich der EB Technik, die keine Möglichkeit besitzen, ein Referenzsignal oder einen Haustakt aufzunehmen. Es war uns also nicht möglich, die Kameras per Blackburst oder Tri-Level-Synchronsignal, das entweder durch den Bildmischer selbst oder in der Regel durch einen externen Generator per Genlock erzeugt wird, framegenau zu synchronisieren. Wir folgten hier den Vorgaben über den von der *Society of Motion Picture and Television Engineers* definierten SMPTE274M-Standard <sup>17</sup>.

Da wir in unserem Fall aber nicht in Echtzeit übertragen mussten, konnten wir von den Möglichkeiten der Resynchronisation jedes Eingangssignals Gebrauch machen. Dabei wird eine Latenz von bis zu drei Frames am Eingang des Bildmischers in Kauf genommen, der Frame wird kurzzeitig gespeichert, bis an jedem Eingang ein Frame anliegt, um dann die Frames synchron auf den Bildmischer zu geben <sup>18</sup>. Dies machte einen Verzicht auf Kameras mit Tri-Level-Synchronisation oder per Black Burst möglich und brachte Entlastung im Auf- und Abbau, da pro Kamera eine BNC-Verbindung zwischen Bildmischer und Kamera weniger verlegt werden musste. Der Wegfall eines Haustaktes erleichterte uns ebenso den Umgang mit Consumer-Bildquellen. Per HDMI oder SDI können somit unsynchronisierte Signale aus Laptops oder anderen Endgeräte Bildsignale an den Bildmischer geschickt werden, was mehr Kreativität in der Produktion zulässt. Andere Quellen können dadurch flexibel und schnell eingebunden werden, solange sie der Auflösung und Framerate der Produktion entsprechen.

Hinsichtlich des gewählten Produktionsformat lässt sich festhalten, dass unser Programm in einer Auflösung und Framerate von 1920 p x 1080 p mit 50 fps erstellt und per Recorder

---

<sup>17</sup> (SMPTE, 2004)

<sup>18</sup> (Blackmagic Design Pty. Ltd., 2021)

aufgezeichnet wurde, ehe es für das Sendeformat in 1280 p x 720 p mit 25 fps umgewandelt wurde. Die Aufzeichnung in höherer Auflösung bedeutete keinen höheren Produktionsaufwand.

Bezüglich des Farbraums waren wir auf den Standard BT.709 beschränkt, der für alle Produktionen, die in Standard Dynamic Range umgesetzt werden, vorgegeben ist. Eine Unterstützung von Rec.2020-Farbräumen ist grundsätzlich nur in Videouploads im definierten HDR-Modus möglich. Bei Livestreamings jedoch wird von BT.709 als Farbraum ausgegangen und entsprechend matriziert. Auch wenn der sRGB-Farbraum dem BT.709-Farbraum sehr ähnlich ist, wird er nicht empfohlen und es erfolgt eine Umwandlung zu BT.709 TRC <sup>19</sup>. Der Farbraum BT.709 ist Teil der ITU-Empfehlung ITU-R BT.709.6, die von der Internationalen Fernmeldeunion herausgegeben wird, die unter anderem die Standards zur Übertragung von Rundfunk und Fernsehen bündelt und definiert <sup>20</sup>. Somit wäre eine spätere Verwertung des Materials in Fernsehen möglich gewesen. Die Mischer der Blackmagic-Design-Reihe *ATEM* unterstützen den entsprechenden Farbraum nach eigenen Angaben des Herstellers <sup>21</sup>. Mit acht möglichen Eingängen – vier davon als SDI nach den Richtlinien SMPTE 259M, 292M und 424M konform ausgeführt – stellte sich die *Blackmagic Design ATEM Television Studio HD* als gute Wahl heraus. Die vier anderen Eingänge standen als HDMI zur Verfügung, über welche Consumer-Geräte angebunden werden konnten, was die Anbindung einer Zuspiegelung, zum Beispiel über *Adobe Premiere Pro* gewährleistete. Ausgangsseitig schlossen wir unser Programm an den Wandler von SDI zu USB3.0 an. Ebenso konnten wir den Monitor der Bildregie mit einem Multiview und einem AUX per SDI versorgen.

#### 4.7.2 Anforderungen an die Kamera und Kameraumgebung

Aufgrund der wechselnden Lokalitäten und der damit unterschiedlichen Lichtsituation sowie dem geringen Platzangebot in zwei unserer Livestreams mussten die Kameras klein, mobil und leicht in der Handhabung sein. Gleichzeitig sollten sie mit schlechten Lichtverhältnissen umgehen können und dabei den technischen Spezifikationen unseres Produktionsformats und des Sendeformats entsprechen. Erwünscht war zudem ein intelligenter Autofokus.

---

<sup>19</sup> (Google LLC, 2021)

<sup>20</sup> (ITU, 2015)

<sup>21</sup> (Blackmagic Design Pty. Ltd., 2021)

Infolge der Fusionierung einiger Funktionen innerhalb des Teams, erforderte die Besetzung dieser Position nicht mehr nur einen Kameramann, sondern optimalerweise einen Techniker/Ingenieur, der über ausgewiesene Kompetenzen im Bereich der audiovisuellen Technik verfügt. Beim Auf- und Abbau und bei der Einrichtung der Technik mussten alle Produktionsbeteiligten alle Bereiche in den Grundkenntnissen verstehen und beherrschen können. Dies war auch deshalb erwünscht, um bei einem Aus- oder Krankheitsfall schnell reagieren und Positionen tauschen zu können. Da Bedienung eines manuellen Fokus viel Erfahrung voraussetzt, die nur von erfahrenen Kameramännern und Kamerafrauen verlangt werden kann, griffen wir auf moderne Kameras mit automatischer Gesichtserkennung und guten Trackingeigenschaften zurück. Um sich an die stark variierenden Lichtverhältnisse anpassen zu können, mussten unsere Kameras über einen variablen ND-Filter verfügen. Bedingt durch die geringe Kameraanzahl mussten unsere Kameramänner mit der gleichen Optik sowohl eine Totale als auch Close-Up-Schüsse vom Geschehen abdecken können. Zoomoptik mit schnellen Fahrten und hohem Brennweitenbereich stellte ebenso ein wichtiges Kriterium dar. Trotz des Wunsches, eine EB-Kamera aufgrund ihrer geringen Größe und der dadurch leichten Platzierbarkeit zu nutzen, sollte die Möglichkeit einer Hinterkamerabedienung per LANC möglich sein, um Zoom- und Fokusfahrten ohne Wackler, dafür ergonomisch aus der Hand des Statives zu bedienen. Bei Schüssen, die aus der Hand erfolgen sollten, baten die Kameramänner um eine leistungsstarke Bildstabilisation. Dementsprechend ausgestattet sind EB-Kameras von Sony, wir nutzten eine *Sony PXW-Z150*. Die Optik war mit einem 12-fach-Zoom von 29 mm bis 348 mm (35 mm Kleinbildformat) bei einem durchschnittlichen Abstand von 6 m zwischen Kamera und Protagonisten ideal. Der 1“-CMOS-Sensor in 4:2:2 und 10bit Farbtiefe entsprach der Anforderung zu unserem BMD-ATEM-Bildmischer mit kompatiblen SDI-Ausgänge nach SMPTE ST 424 / ST 292-1 / ST 259-Standards <sup>22</sup>. Die Kameras waren sehr kompakt und leicht mit einem Gewicht von 2,5 kg inkl. Akku, einem *Sony NP-F970*. Um sicherzugehen, dass wir die Entfernungen zwischen Bühne und Regie würden abdecken können, testeten wir die Reichweite des SDI-Signals. Die Angaben der Hersteller variieren sehr stark und sind stets abhängig von der Art und Qualität des Koaxialkabels sowie der verbauten Stecker und Buchsen. Zu beachten ist, dass Qualität

---

<sup>22</sup> (Sony Corporation, 2016)

und Länge Einfluss auf die frequenzabhängige Dämpfung nehmen können. 90 Meter sind in der Regel machbar, ggf. auch mehr <sup>23</sup>.

Nach einigen Tests arbeiteten wir mit einem zertifizierten Belten HDTV-Koaxialkabel für 120m. Verwendet wurde jedoch in der Produktion eine 100 m-Trommel des Kabels. Da wir bei vier Kameras mit zwei Close-Ups nur schwer auf eine Kamera hätten verzichten können, betrieben wir die Kamera mit einem Netzteil und verwendeten zusätzlich einen NP-F970-Akku, um im Bedarfsfall einen Ausfall des Netzteils über die gesamte Sendung überbrücken zu können. Im ersten Livestream mit Annique Göttler trat dieser Fall ein, als uns kurz vor der Sendung ein Netzteil der Kamera 2 ausgefallen war.

Das geringe Gewicht der Kameras bedeutete allerdings auch einen Nachteil: Aufgrund fehlender Trägheit ließen sich kleine Ruckler und stockende Bewegungen bei Kamerafahrten mit den Stativen nur schwer vermeiden. Daher wählten wir alternative Stative und ölgedämpfte Videoköpfe der Marke *Sachtler*, die sich bei unserem Test mit Produkten aus dem unteren Preissegment (<1000€) gegenüber der Marke *Manfrotto* durchsetzten. Die besten Ergebnisse, im Verhältnis zu Wirtschaftlichkeit, Portabilität und des geringen Stativgewichts von 4,4 kg, erzielten wir mit den *ACE M*-Stativen der Marke *Sachtler*. Diese sind für ein Gesamtgewicht der Kamera von 2,5 kg ideal und bis zu einer Maximallast von 4 kg ausgelegt <sup>24</sup>.

Als Hinterkamerabedienung standen uns die *Manfrotto LANC-Controller Typ MVR901ECLA* zur Verfügung und wurden vorwiegend für dynamische Zoomfahrten genutzt. Unter LANC versteht man eine bidirektionale Verbindung mittels seriellen Protokolls zwischen zwei Geräten. Hier können Daten wie Aufnahmeposition, Record, Play, Timecode etc. abgerufen und kontrolliert werden. Das offene LANC-Protokoll unterstützt auch das Zoomen mit variablen Zoomgeschwindigkeiten, die weiche, ästhetische Zoomfahrten dem Kameramann über die Zoomwippe am LANC-Controller erlauben. Das Zoomen in den Telebereich und in den Weitwinkelbereich wird in jeweils acht Stufen plus Halt unterteilt <sup>25</sup>.

---

<sup>23</sup> (BET-Fachwörterbuch, 2020)

<sup>24</sup> (Sachtler, 2011)

<sup>25</sup> (Boehmel, 2020)

### 4.7.3 Tally – Aufnahmelicht-/Rotlichtsteuerung

Ein Tally-System ist in der Liveproduktion mit Kamera eine wichtige Voraussetzung zum schnellen und effektiven Arbeiten mit handgeführten Kameras. Wir arbeiteten mit einem System mit Grünzeichen für Preview und Rotzeichen für Programm. Die Kameramänner erhielten über das „Grün“ die Warnung für die Schnittvorbereitung, so dass das Achtungszeichen nicht mehr mündlich über das Intercom mitgeteilt werden musste. Das führte zu einer klareren Kommandostruktur in der Kommunikation zwischen Bildregie und Kamera. Ebenso vermittelte das klassische Rotlicht die Information für Kameramann und Protagonist, dass dieses Bild gerade gesendet wird.

In Verbindung mit dem BMD ATEM-Bildmischer war es nicht leicht, eine Lösung zu finden, die einfach im Aufbau ist, zuverlässig funktioniert und sich per IP steuern lässt. Das *FlexTally* von *Cerevo*, ein drahtloses Tally-System, das aufgrund des 433 MHz-Bands anmeldefrei nutzbar ist, konnte ich vorab testen. Es verfügt über eine Basisstation, die über IP mit dem BMD ATEM-Bildmischer verbunden wird. Die Steuerung erfolgt über eine Web-Oberfläche über per Browser. Dort wird mittels IP-Verbindung über Ethernet eine Verbindung zwischen Basisstation und Bildmischer eingerichtet. Die Basisstation überträgt anschließend das Signal drahtlos an die Tally-Leuchten. Diese können per integriertem Akku (max. sechs Stunden Laufzeit bei geringster Helligkeitsstufe), per USB-Ladegerät oder Powerbank betrieben werden <sup>26</sup>. Hierbei ist wichtig, dass der Bildmischer eine statische IP-Adresse erhält, da sonst bei einem Neustart des Routers oder des Bildmischers ein Verlust der Verbindung von Basisstation zum Bildmischer gedroht hätte.

Für ein mobiles Setup erwies sich die Anwendung des Tally-Systems auch mit Blick auf den Akku-Betrieb über die Dauer einer Sendung als einfach, zuverlässig und schnell in Auf- und Abbau.

### 4.7.4 Wandlung des SDI-Signals zur Softwareanbindung

Eine wichtige Schlüsselkomponente im Systemdesign bildet die Wandlung des unkomprimierten digitalen Videosignals von SDI oder HDMI zu USB (Universal Serial Bus) zur

---

<sup>26</sup> (Cerevo, 2021)

Anbindung an einen PC oder Laptop an die Streaming Software, zum Encodieren von Bild und Ton. Da wir mit CMOS-Sensoren auf Grundlage der Bayer-Matrix an unseren Kameras arbeiteten, wurde die Produktion in 4:2:2 nach dem YCbCr-Modell mit 8-bit Farbtiefe geschossen. Nicht möglich war aufgrund der eingesetzten EB-Kameras die Nutzung von 3-Sensor-Kameras mit Prisma. Ein 4:4:4 -Signal mit drei getrennten Bildsensoren für Rot, Grün und Blau hätte eine bessere Farbauflösung und Detailauflösung liefern können, da auf eine Interpolation der Pixel verzichtet werden kann <sup>27</sup>. Dies ist aber aufgrund des Zielsignals für die Social-Media-Plattformen mit einem viel höheren Aufwand verbunden, der nicht zuletzt durch die deutlich höheren Kosten für uns unwirtschaftlich geworden wäre.

