



Bachelorarbeit

im Studiengang Audiovisuelle Medien
mit dem Titel

**Konzepterstellung zur Erneuerung von hochwertigen und multifunktionalen
Tonbearbeitungsräumen und Tonregionen des Südwestrundfunks**

vorgelegt von

Robin Schubert
Matrikelnummer 40973

an der Hochschule der Medien Stuttgart am 29.04.2025

zur Erlangung des akademischen Grades eines Bachelor of Engineering (B. Eng.)

Erst-Prüfer: Prof. Oliver Curdt

Zweit-Prüfer: Fabian Vossler

Ehrenwörtliche Erklärung

„Hiermit versichere ich, Robin Schubert, ehrenwörtlich, dass ich die vorliegende Bachelorarbeit mit dem Titel: „Konzepterstellung zur Erneuerung von hochwertigen und multifunktionalen Tonbearbeitungsräumen und Tonregionen des Südwestrundfunks“ 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. Ebenso sind alle Stellen, die mit Hilfe eines KI-basierten Schreibwerkzeugs erstellt oder überarbeitet wurden, kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden.

Aus Gründen der besseren Lesbarkeit wurde in dieser Arbeit auf die gleichzeitige Verwendung männlicher, weiblicher und diverser Sprachformen verzichtet. Sämtliche Personenbezeichnungen gelten selbstverständlich für alle Geschlechter gleichermaßen.

Einige in dieser Arbeit enthaltene Informationen basieren auf persönlichen Gesprächen mit Mitarbeitenden des SWR. Auf eine explizite Zitierung einzelner Aussagen wurde in einigen Fällen verzichtet, um den Lesefluss und die Übersichtlichkeit zu wahren, insbesondere weil die Informationen informell und von unterschiedlichen Personen stammen.

Ich habe die Bedeutung der ehrenwörtlichen Versicherung und die prüfungsrechtlichen Folgen (§ 24 Abs. 2 Bachelor-SPO) einer unrichtigen oder unvollständigen ehrenwörtlichen Versicherung zur Kenntnis genommen.“

Stuttgart, 29.04.2025

Ort, Datum

Robin Schubert

Unterschrift

Vorwort

An dieser Stelle möchte ich mich herzlich bei allen Personen bedanken, die mich während der Erstellung meiner Bachelorarbeit unterstützt haben.

Mein besonderer Dank gilt meinem Betreuer Fabian Vossler vom SWR sowie meinem betreuenden Professor Oliver Curdt für ihre wertvolle Unterstützung und die Bereitschaft, diese Arbeit zu begleiten. Der intensive und konstruktive Austausch mit Fabian Vossler hat entscheidend dazu beigetragen, komplexe Sachverhalte richtig einzuordnen und ein tiefgreifendes Verständnis der Strukturen und Anforderungen des SWR zu entwickeln.

Ein herzliches Dankeschön richte ich auch an alle Mitarbeiterinnen und Mitarbeiter des SWR, die mir jederzeit bei empirischen Untersuchungen sowie offenen Fragen mit Rat und Tat zur Seite standen. Besonders danke ich Marc-Oliver Brehm, Manuel Michalski, Christian Eickhoff, Moritz Otto, Bernhard Rebmann und Ully Schreiner.

Mein besonderer Dank gilt zudem Meinhard „Obi“ Jenne, der mit seiner Teilnahme an der empirischen Untersuchung zur Latenzakzeptanz entscheidend zu einer fundierten Einschätzung für den SWR beigetragen hat.

Ebenfalls bedanken möchte ich mich bei Florian Simon von MAGIX, Jonathan Baker und Martin Kraut von Qvest und Michael Thumm von den Bauer Studios für ihre wertvolle telefonische Unterstützung und die hilfreichen Beiträge bei Entscheidungsfindungen rund um die Themen Latenz, DAW und Controlling.

Nicht zuletzt bedanke ich mich herzlich bei Lars Hoffmann, der die Erstellung dieser Arbeit überhaupt erst ermöglicht hat.

Kurzfassung

Die vorliegende Arbeit befasst sich mit der Konzeption zur Erneuerung von Tonregien und digitalen Schnittplätzen beim Südwestrundfunk (SWR) am Standort Stuttgart im Rahmen eines umfassenden Modernisierungsprozesses. Ziel ist es, ein zukunftssicheres Studiokonzept zu entwickeln, das moderne digitale Audiotechnik nutzt, multifunktionale Arbeitsweisen unterstützt und Latenzprobleme beim Monitoring minimiert.

Basierend auf der Analyse bestehender Workflows und technischer Grundlagen (u. a. DAWs, Audio-over-IP, Latenzmanagement) wurden praxisnahe Tests durchgeführt. Dabei zeigte sich, dass rein softwarebasierte Setups zwar grundsätzlich funktionieren, bei Live-Aufnahmen jedoch hörbare Latenz und Stabilitätsgrenzen aufweisen. Deutlich bessere Ergebnisse erzielte ein hybrides System mit DSP-gestütztem Monitoring.

Das empfohlene Konzept kombiniert MAGIX Sequoia für Schnitt und Postproduktion mit modernen Controllern, Dante-Audiovernetzung sowie einem optimierten Monitoring- und Ergonomiekonzept. Die Lösung soll nicht nur technische Anforderungen erfüllen, sondern auch Flexibilität, Produktionsqualität und Standardisierung fördern. Zudem berücksichtigt sie Personalaustausch und die Integration in bestehende Produktionsstrukturen des SWR.

Abstract

This thesis presents a concept for the renewal of audio control rooms and digital editing suites at Südwestrundfunk (SWR) as part of a broader modernization process. The aim is to develop a future-proof studio setup that leverages modern digital audio technology, supports multifunctional workflows, and minimizes latency issues during monitoring.

Based on the analysis of existing workflows and technical foundations (including DAWs, audio-over-IP, and latency management), practical tests were conducted. These revealed that purely software-based setups generally function but exhibit noticeable latency and stability limits during live recordings. Significantly better results were achieved with a hybrid system featuring DSP-assisted monitoring.

The recommended concept combines MAGIX Sequoia for editing and post-production with modern controllers, Dante audio networking, and an optimized monitoring and ergonomics concept. The proposed solution is designed not only to meet technical requirements but also to promote flexibility, production quality, and standardization. It additionally considers personnel interchange and integration into the SWR's existing production structures.

Inhaltsverzeichnis

Ehrenwörtliche Erklärung.....	II
Vorwort	III
Kurzfassung.....	IV
Abstract.....	IV
Inhaltsverzeichnis	V
Abbildungsverzeichnis	VII
Tabellenverzeichnis	VIII
Abkürzungsverzeichnis.....	IX
1 Einleitung	1
2 Hintergrund und Motivation.....	3
2.1 Der Südwestrundfunk (SWR) - Organisation und Kontext.....	3
2.2 Aktuelle Studiosituation und Handlungsbedarf	4
2.3 Zielsetzung der Studioerneuerung.....	6
3 Theoretische Grundlagen	8
3.1 Digital Audio Workstations im Rundfunk.....	8
3.1.1 MAGIX Sequoia – Schnittsystem beim SWR.....	8
3.1.2 Avid Pro Tools – Industriestandard für Mixing	9
3.2 Mischpult vs. Controller.....	11
3.3 Audio-over-IP und Vernetzung.....	14
3.4 Latenz in der Audioverarbeitung und Wahrnehmung	16
4 Konzeptionelle Überlegungen.....	20
4.1 Anforderungen an multifunktionale Tonstudios.....	20
4.2 Geplantes technisches Setup	21
5 Empirische Untersuchung.....	24
5.1 Versuchsaufbau und Methodik (inkl. Ergebnis)	25
5.1.1 Test 1: Sequoia im Alleinbetrieb	26

5.1.2	Test 2: Hybrid-Setup mit Pro Tools (HDX)	29
5.1.3	Test 3: Sequoia mit DADman und TotalMix	32
5.1.4	Test 4: Latenzausgleich beim Overdub	34
5.2	Diskussion der Ergebnisse	37
6	Monitoring-Konzept.....	42
7	Ergonomie	46
8	Schlusskonzept.....	50
9	Projektplanung und Umsetzung.....	54
10	Fazit und Ausblick.....	56
	Literaturverzeichnis.....	X

Abbildungsverzeichnis

Abbildung 1: Latenztest Meinhard "Obi" Jenne am 03.04.2025	18
Abbildung 2: Testaufbau „Latenzmessung über ein Mikrofon“	27
Abbildung 3: Testaufbau „Latenzmessung über zwei Mikrofone“	30
Abbildung 4: Testaufbau „Latenzmessung Overdub“	35
Abbildung 5: Mögliche Varianten der Faderanordnung (Ergonomie)	47

Tabellenverzeichnis

Tabelle 1: Auswertung Latenztest 1 – Sequoia im Alleinbetrieb.....	28
Tabelle 2: Auswertung Latenztest 2 – Hybrid-Setup mit Pro Tools (HDX)	31
Tabelle 3: Auswertung Latenztest 3 – Sequoia mit DADman und TotalMix	33
Tabelle 4: Auswertung Latenztest 4 – Latenzausgleich beim Overdub	36
Tabelle 5: Gegenüberstellung der Testergebnisse aus Kapitel 5.1	38
Tabelle 6: Monitoring-Konzept.....	42