Unser Bildsignal wurde aus dem Mischer in 4:2:2 mit 8 bit per SDI mit Level B angeliefert und musste per USB-Verbindung mit dem Rechner verbunden werden, wofür es verschiedene Möglichkeiten gibt: Da wir das Sendeformat von 1280 p x 720 p gewählt haben, konnte hier eine USB 2.0-Verbindung genutzt werden, da diese mit maximal 480 Mbit/s für eine Videodatenrate im Zielformat 1280 p x 720 p mit 25 fps ausreicht. Hierfür griff ich auf den Web-Presenter von *BMD* zurück. Dieser nimmt das unkomprimierte SDI-Signal auf, konvertiert es von Full HD in HD mit 720 p und passt ebenfalls – wenn nötig – die Bildwiederholrate mit entsprechenden Algorithmen an, wie zum Beispiel mit einem 3:2-Pulldown oder -Pullup. Da im System das Sendeformat mit 25 fps vorgegeben war, wurde von 50 fps auf 25 fps konvertiert, wobei jedes geradzahlige Bild übersprungen und das ungeradzahlige Bild doppelt gespielt werden musste – die einfachste Umwandlung, die auch keine Artefakte erzeugt. Die Ausgabe über USB erfolgte im UVC-Standard, der sogenannten *USB Video Class*. Dieser Standard ermöglicht das Streamen von Videosignalen auf einem Rechner ohne speziellen Treiber <sup>28</sup>. UVC ist ein einheitlicher Standard mit einem vorinstallierten Treiber durch das Betriebssystem sowohl bei *Windows* und *Linux* als auch bei *Mac OS* und ermöglicht den „Plug and Play“-Betrieb <sup>29</sup>. Somit lag ein passendes Video-Signal mit Programmtone zum Encodieren für den Livestream an *OBS* an.

Da im Systemdesign das gewünschte Equipment durch Lieferengpässe nicht verfügbar sein kann und infolgedessen andere Komponenten verwendet werden oder schlimmstenfalls

---

<sup>27</sup> (BET-Fachwörterbuch, 2018)

<sup>28</sup> (Blackmagic Design Pty. Ltd., 2021)

<sup>29</sup> (Walke-Chomjakov, 2018)

Workarounds gebaut werden müssen, ist die Variantenbildung an systemkritischen Bereichen elementar. Wichtige Schlüsselkomponenten, die als Bottleneck bzw. Engpass im System als kritische Komponente das gesamte System gefährden können, sollten hier durch verschiedene Varianten abgedeckt werden. Dies kann entweder als Prinzipvariante mit einem anderen Lösungsansatz im Workflow oder als Detailvariante mit verschiedenen Komponenten, die zum gleichen Ergebnis in der Kette führen, geschehen<sup>30</sup>. Beides wurde in der Planung für dieses Livestreaming-Setup bedacht. Nicht umsonst, wie sich später herausstellen sollte, denn die ursprünglich geplante Version konnte aus folgenden Gründen nicht umgesetzt werden:

Aufgrund der Corona-Pandemie kam es zu Lieferschwierigkeiten; zum einen, weil die Industrie in China zum Stillstand gekommen war und zum anderen, weil das Livestreaming eine der wenigen Optionen war, die trotz des Lockdowns in Deutschland Veranstaltungen möglich gemacht hat. Da alle Livestream-relevanten Produkte der Firma *BMD* vergriffen waren (den Web-Presenter erhielt ich letztlich kurz vor der ersten Sendung) und ich aus Gründen der Betriebssicherheit eine Alternative einplanen musste, wählte ich einen Full HD fähigen Konverter. Bei einer Datenrate von 1,5 Mbit/s bei 1920 p x 1080 p mit 25 fps wurde eine schnellere USB-Verdichtung mit mindestens USB3.0-Standard nötig, die Datenraten von bis zu 5 Gbit/s verarbeiten und eine Auflösung vom 1920 p x 1080 p mit bis zu 60 fps gewährleisten kann. Zur Verfügung stand hierzu ein *Marshall VAC-23SHU3* SDI/HDMI zu USB-Konverter. Dieser unterstützt eingangsseitig 4:2:2 mit bis zu 10 bit als SDI SMPTE 296M, 274M, 259M, 425M, in A- oder B-Level, und ausgangsseitig USB3.0 sowie zur Videoübertragung an den Rechner das UVC Protokoll<sup>31</sup>.

#### 4.7.5 Programmaufzeichnung - MAZ

Sowohl Künstler als auch Produktionsteam äußerten den Wunsch, das Programm vor Ort aufzuzeichnen. Da die Künstler unentgeltlich arbeiteten, lag nahe, dass das produzierte Programm ihnen später für Werbung und Demos zur Verfügung gestellt werden würde. Außerdem war es uns im Team wichtig, das Material aller Sendungen in Ruhe auszuwerten.

---

<sup>30</sup> (Prof. Dr. Reinhard Haberfellner, 2018, S. 59-60)

<sup>31</sup> (Marshall, 2021)

Um unabhängig von Sendeformat und Bitrate zu sein, wollten wir das Programm in guter Qualität vor der Kompression für den Stream aufzeichnen. So hätten wir auch bei einem eventuellen Ausfall oder Frame-Drops keine Bildaussetzer in der Aufzeichnung gehabt und hätten im Nachhinein eine korrigierte Fassung hochladen können. Somit hatten wir für die Produktion ein Backup geschaffen. Diesbezüglich habe ich nach einem Hardware-Recorder gesucht, der sich per Bedienpanel steuern, sich aber auch per Netzwerk bedienen lässt und in den Workflow unserer Bildtechnik passt. Die Steuerung per Netzwerk sollte ähnlich zum Bildmischer eine dezentrale Platzierung, nicht in Griffweite der Regie, zulassen und trotzdem bedienbar bleiben.

Als hilfreiche Ergänzung zu unserem Bildmischer hat sich der *BMD HyperDeck Studio Mini* angeboten, der unser 3G-SDI-Signal inklusive eingebettetem Programmtton aufnehmen konnte. Mögliche Formate für unser 4:2:2 8bit Signal waren *AVID DNxHD*, *AVID DNxHR*, *Apple ProRes QuickTime* und H.264. Audio wurde hier wahlweise in AAC mit maximal 320 kbit/s oder per PCM mit maximal 48 kHz und 24 bit aufgezeichnet. Als Speichermedium kamen zwei SD-Karten im UHS I oder II Standard zum Einsatz, die eine Schreibgeschwindigkeit von bis zu 300 MB/s zulassen. Auch der erforderliche SDI-Standard aus der Produktionskette wurde wiederum eingehalten, so dass SMPTE 259M, 292M, 296M und 425M abgedeckt waren<sup>32</sup>. Wir hatten uns aufgrund der Zielplattform *YouTube* und mit dem Wissen, dass das Material nicht weiter editiert oder einem Color-Grading unterzogen werden würde, dafür entschieden, in einem komprimierten Format aufzuzeichnen. Um die Speichermenge im überschaubaren Rahmen zu halten, encodierten wir mit H.264, mit 1080 p und 25 fps bei 45 Mbit/s. Das Audio wurde in einem MOV-Container abgelegt und per PCM mit 48 kHz und 24 bit aufgezeichnet. So erhielten wir mit unserem Programm eine gut zu verwaltende Datenmenge in sehr guter Qualität mit Headroom, die in etwa 3,75-mal so hoch ist, wie die maximale akzeptierte durchschnittliche Upload-Bitrate von YouTube, die bei unserem Produktionsformat bei 12 Mbit/s liegt<sup>33</sup>.

In der Steuerung der MAZ konnten wir auf die Netzwerkkompatibilität des Geräts zurückgreifen. Zum einen war eine Steuerung per Software über das Remotenetzwerk möglich – hier gibt es ähnlich zum *ATEM*-Bildmischer eine eigene Oberfläche für alle

---

<sup>32</sup> (Blackmagic Design Pty. Ltd., 2021)

<sup>33</sup> (Google LLC, 2021)

Kontrollen – zum anderen war für unseren Workflow die Integration in die Bildmischersteuerung interessant. So ist es möglich, per Macro-Befehle der *ATEM*-Bedienoberfläche im Bildmischer die Aufzeichnung und Wiedergabe zu steuern, oder automatisiert per Schnittbefehl auf das Bild der MAZ den Start der Wiedergabe auszulösen. So war hier wiederum die Arbeit eines MAZ-Operators hinfällig, da der Start zum Abspielen einer MAZ automatisiert passierte, sobald das Bild des *HyperDeck Studio Mini* auf das Programm geschnitten wurde. Ebenso wurde anschließend der Ton der Aufzeichnung automatisch per Audio-Follow-Video-Funktion ein- und ausgeblendet. Die Einbindung eines *HyperDeck Studio Mini* erfolgte per IP-Verbindung zum *ATEM*-Bildmischer. Insgesamt lassen sich so an einem Bildmischer bis zu vier Hardware-Recorder von *BMD* integrieren und steuern.

#### 4.8 Aufbau der Tontechnik mit binauralem Sendeton

Im Folgenden betrachte ich das Gewerk Ton. Dieses Gewerk muss im Design zunächst als Blackbox angenommen werden, um die besonderen Anforderungen der Umgebung klar zu definieren. Danach gehe ich in Stufen tiefer, in die einzelnen Lösungsansätze einzelner Subsysteme, um diese zu erarbeiten. Nach dem systemhierarchischen Denken empfiehlt sich zunächst die Definition des Zielformats im Bereich Sendeton und die damit umzusetzenden Anforderungen an das Übersystem Ton <sup>34</sup>. Davon ausgehend gehe ich in den kommenden Punkten auf die verschiedenen Subsysteme ein, die zur Umsetzung des Ziels entsprechend entworfen werden. Dies erfolgt in Form einer Abgrenzung der einzelnen Systemkomponenten in der Tontechnik und deren Zusammenspiel.

Das Sendekonzept sah vor, eine besonders intime Stimmung und Nähe zum Künstler aufzubauen. Im Bereich Bild haben wir dies durch Location, Ausleuchtung und Perspektiven der Kameras geschaffen, indem wir die Künstler in ihrer vertrauten Umgebung gezeigt haben, in der sie sich wohlfühlen (Annique Göttler im Konzertsaal der HMDK, in dem sie unzählige Konzerte gespielt hat, zuletzt ihren Abschluss zum Konzertexamen; Tabea Booz in ihrem Lieblingscafé im Stadtteil Stuttgart-West, in dem sie lebt). Dies hat von Anfang an eine lockere, angenehme Atmosphäre erzeugt. Um diese Situation auch im Ton zu transportieren, entschieden wir uns für immersives Audio. Den Zuschauer wollten wir so stärker emotional

---

<sup>34</sup> (Prof. Dr. Reinhard Haberfellner, 2018, S. 40-42)

berühren und näher an das Geschehen bringen. Mit Ausnahme des Konzertsaals der HMDK waren alle anderen Locations akustisch nicht aufbereitet und ungeeignet für die akustische Aufzeichnung von Musik. Durch entsprechende Mikrofonierung und Mischung des binauralen Sendetons konnten wir einen künstlichen Raum erzeugen, die ungenügende Akustik der realen Räumlichkeiten ausblenden und durch einen passenden, synthetischen 3D-Raum mit Hall und Reflektionen ersetzen.

Mit Blick auf die Zielplattformen *YouTube* und *Facebook* entschieden wir uns bewusst für eine binauralisierte Mischung um die Zuschauer mit einem Stereosignal für Kopfhörer zu erreichen. Denn uns erschien die Option, einen Kopfhörer zu nutzen, für die meisten Zuschauer und Zuhörer realistischer als ein Mehrkanal-Lautsprechersystem. Ebenso konnten wir so spontane Zuschauer erreichen, die in Bus, Bahn oder anderweitig mobil unterwegs waren. Zudem hätten wir bei einer spontanen Änderung der Zielplattform davon ausgehen können, dass Stereo kein Problem in der Übertragung darstellen und grundsätzlich von den Plattformen akzeptiert werden würde.

An diesen Ton-bedingten Vorgaben mussten sich die kommenden Subsysteme orientieren, um technisch und künstlerisch eine Binauralisierung zu ermöglichen und sinnvoll zu gestalten. Ebenso musste ein Workflow mit entsprechend einfacher und schneller Bedienbarkeit des Systems entworfen werden, damit der Ton von einer Person würde betreut werden können.

#### 4.8.1 Mikrofonierung der Künstler

Bei der Wahl der passenden Mikrofonierung und der Mikrofone selbst für die Künstler\*innen war in erster Linie eine nahe, direkte Mikrofonierung erforderlich, um den Raum bestmöglich auszublenden. Neben einer guten Abbildung der einzelnen Instrumente und der Stimme war eine hohe Signaltrennung der Schlüssel zu einer flexiblen und sauber klingenden Binauralisierung mit guter Lokalisation und Ortung der einzelnen Quellen in der fertigen Mischung. Für die akustischen Instrumente planten wir daher ausschließlich Druckgradientenempfänger ein. So wurde beispielsweise der Kontrabass mit einem *DPA4099* Mikrofon mit Supernierencharakteristik nahe am Steg abgenommen, um die Stimme von Tabea Booz bestmöglich auszublenden. Bei elektronischen Instrumenten wie E-Bass und Drum-Computer haben wir auf einen akustischen Verstärker oder Monitorlautsprecher

verzichtet und direkt per DI-Ausgang oder mit einer Symmetrierung per DI-Box abgenommen, um zusätzliche Schall- und potenzielle Störquellen im Sendeton zu vermeiden. Für das Monitoring bedeutete dies, dass für den Livestream mit Tabea Booz ein In-Ear-Monitoring benötigt wurde.

Zur Übertragung der Stimme und des Gesangs probierten wir verschiedene Möglichkeiten aus: sowohl Kopfbügelmikrofone/Headsets mit verschiedenen Richtcharakteristiken als auch Handmikrofone. Unsere Wahl fiel aufgrund klanglicher Aspekte auf Handheld-Mikrofone, da mit einer Hypernierenrichtcharakteristik mit einem Kleinmembrankondensatormikrofon die besten Ergebnisse in den Bereichen Klang und Signaltrennung zu anderen Quellen im Raum geschaffen werden können. Damit hatten wir optimale Voraussetzungen und eine hohe Flexibilität zur Binauralisierung.