Abkürzungsverzeichnis

A/D	<i>Analog-to-Digital (Wandlung)</i>
AES/EBU	<i>Audio Engineering Society / European Broadcasting Union</i>
AES67	<i>Audio Engineering Society Standard 67</i>
AoIP	<i>Audio-over-IP</i>
ASIO	<i>Audio Stream Input/Output</i>
AUX	<i>Auxiliary</i>
Comp	<i>Compressor</i>
CPU	<i>Central Processing Unit</i>
D/A	<i>Digital-to-Analog (Wandlung)</i>
DAW	<i>Digital Audio Workstation</i>
DigAS	<i>Digitaler Aktualitätenspeicher (Redaktions- und Produktionssystem)</i>
DSP	<i>Digital Signal Processing</i>
DSZ	<i>Digitales Schnittzentrum</i>
EQ	<i>Equalizer</i>
EUCON	<i>Extended User Control (Steuerungsprotokoll von Avid)</i>
GPI	<i>General Purpose Input</i>
HDSPe	<i>Hammerfall DSP enhanced (Produktionsserie von RME)</i>
HDX	<i>High Definition Expansion (Avid DSP-Technologie)</i>
KVM	<i>Keyboard, Video, Mouse</i>
MADI	<i>Multichannel Audio Digital Interface (Protokoll)</i>
MIDI	<i>Musical Instrument Digital Interface (Protokoll)</i>
MoIP	<i>Media-over-IP</i>
PC	<i>Personal Computer</i>
PFL	<i>Pre-Fader Listening</i>
PPM	<i>Projektportfoliomanagement</i>
PuS	<i>Produktions- und Sendestudios (Abteilung im SWR)</i>
SADiE	<i>Studio Audio Disk Editor</i>
SDR	<i>Süddeutscher Rundfunk</i>
Smp	<i>Samples</i>
SMPTE	<i>Society of Motion Picture and Television Engineers</i>
SSD	<i>Solid State Drive</i>
SWF	<i>Südwestfunk</i>
SWR	<i>Südwestrundfunk</i>
USB	<i>Universal Serial Bus</i>

1 Einleitung

Die Medienbranche befindet sich in einem kontinuierlichen technologischen Wandel, der auch etablierte Rundfunkanstalten betrifft. Der Südwestrundfunk (SWR) als zweitgrößter Sender der ARD steht ab dem Jahr 2025 vor der Aufgabe, mehrere seiner Tonbearbeitungsräume und Tonregien technisch auf den neuesten Stand zu bringen. Neben den sogenannten digitalen Schnittzentren (DSZs), die primär für die Nachbearbeitung von Tonaufnahmen genutzt werden, umfasst diese Modernisierung auch das Podcast- und das Kammermusik-Studio. Zukünftig sollen die Studios multifunktional genutzt werden, wodurch Synergien geschaffen sowie Räume flexibler ausgelastet werden können. Gleichzeitig ergibt sich hieraus eine Reihe von Herausforderungen, insbesondere im Hinblick auf die technische Ausstattung und die latenzfreie Audioübertragung beim Aufnahmeeinsatz.

Die Erneuerung der Studios ist nicht allein ein Selbstzweck, sondern wird durch konkrete Anlässe getrieben. Zum einen steht am SWR-Standort Stuttgart eine umfangreiche Gebäudesanierung bevor, welche die Studioebene betrifft. In diesem Zuge müssen die dort befindlichen Tonbearbeitungsräume neu aufgebaut werden, was die Möglichkeit bietet, veraltete Technik zu ersetzen und das Studio-Layout zu überdenken. Zum anderen entspricht das Vorhaben dem Selbstbild des SWR als innovative und wandlungsfähige Rundfunkanstalt: „Veränderung ist dem SWR in die Wiege gelegt und Teil unserer DNA“, betonte SWR-Intendant Kai Gniffke anlässlich des 25-jährigen Jubiläums der Anstalt (SWR, 2023). Dies unterstreicht die Bereitschaft, neue Wege zu gehen und technische Innovationen – wie den vermehrten Einsatz von Software-basierten Workflows und Audio-over-IP – proaktiv anzugehen.

Bislang sind die DSZs rein für die Bearbeitung bereits aufgezeichneter Inhalte ausgelegt. In solchen Umgebungen ist eine Latenz unkritisch, da keine Musiker oder Sprecher ein direktes Kopfhörersignal abhören müssen. Soll jedoch ein Studio multifunktional sowohl für Schnitt als auch für Aufnahmen genutzt werden, steigen die Anforderungen: Beim Einsprechen von Moderationen oder Einspielen von Musik in denselben Räumen muss eine nahezu latenzfreie Abhörmöglichkeit gewährleistet sein, damit sich die Ausführenden selbst ohne wahrnehmbare Verzögerung hören können. In klassischen Aufnahmestudios wird dies oft durch direkte analoge Signalwege oder dedizierte Hardware-Mischpulte erreicht. In einem modernen digitalisierten Studio stellt dies jedoch eine Herausforderung dar, da Digital Audio Workstations (DAWs) und Audio-Interfaces inhärente Latenzen durch die Wandlung zwischen analogem und digitalem Signal sowie Pufferung mit sich bringen. Vor

diesem Hintergrund stellt sich die zentrale Frage: Wie kann der SWR seine Tonbearbeitungsräume so erneuern, dass sie höchsten Qualitätsansprüchen genügen, multifunktional nutzbar sind und eine latenzarme Aufnahmeumgebung ermöglichen?

Unter dieser Problematik sind einige Unterfragen zu klären, unter Anderem: Soll weiterhin ein Hardware-Mischpult in die Regie integriert werden, oder kann dieses durch eine Kombination aus DAW-Software und Controller ersetzt werden? Welche technischen Lösungen gibt es, um die Latenz beim Monitoring auf ein für Musiker nicht wahrnehmbares Maß zu reduzieren? Wie lassen sich neue Audio-over-IP-Technologien nutzen, um Flexibilität in der Studiovernetzung zu erreichen? Welche Arbeitsabläufe und Schnittstellen zu anderen Produktionsbereichen sind zu berücksichtigen? Und: Welche Aspekte der Ergonomie und des Monitorings müssen beachtet werden, damit in den Studios effizient gearbeitet werden kann?

Diese Bachelorarbeit verfolgt das Ziel, ein technisches Konzept für die geplante Studioerneuerung zu entwickeln. Dieses Konzept soll auf fundierten Recherchen, Mitarbeiterkommunikation und empirischen Tests basieren und letztlich als Entscheidungsgrundlage für den SWR dienen. Hierbei wird einerseits theoretisch beleuchtet, welche Optionen zur Verfügung stehen, wie den Einsatz eines Mischpults entgegen einem reinen DAW-Setup, verschiedene Monitoring-Möglichkeiten und Vor- und Nachteile spezifischer Systeme. Andererseits werden praktische Tests durchgeführt, um die Leistungsfähigkeit konkreter Lösungsansätze zu prüfen – insbesondere im Hinblick auf Latenz und Systemstabilität. Die Ergebnisse dieser Untersuchungen fließen direkt in die Konzeptempfehlung ein. Darüber hinaus dokumentiert die Arbeit die bereits im SWR getroffenen Vorentscheidungen.

Zunächst werden in Kapitel 2 der SWR und der Projektkontext vorgestellt. Kapitel 3 erläutert die technischen und theoretischen Grundlagen, die für das Verständnis der Problemstellung wichtig sind. In Kapitel 4 werden darauf aufbauend die konzeptionellen Überlegungen dargelegt, die im Vorfeld der empirischen Untersuchung getroffen wurden. Den Hauptteil bildet Kapitel 5 mit der Darstellung und Interpretation der durchgeführten Tests und deren Ergebnisse. In Kapitel 6 werden Monitoring-Konzepte für die DSZs und diese Resultate kritisch diskutiert und interpretiert sowie in Kapitel 7 ergonomische Aspekte der Studioerneuerung behandelt, bevor in Kapitel 8 das konkrete Technikkonzept vorgestellt wird. Kapitel 9 beleuchtet organisatorische Aspekte der Projektumsetzung wie den Zeitplan und die bereichsübergreifende Zusammenarbeit. Abschließend fasst Kapitel 10 die wichtigsten Erkenntnisse zusammen und gibt einen Ausblick auf zukünftige Schritte und Überlegungen.

2 Hintergrund und Motivation

In diesem Kapitel wird zunächst ein Überblick über den Südwestrundfunk gegeben, um den organisatorischen Rahmen der vorliegenden Arbeit zu verdeutlichen. Dabei werden historische Entwicklungen und die technische Infrastruktur dargestellt, um nachvollziehbar zu machen, weshalb eine Erneuerung der Tonbearbeitungsräume notwendig und sinnvoll ist. Anschließend folgt eine detaillierte Beschreibung der aktuellen Studiosituation am Standort Stuttgart, einschließlich der daraus resultierenden Herausforderungen und des dringenden Handlungsbedarfs. Dies bildet die Grundlage für die Zielsetzungen und Anforderungen der geplanten Modernisierung.

2.1 Der Südwestrundfunk (SWR) - Organisation und Kontext

Der Südwestrundfunk (SWR) ist die öffentlich-rechtliche Rundfunkanstalt für die Länder Baden-Württemberg und Rheinland-Pfalz. Er ging 1998 aus der Fusion des Südwestfunks (SWF) und des Süddeutschen Rundfunks (SDR) hervor (SWR, 2023).

Mit Hauptstandorten in Stuttgart, Baden-Baden und Mainz ist der SWR heute die zweitgrößte Rundfunkanstalt innerhalb der ARD. Er betreibt mehrere Hörfunk- und Fernsehprogramme und verfügt über umfangreiche Produktionsressourcen in Radio, Fernsehen und Onlinemedien (SWR, o. D.). Innerhalb des SWR gibt es verschiedene Abteilungen, die für die Produktion von Audioinhalten zuständig sind. Im Bereich Hörfunk fallen darunter sowohl der Sendebetrieb und die Außenübertragung als auch die Produktion und Postproduktion im Hörspiel, Beitragsschnitt und der Produktion von Musikaufnahmen. Die hier betrachteten Tonbearbeitungsräume und Regien gehören organisatorisch zur Hauptabteilung „Produktions- und Sendestudios“ (PuS) am Standort Stuttgart. Vergleichbare Einrichtungen existieren auch in den Studios in Mainz und Baden-Baden.