Für eine gute Bewegungsfreiheit auf der Bühne und im Talk haben wir Funkmikrofone verwendet. Zu bedenken galt, dass das System – da wir mit nur einem Toningenieur auskommen mussten – über unser Netzwerk konfigurierbar und überwachbar sein sollte, um eventuelle Probleme wie Signalstörungen, Batterieanzeige, Signalstärke und Pegel von der Position des Toningenieurs am Arbeitsplatz in der Regie überwachbar zu halten. Ebenso sollte das System eine analoge Übertragungsstrecke nutzen, um zusätzliche Latenz im System zu vermeiden. Zum einen erzeugen wir bereits durch die Binauralisierung eine Latenz. Zum anderen brachten wir In-Ear-Monitoring zum Einsatz, wobei jegliche zusätzliche Latenz neben der Wandlung am Audiosystem zu vermeiden ist. Unser *MIDAS M32C* hat inklusive der AD- und DA-Wandlung 4 ms Latenz an den Stageboxen erzeugt. Das uns zur Verfügung stehende digitale Funksystem *ULX und Axxient ADX* von *Shure* gilt als System mit geringsten Latenzen und hätte 2 bis 3 ms zusätzliche Latenz erzeugt<sup>35</sup>. Das ist für den Betrieb im In-Ear-Monitoring grenzwertig, zumal hier insgesamt knapp 6 bis 7 ms Latenz entstanden wären. Wenn nun obendrein digitale Effektpedale (Amp-Simulationen durch die Backline des Musikers) hinzukommen, kann es schnell zu einer zu hohen Latenz kommen, die als störend empfunden wird, wenn das Monitorsignal trocken oder mit einfachem Hall auf die Ohren der Musiker gegeben wird<sup>36</sup>. Ich habe mich unter den mir zur Verfügung stehenden Systemen und zu beachtenden Anforderungen für die *Sennheiser ew500g4* aus dem Jahr 2020 mit *Neumann*

---

<sup>35</sup> (Kaminski, 2020)

<sup>36</sup> (Georgios Marentakis, 2012)

KK205 *Handsendern* entschieden. Die Kleinmembrankondensatorkapseln in Hypnierencharakteristik mit einem gleichmäßigen Aufnahmeverhalten, geringem Übersprechungseffekt und guter Dämpfung von Umgebungsgeräuschen erleichterten die Binauralisierung. Das System baut auf eine moderne, aber analoge Funkstrecke auf, die durch einen rein analogen Übertragungsweg keine zusätzlichen Latenzen erzeugt. Allerdings ließen sich alle wichtigen Parameter per Netzwerk steuern und überwachen und auch die Frequenzbereichsüberwachung war ebenfalls über Netzwerk einsehbar und kontrollierbar.

Abschließend ist mit dem Hygienekonzept eine weitere Besonderheit zu nennen. Wir teilten keine Mikrofone zwischen Künstler\*innen und Moderation und statteten sie jeweils mit einem Mikrofon für den Talk bzw. den Gesang aus. Der Sender wurde vor und nach der Produktion gereinigt und desinfiziert, ebenso wurden der Schaumstoff des Popschutzes desinfiziert und ausgelüftet. Auch deshalb war eine Überwachung des Senders per Netzwerk erforderlich und eine wichtige Anforderung des Systems.

#### 4.8.2 Wahl des digitalen Audiosystems – Blackbox Betrachtung

Um zunächst zu bestimmen, welche Form des digitalen Audiosystems sich am besten für das System eignet, müssen die Anforderungen klar definiert sein. Hier kann im Sinne des Systemdenkens von einer Blackbox bzw. einer Greybox ausgegangen werden. Um die Komplexität des Systems zu vereinfachen, werden die Details des Teilsystem „Audiosystem“ zunächst ausgeblendet, die äußeren Umstände analysiert und genauer betrachtet. Da Teilstrukturen und Anforderungen an das Audiosystem bereits bekannt sind, wird hier von einer Greybox gesprochen<sup>37</sup>. Eine grobe Aufteilung der benötigten Positionen der Inputs und Outputs ist ebenfalls schon bekannt. An der Bühne und in der Regie müssen Eingänge und Ausgänge zum System vorhanden sein, ebenso muss die DSP-Einheit zur Audibearbeitung und für das Routing unabhängig der Positionen in der Regie und der Bühne platzierbar sein. Dies ist wichtig, um die gewünschte Flexibilität zu halten und sich auf verschiedenste Räumlichkeit schnell anpassen zu können.

---

<sup>37</sup> (Prof. Dr. Reinhard Haberfellner, 2018, S. 32-33)

Als nächstes wird bei der Wahl des richtigen Audiosystems die grundlegende Art bestimmt. Digitale Audiosysteme lassen sich mit dem OSI-Schichtenmodell einteilen und anhand dessen auch ihre Flexibilität bzw. Vor- und Nachteile betrachten. Im Layer 1 werden Systemkomponenten mit Hilfe von physikalischen Netzwerkkomponenten direkt miteinander verbunden. In diesem Fall wird auf Hardware wie Cat.5-Kabel zurückgegriffen und eine Steckverbindung aus dem Netzwerkstandard genutzt. Hier kommen Verbindungen wie RJ45 oder die robustere Version EtherCON mit Schutztüle und zusätzlicher Schirmung zum Einsatz. Vertreter dieser Systeme sind Audionetzwerke wie AES50, SuperMAC, HyperMAC, Rocknet300, Optocore oder auch MediorNet, die in der Regel in einer Kettenanordnung, dem sog. „Daisy-chain-Verfahren“ aufgebaut sind. Vorteil dieses Aufbaus ist eine einfache Struktur, die einen schnellen Auf- und Abbau ermöglicht. Außerdem haben diese Systeme eine meist sehr geringe Latenz und sind durch ihre klare Struktur und proprietäre Technik weniger anfällig für fehlerhafte Aufbauten und Fehlbedienung. Als statischer Netzwerkaufbau kann der kettenförmige Aufbau auch ein Nachteil sein, da er keine dynamischen Änderungen zulässt: Komponenten werden fest verbunden und können im laufenden Betrieb nicht getauscht oder verändert werden. Ebenso sind alle Komponenten an einen Hersteller des proprietären Systems gebunden <sup>38</sup>.

Layer-2-Systeme nutzen nicht nur die physikalische Verbindung, sondern auch Ethernet-Komponenten, um Netzwerkstrukturen zu ermöglichen, indem zum Beispiel Switches für sternförmige Aufbauten zur Datenverteilung genutzt werden können. Ein Vorteil der Layer-2-Struktur der auch „Audio-Over-Ethernet“ genannten Systeme ist ein möglicher flexiblerer Aufbau und die dadurch verbesserte Anpassungsfähigkeit an die Produktionsstrukturen, so dass ein sternförmiger oder baumartiger Aufbau möglich ist. Hier kann auf gängige Netzwerkkomponenten zur Verteilung zurückgegriffen werden ebenso wie auf Switches aus dem IT-Bereich. Im Havariefall wäre ein schneller Austausch möglich und durch den Einsatz von handelsüblichen IT Produkten wird es dank der Konkurrenz auf dem Markt kostengünstiger <sup>39</sup>. Wichtige Vertreter der Layer-2-Netzwerke sind CobraNet, EtherSound und SoundGrid.

---

<sup>38</sup> (Hildebrand, 2018, S. 13)

<sup>39</sup> (Hildebrand, 2018, S. 15)

Für die Planung dieses Systems habe ich außerdem Systeme mit Layer-3-Struktur betrachtet, z.B. das Audio-Over-IP, wobei auf Basis des Internet-Protocol ein System genutzt wird, das auf bestehende IP-Netzwerke zurückgreift und sich ebenso Ethernet-Struktur und Komponenten mit anderen IP-Systemen teilen kann. Es wird noch flexibler in Auf- und Abbau, beliebige Netzwerkstrukturen wie Stern, Baum oder Netze werden möglich. Ebenso können Ketten- und Ringstrukturen etabliert werden. Die Größe des Netzwerks ist ebenfalls nicht beschränkt, so lange Datenraten und Kapazitäten vorhanden sind und von den Audiogeräten unterstützt werden. Nachteil ist hier ein hoher Aufwand in Planung, Aufbau und Erstinbetriebnahme. Darüber hinaus sollte das Netzwerk vor der Nutzung getestet werden, da ein stabiler Betrieb von Hersteller nicht gewährleistet werden kann. Dies liegt darin begründet, dass für den Netzwerkaufbau Geräte von Drittanbietern genutzt werden können, die unter Umständen nicht zertifiziert sind. Hinzu kommt, dass die Latenz im System größer ist als in Layer 1- und 2-Systemen. Zusätzlich ist die Latenz variabel und hängt von Größe und Anzahl der verbundenen Geräte ab und kann erst während des Betriebs gemessen werden. Wichtige Vertreter sind *Audinate Dante*, *Livewire* oder *Q-LAN* <sup>40</sup>.

Im benötigten System für unsere *Stayin' Live*-Produktion waren die Positionen fest definiert: Eine Stagebox wurde an der Bühne benötigt, eine in der Regie. Zur Steuerung und für die Audibearbeitung des Systems war ein Mischpultsystem mit passender DSP notwendig, welches sich per IP-Remote steuern ließ. Hinsichtlich des Audiosystems war daher eine flexible und skalierbare Steuerung nötig, wobei die Audiostruktur statisch sein und sich die Dimensionierung nur an einer maximalen Kanalzahl für Eingänge und Ausgänge orientieren sollte. Entgegen meiner ursprünglichen Überlegung, ein Layer-3-System zu verwenden, entschied ich mich mit dem *SuperMAC Protocol* von *Sony* für ein Layer-1-System, das kostengünstiger war und alle Voraussetzungen erfüllte: Wir benötigten ein System, das 48 kHz Samplerate mit 24 bit für einen sendefähigen Ton unterstützt; außerdem ein System mit minimaler Latenz für ein In-Ear-Monitoring und Flexibilität im Umgang mit der rechenintensiven Binauralisierung, die zusätzlich eine hohe Latenz in das System brachte. *SuperMAC*, bzw. *AES50* wurde ursprünglich von *Sony* entwickelt und unterstützt DSD wie auch das von uns genutzte PCM-Verfahren und ermöglicht unter Nutzung von kostengünstigen Cat5-

---

<sup>40</sup> (Hildebrand, 2018, S. 17)

, Cat6- oder Cat7-Kabeln Übertragungslängen von bis zu 100 m. Mit Latenzen von 50 bis 80 µs können 48 Kanäle bidirektional in 48 kHz und 24 bit übertragen werden <sup>41</sup>.

Das hier zur Verfügung stehende *MIDAS-M32C*-Audiosystem deckte alle unsere Anforderungen. Neben den genannten technischen Eckdaten des AES50-Netzwerks standen uns 32 Mischkanäle und 25 Busse zur Verfügung. Viel wichtiger war jedoch, dass die *M32 Compact* – eine DSP-Einheit mit nur einer Höheneinheit und 19“-Rackformat – ein sehr platzsparendes System ist. Die Steuerung erfolgte über einen Ethernet-Port per IP-Remote-Software an PC, Mac oder Tablet. Zusätzlich konnte ein Touchscreen für die Software genutzt werden. Für die Mischung und Ergonomie im Arbeiten konnten wir eine Mischeinheit mit motorisierten Fadern per *Midi Over Ethernet Protocol* an den Mischer anbinden. Hierbei war die Fader-Unit als Remote-Controller über eine IP-Verbindung direkt mit der DSP-Einheit verbunden und nicht über einen PC angebunden. Durch diese unabhängige Verbindung hätte das *M32C*-System auch ausschließlich per Fader-Unit oder Remote Software gesteuert. Diese Flexibilität konnte uns kein anderes zur Verfügung stehendes System bieten. Bei einem zuvor angedachten *Yamaha*-System hätten wir stets mindestens ein Mischpult (*Yamaha QL1* oder *CL3*) im Touring-Case nutzen müssen. Dieses zu transportieren und auf- und abzubauen wäre ein unnötiger Kraft- und Zeitaufwand gewesen. Obendrein wäre in zwei Locations ein Platzproblem aufgetreten und unser angedachter Regie-Aufbau hätte nicht den Anforderungen des Hygienekonzepts entsprechen können. Ein *Yamaha*-System mit *Audinate Dante* wäre überdies deutlich teurer gewesen:

Systemkomponente	Produkt	Preise inklusive MwSt.
Midas M32 System	Verwendete Komponenten	
DSP-Einheit	MIDAS M32C	609 €
Fader Unit – Remote Controller	Behringer X-Touch	398 €
Touchscreen	ProLite T2435MSC-B2	314 €
Stagebox (Bühne)	MIDAS DL251	3490 €
Stagebox (Regie)	MIDAS DL16	749 €
USB Audio Interface	Inklusive (M32C)	0 €
	Summe:	<b>5560 €</b>
Yamaha System	Alternative Komponenten	
DSP-Einheit und Mischpult	Yamaha QL1	7890 €
Fader Unit – Remote Controller	(Nicht nötig, durch Mischpult)	0 €
Touchscreen	(Nicht nötig, durch Mischpult)	0 €
Stagebox (Bühne)	Yamaha RIO 1608-D2	4480 €

<sup>41</sup> (Sony Pro - Audio Lab, 2004)

Stagebox (Regie)	Yamaha RIO 1608-D2	4480 €
USB Audio Interface	RME Digiface Dante	1299 €
	Summe:	<b>18149 €</b>

Tabelle 4 - Preisvergleich - Neupreise abgerufen am 28.01.2021 von Thomann.de

In der Rechnung wurden die verwendeten Kabel ausgespart, da diese für beide Systeme dieselben gewesen wären und im Testaufbau kompatibel waren. Als Interface im *Yamaha*-System könnte eine *Audinate Dante-Virtual-Soundcard* nicht genutzt werden, da der Rechner mit der Binauralisierung einen hohen Rechenaufwand zu bewerkstelligen hat. Die zusätzliche Rechenleistung zur Wandlung müsste außerhalb des Rechners in einem Hardware-Interface geschehen. Da auch im *M32*-System ein Hardware-Interface zum Einsatz kommt, nutzte ich bei dem getesteten *Yamaha*-System ebenfalls ein Hardware-Interface von *RME*. Der enorme Preissprung von 12589 € rechtfertigt eine mögliche redundante Verkabelung im *Yamaha*-System aus meiner Sicht nicht, da im Zweifel sogar zwei vollwertige *M32*-Systeme gemietet oder gekauft werden könnten, um eine vollständige Hardware-Redundanz zu rechtfertigen. Ebenso ist keine flexible und skalierbare Steuerung möglich, da mit dem *Yamaha QL1* ein statisches Mischpult zum Einsatz kommt; das günstigste Mischpult, das zu unseren Systemanforderungen passt und verwendbar wäre.