Der SWR hat in den letzten Jahrzehnten mehrfach seine technische Infrastruktur erneuert, um mit der Entwicklung Schritt zu halten. Beispielsweise wurden analoge Bandmaschinen durch digitale Schnittsysteme ersetzt und klassische analoge Studiomischpulte wichen zunehmend digitalen Konsolen und computergestützten Workflows. Seit vielen Jahren nutzt der SWR im Bereich der Audioaufnahme und Postproduktion vor allem die DAW-Software „MAGIX Sequoia“. Sequoia hat sich insbesondere für Musik- und Wortproduktionen bei vielen Rundfunkanstalten als Standard etabliert, dank der umfassenden Funktionen im klassischen Musikschnitt und bei der Bearbeitung von Audioaufnahmen. Daneben kommt in bestimmten Bereichen wie dem Hörspiel oder Bildproduktionen auch Avid Pro Tools zum

Einsatz. In einigen Regionen sind DAW-Controller der Avid S-Serie im Einsatz – z. B. eine Avid S6 Konsole in der Hörspielproduktion oder Avid S1 in anderen Studios. Diese Geräte fungieren als physische Kontrolloberflächen für die DAW und bieten viele der haptischen Vorteile eines klassischen Mischpults. Ihre Präsenz im SWR bedeutet, dass Know-how für die Bedienung und Wartung vorhanden ist und Synergieeffekte bei einer Erweiterung der Systeme genutzt werden können. Auch klassische Hardware-Mischpulte wie die digitalen Broadcast-Pulte von LAWO sind in manchen Regionen in Betrieb, insbesondere in den Sendestudios.

Zusammenfassend bildet der SWR als Organisation den Rahmen, in dem die vorliegende Arbeit angesiedelt ist. Die Modernisierung der Tonbearbeitungsräume ist Teil der kontinuierlichen Bestrebungen des Senders, seine Produktionsmittel zu optimieren und den technologischen Fortschritt zu nutzen, um effizienter und qualitativ hochwertiger zu arbeiten. Im nächsten Abschnitt wird auf die konkrete Studiosituation und den daraus resultierenden Handlungsbedarf eingegangen.

2.2 Aktuelle Studiosituation und Handlungsbedarf

Die drei Tonbearbeitungsräume (DSZs) am Standort Stuttgart, die Gegenstand der Erneuerung sind, befinden sich aktuell in der 4. Ebene des Gebäudekomplexes B, die aus brandschutz- und haustechnischen Gründen kernsaniert werden muss. Im Zuge dieser Sanierung sollen die Studios in die 5. Ebene verlegt und anschließend neu eingerichtet werden. Eine räumliche Neuordnung der Technikräume ist dabei ebenfalls notwendig. Somit ist der Umzug ein guter Zeitpunkt, ein neues technisches Konzept umzusetzen und die Studios zukunftsfähig zu machen.

Die DSZ-Räume wurden zuletzt vor ca. 15–20 Jahren grundlegend ausgestattet. Dementsprechend stoßen einige Komponenten an das Ende ihres Lebenszyklus: Die Technik in den Räumen ist nicht mehr auf dem aktuellen Stand. Dies betrifft nicht nur die DSZs für das reine Nacheditieren von Musikaufnahmen, sondern auch das Kammermusik-Studio und das Podcast-Studio, in denen Musik- und Sprachaufnahmen stattfinden. Hier müssen die tontechnischen Komponenten ebenfalls modernisiert werden und die Studios weisen bei der notwendigen Ausstattung einige Parallelen zu den DSZs auf. So ist auch hier die Hardware in den Räumen veraltet.

Im Podcast-Studio beeinträchtigen Mikrofonvorverstärker und Wandler aus den frühen 2000er Jahren die Signalqualität erheblich. Dies führt zu einem schlechten Signal-Rausch-

Abstand, insbesondere bei der Verwendung von dynamischen Mikrofonen, die aufgrund des geringeren Übersprechens im Podcast bevorzugt werden. Die Phantomspeisung des veralteten DHD-Mischpults verursacht Störgeräusche, welche auf andere Kanäle einstreuen und dadurch die Nutzung der Kanäle einschränkt. Somit ist eine Nachbearbeitung der Signale unumgänglich. Routingmöglichkeiten sind am Mischpult nur schwer und komplex umsetzbar.

Auch im Kammermusik-Studio ist zwar keine gebäudetechnische Sanierung geplant, allerdings ist die Technik nicht mehr auf dem aktuellen Stand. Ein veraltetes Yamaha DM-2000-Pult mit eingeschränkten Routingmöglichkeiten, bei dem u. a. keine Klartextnamen verwendet werden können, analoge Mikrofonvorverstärker, die kein Abspeichern der Einstellungen für Folgeproduktionen ermöglichen sowie die Beschränkung, dass der Aufnahmeraum derzeit nur 16 Signale aufnehmen kann, machen komplexere Produktionen in der heutigen Zeit schwer umsetzbar.

Zudem sind gewisse Räume in ihrer Funktionalität spezialisiert – es gibt reine Schnittplätze ohne Aufnahmemöglichkeit und separate Aufnahmeräume mit Regien für z. B. Sprachaufnahmen. Diese Trennung führt teils zu einer ineffizienten Raumnutzung. Um eine höhere Auslastung und effizientere Nutzung zu ermöglichen, soll deshalb ein DSZ in die Kammermusik integriert werden, während die anderen beiden DSZs als reine Nachbearbeitungsstudios vermutlich in die 5. Ebene verlegt werden. Dabei soll eins der geplanten DSZs mit einer 7.1.4-Abhöreinrichtung ausgestattet werden, um den aktuellen Anforderungen und Entwicklungen im Bereich 3D-Audio gerecht zu werden.

Aktuell sind die DSZ-Räume mit Sequoia-Workstations ausgestattet und nutzen MADI-basierende Audiointerfaces von RME für die Anbindung an das zentrale Audiorouting. In reinen Schnitt-Betriebsszenarien wird das zu editierende Material meist von anderen Quellen zugeliefert, z. B. über mitgebrachte Festplatten von Tonmeistern, und es finden keine simultanen Aufnahmen statt. Daher wird die Latenz durch großzügige ASIO-Puffer gelöst – so laufen z. B. die Systeme mit Puffergrößen von 256 oder 512 Samples, was zwar eine hohe Latenz während des Editierens verursacht, aber maximale Stabilität gewährleistet. Beim Integrieren des DSZs in die Kammermusik und beim Podcast-Studio muss allerdings ein Weg gefunden werden, dort während einer Aufnahme die Latenz so weit zu senken, dass ein Sprecher oder Musiker sich in Echtzeit auf dem Kopfhörer hört, ohne jedoch die Systemstabilität zu gefährden oder auf die im SWR etablierte DAW „Sequoia“ verzichten zu müssen. Interne Diskussionen im Vorfeld der Arbeit zeigten verschiedene Meinungen und bereits Tendenzen: Einige Techniker sprachen sich dafür aus, zumindest im geplanten Kammermusik-/Podcast-Studio ein kleines digitales Mischpult vorzusehen, um konservativ die Latenzfrage zu lösen. Dort könnte dann das

Kopfhörermonitoring analog oder via Pult-DSP umgesetzt werden. Andere plädierten für eine einheitliche Lösung in allen Räumen mit DAW-Controllern, um die Bedienkonzepte umfassend konsistent zu halten. Diese Überlegungen galt es durch objektive Untersuchungen zu untermauern.

2.3 Zielsetzung der Studioerneuerung

Aus dem bisher Gesagten lassen sich die Ziele der Studioerneuerung beim SWR wie folgt zusammenfassen:

Die Infrastruktur soll auf einen aktuellen Stand gebracht werden, um bis spätestens 2027 über zukunftssichere Studios zu verfügen. Geplant ist, das technische Konzept bis Ende 2025 auszuarbeiten und 2026 mit der Umsetzung zu beginnen. Dabei werden zentrale Komponenten wie die Steuerung der Workstations, Audio-Interfaces und Abhöranlagen erneuert.

Durch den Fokus auf DAW-gestützte Workflows sollen die Studios digitalen Anforderungen entsprechen und effizient genutzt werden. So lässt sich die gesamte Produktionskette innerhalb der Digital Audio Workstation abwickeln, wodurch Projekte studiounabhängig und ohne externes Mischpult schnell aufgerufen, bearbeitet sowie Korrekturen und Automationsdaten unmittelbar gespeichert werden können.

Das DSZ in der Kammermusik sowie das Podcast-Studio sollen multifunktional werden, d. h. sowohl als komfortabler Schnittplatz für die Postproduktion dienen können, als auch als Regie für Aufnahmen. Dies erhöht die Flexibilität, erfordert aber einen modularen Aufbau. So muss z. B. die Verkabelung die Nutzung von Aufnahmeräumen mit Mikrofonen zulassen, und die Einrichtung muss sowohl für die längere Arbeit am Bildschirm als auch für die Betreuung von Aufnahmesessions geeignet sein.

Ein zentrales Qualitätsziel ist die Minimierung von Latenzen im gesamten Audio-Signalfluss, um Echtzeit-Monitoring zu ermöglichen. Gleichzeitig darf dies nicht zulasten der Stabilität gehen – das System muss auch bei der Aufnahme von 32 Spuren und DSP-Belastung zuverlässig laufen. Der SWR strebt hier einen stabilen und flexiblen Workflow an, der heutigen und zukünftigen Anforderungen genügt. Konkret soll die Lösung auch mit hohen Samplingraten bis 96 kHz und großen Spurenzahlen performant sein, ohne hörbare Verzögerungen für die Künstler zu verursachen.

Die neuen Studios müssen außerdem in die bestehende SWR-Struktur passen. Dazu gehört v. a. die Kompatibilität mit vorhandenen Systemen, wie z. B. dem zentralen Synchronisations-Takt oder eventuell vorhandenen zentralen Speichersystemen. Ebenso ist zu beachten, dass der SWR bereits bestimmte Software wie Sequoia und Pro Tools lizenziert hat und ein Wechsel auf komplett neue Software unwahrscheinlich ist. Vielmehr geht es darum, mit den vorhandenen Tools bessere und flexiblere Workflows zu ermöglichen.

Auch aus menschlicher Sicht ist das Ziel, das Konzept bestmöglich in den SWR zu integrieren. In der Praxis zeigt sich, dass Mitarbeitende im SWR auch längere Zeiträume von 2-3 Monaten nicht im Studio sind und anschließend schnell wieder in die Produktionsabläufe eingebunden werden müssen. Deshalb ist es essenziell, dass die Bedienung in den neuen Studios intuitiv und unkompliziert bleibt, um intensive Einarbeitungsphasen zu vermeiden und sich an die realen Bedürfnisse der Nutzer bestmöglich anzupassen.