Somit war für mich in der Planung das Audiosystem gesetzt: Zum Einsatz kam das *M32C* mit beschriebenen Komponenten und bildete die Kerninfrastruktur für unser Audio- und Kommunikationssystem in der *Stayin' Live* Produktion.

#### 4.8.3 Intercom und Abhörsituation der Produktion

Im Rundfunk- und Fernsehbetrieb ist ein Kommunikationssystem zur Umsetzung der eigentlichen Produktion unersetzlich. Hierfür wird meist ein zusätzliches System für die Übertragung von Sprachkommandos aufgebaut. Dieses System hat keine primäre Funktion, ist losgelöst vom Sendeton und steht in keiner direkten Verbindung zum erzeugten Produkt in Bild und Ton. Es dient rein zur Kommunikation aller Produktionsbeteiligten. Das Zentrum bildet eine Basisstation, die von Inspizienten bzw. Regisseuren genutzt wird. Von dieser Station gehen kabelgebundene oder drahtlose Hör-Sprech-Stellen zu allen Gewerken wie Kameraleute, Tonregie, Bildregie, Setrunner oder anderen Beteiligten der Produktion. Es gibt einfache analoge Systeme – auch Partyline genannt – wobei alle Teilnehmer zu allen

Teilnehmern und umgekehrt sprechen. Durch eine Simplex-Verbindung kann stets nur eine Person sprechen, andere Teilnehmer müssen warten und können nicht gleichzeitig sprechen, solange die Leitung genutzt wird. Moderne Systeme, auch digitale Duplex-Systeme, lösen dieses Problem, wie zum Beispiel digitale Systeme von *ClearCom*, das sog. *FreeSpeak*. Diese Systeme unterstützen hierarchische Strukturen und mehrere Gesprächsgruppen, in die jeder Teilnehmer mit entsprechender Rechehierarchie sprechen kann <sup>42</sup>.

Bei der Systemgestaltung wurde zügig beschlossen, dass auf ein solch dediziertes Intercom-System verzichtet werden kann. Die Systemkapazitäten, die das Audiosystem M32 mit sich bringt, war ausreichend für alle Funktionen und Anforderungen unserer Produktion ein.

Jeder an der Produktion Beteiligte sollte über die Möglichkeit des Sendens und Empfangens von Kommandos verfügen, wobei dennoch Hierarchien möglich sein sollten. Die Bildregie sollte beispielsweise von jedem Intercom-Teilnehmer (Kammerleute, Toningenieur, MAZ, Moderatorin) hörbar sein. Ein Kameramann sollte bei seinen Kommandos aber nur von der Bildregie gehört werden, um unnötige Kommunikation an Toningenieur und MAZ zu vermeiden. Ebenso sollten Anweisungen des Toningenieurs bei Bedarf nur von der Bildregie, MAZ und Moderatorin zu hören sein. Um diese Struktur umzusetzen, erhielt jeder Teilnehmer einen eigenen Eingang am System, der individuell auf die Busse matriziert werden konnte. Ebenso bekam jeder Teilnehmer je nach Kopfhörersituation in Stereo- oder Mono-Konfiguration einen eigenen Buss zum Empfangen von Kommandos (die Intercom-Teilnehmer in der Regie in Stereo; Kameraleute und Moderatorin in Mono). Um die Kanäle schaltbar zu machen, setzten wir Hardware-Taster nach den Mikrofonen als Fuß- oder Handtaster ein. Die verwendeten Taster schalten das Audiosignal ohne Knacken und geben Phantomspeisung, falls benötigt, durchgehend an die Mikrofone weiter. Hier hatte ich *Rolls MM 11 Pro* verbaut, robuste Schalter aus Metall, die die Anforderungen vollumfänglich erfüllten.

Beim Aufbau der Signalkette wurde das analog geschaltete und getastete Eingangssignal auf einen Eingang der DSP-Einheit geroutet. Dort wurde es per Hochpass bei 100 Hz mit 12 db/oct. von Störfrequenzen im tieffrequenten Bereich befreit. Danach stand ein Gate, das unnötige Signalübertragung vermied. Es folgte ein starker Kompressor mit einer Kompressionsverhältnis von 5 bei einem Threshold von -40 dBfs. Hierdurch konnte eine

---

<sup>42</sup> (Grzesinski, 2020)

Sprachverständlichkeit, besonders bei den Kameralenten, bei lauter und auch bei sehr leisen Kommandos gewährleistet werden. Der verlorene Pegel wurde mit 15 dB Make-Up-Gain aufgeholt. Nach einer Entzerrung mit einem parametrischen Equalizer zur besseren Sprachverständlichkeit ging das Signal auf die Busmatrix und wurde dort an entsprechende Empfänger im Post-Fader-Modus geroutet. Dies hatte zur Folge, dass bei Pegelschwankungen das Eingangssignal über den Fader für alle Empfänger gleichzeitig angepasst werden konnte. Im Ausgang konnte das Signal flexibel geroutet, Talk-Wege an die Künstler\*innen für In-Ear-Monitoring eingeschliffen und Anweisungen des Toningenieurs oder der Bildregie an die Künstler\*innen oder die Moderatorin weitergegeben werden. Die Gesprächsgruppen bewegten sich in der Produktion nur zwischen zwei Zuständen: Beim Rounting-Setup für die Probe konnten bei Bedarf Regie, Toningenieur und Moderation mit den Künstler\*innen kommunizieren. Das zweite Setup hingegen beschränkte die Kommunikation nur auf die Regie, um Künstler\*innen und Moderation in der Sendung nicht zu stören. Hier wurde die Dauerschaltung auf deren Busse deaktiviert, so dass nur die Regie auf Wunsch zu Moderation und Künstler\*innen sprechen konnte. Diese Setups wurden in Teilmatrizen, sog. Snippets, gespeichert und waren per Knopfdruck sofort schaltbar. Die Snippets konnten per Remote-Control-Software oder Fader-Unit von Toningenieur oder Regie abgerufen werden.

Dieses Setup hatte für uns neben der beschriebenen Flexibilität und des Wegfalls zusätzliche Equipments einen entscheidenden Vorteil: Die Kommunikationsbarriere zwischen Künstler\*innen und Produktion war nicht mehr vorhanden, so dass es dem Team möglich war, schnell und unkompliziert Anweisungen und Hinweise mit den Künstler\*innen über deren In-Ear-Monitoring auszutauschen, da alle auf die gleichen Ressourcen des Audiosystems Zugriff hatten. Gerade bei einem kleinen Team wie unserem mit vielen Doppelfunktionen war dies sehr hilfreich und beschleunigte das Arbeiten.

Gelöst war hiermit auch die Abhörsituation für die Regie, da die Räumlichkeiten oftmals direkt nebeneinander lagen, Regie und Bühne nur durch eine Wand getrennt waren und somit auf die Nutzung von Abhörlautsprechern verzichtet werden musste. Stattdessen erhielten alle Beteiligten in der Regie die binauralisierte Mischung auf ihre Kopfhörer. Diese konnte über die Busmatrize mit den Kommandos im jeweils benötigten Verhältnis gemischt werden. Die Regie wünschte Sendeton und Kommandos in gleicher Lautstärke, der Toningenieur erhielt

auf seinem Kopfhörer zur Mischung den Sendeton 10 db lauter als die Kommandos. Jeder konnte also seine Abhörsituation frei gestalten und mischen.

## 4.9 IP-basierte Steuerung der Bild- und Tontechnik

Die Bedienung des Systems beanspruchte ein hohes Maß an Flexibilität bzw. Systemagilität: Infolge der Pandemie und der damit einhergehenden Personenzahlbeschränkungen pro Raum mussten wir unser Personal stark reduzieren. Nachdem die Bereiche so vereinfacht, automatisiert und zusammengelegt worden waren, musste das Gesamtsystem kontrollierbar gemacht werden, ohne auf die erreichte Flexibilität verzichten zu müssen. Alle Arbeitsplätze mussten in ihrer Größe skalierbar sein, um die gesamte Regie mit drei Personen in allen geplanten Räumlichkeiten unterbringen zu können. Bei drei Personen gab es personell keinen Spielraum für Ausfälle und schnellen Ersatz. Hier bestand die Anforderung an das Netzwerk, dass jede/s Gewerk/Position per Desktopoberfläche von jedem Arbeitsplatz im Notfall kontrolliert und übernommen werden könnte, um zumindest im Notbetrieb sendefähig zu bleiben. Dies bedeutete für uns notwendigerweise das Verlassen des systemhierarchischen Denkens und die Einführung einer Kontrollebene ohne Stufen im Bereich der Steuerung <sup>43</sup>. Obwohl im technischen Aufbau, Signalverlauf und den Arbeitsbereichen klare Systemstrukturen mit Über- und Untersystemen vorhanden sind, ist der Bereich der Steuerung herausgelöst. Hier erfolgt alles über IP-basierte Hardware- und Softwarelösungen und bewegt sich auf einer Ebene, auf der alle Teilnehmer dieselben Zugriffsrechte besitzen. Um diese Agilität zu gewährleisten und eine flexible Anbindung zu schaffen, ist ein dynamisches Netzwerk notwendig und kann als dynamisch, kompliziertes System betrachtet werden <sup>44</sup>. Die Netzwerkgeräte der Ton- und Bildtechnik sind statisch vorhanden, jedoch sind die Teilnehmer am Netzwerk dynamisch und ändern sich in Anzahl und über den zeitlichen Verlauf der Produktion. PCs, Laptops und Tablets melden sich dynamisch im Netzwerk an und ab und greifen somit beliebig auf das System zur Steuerung zu. Im Entwurf und der Ausführung des Systems äußert sich dies über statische IP-Adressierung der Ton- und Bildtechnik, da sich diese nicht verändert, sowie über dynamische IP-Adressierung der

---

<sup>43</sup> (Prof. Dr. Reinhard Haberfellner, 2018, S. 40-42)

<sup>44</sup> (Prof. Dr. Reinhard Haberfellner, 2018, S. 35)

Steuerungsgeräte der Anwender mit ihrem Endgerät per DHCP. Hinzu kommen Hardwaresteuerung für Bild und Ton wie Fader-Units und eine Bildmischeroberfläche, die ebenso per DHCP am Netzwerk teilnehmen.

Für die Umsetzung wurde ein Gigabit-Netzwerk mit passendem Router und Netzwerk-Switches eingeplant. Im Router war ein Access-Point verbaut, der einen drahtlosen Zugriff per 2,4 GHz- und 5 GHz-WLAN möglich macht. Alle IP-fähigen Geräte wurden mit diesem Netzwerk verbunden und auf Nutzer-Seite mit DHCP versorgt. Die gesamte Steuerung des Systems lief somit über einen Router in einem LAN ab und erzeugte die gewünschte Agilität des Netzwerkes. In der Regie konnte die Steuerung dynamisch erfolgen, so dass zum Soundcheck und zur Einrichtung per WLAN mobile Geräte wie Tablets oder Laptops mit zur Bühne genommen werden konnten, um vor Ort am Monitoring Einstellungen gemeinsam mit den Künstler\*innen vorzunehmen bzw. Kameraeinstellungen und Bildmischereinstellungen nebeneinander zu vergleichen und zu ändern.

An dieser Stelle möchte ich auf eine Gefahr hinweisen, die dieser Aufbau mit sich brachte: Die gesamte Steuerung des Systems hing an einem Router, an einem Hardwaregerät. Fällt dieser aus, wäre davon auch die Steuerung einiger Geräte betroffen. Anmeldungen neuer Teilnehmer im Netzwerk per DHCP wäre in diesem Fall ebenfalls nicht möglich. Deshalb stand ein zweiter Router mit gespiegelten Einstellungen bereit und hätte im Ernstfall für den ersten Router übernehmen können. Somit hätte ein Bottleneck der Produktion vermieden werden können, der im Ernstfall zum Sendeabbruch geführt hätte. Die Übernahme des zweiten Routers kann per *Virtual Router Redundancy Protocol (VRRP)* erfolgen, wobei die Gateway-Adresse bestehen bleibt und sich für die Netzwerkteilnehmer nichts ändert <sup>45</sup>. Diese Redundanz im Netzwerk ist von entscheidender Wichtigkeit. Die konsequente Zusammenlegung des Steuernetzwerkes für alle Gewerke in der Produktion schafft zwar die gewünschte Flexibilität und Agilität im System, gefährdet aber bei einem Ausfall die gesamte Sendung. Dies muss den Verantwortlichen in der Systemplanung stets bewusst sein. Eine entsprechende Redundanz an Hardware ist daher zwingend erforderlich. Daraus lässt sich schlussfolgern: Je weniger Systemabgrenzung und Unterteilung in Untersysteme erfolgt,

---

<sup>45</sup> (Markus Nispel, 2010)

desto anfälliger wird das Gesamtsystem, da sich Probleme und Ausfälle über einen größeren Bereich der Produktion erstrecken und negativ auswirken können.

## 5. Aufbau des Audiosystems mit binauralisierten Sendeton

### 5.1 Systemanforderungen der Live-Binauralisierung

Im Bereich der Postproduktion gibt es bereits viele Ansätze von nichtlinearen Medien um eine Binauralisierung für gut funktionierende Workflows für immersives Audio vorzunehmen. Dies gilt auch für den Teilbereich der Binauralisierung von 3D-Mischungen für Kopfhörer. Das gilt allerdings nicht für Live-Binauralisierung im Rundfunk und Fernsehen. In der Postproduktion kann per Software im Rechner eine Mischung im 3D-Raum mit Mehrkanalbussen und Tools zur Binauralisierung in einer DAW vorgenommen werden. Die rechenintensiven Plugins benötigen Zeit bzw. erzeugen hohe Latenzen und beanspruchen die CPU. Wenn hier ein Rechenprozess mehrere Millisekunden oder sogar ganze Frames benötigt, ist das in der Postproduktion kein Problem. Anders verhält es sich im Umfeld einer Fernsehsendung bzw. eines Livestreams. Hier müssen Bild und Ton stabil und synchron laufen. Der Ton darf in der Latenz nicht variieren und muss zuverlässig über den Zeitraum der Produktion ohne Signalstörungen oder gar Ausfälle zur Verfügung stehen. Neben einem sauberen Signal sind Zuverlässigkeit und Betriebssicherheit die wichtigsten Parameter, um einen sauberen Sendungsablauf und stressfreies Arbeiten in der Produktion zu gewährleisten.

Neben einem sicheren Workflow der Binauralisierung sollte ebenso eine gute Bedienbarkeit möglich sein. Im Livebetrieb muss ein schneller Zugriff auf alle Parameter möglich sein; dies gilt auch für das Positionieren der Quelle im 3D-Raum, ebenso sollen Eingriffe in Klang, Entzerrung und Effekte wie Hall unkompliziert möglich sein, da diese als gestalterische Mittel in der Musikproduktion von besonderer Wichtigkeit sind. Hierbei funktioniert ein direkter Zugriff auf Parameter nur, wenn Hardwarekomponenten zur Steuerung vorhanden sind oder die Arbeit automatisiert werden kann, um Arbeitsschritte zu vereinfachen und auf notwendige, gestalterische Elemente zu beschränken. Nur so kann ein schnelles Reagieren auf unvorhergesehene Situation im Livebetrieb realisiert werden, da sich die Konzentration des Toningenieurs während der Sendung auf die technisch und gestalterisch wichtigen Komponenten der Mischung richtet.

Es wurde schnell klar, dass weder ein reiner Mischpultbetrieb noch eine komplette Mischung „In the Box“, also im Rechner, all diese Anforderungen umsetzen könnte. Deshalb bot sich ein Hybridsystem an. Zum einen ein digitales Mischpult mit diversen Routing-Möglichkeiten, für Sendeton, Monitoring und Intercom. Mit geringen Latenzen für Monitormischungen per In-Ear für die Künstler samt haptischem Feedback und direktem Zugriff auf Fader und Regler bzw. passenden Automationen, um schnell auf alle Situationen reagieren zu können. Und zum anderen ein flexibles, intuitives System im Rechner, mit dem eine Binauralisierung der einzelnen Kanäle möglich wird. Voraussetzung hierfür ist die Option für ein flexibles 3D-Panning der Quellen am Rechner. Zusätzlich muss ein 3D-Raum mit passendem Hall berechnet werden, um die dreidimensionale Räumlichkeit und Wahrnehmung zu unterstreichen.

Unter all diesen Gesichtspunkten habe ich die nachstehende Systemstruktur für den Workflow und einen passenden Signalfluss im Mischpult entwickelt, und die Binauralisierung durch einen externen Rechner in dieses System integriert.

## 5.2 Signalstruktur im Hybrid-Audiosystem

Das nun entstandene, komplexe Hybrid-Audiosystem lässt sich zu genauerer und besserer Betrachtung in verschiedene Untersysteme unterteilen. So kann im Sinne des Systemdenkens leichter auf Probleme eingegangen werden. Dadurch können Lösungen und alternative Herangehensweisen überschaubar gestaltet werden, ohne dabei andere Systeme zu verändern oder ebenfalls neu planen zu müssen. Obwohl hier die Systeme auf die gleiche Hardware zurückgreifen, ist eine Trennung der Bereiche in Sendeton mit Binauralisierung, In-Ear-Monitoring und Produktionskommunikation ratsam. Hier kann von „System von Systemen“ (SvS) gesprochen werden, da die Systeme zusammenhängen, aber nicht vollständig voneinander abhängig sind. Jedes der drei beschriebenen Systeme <sup>46</sup> greift auf ein übergeordnetes System „Tontechnik“ zurück, verfolgt aber einen ganz eigenen Zweck mit jeweils anderem Ziel <sup>47</sup>.

---

<sup>46</sup> Siehe Kapitel 5.2 Signalstruktur im Hybrid-Audiosystem inklusive 5.2.1, 5.2.2, 5.2.3

<sup>47</sup> (Prof. Dr. Reinhard Haberfellner, 2018, S. 31)

### 5.2.1 Signalfluss und Signalgestaltung des Sendetons

Beginnend am Mikrofon werden die Signale abgenommen und bis zur Stagebox geführt. Dies erfolgt per symmetrischem Signalweg über eine XLR-Steckverbindung. An der Stagebox werden die Signale analog verstärkt und mittels Analog-Digital-Wandler (AD-Wandler) digitalisiert und per AES50-2011 dem Audionetzwerk und damit dem Audiosystem zur Verfügung gestellt. Zusätzlich können an der Stagebox die benötigten Mikrofone mit 48 V Betriebsspannung versorgt werden. Unsere AD-Wandler hatten einen ausreichenden Dynamikumfang von 107 dB mit einer maximalen Vorverstärkung von +45 dB, was für unsere Produktion einen ausreichenden Rahmen darstellte <sup>48</sup>. Von den Stageboxen, eine an der Bühne und eine in der Regie, liefen die Signale per AES50 über eine Cat5e-Leitung, wobei eine maximale Leitungslänge von 100 m möglich war. Genutzt wurden 75 m, um bei elektromagnetischen Störungen bzw. Einstreuungen und einem eventuell beschädigtem Leitungsschirm genügend Headroom in Reserve zu haben. Da die AES50-Leitung sternförmig zur DSP-Einheit M32C lief, hatte jede Stagebox ihren eigenen Anschluss. Um Systemengpässe bzw. Störanfälligkeit zu minimieren, verzichteten wir auf einen Daisy-Chain-Aufbau <sup>49</sup>.

An der DSP-Einheit anliegend wurden die Inputs entsprechend der Untersysteme gesplittet. Für den Sendeton liefen die Signale von dort auf das USB-Interface und wurden an den Rechner weitergeleitet. Im Rechner erfolgte die Übernahme der Signale per dediziertem ASIO-Treiber für das M32C-Interface. Als stabile Puffergröße haben sich in dieser Produktionsumgebung nach einigen Tests 512 Samples als zuverlässig erwiesen. Nach der Binauralisierung lag jede Quelle als binauralisiertes Stereosignal an und wurde per Interface zurück in das Mischpult gegeben. Von der Bühne kamen in unserer Produktion acht Inputs, die zur Binauralisierung in den Rechner geleitet wurden. Daraus folgten 16 Kanäle, sprich acht Stereo-Quellen zurück in das Mischpult. Diese wurden im Mischpult auf 16 Kanäle aufgelegt. Jeweils zwei Kanäle wurden anschließend verlinkt, um diese wie einen Stereokanal zu betrachten und zu bearbeiten. Hier konnte nun im gewohnten Umfeld gemischt und Balancen per Fader gefahren werden. Die binauralisierten Quellen wurden auf einen Programm-Bus geroutet und dort summiert. Jener Bus konnte dort zusätzlich dynamisch aufbereitet werden, um die Lautstärke auf ein sendefähiges Niveau mit beschränktem Dynamikumfang zu bringen.

---

<sup>48</sup> (MUSIC Group IP Lt., 2015)

<sup>49</sup> Siehe Kapitel 4.8.2 Wahl des digitalen Audiosystems – Blackbox Betrachtung

Ebenso wurde ein Limiter eingeschliffen, um im Grenzbereich per „TruePeak-Analyse“ Spitzen einzufangen und ein Zerren des Programmsignals zu vermeiden.

Für die Signalgestaltung des Sendetons in der Übertragung habe ich mich in der Planung an der Größe der „Loudness Units“ orientiert. Definiert durch die ITU bildet sie in der Sendeumgebung des Hörfunks, im Fernsehen und auch auf den digitalen Plattformen wie YouTube, Spotify, Amazon Music und iTunes eine feste Größe, um die empfundene Lautstärke von Programmen aufeinander abzustimmen. Angestrebt wurden eine Sendelautstärke von -23 LUFS und eine Signalspitze von -1 dBTP. Da wir *YouTube* als primäre Streaming-Plattform festgelegt hatten, liegt es nahe, sich bzgl. der Ziellautheit an den Normen von *YouTube* orientieren. Zwar hat *YouTube* diesbezüglich keine genauen Angaben veröffentlicht, aufgrund von Test-Uploads vermutet die Mastering-Community allerdings, dass sich die integrierte Lautheit bei -14 LUFS bewegt. Ein anderes Beispiel: Spotify erklärt auf seiner Website, dass auch eigene Algorithmen zur Analyse der Lautheit verwendet werden, die höchstens als „ähnlich“ zur Größe „Loudness Unit“ der ITU bezeichnet werden kann und nennt diese Größe „ReplayGain“<sup>50</sup>. Diese Größen sind auf den von uns genutzten Plattformen nicht weiter technisch begründet oder detaillierter beschrieben, so dass ich mich während der Sendungsplanung bewusst gegen diesen wagen Wert von -14 LUFS entschied. Da wir in unseren Sendungen sowohl klassische Musik als auch Sprechkunst übertragen haben, beläuft sich der gewünschte Dynamikumfang auf mindestens 30 dB und ist damit so groß, dass wir ohne starke Kompression einen Zielwert von -14 LUFS nicht hätten bewerkstelligen können. Hinsichtlich der Binauralisierung und einer sauberen räumlichen Darstellung war es uns wichtig, die Dynamik nach der Binauralisierung nicht zu stark zu komprimieren oder anderweitig zu bearbeiten, um ein angenehmes Hörerlebnis zu gewährleisten. Wir orientierten uns daher an der EBU-Empfehlung R128, die nach Definition der „Loudness Units“ durch ITU-RBS.1770 einen Zielwert von -23 LUFS vorgibt. Der maximale „TruePeak“ darf demzufolge bei höchstens -1 dBTP liegen<sup>51</sup>. Somit verfügten wir aus gestalterischer Sicht über genügend Dynamikumfang und erhielten uns die Möglichkeit, die Sendung im Nachhinein auch anderweitig zu verwerten oder auf anderen Plattformen zu senden.

---

<sup>50</sup> (Spotify AB, 2021)

<sup>51</sup> (EBU, 2020)

Nachdem der Sendeton die passenden Zielwerte erreicht hatte, wurde das Signal aus der DSP-Einheit an unsere Stagebox in der Regie per AES50 übergeben. Dort gaben wir das Signal per DA-Wandler als symmetrisches Signal per XLR an den analogen Eingang des Bildmischers weiter, wo das finale Einbetten des Sendetons in den Programmfeed erfolgte.

### 5.2.2 Signalfluss des In-Ear-Monitoring-Systems

Um die Musiker mit passendem Monitoring zu versorgen, wurden die ankommenden Signale an der DSP direkt auf Eingangskanäle des Mischpultes aufgelegt, in unserem Fall umfasste das acht Monokanäle in einer Standardkonfiguration der Kanalzüge. Jeder Kanalzug umfasste Trimm, Hochpass Filter, Gate, Equalizer, Kompressor, Einschleifpunkt für Effekte, Output-/Busrouting und Fader-Einheit. Hier wurde eine passende Stereomischung für alle beteiligten Künstler\*innen gemischt, mit eventuell gewünschten Effekten wie Hall oder Delay versehen und per Busmatrize auf die passenden Ausgänge geroutet. An dieser Stelle konnten wir zudem die Kommandos einbinden, um während der Probe Anweisungen von Toningenieur und Bildregie an die Künstler\*innen direkt in die In-Ear-Mischungen individuell einbinden zu können. Ebenso konnten hier innerhalb des Sendebetriebs kurze Anweisungen von der Regie zum Programmablauf weitergegeben werden. Durch die Verwendung von Snippets konnten wir per Knopfdruck entsprechende Schaltungen kurzzeitig abrufen.

Über das AES50-Netzwerk wurden die Signale von der DSP-Einheit zurück an die Stagebox an der Bühne geschickt, auf die entsprechenden Digital-Analog-Wandler gegeben und per XLR-Kabel an die verwendeten In-Ear-Kopfhörerverstärker geschickt. Da beim analogen Kopfhörersystem von *Fischer Amps*, den sog. *In-Ear-Sticks*, mittels Adapter zwei symmetrische Line-Signale auf ein dreipoliges Mikrofonkabel mit XLR-Steckverbindung zusammengefasst werden kann, verkürzten sich nicht nur Auf- und Abbauzeiten. Auch kann der Kopfhörerverstärker mit entsprechender Eingangsschaltung das Stereosignal aus einem Kabel gewinnen. Somit müssen für ein Stereosignal an die Künstler\*innen nur ein statt zwei Kabel gezogen werden. Wir erzielten eine Latenz von unter 3 ms von AD- zu DA-Wandlung und konnten den Künstler\*innen damit ein bestmögliches Monitoring gewährleisten. Um auf verschiedene Mischsituationen zwischen den Programmpunkten eingehen zu können, nutzten wir die Vorzüge von Snippets, so dass wir Zustände des gesamten Systems oder auch

nur Ausschnitte aus dem Monitoring speichern und automatisiert oder teilautomatisiert abfahren konnten. Daher war es möglich, das Routing für einzelne Kanäle sowie Effekte und deren Parameter zu verändern, sie an- oder auszuschalten. Ebenso konnten ungenutzte Mikrofonwege in einigen Titeln temporär abgeschaltet werden, wodurch störendes Übersprechen vermieden werden konnte. Auch an dieser Stelle waren die Snippets zur Automation ein wichtiges Tool, um die Doppelbelastung des Toningenieurs möglichst niedrig zu halten und technische Abläufe weitestgehend zu automatisieren. Dies gab dem Toningenieur die Freiheit, sich vermehrt auf die gestalterische Arbeit im Mischen zu konzentrieren.