Ein implizites Ziel – gerade auch dieser Arbeit – ist es, die Entscheidungsfindung transparent zu machen und zu dokumentieren. Die verschiedenen Ansätze wie die Entscheidung zwischen Mischpult und Controller oder unterschiedliche Monitoring-Modelle sollen verglichen und die Gründe für die letztliche Entscheidung nachvollziehbar festgehalten werden. Dadurch entsteht für die Abteilung und die Leitung eine fundierte Basis, um die Investitionen und Änderungen zu begründen.

Zusammengefasst soll am Ende ein Konzept stehen, das effiziente, flexible und zukunftssichere Studios beschreibt, in denen hochqualitative Audioproduktionen durchgeführt werden können. Die nachfolgenden Kapitel schaffen zunächst das theoretische Fundament in Kapitel 3 und erläutern im 4. Kapitel die initialen konzeptionellen Überlegungen, bevor die praktischen Untersuchungen und darauf aufbauenden Lösungen in Kapitel 5 präsentiert werden und Lösungen zum Monitoring und zur Ergonomie im 6. und 7. Kapitel diskutiert werden.

3 Theoretische Grundlagen

Im folgenden Kapitel werden die wesentlichen technischen und theoretischen Grundlagen dargelegt, die für das Verständnis der Konzepterstellung relevant sind. Zunächst wird auf die Rolle von Digital Audio Workstations (DAWs) im Rundfunk eingegangen und speziell die im SWR verwendeten Systeme „Sequoia“ und „Pro Tools“ beleuchtet. Anschließend werden die beiden Herangehensweisen in Studio-Regien mit einem Mischpult gegenüber einem Controller diskutiert. Daraufhin folgt ein Überblick über Audio-over-IP-Technologien und deren Bedeutung für moderne Studioumgebungen. Schließlich wird das Thema Latenz in der Audioproduktion erläutert, sowohl hinsichtlich technischer Ursachen als auch der Wahrnehmung durch den Menschen. Diese Grundlagen bilden den Rahmen, innerhalb dessen die weiteren Überlegungen stattfinden.

3.1 Digital Audio Workstations im Rundfunk

Die Einführung von Digital Audio Workstations (DAWs) seit den 1990er Jahren hat die Audioproduktion grundlegend verändert. Statt linearer Schnitte auf Tonband oder rein analogen Mischungen erlauben DAWs eine non-lineare Bearbeitung mit nahezu unbegrenzten Möglichkeiten in Spurenzahl, Bearbeitungsschritten und dem Rückgängigmachen von Arbeitsschritten (Beacham, 2017). Im Rundfunkbereich wurden Systeme wie Fairlight MFX-3 oder MAC PowerPC basierende Schnittsysteme von Sonic Solutions zunehmend durch spezialisierte Systeme wie SADiE abgelöst. Später kamen verbreitete DAWs wie Pro Tools hinzu, insbesondere wenn eine Verzahnung mit Musik- und Mischproduktionen stattfand. Seit 2006 kommen beim SWR die DAWs Sequoia von MAGIX sowie seit den 90ern Pro Tools von Avid für die Aufnahme und Produktion von Sprache und Musik zum Einsatz. Im folgenden Kapitel werden beide Programme näher vorgestellt und in Bezug auf ihre Anwendung beim SWR eingeordnet.

3.1.1 MAGIX Sequoia – Schnittsystem beim SWR

MAGIX Sequoia ist die DAW, die beim SWR in der Tonbearbeitung vorherrschend genutzt wird. Sequoia ging aus der Software Samplitude hervor und wurde als High-End-Lösung für Mastering, Broadcast und Postproduktion etabliert. Ihre Stärken liegen u. a. in der objektbasierten Bearbeitung, d. h. Audioobjekte mit individuellen Effekten, dem Vierpunkt-

Schnitt – ein Feature, das insbesondere für klassische Musikproduktion sehr effizient ist – sowie einem präzisen Crossfade-Editor. Zudem unterstützt Sequoia hohe Kanal- und Spurenzahlen von bis zu 512 I/O-Kanäle und 32-Bit Fließkomma-Audio, was für höchste Klangqualität und internen Headroom sorgt. (Sequoia Pro 17 – The engineers' choice, o. D.)

Im SWR-Kontext wird Sequoia vor allem für die Musikproduktion verwendet. So wird Sequoia z. B. bei Mitschnitten von Konzerten oder klassische Aufnahmen, Hörspiel- und Podcast-Produktionen sowie schnitttechnischen Beiträgen wie der Zusammenstellung von O-Tönen, Atmo und Musik zu sendefertigen Beiträgen eingesetzt. Die Benutzer im SWR sind mit Sequoia langjährig vertraut, was ein wichtiger Faktor ist.

Eine Besonderheit von Sequoia ist, dass es traditionell auf native Verarbeitung setzt, d. h. es benötigt keine dedizierte DSP-Hardware und läuft auf Standard-PCs mit Windows-Betriebssystem. Dadurch ist es flexibel einsetzbar, allerdings teilt es sich die CPU-Ressourcen mit dem Betriebssystem, was bei sehr geringen Puffergrößen für eine geringe Latenz kritisch werden kann. Sequoia bietet einen Modus für direktes Hardware-Monitoring über ASIO Direct Monitoring, bei dem die Software das Monitoring an das Audio-Interface delegiert (F. Simon, persönliche Kommunikation, 20. Februar 2025). Dieser Modus spielt im Kontext unserer Untersuchungen eine Rolle, denn er soll theoretisch latenzfreies Abhören ermöglichen, hat aber in der Praxis Einschränkungen gezeigt (siehe empirischer Teil).

Im Hinblick auf die Integration in größere Systeme kann Sequoia mittels standardisierter Schnittstellen mit externer Hardware kommunizieren, z. B. durch MIDI oder EUCON, ein Steuerungsprotokoll von Avid für DAW-Controller. Der SWR hat diesen Weg bereits beschritten – in manchen Studios wie dem Hörspielstudio steuert eine Avid S6-Konsole parallel Sequoia und andere Systeme.

Zusammenfassend lässt sich sagen, dass Sequoia als zentraler Bestandteil des Workflows im SWR gesetzt ist. Ihre Stabilität und Leistungsfähigkeit in den bisherigen Anwendungen sind bewährt. Die Herausforderung besteht nun darin, Sequoia in einen Live-Aufnahme-Kontext einzubinden. Daher wird in der Konzeptentwicklung geprüft, wie Sequoia mit anderen Komponenten kombiniert werden kann, um die Anforderungen zu erfüllen.

3.1.2 Avid Pro Tools – Industriestandard für Mixing

Avid Pro Tools ist weltweit eine der am weitesten verbreiteten DAWs, insbesondere in Tonstudios für Musikproduktion, Filmtoun und Broadcast-Postproduktion. Während Sequoia

im deutschsprachigen Rundfunk stark vertreten ist, gilt Pro Tools international als Standard in professionellen Tonstudios. Pro Tools zeichnet sich durch seine robuste Aufnahme- und Mixing-Engine aus, die insbesondere in Verbindung mit Avids eigener Hardware Pro Tools HDX niedrige Latenzen und hohe Zuverlässigkeit ermöglicht (Beacham, 2017).

Ein Alleinstellungsmerkmal war lange der Einsatz von HDX-DSP-Karten, die Plugins und Mixer berechneten, sodass die CPU entlastet wurde und konstante Performance erzielt werden konnte – ein Vorteil, der vor dem Aufkommen leistungsfähiger CPUs entscheidend war. Auch heute noch bietet das HDX-System den Vorteil, dass Monitoring nahezu latenzfrei möglich ist, da das Signal direkt auf der DSP-Hardware gemischt wird (Camou, 2013).

Im SWR wird Pro Tools bisher punktuell genutzt. Beispiele sind die Fernseh-ton-Postproduktion, wo Pro Tools-Projekte mit Video-Schnittsystemen austauschbar sein müssen, und manche Musikaufnahmen, bei denen externe Produzenten oder Tonmeister auf Pro Tools setzen. Außerdem besitzt der SWR bereits einige Pro Tools Lizenzen und Hardware. So steht in einigen Studios etwa ein Avid MTRX Interface zur Verfügung – eine modulare Audio-Kreuzschiene, die MADI, Dante und Analog verbinden kann – und zumindest eine HDX-Hardware, die in einem Studio verwendet wird.

Für die vorliegenden Konzeptuntersuchungen ist Pro Tools vor allem in Hinblick auf das Monitoring interessant. Avid hat in den letzten Jahren den sog. Hybrid Engine Modus eingeführt, der es erlaubt, in Pro Tools DSP und native Verarbeitung zu kombinieren. Im Aufnahmebetrieb kann ein Track in den DSP-Modus geschaltet werden, sodass dessen Monitoring und Basisplugins auf der HDX-Hardware quasi latenzfrei laufen, während andere Spuren und Prozesse nativ berechnet werden, um die DSP nicht zu überlasten (Lebowski, 2021). Dieses Feature könnte genutzt werden, um in einem ansonsten nativen Workflow mit Sequoia gezielt die Monitoring-Pfade auszulagern. Ein denkbare Szenario – welches später praktisch getestet wurde – ist, Pro Tools als vorgeschalteten „Monitor-Mixer“ zu verwenden: Die Mikrofoneinspeisungen laufen parallel in Pro Tools für das Monitoring und in Sequoia für die Aufnahme. Pro Tools übernimmt mit seiner DSP-Power das Erstellen des Kopfhörermixes ohne wahrnehmbare Latenz, während Sequoia im Hintergrund aufnimmt. Die Performer hören sich über Pro Tools ab und nehmen dessen Anwesenheit dabei nicht wahr, während die Tontechniker gleichzeitig in Sequoia arbeiten und nach der Aufnahme sämtliche Spuren dort vollständig zur Verfügung haben. Dieses Prinzip erfordert natürlich, dass beide Workstations synchron und zuverlässig laufen und miteinander kommunizieren. Auch müssen potenzielle Probleme wie doppelt hörbare Signale oder Latenzunterschiede zwischen den Systemen gelöst werden. Pro Tools bringt ebenfalls eine breite Unterstützung für Controller mit. Avids eigene Konsolen der S-Serie sind primär für

Pro Tools entwickelt und erlauben eine tiefe Integration mit Pegelanzeigen, Plugin-Control sowie Multi-DAW-Control mit der Möglichkeit, zwischen Workstations umzuschalten.