### 5.2.3 Signalfluss des Kommunikationssystems

Für die Kommunikation wurden neben den Mikrofonen auf der Bühne auch die Kommandomikrofone von Toningenieur, Bildmischer, Streaming-Techniker und den Kameraleuten auf die Eingangskanäle des Mischpults in der DSP Einheit aufgelegt. Die Eingänge und Ausgänge mit entsprechender AD- und DA-Wandlung wurden sowohl an den Stageboxen der Regie als auch an der Bühne umgesetzt, so dass die Signale dem gesamten Audiosystem zur Verfügung standen, flexibel und schnell geroutet und beliebig genutzt werden konnten. Die Kanalstruktur entsprach hierbei der gleichen Struktur wie einem In-Ear-Monitoring-Kanal <sup>52</sup>.

## 5.3 Aufbau und Struktur der „Digital Audio Workstation“ (DAW)

### 5.3.1 Anforderung und Wahl der passenden DAW

Für die Umsetzung einer Binauralisierung sowie der daraus resultierenden binauralen Mischung werden zunächst einzelne Kanäle benötigt, ferner eine DAW, die ein flexibles Routing der Ein- und Ausgangskanäle zulässt. Um den beschriebenen Workflow der Binauralisierung aller Einzelkanäle vorzunehmen, muss es möglich sein, einzelne Kanäle aus dem Audiointerface aufzunehmen, diese zu bearbeiten und per Audio-Plugin zu

---

<sup>52</sup> Siehe Kapitel 5.2.2 Signalfluss des In-Ear-Monitoring-Systems

binauralisieren und im Ausgang als Stereokanal, getrennt vom Master, an das Audiointerface zu übergeben, um schließlich die Signale dem Audiosystem wieder zurückzugeben.

Die Integration einer passenden Software-Erweiterung, eines Plugins, zur Binauralisierung muss zuverlässig möglich und in hoher Zahl verfügbar sein, da jedes Signal eine einzelne Binauralisierung erfährt und daher einen getrennten Kanal benötigt.

Da die DAW im Livebetrieb zum Einsatz kommt, muss ein stabiler Betrieb unter allen Umständen durchführbar sein. Die Systemsicherheit ist hier im Gegensatz zur Postproduktion bzw. Produktion von nichtlinearen Medien neben gestalterischen Merkmalen und der Qualität die wichtigste Bedingung<sup>53</sup>. Die DAW sollte daher sowohl in der Aufnahme als auch im Wiedergabemodus das gleiche Verhalten aufweisen. Weder klangliche noch technische Veränderungen zwischen diesen zwei Zuständen sollten wahrnehmbar sein, um sowohl in der Sendung als auch im Probenbetrieb immer gleiche Voraussetzungen vorzufinden.

Uns standen verschiedene DAWs zur Verfügung, die diese Kriterien erfüllten. Zur Auswahl standen *REAPER*, *AVID ProTools* und *Magix Sequoia15* mit ihren aktuellen Versionen (Stand Mai 2020). Nach einigen Tests stellten wir die höchste Betriebsstabilität mit unserem *Plugin dearVR Pro* mit der DAW *Magix Sequoia15* fest. Hier konnten wir das Plugin live rechnen lassen. Das Verhalten der DAW hat sich sowohl im Pause- als auch im Record-Modus nicht wahrnehmbar geändert, wodurch das System berechenbar blieb.

### 5.3.2 Anforderungen und Wahl des Plugins zur Binauralisierung

Um die Einbindung in möglichst vielen DAWs zu ermöglichen und somit einen flexiblen Einsatz in verschiedenen Systemen fortzuführen, war uns ein Plugin wichtig, das eine Positionierung im 3D-Raum und das Verhalten mit einem künstlich erzeugten 3D-Raum ermöglicht. Außerdem sollte direkt ein binauralisiertes Signal ausgegeben werden, um auf die Verwendung von mehrkanaligen Bussen für Ambisonic-Formate verzichten zu können. Die Verwendung von Mehrkanalbussen für 3D-Audio-Systeme und Ambisonics wird noch nicht von allen DAWs unterstützt, ebenso ist das Rechnen von 16 oder gar 64 Kanalbussen für Ambisonic höherer Ordnung sehr rechenaufwendig. Meine Planung erforderte jedoch eine

---

<sup>53</sup> Siehe Kapitel 5.1 Systemanforderungen der Live-Binauralisierung

flexible und einfache Nutzung der Technik und sollte für die Nutzer eine einfache Bedienbarkeit gewährleisten. Die Leistungsanforderungen an Software und Rechner sollten ebenso in einem überschaubaren Rahmen bleiben, um mit mobilen Laptops die Produktion sicher und schnell bewerkstelligen zu können. Notwendigerweise musste ein All-in-One-Plugin genutzt werden. Die erzeugten Stereosignale waren überdies auch von einer DSP-Einheit bzw. dem Mischpult gut zu verarbeiten, wohingegen mehrkanalige Formate von digitalen Mischpulten nur in einer sehr hochpreisigen Klasse unterstützt werden und die hierfür benötigten Geräte nicht klein genug und transportabel sind, wie es das geplante System vorsah.

Um die Bedienbarkeit und den Workflow zu beschleunigen, war es aus mehreren Gründen wichtig, nur mit Mono- und Stereokanälen in der DAW zu arbeiten. Hätten wir trotzdem auf eine DAW mit mehrkanaligem Bus-Setup für Ambisonic Formate gesetzt, wäre die Zahl der möglichen DAWs stark geschrumpft. Eine moderne DAW ist so komplex, dass ein schnelles Einarbeiten innerhalb von wenigen Stunden nicht möglich ist. Das Erlernen einer DAW dauert Tage, will ein sicheres Verständnis im Umgang – im Besonderen bei einer ersten Anwendung während einer Liveproduktion – sichergestellt werden. Daher war es elementar, dass das genutzte Plugin auf möglichst vielen DAWs zuverlässig funktioniert, sodass der Toningenieur eine DAW, die er sicher beherrscht, nutzen kann. Da Stereobusse von allen professionellen DAWs unterstützt werden, ist die Wahrscheinlichkeit groß, dass der Toningenieur die verwendete DAW bereits kennt oder im Zweifelsfall eine ihm bekannte DAW unkompliziert einbinden kann.

Für ein dreidimensionales, räumliches Hörereignis sollte das Plugin über einen integrierten dreidimensionalen Hall mit entsprechenden Reflexionsmustern verfügen. Aufgrund der Abnahme der Instrumente im Nahfeld waren alle Signale sehr direkt und trocken, weshalb sie aus gestalterischer Sicht einen akustischen Raum brauchten. Damit die binauralisierten Signale in ihrer Räumlichkeit erhalten werden konnten, stellte der klassische Weg – dem Signal einen Stereohall hinzuzumischen – keine Option dar. Bei einem Plugin mit bereits eingebautem Werkzeug zum Verhalten des Signals dagegen bleibt die richtige Lokalisation nach dem Berechnen des 3D-Raums und dem Hinzumischen zum Signal erhalten; eine Richtungswahrnehmung und passende Reflexionsmuster sind dafür essenziell. Erst dann kann eine Binauralisierung zu einem zweikanaligen Signal erfolgen.

Ein Plugin, das all diese Anforderungen erfüllt und zusätzlich eine sehr schnelle und intuitive Bedienung ermöglicht, ist *dearVR Pro* von *Dear Reality*, einem deutschen Entwickler aus Düsseldorf. Hier erfolgt im Plugin die Wandlung des Eingangssignals in ein mehrkanaliges Ambisonic-Format und kann mit einem dreidimensionalen Hall sowie Reflexionsmustern versehen werden. Ausgangsseitig besteht die Möglichkeit der Ausgabe in verschiedene Ambisonic-Formate, wie zum Beispiel Ambisonics 3. Ordnung. Möglich sind auch statische, vordefinierte 3D-Audio-Lautsprecher-Konfigurationen, zum Beispiel mit 7.1.4 oder 13.1. Abgesehen von diesen Möglichkeiten gab es alternativ die direkte Binauralisierung in ein zweikanaliges Format. Hier kam eine Head-Related-Transfer-Function (HRTF) zum Einsatz, die von *Dear Reality* vorgegeben wurde und auch nicht veränderbar war <sup>54</sup>.

Im *Stayin'Live Stuttgart*-Projekt habe ich mich in der Planung für dieses Plugin entschieden, da nicht nur alle technischen Anforderungen, sondern auch ein sehr guter Support durch die Entwickler mit Achim Fell gegeben war. Das Plugin wurde den Systemanforderungen gerecht.

## 6. Postproduktion: Optimierung des Systems

Im folgenden Abschnitt möchte ich auf Optimierungen des Systems eingehen. Nach den durchgeführten Testläufen, der Erfahrung in Sendungsabläufen im Livestreaming und der Auswertungen im Team möchte ich in diesem Abschnitt meine Optimierungen und Veränderungen im Systemdesign beschreiben, die sich über die vier Sendungen ergeben haben.

### 6.1 Optimierung im Stream-Encoding per Hardware-Encoder

Nachdem wir unsere Sendungen ausgewertet hatten, fiel uns ein Problem im Bild auf: Während den Livesendungen fiel in allen Sendungen nach ca. 30 Minuten Sendung für ein paar Sekunden regelmäßig die Bildwiederholrate, wodurch es zu Drops auf bis zu 12 fps kam. In allen Tests, die vor den Sendungen abgelaufen waren, war dieses Problem nicht aufgetreten, da wir nie vier bis fünf Stunden am Stück einen Test-Stream encodierten hatten.

---

<sup>54</sup> (Dear Reality GmbH, 2020)

Außerdem war OBS (Version 25.7) an den Produktionstagen stets ohne Unterbrechung gelaufen; einmal zur Probe gestartet, streamte die Software bis zum Ende der Sendung pausenlos an *Restream.io*. Diese Situation konnten wir so in Tests zuvor nicht herstellen. Wir konnten die Situation im Nachhinein jedoch wiederherstellen und Hardwareprobleme ausschließen, da wir Signalwandler von SDI zu USB und ebenso Rechner getauscht hatten. Das Problem tauchte nach einigen Sendestunden dennoch immer wieder auf. Ebenso hatten wir verschiedene Auflösungen von 1080 p bis 480 p und verschiedene Codierverfahren auf *OBS* verwendet und gewechselt. Trotzdem war das Problem geblieben. Inzwischen gibt es die *OBS*-Version 26.1.1, bei der dieses Problem gelöst ist <sup>55</sup>.

Bei einer solch kritischen Komponente im System, worauf man bei Software keinen Einfluss hat und die die Qualität des Produktes grundlegend beeinträchtigt, wäre eine Hardwarelösung von Vorteil gewesen. Daher setzten wir im System ab September 2020 einen Hardware-Encoder ein. Der *Magewell Ultra Stream SDI* erhält das Signal per SDI und gibt per RJ45-Anbindung an das Netzwerk den encodierten RTMP-Stream direkt an *Restream.io* weiter. Ebenso kann parallel eine Aufzeichnung des encodierten Programms per MOV-Container mit H.264 auf eine HDD oder SSD per USB3.0 erfolgen <sup>56</sup>. Die Steuerung erfolgte in unserem Fall über eine Web-Oberfläche, die per IP angesprochen wurde und die letzten sieben Sendungen sehr zuverlässig arbeitete. Ebenso konnte die Auflösung auf 1920 p x 1080 p mit 25 fps bei 8 Mbit/s erhöht werden.

## 6.2 Wechsel des SDI-USB-Konverter

Der verwendete Konverter der Firma *Marshall* hatte im Video zuverlässig funktioniert und erwies sich als einfach und schnell in Bedienung und Einrichtung. Allerdings waren Probleme im Ton per SDI aufgetaucht, weshalb es zu starken Verzerrungen des Audiosignals bei der Wandlung gekommen war. Das Problem wurde per Ausschluss ermittelt, indem das Audiosignal – unser Programmtone hinter dem Bildmischer – per *BMD Web Presenter* abgehört und geprüft wurde. Hier war das Signal sauber und konnte mit richtigem Pegel abgehört werden. Bei Verwendung des beschriebenen *Marshall* <sup>57</sup> kam allerdings ein stark verzerrtes

---

<sup>55</sup> (OBS, 2021)

<sup>56</sup> (Magewell Electronics, 2020)

<sup>57</sup> Siehe Kapitel 4.7.4 Wandlung des SDI-Signals zur Softwareanbindung

Signal in OBS an, das nicht sendefähig gewesen wäre. In den Sendungen schickten wir den Ton per Audio-Interface (*RME MADIface Pro*) über zwei analoge, symmetrische XLR-Verbindungen von der Regie-Stagebox direkt an den Streaming-Rechner in *OBS*, um dort die Synchronisation erneut flexibel anzupassen. Da sich dies aber nach einigen Sendungen als überflüssig herausgestellt hatte, konnten wir auf den Aufbau der analogen Leitung und des Audio-Interface verzichten und ersetzen den *Marshall Konverter* mit dem *BMD Web Presenter*. Dies brachte drei Vorteile: Zunächst reduzierte sich die Zahl der zu betreuenden Geräte durch den schnelleren Auf- und Abbau des Systems. Des Weiteren konnten wir die zu verarbeitende Datenmenge und Rechenprozesse im Streaming-Rechner minimieren, da die Konvertierung von 1920 p x 1080 p in 1280 p x 720 p im *BMD Web Presenter* erfolgte, wozu nur noch eine USB 2.0-Verbindung benötigt wurde. Als dritter Vorteil ist die Skalierung der Bild-Auflösung durch den *Web Presenter* erfolgt. Die Algorithmen zur Bildskalierung in *OBS* in das erforderliche Sendeformat von Full HD zu HD erzeugten sichtbare Artefakte. Die Algorithmen *Bicubic* lieferten zwar akzeptable Ergebnisse, waren aber rechenintensiv. *Bilinear*, *Area* oder *Lanczos* waren deutlich schlechter im Vergleich zum hardwareskalierten Bild mit dem *BMD Web Presenter*.

### 6.3 Einbindung einer individuellen HRTF

Dass die Zielplattform YouTube keine Einbindung einer eigenen, individuellen HRTF des Zuschauers zulässt, habe ich zunächst hingenommen, was das System in der Planung und Umsetzung mit einer Standard-HRTF in unserem *Plugin dearVR Pro* zur Binauralisierung vereinfachte und überschaubar machte. Ebenso war durch den zuverlässigen Umgang mit Stereokanälen in der DAW ein sicherer Sendungsablauf gewährleistet, der zuverlässig über die Produktion funktionierte und für den Toningenieur einen bekannten Workflow bedeutete.

Bei der Auswertung kam jedoch die Frage auf, ob das Ergebnis der Binauralisierung aus klanglicher Sicht noch optimiert werden könnte. Im Raum stand, ob durch die Einbindung einer anderen HRTF in die Signalkette ein besseres Hörerlebnis zu erreichen wäre. Daher baute ich den Workflow im Nachhinein erneut auf erstellte eine Mischung. Zusätzlich spiegelte ich die Eingänge und stellte bei gleicher Plugin-Einstellung den Ausgang auf

Abisonics dritter Ordnung um. Die Binauralisierung nahm ich mit dem *sparta ambiBIN*, einem weiteren Plugin von Aalto, einem binauralen Ambisonic-Decoder vor, der die Integration einer individuellen HRTF per Spatially Oriented Format for Acoustics (SOFA) zulässt. Das Plugin wurde direkt hinter das *dearVR Pro*-Plugin in der Kette gesetzt, so dass eine Einbindung einer anderen als die Standard-HRTF möglich wäre. In der verwendeten DAW *Sequoia15* war eine Mehrkanaleinbindung von Ambisonic der dritten Ordnung möglich. Leider haben die Plugins zur Binauralisierung nur 8 der 16 Kanäle lesen können. Getestet wurden die Plugins von Aalto und IEM zur Binauralisierung von Ambisonics höherer Ordnungen.

Solche Probleme können aktuell in DAWs bei Verwendung von Drittanbieter-Plugins in Bezug auf Ambisonic Formate auftreten, da sich diese Formate derzeit erst etablieren. Es fehlen noch klare Softwarestandards, die Hersteller-übergreifend Anwendung finden. Auch wenn das Ergebnis in Ambisonics der ersten Ordnung und einer Binauralisierung durch *ambiBIN* vielversprechend klingt, war eine technisch saubere Umsetzung in diesem System mit *Sequoia15* nicht möglich. Hier müsste auf eine andere DAW umgestiegen werden, wie zum Beispiel *REAPER*. Dann müsste das gesamte System aber erneut getestet werden.

## 7. Fazit

Im Fazit möchte ich kurz auf die erfolgreiche Umsetzung des Systemdesigns, die gewonnenen Erkenntnisse und den daraus resultierenden Anregungen für weitere Projekte eingehen.

Die erfolgreiche Planung, mit passendem Systemkonzept auf Grundlage des hier beschriebenen Systemdesigns lässt sich an 4 erfolgreichen Sendungen betrachten. Das System hat sich für alle 4 Genres als kompatibel erwiesen, von Seiten der Künstler gab es keine Wünsche für Veränderungen, ebenso wurde in der Produktion selbst kein Mangel, technischer Defekt oder Ausfälle festgestellt.

Auch konnte flexibel auf äußere Umstände spontan reagiert werden. Ein Beispiel: Da durch mangelndes Budget keine Business Leitung eines Providers angemietet werden konnte, oder die Verwendung einer *Dark Fiber* ausgeschlossen war, war die Produktion immer auf den Internetanschluss mit entsprechendem Datenlink nach Tarif des jeweiligen Veranstaltungsortes angewiesen. Im *Mókuska Caffè* zu unserer 2. Sendung kam es zum Ausfall

des Internets. Nach 2 Stunden Probe mit entsprechendem Test-Livestream kam es zu einem Drop der Upload-Geschwindigkeit. Diese fiel auf unter einen Mbit/s. Somit war ein Senden von Bild und Ton unmöglich. Hier kam unser Notkonzept erfolgreich zum Einsatz und die Sendung wurde um 17 Uhr aufgezeichnet. Die Aufzeichnung wurde pünktlich zum Sendungszeitpunkt an einem anderen Ort vom Band ausgestrahlt. Durch dieses Notkonzept mit „Point of no Return“, dass bis 16 Uhr ein Sendefähiges System stehen muss oder andernfalls spätestens 17 Uhr aufgezeichnet wird, hat sich als richtig erwiesen.

Durch die Anpassung des Systems, indem Arbeitsbereiche zusammengelegt, Aufgaben sinnvoll verteilt und teilweise automatisiert wurden, konnte Personal eingespart werden. Aufgrund der Automatisierung und Vorbereitung der Arbeitsbereiche konnte sich das Personal auf die kreativen Aufgaben und die Überwachung konzentrieren. Um einen technisch und gestalterisch professionellen Livestream zu ermöglichen, waren daher fünf Personen nötig: ein Toningenieur für die Mischung und damit die klangliche Gestaltung der Livemischung des Sendetons und der Monitormischung der Künstler, ein Bildmischer/eine Bildregie für den Bildschnitt, Kamerakoordination und Bildgestaltung sowie ein Livestream-/MAZ-Techniker für die Einrichtung der Livestream-Technik vor der Sendung, die Überwachung der Livestream-Technik und für die Bedienung der Grafik- und Bildeinspielung während der Sendung. Im Sinne eines ästhetischen Bildschnitts mit mehreren statischen und zwei handgeführten Kameras waren außerdem zwei Kameraleute unverzichtbar.

Zum Zeitpunkt der Produktion waren die geltenden Hygieneauflagen und Personenbeschränkungen damit eingehalten.

Aus Sicht der Systemgestaltung ist im Bereich Planung und Umsetzung Folgendes festzuhalten: Auch wenn die Gewerke der Produktion innerhalb ihrer Bereiche technisch, strukturell und aufgrund ihres Umfeldes klar voneinander getrennt werden können, gilt das nicht in Bezug auf ihre Bedienung und Anwendung im Workflow. Im Sinne der Systembetrachtung kann eine Unterteilung der Bereiche Ton-, Licht-, Bild- und Sendetechnik nur in ihren Grundaufgaben und der Technik selbst erfolgen. Hierarchien sind noch immer klar erkennbar und ein systemhierarchisches Denken kann zur Konzeptionierung, Planung und Gestaltung eines Systems angewendet werden. Dies funktionierte aber aufgrund fehlender Erkennbarkeit der Hierarchien im Steuerungskonzept des Workflows und der Steuerungstechnik nicht mehr. Alle Regie-Beteiligten hatten über das Steuerungsnetzwerk

Zugriff auf alle Positionen und Arbeitsplätze, um auch im Notfall übernehmen zu können oder um in der Einrichtung und im Aufbau flexibel ihren Kollegen helfen zu können. Dies lässt sich auch auf das Produktionsteam selbst übertragen, da jeder im Team alle Position kennen und beherrschen musste, da keine Redundanz im klassischen Sinne vorhanden war. Dies hatte im Gegenzug den großen Vorteil, dass auf zusätzliche Technik und unnötigen Platz verzichtet und damit eine langsamere Produktion, die weniger agil gearbeitet hätte, vermieden werden konnte. Statt doppelter Technik bzw. Personal als Redundanz wurden wichtige Schlüsselkomponenten gedoppelt, wie der Router/Switch für die Steuerung des Systems. Eine Redundanz in den Gewerken bestand in redundanten Arbeitsplätzen und multifunktionalem Personal. Das Team investierte viel Zeit in der Vorbereitung, um alle Beteiligten mit allen Gewerken einzuarbeiten. Da sie alle mit Schwerpunkt auf audiovisuelle Medien und Technik ausgebildet werden, war ein grundlegendes Verständnis für alle Gewerke vorhanden und somit auch ein Einlernen auf Details der Produktion in Bild-, Ton-, Licht- und Stream-Technik kein Problem.

Hier sehe ich klar die Zukunft und auch den Sinn dieses Systemkonzepts. Um schnell und flexibel verschiedenste Produktionen umzusetzen, anpassbar an verschiedenste Produktionsumstände und Gegebenheiten zu sein, muss eine Einschränkung auf das Nötigste möglich sein, ohne auf gestalterische und technische Qualität zu verzichten. Technik und Personal müssen multifunktional einsetzbar sein, denn je weniger Personal vorhanden ist, desto größer und anspruchsvoller werden die Aufgaben für alle Beteiligten. Ein übergreifendes Steuerungskonzept mit einem flexibel und breit ausgebildeten Personal ist essenziell, um eine Produktion sicher und zuverlässig umzusetzen. In der Zukunft wird Systemsicherheit nicht mehr nur mit technischer Redundanz gleichgesetzt sein, sondern sich auch auf die mögliche Automatisierung von technischen Abläufen beziehen, um das Personal dahingehend zu entlasten, dass die Konzentration komplett auf gestalterische Aufgaben gerichtet werden kann. Auch personelle Redundanz wird nicht mehr durch eine Doppel-Besetzung geschaffen werden, sondern durch die Reduktion auf Personal, das alle Gewerke in den Grundlagen so weit beherrscht, dass im Notfall eine Sendefähigkeit gewährleistet werden kann. Dennoch sollten alle Beteiligten auf ein Gewerk spezialisiert sein. Klassische Systemstrukturen werden auf technischer Seite weiter bestehen und Hierarchien zur Planung und Abgrenzung der einzelnen Bereiche erhalten bleiben. Auf Anwenderseite wird es mit

entsprechendem Steuerungssystem mit diesem System keine Hierarchie mehr geben, da einer der Knackpunkte darin besteht, dass alle Beteiligten gemeinsam auf demselben Level der Zuständigkeit und des Zugriffs arbeiten, auch wenn feste Positionen mit jeweiliger Spezialisierung vorhanden sind. Hierfür werden die Ansprüche an Techniker und Ingenieure an eine vielfältige Ausbildung und breite Kenntnisse größer werden. Außerdem wird ein vertrauter Umgang mit Netzwerktechnik eine wichtige Voraussetzung für zukünftige Arbeitsplätze sein.

## Tabellenverzeichnis

Tabelle 1 - Leistungsbedarfsrechnung der Produktion.....	21
Tabelle 2 - Quellen und Senken in der Bildtechnik.....	23
Tabelle 3 - Quellen und Senken in der Tontechnik.....	24
Tabelle 4 - Preisvergleich - Neupreise abgerufen am 28.01.2021 von Thomann.de .....	45

## Abbildungsverzeichnis

Abbildung 1 – Systemdarstellung Stayin' Live Stuttgart.....	70
---	----

## Literaturverzeichnis

Bernstädt, H. (6. August 2019). *Phasenanschnittsdimmer - Warum summt der Dimmer?*

Abgerufen am 23. Januar 2021 von production-partner.de: <https://wiki.production-partner.de/licht/warum-summt-der-dimmer/#Top>

BET-Fachwörterbuch. (16. Juni 2018). 4:2:2. Abgerufen am 26. Januar 2021 von bet.de:

<https://www.bet.de/lexikon/4zu2zu2/>

BET-Fachwörterbuch. (21. September 2020). 3G-SDI. Abgerufen am 25. Januar 2021 von

bet.de: <https://www.bet.de/lexikon/3gsdi/>

Blackmagic Design Pty. Ltd. (2021). *ATEM Television Studio*. Abgerufen am 25. Januar 2021

von blackmagicdesign.com:

<https://www.blackmagicdesign.com/de/products/atemtelevisionstudio>

Blackmagic Design Pty. Ltd. (2021). *ATEM Television StudioHD - Technische Daten*.

Abgerufen am 25. Januar 2021 von blackmagicdesign.com:

<https://www.blackmagicdesign.com/api/print/to-pdf/de/products/atemtelevisionstudio/techspecs/W-APS-08?filename=atem-television-studio-hd-techspecs.pdf>

Blackmagic Design Pty. Ltd. (2021). *Blackmagic Web Presenter*. Abgerufen am 26. Januar

2021 von blackmagicdesign.com:

<https://www.blackmagicdesign.com/de/products/blackmagicwebpresenter/techspecs/W-WPR-01>

Blackmagic Design Pty. Ltd. (2021). *Technische Daten - HyperDeck Studio Mini*. Abgerufen am 26. Januar 2021 von blackmagicdesign.com:

<https://www.blackmagicdesign.com/api/print/topdf/de/products/hyperdeckstudiomini/techspecs/W-HYD-07?filename=hyperdeck-studio-mini-techspecs.pdf>

BMVI. (2021). *Das Breitbandförderprogramm des Bundes*. Abgerufen am 24. Januar 2021 von Bundesministerium für Verkehr und digitale Infrastruktur:

<https://www.bmvi.de/DE/Themen/Digitales/Breitbandausbau/Breitbandfoerderung/breitbandfoerderung.html>

Boehmel, M. (08. November 2020). *So funktioniert das SONY LANC(tm)-Protokoll*. Abgerufen am 25. Januar 2021 von boehmel.de: <http://www.boehmel.de/lanc>

Bundesministerium für Wirtschaft und Energie. (2020). *Kultur- und Kreativwirtschaft*.