Für den SWR, der bereits Avid-Controller verwendet, wäre es insofern vorteilhaft, wenn Pro Tools Teil des Konzepts ist, da die vorhandene Bedienoberfläche dann optimal genutzt wird. Eine wichtige Überlegung ist jedoch: Will man eine zweite DAW parallel betreiben? Dies erhöht die Komplexität für den Bediener, und es bedarf klarer Regeln, wann welche DAW „führend“ ist. Mitarbeiter, die bislang nur Sequoia nutzen, müssten im Umgang mit Pro Tools zumindest mit den Grundfunktionen und dem Erstellen eines Kopfhörermixes geschult werden. Abschließend sei betont, dass die Stärken von Pro Tools und Sequoia sich in gewisser Weise komplementär verhalten: Sequoia ist sehr stark in Schnitt, Objekt-Editing und Mastering, Pro Tools ist sehr stark im Live-Tracking, Mixing und mit DSP-Hardware praktisch ohne Latenz im Monitoring. Beide Systeme parallel einzusetzen, könnte also eine Strategie sein, die das Beste aus beiden Welten nutzt – allerdings um den Preis erhöhter Systemkomplexität. Diese Abwägung wird im weiteren Verlauf der Arbeit eine Rolle spielen.

3.2 Mischpult vs. Controller

Eine grundlegende strategische Entscheidung bei der Studioerneuerung ist die Frage, ob in Aufnahmerräumen ein klassisches Mischpult als Hardware zum Einsatz kommen soll oder ob ein sogenanntes DAW-Controller-basiertes Setup gewählt wird. Beide Ansätze haben spezifische Vor- und Nachteile, die im Folgenden beleuchtet werden, insbesondere im Hinblick auf die Anforderungen des SWR-Projekts.

Bei einem physischen Audiomischpult werden alle Audioquellen, d. h. Mikrofone, Zuspeler oder DAW-Ausgänge an ein Hardwaremischpult angeschlossen. Im einfachsten Fall könnte dies ein analoges Pult sein, wahrscheinlicher wäre heutzutage jedoch ein digitales Broadcast-Pult wie ein LAWO-Pult aus der mc²-Serie. Ein Mischpult übernimmt die Mischung der Signale in Hardware und bietet i. d. R. direkte Abhörwege mit minimaler Latenz, da das Signal innerhalb des Pults oder seines DSPs verarbeitet wird.

Ein Mischpult arbeitet eigenständig und stabil, unabhängig von einem PC-Betriebssystem. Es ist für den Live-Betrieb konzipiert und daher nahezu ausfallsicher. Bediener schätzen an Mischpulten die haptische Kontrolle und oft intuitive Bedienung. Für Monitoring-Zwecke ist ein Pult ideal, da man analoge oder digitale Direktausspielwege nutzen kann – der Sprecher oder Musiker kann sein Signal praktisch latenzfrei über das Pult auf den Kopfhörer abhören.

Auch mehrere unabhängige Kopfhörermischungen für verschiedene Musiker lassen sich an vielen Pulten einfach durch AUX-Sends erstellen. In laufenden Produktionen sind Pulte zudem robust gegenüber hohen Lasten – es gibt keine „Buffer Underruns“ (das Problem, dass der Audiopuffer nicht schnell genug mit neuen Daten gefüllt wird), solange die DSP des Pultes nicht überfordert ist, was in typischen Anwendungen selten der Fall ist. Die Wartung ist außerdem deutlich weniger komplex im Vergleich zu Controllern, bei denen häufigere Software- und Hardwareupdates zu Fehlerquellen werden.

Allerdings hat ein fest installiertes Mischpult v. a. hinsichtlich Flexibilität und Zukunftssicherheit einige Nachteile. Es ist ein beträchtliches Investment, insbesondere bei der Erstinvestition und der Wartung. Zudem sind Einstellungen wie Mixing-Automationen oder Routing vergleichsweise nicht so portabel: Während in einer DAW Projekte einfach zwischen Studios kopiert werden können, muss man bei Pulten auf Kompatibilität achten, bzw. die Pultsession separat laden. Für Postproduktions-Workflows ist ein Pult mitunter hinderlich: Wenn z. B. ein Mix auf dem Pult gemacht wurde, muss für spätere Änderungen entweder wieder dasselbe Pult genutzt werden, wodurch der Raum blockiert wird, oder man muss die Mischung in der DAW reproduzieren. Ein weiterer Aspekt ist der Platzbedarf: Mischpulte – vor allem wenn sie viele Kanäle bieten sollen – beanspruchen viel Raum und dedizierte Möbel/Installation.

Im SWR wurden in der Diskussion vor allem die Punkte Latenz und Bedienung als Pro-Argumente für ein Mischpult genannt: Gerade für spontane Aufnahmen oder Selbstfahrer-Situationen ist ein vertrautes Pult mit physikalischen Knöpfen möglicherweise effizienter. Dagegen spricht aber, dass in der Postproduktion mit Sequoia dann ein Medienbruch entstünde – man würde quasi zwei Systeme parallel fahren müssen. Sequoia wäre für die Aufnahme und das Pult für die Abhöre da, mit dem Problem, dass z. B. Automationen oder Mix-Einstellungen aus dem Pult nicht in Sequoia vorhanden wären.

Ein DAW-Controller ist im Prinzip ein Mischpult-ähnliches Bediengerät, das jedoch keinen eigenständigen Audioweg hat, sondern lediglich die Steuerung einer Software übernimmt. Beispiele sind die Avid S1/S4/S6, aber auch kleinere Geräte wie Presonus Faderport oder die früher verbreiteten Mackie Control. Mit einem Controller steuert man die DAW direkt: Bewegt man einen Fader am Controller, bewegt sich der virtuelle Fader in der Software, drückt man Solo/Mute, wird dies im Projekt ausgeführt.

Der große Vorteil liegt in der Integration und Automatisierung. Alle Mix-Einstellungen werden im DAW-Projekt gespeichert. Wenn ein Projekt an einem anderen Arbeitsplatz geöffnet wird, mit oder ohne Controller, sind alle Parameter vorhanden. Änderungen können erfolgen, ohne dass man an ein bestimmtes Studio gebunden ist – man könnte

theoretisch eine Mischung mobil auf dem Laptop vorbereiten und im Studio finalisieren. Zudem sind Controller-Setups oft platzsparender und modularer. Ein Gerät wie die Avid S4 kann z. B. mit 8 oder 16 Fadern konfiguriert werden, je nach Platz. Zusätzliche Module wie Displays oder Knobs sind optional. Sollten sich Anforderungen ändern, kann der Controller erweitert oder umgestellt werden. Ein weiterer Vorteil ist die Multifunktionalität. Ein guter Controller kann zwischen verschiedenen Workstations umschalten. Die Avid S6 etwa ermöglicht es, simultan mehrere DAWs zu kontrollieren – beispielsweise könnte auf Faderbank 1-8 Sequoia liegen und auf 9-16 Pro Tools oder DADman, was für einen Anwendungsfall Fall interessant wäre.

Der Hauptnachteil von einem Controller ist jedoch die Abhängigkeit der Leistungsfähigkeit von der jeweiligen Software, die er steuert. In Bezug auf Latenz bedeutet dies: Ein Controller selbst reduziert keine Latenz, er stellt nur eine Fernbedienung für die DAW dar. Wenn also eine DAW mit 5 ms Latenz arbeitet, kann der Controller nichts daran ändern; im Gegensatz kann ein Mischpult eigenständig latenzfrei routen. Auch gibt es bei reinen Controller-Lösungen potenziell mehr Abhängigkeiten von Netzwerkverbindungen und Treibern. Sollte die DAW abstürzen, ist der Controller nutzlos – wohingegen ein Mischpult oft unabhängig weiter Audio vom Band wiedergeben könnte. In Live-Situationen sehen manche Toningenieure daher reinen Software-Lösungen mit Skepsis entgegen. Ein weiterer Punkt ist die Lernkurve: Ein universeller Controller hat oft Layer und Seiten, um die vielen Funktionen der DAW abzubilden. Dies kann weniger intuitiv sein als ein analoges Pult, bei dem es pro Kanal einen EQ mit festen Reglern gibt. Zwar versucht man, mit Touchscreens und Displays bei modernen Controllern Abhilfe zu schaffen, aber ein gewisses Umschalten zwischen verschiedenen Steuerungsebenen bleibt.

Für die SWR-Studios spricht zugunsten der Controller-Lösung, dass man bereits im Haus positive Erfahrungen damit hat. Die Avid S-Serie ist eine "etablierte Marke" und vertraut in der Wartung und Bedienung. Außerdem ist der Trend im Studiobau allgemein zu beobachten, dass Software-basiert gearbeitet wird und große Konsolen eher in speziellen Regien wie den Übertragungsregien für Orchester vorbehalten bleiben. Flexibilität und Portabilität wurden ebenfalls als Pluspunkte erkannt: Projekte lassen sich leicht in andere Räume mitnehmen, wenn identische Software vorhanden ist. Zudem spart der Verzicht auf ein Mischpult Kosten und bietet ergonomische Vorteile, da Bedienelemente beim Controller modular aufgebaut sind und somit Module flexibel gekauft und angeordnet werden können. Eine mögliche Hybridlösung wäre beides zu kombinieren: Ein kleines Hardware-Mischpult nur für grundlegende Dinge wie der Vorverstärkung oder dem Erstellen von Monitor- und Kopfhörermischungen und daneben ein Controller für die Feinmischung und Automation in der DAW. Dies wurde in internen Sitzungen als „Best of both worlds“ diskutiert: Dabei würde

das Pult im Prinzip wie ein großes Audio-Interface fungieren, das Eingänge mischt und latenzfrei verteilt, während die Feinheiten in Sequoia via Controller gemacht werden. Der Vorteil wäre, dass Musiker sich latenzfrei abhören können und trotzdem alle Automationen in der DAW editierbar bleiben. Der Nachteil liegt allerdings ebenso auf der Hand: man benötigt eine doppelte Ausstattung mit Mischpult und Controller und entsprechend mehr Platz sowie eine aufwendige Synchronisation beider Welten. Insgesamt lässt sich festhalten, dass die Entscheidung Mischpult vs. Controller stark von den Prioritäten abhängt: Stehen die Latenzfreiheit und Standalone-Stabilität an oberster Stelle, hat ein Mischpult Vorteile. Geht es um Integration, Recall und maximale Flexibilität, ist ein Controller/DAW-Setup einem Mischpult überlegen. Das Konzept in dieser Arbeit wird versuchen, die Latenzproblematik auch ohne ein Hardware-Mischpult in den Griff zu bekommen, um die Vorteile der integrativen Lösung eines Controllers zu realisieren.