(BMWi, Hrsg.) Abgerufen am 01. Februar 2021 von bmwi.de:

<https://www.bmwi.de/Redaktion/DE/Artikel/Branchenfokus/Wirtschaft/branchenfokus-kultur-und-kreativwirtschaft.html>

BW, L. (26. Mai 2020). *Zweite Verordnung der Landesregierung zur Änderung der Corona-Verordnung*. (L. B. Württemberg, Hrsg.) Abgerufen am 22.. Januar 2021 von Baden

Württemberg: [https://www.baden-wuerttemberg.de/fileadmin/redaktion/dateien/PDF/Coronainfos/200526\\_Zweite\\_VO\\_der\\_LReg\\_zur\\_Aenderung\\_der\\_CoronaVO.pdf](https://www.baden-wuerttemberg.de/fileadmin/redaktion/dateien/PDF/Coronainfos/200526_Zweite_VO_der_LReg_zur_Aenderung_der_CoronaVO.pdf)

Cerevo. (2021). *Specifications*. Abgerufen am 25. Januar 2021 von flextally.cerevo.com:

<https://flextally.cerevo.com/en/#spec>

Dear Realiy GmbH. (März 2020). *dearVR Pro - USER MANUAL v1.4.1*. Abgerufen am 30. Januar 2021 von dearvr.com:

[https://cdn.shopify.com/s/files/1/0120/4748/2938/files/dearVRpro\\_Manual\\_v1.4.1\\_rev4.pdf?v=1591283987](https://cdn.shopify.com/s/files/1/0120/4748/2938/files/dearVRpro_Manual_v1.4.1_rev4.pdf?v=1591283987)

EBU. (2020). *R128 LOUDNESS NORMALISATION AND PERMITTED MAXIMUM LEVEL OF AUDIO SIGNALS*. Abgerufen am 29. Januar 2021 von European Broadcast Union - ebu.ch: <https://tech.ebu.ch/docs/r/r128.pdf>

Georgios Marentakis, C. K. (2012). *Latency Tolerance Enhancement in In-Ear Monitoring Systems*. (K. G. Institut für Elektronische Musik und Akustik, Hrsg.) Abgerufen am 27. Januar 2021 von dega-akustik.de: [https://pub.dega-akustik.de/DAGA\\_2012/data/articles/000239.pdf](https://pub.dega-akustik.de/DAGA_2012/data/articles/000239.pdf)

Google LLC. (2021). *Empfohlene Einstellungen für die Upload-Codierung*. Abgerufen am 25. Januar 2021 von Support.Google: <https://support.google.com/youtube/answer/1722171?hl=de#zippy=%2Cfarbraum>

Google LLC. (2021). *Live-Encoder-Einstellungen, Bitraten und Auflösungen wählen*. Abgerufen am 24. Januar 2021 von Support.Google: [https://support.google.com/youtube/answer/2853702?hl=de&ref\\_topic=9257892#zippy=%2Cp%2Cp-bei-fps%2Ck-p-bei-fps](https://support.google.com/youtube/answer/2853702?hl=de&ref_topic=9257892#zippy=%2Cp%2Cp-bei-fps%2Ck-p-bei-fps)

Grzesinski, C. (2020). *Symrek - Tontechnik für Veranstaltungstechnik in Ausbildung und Praxis* (4. Auflage Ausg.). Stuttgart, Detuschland: Hirzel Verlag. Abgerufen am 28. Januar 2021

Heuzeroth, T. (21. Juli 2019). *So schnell ist das Internet in Ihrem Bundesland*. Abgerufen am 24. Januar 2021 von WELT: <https://www.welt.de/wirtschaft/webwelt/article197193483/Internetgeschwindigkeit-Diese-Bundeslaender-sind-die-schnellsten.html>

Heuzeroth, T. (15. November 2020). *Empfang, Auflösung, Streaming – das sind Deutschlands neue Fernseh-Vorlieben*. Abgerufen am 24. Januar 2021 von WELT: <https://www.welt.de/wirtschaft/webwelt/article220173544/Fernsehkonsum-Deutsche-Haushalte-haben-im-Schnitt-1-57-Fernseher.html>

Hildebrand, A. (26. Januar 2018). *Overview on IP Audio Networking*. Abgerufen am 28. Januar 2021 von hs-rm.de: [https://www.hs-rm.de/fileadmin/persons/khofmann/Gastvortraege/Vortragsfolien/20180126-Hildebrand-IP\\_Audio\\_Standards.pdf](https://www.hs-rm.de/fileadmin/persons/khofmann/Gastvortraege/Vortragsfolien/20180126-Hildebrand-IP_Audio_Standards.pdf)

- ITU. (Juni 2015). *Recommendation ITU-R BT.709-6*. Abgerufen am 25. Januar 2021 von ITU.int: [https://www.itu.int/dms\\_pubrec/itu-r/rec/bt/R-REC-BT.709-6-201506-1!!PDF-E.pdf](https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.709-6-201506-1!!PDF-E.pdf)
- Kaminski, P. (05. März 2020). *Digitales UHF-Wireless - Test: Shure Axient Digital*. Abgerufen am 27. Januar 2021 von Digitales UHF-Wireless: <https://www.production-partner.de/test/test-shure-axient-digital/>
- L MUSIC Group IP. (2016). *Product Information Document - M32C*. (L. M. IP, Hrsg.) Abgerufen am 23. Januar 2021 von midasconsoles.com: [https://mediadl.musictribe.com/media/PLM/data/docs/POBM8/Midas\\_M32C%20POBM8\\_Product%20Information%20Document.pdf](https://mediadl.musictribe.com/media/PLM/data/docs/POBM8/Midas_M32C%20POBM8_Product%20Information%20Document.pdf)
- Magewell Electronics. (27. November 2020). *Ultra Stream SDI Technical Specifications*. Abgerufen am 24. Januar 2021 von magewell.com: [https://www.magewell.com/static/tech-specs/Ultra\\_Stream/UltraStreamSDI.pdf](https://www.magewell.com/static/tech-specs/Ultra_Stream/UltraStreamSDI.pdf)
- Mantei, D. (2019). *Heben Halten Tragen Dossier*. (B. f. Arbeitsmedizin, Hrsg.) Abgerufen am 23. Januar 2021 von Bundesanstalt für Arbeitsschutz und Arbeitsmedizin: [https://www.baua.de/DE/Themen/Arbeitsgestaltung-im-Betrieb/Gefaehrdungsbeurteilung/Expertenwissen/Physische-Belastung/Heben-Halten-Tragen/Heben-Halten-Tragen\\_dossier.html?view=pdfViewExt](https://www.baua.de/DE/Themen/Arbeitsgestaltung-im-Betrieb/Gefaehrdungsbeurteilung/Expertenwissen/Physische-Belastung/Heben-Halten-Tragen/Heben-Halten-Tragen_dossier.html?view=pdfViewExt)
- Markus Nispel, S. K. (26. Juli 2010). *Redundanz, Ausfallsicherheit & Co. – So erreichen Sie High Availability*. Abgerufen am 27. Januar 2021 von IP-Insider.de: <https://www.ip-insider.de/redundanz-ausfallsicherheit-co-so-erreichen-sie-high-availability-a-274602/?p=4>
- Marshall. (2021). *VAC-23SHU3 Manual*. Abgerufen am 26. Januar 2021 von Mashall-USA.com: <http://www.marshall-usa.com/pdf/VAC-23SHU3-manual.pdf>
- MUSIC Group IP Lt. (2015). *IO Interface - PRO SERIES DL251*. Abgerufen am 29. Januar 2021 von musictribe.com: [https://mediadl.musictribe.com/media/PLM/data/docs/POAFN/MIDAS\\_DL251%20POAFN\\_Product%20Information%20Document.pdf](https://mediadl.musictribe.com/media/PLM/data/docs/POAFN/MIDAS_DL251%20POAFN_Product%20Information%20Document.pdf)

OBS. (06. Januar 2021). *OBS Project*. Abgerufen am 24. Januar 2021 von github.com:  
<https://github.com/obsproject/obs-studio/releases>

Petzold, J. (2020). Abgerufen am 22. Januar 2021 von Kulturinsel Stuttgart gemeinnützige GmbH: <https://www.kulturinsel-stuttgart.org/>

Prof. Dr. Reinhard Haberfellner, P. D.-I. (2018). *Systems Engineering - Grundlagen und Anwendung*. Zürich, Schweiz: Orell Füssli Verlag.

Restream Inc. (2021). *Developers Documentation*. Abgerufen am 24. Januar 2021 von [restream.io](https://developers.restream.io/docs#getting-started): <https://developers.restream.io/docs#getting-started>

Sachtler. (2011). *Sachtler ACE M MS - Manual*. Abgerufen am 25. Januar 2021 von [sachtler.com](http://sachtler.com):  
[file:///C:/Users/ARNEMO~1/AppData/Local/Temp/1001\\_documents\\_1\\_system.pdf](file:///C:/Users/ARNEMO~1/AppData/Local/Temp/1001_documents_1_system.pdf)

SMPTE. (2004). *PROPOSED SMPTE STANDARD for Television 1920 x 1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequences for Multiple Picture Rates*. Abgerufen am 25. Januar 2021 von [hifpga.com](http://hifpga.com):  
<https://hifpga.com/upfiles/15925738469707805.pdf>

Sony Corporation. (27. April 2016). *Festspeicher-Camcorder Bedienungsanleitung*. Abgerufen am 23. Januar 2021 von Sony Professional:  
<https://pro.sony/support/res/manuals/4586/45868034M.pdf>

Sony Pro - Audio Lab. (April 2004). *Product Brief*. Abgerufen am 28. Januar 2021 von [aesnashville.org](http://www.aesnashville.org): <http://www.aesnashville.org/PDFs/supermac.pdf>

Spotify AB. (2021). *Spotify FAQ*. Abgerufen am 29. Januar 2021 von [spotify.com](https://spotify.com):  
<https://artists.spotify.com/faq/mastering-and-loudness#how-does-spotify-calculate-loudness>

Thomann GmbH. (23. Juni 2020). *Bedienungsanleitung - Botex UP-2 1Ch Dimmer 10A*. (M. Thomann, Herausgeber) Abgerufen am 23. Januar 2021 von [thomann.de](https://thomann.de):  
[https://im.static-thomann.de/pics/atg/atgdata/document/manual/316126\\_c\\_316126\\_v2\\_de\\_online.pdf](https://im.static-thomann.de/pics/atg/atgdata/document/manual/316126_c_316126_v2_de_online.pdf)

Walke-Chomjakov, I. (02. November 2018). *Das bedeutet USB Video Class auf Ihrer Kamera.*

Abgerufen am 26. Januar 2021 von PC-Welt: <https://www.pcwelt.de/a/das-bedeutet-usb-video-class-auf-ihrer-kamera,3462796>

### Anhang - Abbildung

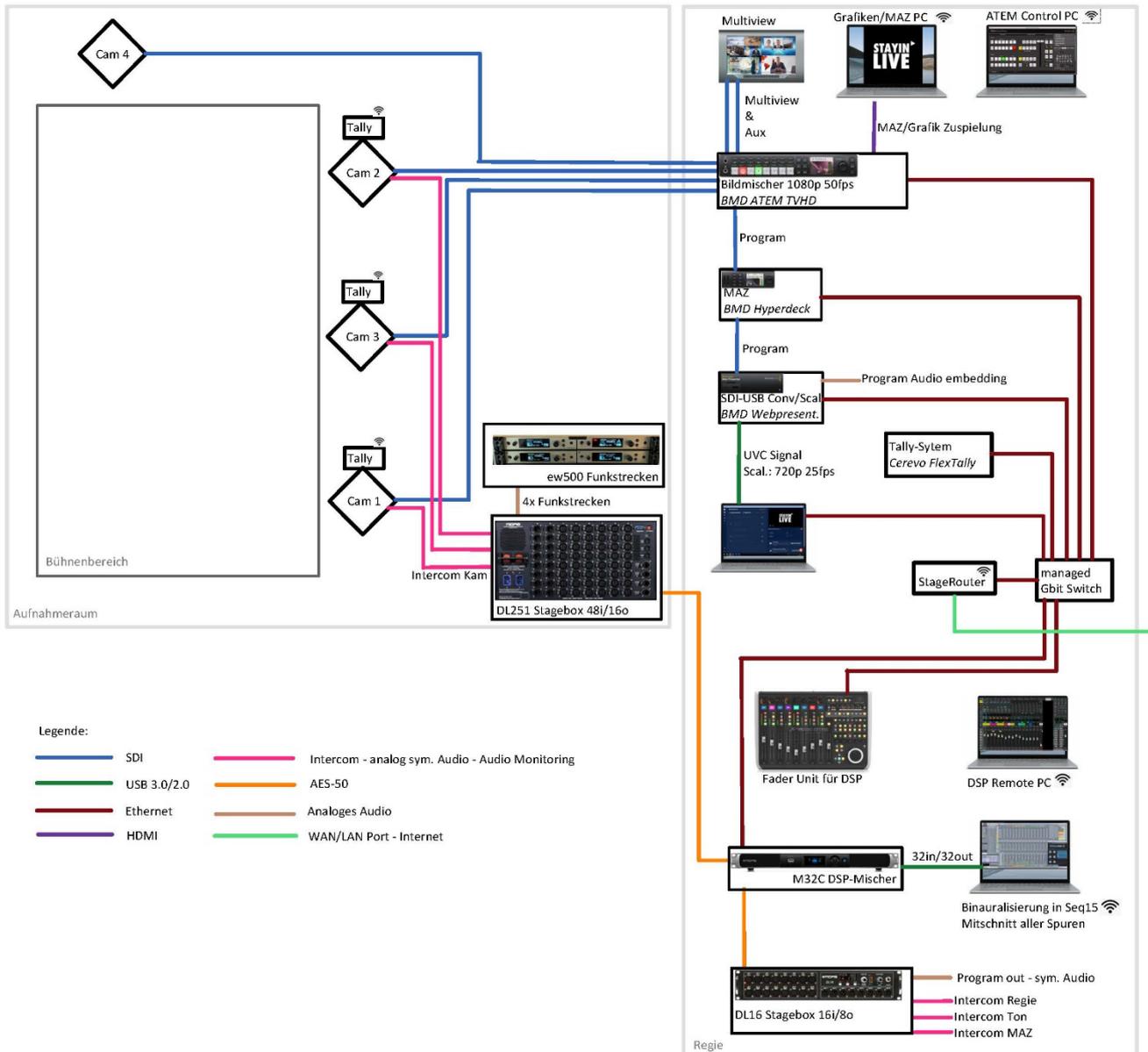


Abbildung 1 – Systemdarstellung Stayin' Live Stuttgart