3.3 Audio-over-IP und Vernetzung

Moderne Studios sind keine Inseln, sondern meist Teil eines größeren Netzwerks von Audiogeräten. Traditionell wurden Audioverbindungen über analoge Multicores oder digitale Punkt-zu-Punkt-Leitungen wie MADI oder AES/EBU realisiert. In den letzten Jahren hat sich jedoch Audio-over-IP (AoIP) als neuer Standard etabliert. AoIP bezeichnet die Übertragung von Audiosignalen über gewöhnliche IT-Netzwerke wie Ethernet, meist unter Verwendung spezifischer Protokolle, die eine synchronisierte und paketbasierte Übertragung ermöglichen. Zu den gängigen AoIP-Protokollen gehören u. a. Dante von der Firma Audinate, SMPTE 2110 von der Society of Motion Picture and Television Engineers (SMPTE) sowie Ravenna - eine offene Technologie, die oft im Broadcast eingesetzt wird und kompatibel mit AES67 Standard ist (Audio over IP: The Growing Trend Amongst High End Integrators - StormAudio, 2025).

Zum einen können durch AoIP-Systeme alle Studios an ein gemeinsames IP-Audionetz angeschlossen werden, anstatt jede Regie einzeln über dedizierte Audioleitungen zu verkabeln. Ein zentrales, softwaregestütztes Routing ermöglicht eine flexible Zuordnung von Quellen und Senken. Dies ist besonders nützlich für den schnellen Austausch von Audiomaterial zwischen verschiedenen Studios oder für das Teilen zentraler Ressourcen wie Mikrofonvorverstärker oder Effektracks.

Außerdem können externe Rechner (z. B. mitgebrachte Laptops von Tonmeistern) mittels AoIP-Protokolle wie Dante von der Firma Audinate über ein USB-Interface ins studiointerne Netzwerk eingebunden werden, wodurch auf zusätzliche analoge Verkabelung verzichtet

werden kann. Insbesondere bei einer 7.1.4-Abhörkonfiguration, wie sie in einem der DSZs geplant ist, bietet Dante einen deutlichen Vorteil gegenüber einer traditionellen analogen Verkabelung (12 analoge Kabel).

AoIP-Systeme zeichnen sich durch eine hohe Skalierbarkeit aus. Ein Standard-Gigabit-Netzwerk genügt bereits für hunderte Audiokanäle gleichzeitig, was die Infrastruktur für zukünftige Erweiterungen oder immersive Audioformate vorbereitet. Durch Umstieg auf 10-Gbit-Netzwerke kann die Kapazität bei Bedarf zusätzlich erhöht werden.

Eine wesentliche Sorge ist häufig die Latenz im IP-Netzwerk. Dante z. B. gewährleistet jedoch sehr geringe und feste Latenzen von standardmäßig etwa 1 ms, konfigurierbar sogar bis hinunter zu 150 μ s (Dante Advanced Configuration, 2015). In der Praxis sind diese Werte vernachlässigbar und stellen keinen limitierenden Faktor für das System dar.

Einige AoIP-Interfaces, wie etwa das Avid MTRX oder der LAWO PowerCore, bieten internes Routing und Mixing an, was die Erstellung von Monitor-Mischungen unabhängig von der DAW ermöglicht. Die Software DADman erlaubt hierbei grundlegende Steuerungsmöglichkeiten über das EUCON-Protokoll, auch wenn aktuell keine vollständige Mischpult-Funktionalität vorhanden ist.

Nicht alle Bereiche des SWR profitieren gleichermaßen von AoIP. In den DSZs beispielsweise macht AoIP aktuell wenig Sinn. Diese arbeiten als isolierte Einheiten, in denen Audiomaterial über Server oder externe Datenträger eingespielt wird. Da keine unmittelbare Vernetzung notwendig ist, besteht keine Notwendigkeit für umfangreiche AoIP-Strukturen. Einzige Ausnahme sind externe Geräte, wie bereits erwähnte Laptops, deren Anbindung über Dante deutlich vereinfacht werden könnte. In der Kammermusik und im Podcast-Studio hingegen könnte AoIP perspektivisch sinnvoll sein, insbesondere, um zukünftige Anbindungen an das hausinterne Audionetzwerk zu ermöglichen. Der konkrete Anwendungsfall dafür existiert aktuell noch nicht, wird jedoch als potenzielle Zukunftsoption gesehen – beispielsweise bei Live-Produktionen, die eine Verbindung zum zentralen Schaltraum erfordern.

Das Großprojekt MoIP (Media-over-IP) im SWR hat zum Ziel, mittelfristig sämtliche Mediendaten wie Audio und Video über IP-basierte Infrastrukturen zu verteilen. Dieses Projekt unterstreicht die strategische Bedeutung der IP-Technologie für den SWR, auch wenn aktuell noch keine konkrete Anbindung der Tonstudios an das zentrale MoIP-System vorgesehen ist. Es wird jedoch für dieses Projekt angedacht, die AoIP-Kompatibilität im Hinterkopf zu behalten, um bei Bedarf nachrüsten zu können.

Zusammenfassend bleibt Audio-over-IP eine wichtige Schlüsseltechnologie im neuen Studiokonzept des SWR. Sie gewährleistet Flexibilität, Skalierbarkeit und einfache Integration externer Geräte, allerdings mit gezielten und pragmatischen Einschränkungen dort, wo keine akuten Anforderungen an Vernetzung bestehen.

3.4 Latenz in der Audioverarbeitung und Wahrnehmung

Latenz bezeichnet die Verzögerung, die zwischen dem Zeitpunkt des Einsprechens/-spielens eines Signals und dem Zeitpunkt des Hörens desselben Signals nach der Verarbeitung auftritt. In einem analogen Signalweg ist Latenz hauptsächlich durch die Laufzeit in Kabeln oder Schaltungen bestimmt und liegt im Mikrosekunden-Bereich – praktisch unmerklich (Smyrek, 2009, S. 460). In digitalen Systemen hingegen entsteht Latenz durch A/D- und D/A-Wandlung, Pufferung, Signalverarbeitung (z. B. durch Plugins) und digitale Übertragungswege. Technisch setzt sich die Gesamtlatenz bspw. wie folgt zusammen:

Moderne Analog-Digital-Wandler weisen Gruppenlaufzeiten auf, die oft einige Dutzend Samples umfassen. Diese liegen typischerweise zwischen 0,25 ms und 1 ms, je nach System und Abtastrate. Digitale Audio Workstations arbeiten zudem mit Verarbeitungspuffern, um die Ein- und Ausgänge sowie die interne Bearbeitung zu managen. Ist beispielsweise eine Puffergröße von 128 Samples eingestellt, ergibt sich bei 48 kHz eine Verzögerung von etwa 2,67 Millisekunden je Richtung, also für den Eingang oder Ausgang allein. Bei einem vollständigen Roundtrip, das heißt dem Signalweg von Eingang über die DAW bis zum Ausgang, verdoppelt sich diese Verzögerung auf etwa 5,3 Millisekunden, hinzu kommen weitere systembedingte Latenzen durch Faktoren wie Computer Hardware, Verarbeitungspuffer, Netzwerk oder Sample Rate Konvertierung (Wang, 2017, S. 37-39).

Auch Software-Plugins und Signalprozessoren können zusätzliche Verzögerungen verursachen, insbesondere bei Prozessoren wie Linearphase-Equalizern oder Lookahead-Kompressoren. In Monitoring-Signalwegen vermeidet man deshalb solche Effekte oder verwendet spezielle Versionen, die keine Latenz erzeugen.

Die Digital-Analog-Wandlung erzeugt ähnliche Verzögerungen wie die Analog-Digital-Wandlung und beträgt typischerweise zwischen 0,25 ms und 1 ms (Wang, 2017, S. 37-39). Netzwerkübertragungen, etwa mittels Dante, verursachen weitere Latenzen von rund einer Millisekunde (Dante Advanced Configuration, 2015), wobei diese entfallen können, wenn

alle Komponenten bereits innerhalb des Audio-Interfaces integriert sind. Summiert man, kommt man selbst bei gut optimierten Systemen schnell auf mehrere Millisekunden.

Ein Beispiel: 64 Samples Puffer (1,33 ms bei 48 kHz) * 2 + Wandlung (z. B. 1,4 ms gesamt für A/D+D/A) $\approx$ 4,1 ms (ohne Plugin-Verzögerungen). 4 ms entsprechen etwa 1,36 Meter Schall-Luft-Weg – das ist in etwa die Verzögerung, die entsteht, wenn ein Musiker 1,4 m von seiner Schallquelle entfernt steht (z. B. ein Sänger, der seine Stimme über Kopfhörer mit 4 ms Verzögerung hört, als stünde er 1,4 m von sich selbst entfernt). Diese Veranschaulichung hilft ein Gefühl zu bekommen, wann Latenz relevant wird.

Die Forschung hat gezeigt, dass die Wahrnehmbarkeit von Latenz stark vom Kontext abhängt. In einer 2007 veröffentlichten Studie von Lester und Boley wurde mit professionellen Musikern ermittelt, wieviel Verzögerung im Monitoring als störend empfunden wird. Ihre Ergebnisse zeigen eine große Spannweite der akzeptierten Latenzen: Unter bestimmten Bedingungen, insbesondere beim In-Ear Monitoring, können bereits Verzögerungen von weniger als 1,4 ms hörbar und störend sein. Demgegenüber tolerierten Musiker in anderen Settings, insbesondere bei der Verwendung von klassischen Wedge-Monitoren (Bodenlautsprechern), Verzögerungen bis zu 42 ms (Lester & Boley, 2007, S. 1). Lester und Boley stellten zudem fest, dass die Empfindlichkeit für Latenz maßgeblich vom Instrument abhängt. Instrumente, die eine unmittelbare und physische Rückkopplung zwischen Aktion und Klang erzeugen (wie z. B. ein Schlagzeug, Saxophon oder Gesang), erfordern deutlich niedrigere Latenzen, da Musiker hier besonders sensibel auf Verzögerungen reagieren. Im Gegensatz dazu zeigten Instrumente mit weniger direkter physischer Kopplung und langsamerem Klangaufbau (wie Keyboard oder Bass) eine höhere Toleranz gegenüber Latenz. Besonders Schlagzeuger und Sänger benötigen eine unmittelbare Rückmeldung, da selbst geringe Verzögerungen ihre Timing-Genauigkeit und Performance spürbar beeinträchtigen können (Lester & Boley, 2007, S. 2-3).

Weiterhin wurde verdeutlicht, dass Musiker beim In-Ear Monitoring grundsätzlich sensibler gegenüber Latenz reagieren, da keine akustische Verzögerung oder Raumhall vorhanden ist, der die Wahrnehmung der Verzögerung mildern könnte. Bei Nutzung von Wedge-Monitoren hingegen ist bereits eine natürliche Schallverzögerung durch die Distanz zwischen Lautsprecher und Musiker vorhanden, weshalb hier höhere Latenzen akzeptiert werden (Lester & Boley, 2007, S. 2-3).

Um für den SWR eine praxisnahe Einschätzung zur akzeptablen Latenz bei Studioaufnahmen zu erhalten, wurde der erfahrene Studiomusiker Meinhard „Obi“ Jenne zu einem Latenztest eingeladen. Als professioneller Schlagzeuger mit umfangreicher Bühnenerfahrung – unter anderem durch zahlreiche Auftritte im Rahmen von SWR-

Produktionen – war seine Rückmeldung besonders wertvoll, um eine realistische Akzeptanzgrenze für Monitoring-Latenzen zu ermitteln. Im Rahmen des Tests erfolgte das Kopfhörermonitoring zunächst mit der minimal möglichen Latenz über ein Setup in Pro Tools unter Verwendung von Pro Tools HDX. Anschließend wurde die Roundtrip-Latenz mithilfe des Plug-ins „Timeadjuster“ schrittweise und präzise auf Sample-Ebene erhöht, während Obi Jenne auf einem Woodblock mit kurzer perkussiver Klangcharakteristik auf ein Kopfhörermonitoring seines eigenen Signals spielte. Die Latenz, die von Obi Jenne subjektiv als noch akzeptabel empfundenen wurde, konnte in einem ergänzenden Klicktest nachgemessen und bestätigt werden. Dieses Verfahren wird in den empirischen Untersuchungen in Kapitel 5 näher erläutert. Als Ergebnis hielten wir fest, dass Latenzen bis zu 3,75 ms (etwa 180 Smp bei 48 kHz) als unkritisch angesehen werden können (M. Jenne, persönliche Kommunikation, 03. April 2025).



Abbildung 1: Latenztest Meinhard "Obi" Jenne am 03.04.2025

DAWs haben in der Regel Funktionen wie den Latenzausgleich (Plugin Delay Compensation) integriert. Das bedeutet, dass bei der Wiedergabe alle Spuren verzögert werden, sodass das Projekt trotz unterschiedlicher latenzverursachender Effekte synchron läuft. Beim Live-Monitoring hätte dieser Ausgleich jedoch zur Folge, dass auch das eigene Signal verzögert wird, um eine Synchronität zu ermöglichen. Es gibt auch die Technik des ASIO Direct Monitorings, bei welchem die DAW das Input-Signal direkt im Interface auf den Ausgang routet und somit die Software umgangen wird. Sequoia nutzt dies im "Hardware

Monitoring"-Modus. Dieser Modus hat bei unseren Tests die Latenz nahezu eliminiert, aber wie in Kapitel 5 zu sehen sein wird, brachte es andere Einschränkungen in Sachen Mischbarkeit.

Interessant ist auch, dass geringe Latenzen teilweise nicht negativ auffallen, da wir aus natürlichen Situationen kleine Verzögerungen gewohnt sind. Beispielsweise hört ein Geiger seinen Ton mit ca. 0,3 ms Verzögerung, dies ist der Weg, den der Schall vom Instrument zum Ohr zurücklegt. Ein Orchestermusiker hört Mitmusiker einige Meter entfernt mit ein paar Millisekunden Verzögerung. Das menschliche Gehirn kann diese kleinen Differenzen integrieren. Erst wenn Latenzen deutlich hörbar werden oder als Echo erscheinen, werden sie störend. Bei Audioaufnahmen im Studio haben Musiker allerdings oft keinerlei andere akustische Referenz als das Kopfhörersignal – weshalb es dort umso entscheidender ist, das Signal unmittelbar auf dem Ohr wahrnehmen zu können.

In der Konzeptentwicklung muss also sichergestellt werden, dass die Summe aller Latenzen im Monitoring-Pfad so gering wie möglich ist. Idealerweise sollte die Kopfhörer-Latenz unter dem ermittelten Wert von 3,75 ms liegen. Das setzt niedrige Puffereinstellungen bzw. spezielle Hardware voraus. Zugleich darf dies nicht auf Kosten von Stabilität gehen: Ein Knacksen oder Dropout (Pufferfehler) ist für eine Aufnahme ebenso unbrauchbar. Es gilt also, einen robusten Kompromiss zu finden, der das Problem umgehen kann. Die folgenden Kapitel (insb. Kapitel 5 und 6) werden sich intensiv mit praktischen Messungen der Latenz befassen und evaluieren, wie sich die verschiedenen Konfigurationen messen. Vorab werden im nächsten Kapitel die konzeptionellen Ideen skizziert, die zu diesen Tests geführt haben.

4 Konzeptionelle Überlegungen

Bevor die empirischen Tests durchgeführt wurden, gab es seitens des Projektteams beim SWR verschiedene konzeptionelle Überlegungen und Hypothesen. In diesem Kapitel werden die wichtigsten davon beschrieben. Diese Überlegungen bilden den “Plan” dessen, was im Hauptteil der Arbeit überprüft und ausgearbeitet wird. Sie umfassen die Anforderungen, erste Entwürfe des technischen Setups und die Auswahl der Komponenten, die in Frage kommen. Einige der hier beschriebenen Punkte knüpfen an die theoretischen Grundlagen an, die in Kapitel 3 gelegt wurden.

4.1 Anforderungen an multifunktionale Tonstudios

Aus der Problemstellung und den Zielen aus Kapitel 1 und 2 lassen sich die Anforderungen an das neue Studio-Setup nochmals konkret zusammenfassen.

Die Minimierung der Latenz beim Einspielen ist für die Aufnahmesituation von hoher Priorität. Das System muss also so gestaltet sein, dass ein Musiker oder Sprecher sich selbst ohne wahrnehmbare Verzögerung hört. Als Richtwert sollte die Gesamtverzögerung vom Mikrofon auf den Kopfhörer unter 3,75 ms liegen.

Des Weiteren soll die Anlage professionellen Ansprüchen genügen. Dazu zählen transparente Wandler mit einem hohen Dynamikumfang und linearen Frequenzgang bei mindestens 48 kHz Samplerate, im besten Fall 96 kHz, um zukunftsicher hochauflösend arbeiten zu können, und hochwertige Abhörmonitore. Darüber hinaus sind mindestens 24 Bit Auflösung Standard, was in allen verwendeten Geräten sichergestellt sein muss.

Auch wenn viel Flexibilität gewünscht ist, darf dies nicht zulasten der Betriebssicherheit gehen. Ein Aufnahmetake darf z. B. nicht durch Software-Aussetzer verloren gehen. Daher muss ggf. mit redundanten Komponenten gearbeitet werden, z. B. einem Recording-PC mit Backup-Aufnahme, oder es müssen bewährte Systeme zum Einsatz kommen, die entsprechende Resilienz haben.

Außerdem sollen die Räume verschiedenste Produktionen ermöglichen. Während an Tag 1 ein Podcast mit einem Sprecher aufgenommen wird, könnte an Tag 2 die Aufnahme eines kammermusikalischen Ensembles mit 16 Mikrofonen und an Tag 3 ein reiner Schnitt eines Hörspiels stattfinden. Das Setup muss diese Bandbreite an Szenarien abdecken. Das bedeutet modulare Verkabelung, ausreichend I/O-Kanäle und eine flexible Software-

Konfiguration, die sich schnell anpassen lässt. Außerdem sollen die Studios inklusiv für Videoproduktionen nutzbar sein, um z. B. Podcasts zusätzlich zu filmen.

Trotz der Vielfalt an möglicher Technik ist gewünscht, dass die Bedienung für die Tontechniker möglichst einheitlich und einfach bleibt. Ideal wäre es, wenn man unabhängig von Schnitt oder Aufnahme im Wesentlichen mit der gleichen Oberfläche und Methodik arbeiten kann. Dies spricht für einen Ansatz mit z. B. einem Controller einer bestimmten Serie in jeder Regie statt für getrennte Systeme. Ein einheitlicher Workflow erhöhe auch die Akzeptanz bei den Mitarbeitern.

Die geplante Lösung soll mindestens 10–15 Jahre im Einsatz bleiben können, ohne grundlegend erneuert werden zu müssen. Deshalb soll Technologie verwendet werden, die auch mit kommenden Entwicklungen mithalten kann. AoIP ist z. B. derzeit Stand der Dinge und bietet einige Vorteile, weshalb es für die Studios mitbedacht werden soll. Ebenso sollte der Controller modular und erweiterbar sein, falls in Zukunft mehr Kanäle oder neue Formate wie 3D Audio vermehrt gewünscht werden.

Nicht zuletzt muss das Konzept realistisch im verfügbaren Budget umsetzbar sein. Anschaffungen mehrerer großer High-End-Mischpulte würden aus Kostengründen z. B. nicht in Frage kommen. Daher sind auch wirtschaftliche Überlegungen Teil der Konzeptfindung: Wo kann vorhandenes Equipment weitergenutzt werden? Welche Lösung bietet Preis-Leistungs-Vorteile?

Diese Anforderungen wurden vor Beginn der technischen Konzeptplanung erhoben und priorisiert. Besonders die Latenzminimierung, die Flexibilität an verschiedene Produktionen, der Kostenfaktor und die einfache Bedienbarkeit sind Kernpunkte, die für moderne und multifunktionale Tonstudios beim SWR unabdingbar sind.

4.2 Geplantes technisches Setup

Auf Basis der Anforderungen wurde ein erstes gedankliches Blockschaltbild des favorisierten Studiokonzepts entwickelt. Im Zentrum steht ein leistungsfähiger PC, auf dem Sequoia als zentrale Plattform für Schnitt, Bearbeitung und Aufnahme betrieben wird. Diese DAW bildet das Herzstück der Produktionsumgebung, da sie im SWR bereits seit vielen Jahren etabliert ist und auf eine fundierte Infrastruktur und Nutzererfahrung, u. a. in den Produktions- und Sendestudios, zurückgegriffen werden kann.

Der Rechner ist mit einem hochwertigen, modularen Audio-Interface als Kreuzschiene verbunden, das sämtliche Ein- und Ausgänge bereitstellt. Als besonders geeignet wurde die Avid MTRX identifiziert, da sie maximale Flexibilität bei gleichzeitig hoher Audioqualität bietet. Durch die modulare Bauweise kann die MTRX individuell mit Mic-Preamps, Line-Outs, MADI- oder Dante-Karten bestückt werden. Die Verbindung zum Rechner erfolgt dabei entweder über DigiLink für die Integration mit Pro Tools HDX oder über Thunderbolt bzw. MADI für den Betrieb mit Sequoia im ASIO-Modus.

Im Zusammenhang mit der MTRX wurde auch die Software DADman als zentrale Steuereinheit berücksichtigt. Sie ermöglicht ein präzises Routing sowie das Erstellen individueller Monitoring-Setups innerhalb der MTRX. Ergänzend kann die Software TotalMix von RME eingesetzt werden, insbesondere wenn RME-Interfaces für das Kopfhörermonitoring verwendet werden sollen. TotalMix bietet umfassende Routing- und Mixing-Möglichkeiten mit nahezu latenzfreier Signalbearbeitung und kann das Monitoring innerhalb des Studios auf flexible Weise ergänzen. Die Kombination aus DADman und gegebenenfalls TotalMix ermöglicht eine zentrale und übersichtliche Verwaltung aller Monitoring- und Routingaufgaben.

Zur Steuerung des Setups kommt eine Avid S-Serie Konsole zum Einsatz, die als zentrales Mischpult fungiert und sowohl die DAWs als auch das Audiointerface ansteuern kann. Diese Controller bieten motorisierte Fader, Drehregler, OLED-Anzeigen und Touchscreens, was eine intuitive Bedienung von Sequoia, DADman und bei Bedarf auch Pro Tools ermöglicht. Aufgrund bestehender Erfahrung und vorhandener Infrastruktur innerhalb des SWR fiel die Wahl auf die Avid S-Serie, da die Avid S4 und S6 auch die Kopplung mehrerer Workstations unterstützt (Avid, 2024, S. 49) – ein wichtiger Faktor bei der möglichen parallelen Verwendung von Sequoia und Pro Tools für Aufnahmen. Alternative DAWs wie Cubase oder Reaper wurden frühzeitig ausgeschlossen, da sie keinen relevanten Mehrwert gegenüber Sequoia und Pro Tools bieten und im Haus keine Erfahrung damit vorhanden ist.

Des Weiteren wurde das Konzept einer parallelen DAW-Nutzung geprüft, bei dem Sequoia die Masterrolle für Schnitt und Aufnahme übernimmt, während Pro Tools als vorgelagertes Mischpult für das Monitoring fungiert. Für ein latenzarmes Kopfhörermonitoring mit Pro Tools ist die Nutzung von Pro Tools HDX vorgesehen. Dieses DSP-basierte Monitoring bietet zusätzliche Sicherheit und Performance. Das Vorgehen hat sich im Hörspielbereich des SWR bereits bewährt, bringt jedoch eine gesteigerte technische Komplexität mit sich, insbesondere durch den parallelen Einsatz unterschiedlicher Betriebssysteme (Windows für Sequoia, macOS für Pro Tools).

Das Gesamtsystem ist als modulares digitales Tonstudiogerüst konzipiert, in dessen Zentrum der Arbeitsplatz mit einer oder zwei DAWs steht. Umgeben ist dieser von modularen Ein- und Ausgangseinheiten, darunter Mikrofoneingänge, Kopfhörerausgänge und Lautsprecherverbindungen. Steuerprotokolle wie EUCON oder MIDI/USB ermöglichen die Integration zusätzlicher Geräte. Für kleinere Produktionsräume (z. B. die DSZs ohne Aufnahmeeinsatz) ist eine vereinfachte Version des Setups vorgesehen, beispielsweise mit einer Avid S1 und einem kompakten Interface. In größeren Studios, wie dem Kammermusik- oder Podcast-Studio, ist dagegen das vollständige Setup mit einer Avid S4 oder S6 vorgesehen, um den hohen Anforderungen an Audioqualität und Monitoring gerecht zu werden. Durch diese Staffelung können hochwertige Komponenten dort eingesetzt werden, wo sie den größten Mehrwert bieten, während gleichzeitig eine einheitliche Benutzeroberfläche für alle Studios erhalten bleibt.

Neben der technischen Ausstattung wurden auch weitere betriebsrelevante Aspekte in die Planung aufgenommen. So ist für die Kommunikation ein Talkback-System vorgesehen, das über ein Mikrofon in der Regie und die DADman-Software in der MTRX gesteuert wird. Dies kann beispielsweise über eine programmierbare Taste auf dem Controller oder ein GPI-Signal erfolgen.

Abschließend wurde die notwendige Rechnerhardware definiert. Vorgesehen sind leistungsfähige, geräuscharme PCs, die im Technikraum platziert und über KVM-Switches (Keyboard, Video, Mouse) von der Regie aus bedient werden können. Diese Praxis ist im SWR bereits etabliert und erlaubt eine lüfterfreie Regieumgebung bei gleichzeitig direktem Zugriff auf die Workstations.

Die detaillierte Betrachtung der Komponenten und Systemarchitektur machte dabei deutlich, dass insbesondere das Thema Latenz eine zentrale Rolle spielt und als entscheidender Faktor bei der Ausgestaltung des Gesamtkonzepts berücksichtigt werden muss. Zum einen bei den Puffereinstellungen und der Leistungsfähigkeit der Workstation, zum anderen bei der Notwendigkeit eines DSP-basierten Monitorings, wie es mit Pro Tools HDX oder TotalMix realisierbar ist.

5 Empirische Untersuchung

Im empirischen Teil der Arbeit wurden mehrere Testreihen durchgeführt, um die zuvor theoretisch diskutierten Lösungen praktisch zu erproben. Der Fokus lag dabei auf der Überprüfung der Latenz und Stabilität in verschiedenen Konfigurationen. Konkret wurden vier Haupt-Testsettings definiert:

Im ersten Test wurde geprüft, inwieweit Sequoia im Alleinbetrieb eine Aufnahme mit vielen Spuren und Plugins handhaben kann, ohne störende Monitoring Latenz oder Fehler zu erzeugen.

Beim zweiten Test wurde Pro Tools dem System als DAW-Mischpult vorgeschaltet – In diesem Szenario wurde das Monitoring in Pro Tools unter Verwendung der HDX-Karten umgesetzt, während die Signale parallel in Sequoia für die eigentliche Mischung und Aufzeichnung geroutet wurden. Untersucht wurde die erzielbare Latenzreduktion durch HDX-DSP und die Systemstabilität bei dieser Kopplung.

Im dritten Test wurde das Monitoring mittels DADman und TotalMix geprüft. Dabei stand im Vordergrund, welche Latenzwerte in der Praxis tatsächlich erreichbar sind, wenn das Signal direkt an der MTRX abgegriffen wird und inwiefern sich DSP-gestützte Plugins in TotalMix auf die Latenz auswirken.

Zuletzt ging es im vierten Test um die Messung der in der Regie auftretenden Latenzen während der Aufnahme sowie deren anschließender Ausgleich. In diesem Szenario wurde überprüft, ob eine Beurteilung der Mischung und des Timings während der Aufnahme möglich ist, wenn dem Musiker aus Sequoia ein Playback abgespielt wird.

Jeder dieser Tests wurde mit einem spezifischen Versuchsaufbau durchgeführt. Im folgenden Kapitel werden zunächst die grundlegenden Rahmenbedingungen und Messverfahren erläutert. Anschließend erfolgt die detaillierte Darstellung der einzelnen Testszenarien sowie die Auswertung der jeweiligen Ergebnisse. Zuletzt werden die Resultate vergleichend gegenübergestellt und im Hinblick auf ihre praktische Relevanz interpretiert.

5.1 Versuchsaufbau und Methodik (inkl. Ergebnis)

Die Tests wurden von Januar bis März 2025 in einem Provisorium im SWR durchgeführt. Als Hardware stand ein Windows 11 PC mit Intel Core i9-9900K Prozessor (8 Kerne bei 3,6 GHz), 32 GB Arbeitsspeicher und einer Samsung SSD mit 6900 MB/s Lesegeschwindigkeit zur Verfügung. Darauf waren Sequoia 17 und TotalMix V1.97 installiert. Ein weiterer MAC Pro (2019) Rechner mit Mac Sonoma 14.7.2, Intel Xeon W Prozessor (16 Kerne bei 3,2 GHz) und 32 GB Arbeitsspeicher diente für die Tests, in denen Pro Tools als Mischpult vorgeschaltet wurde, sowie für die Steuerung der Avid MTRX als zentrale Kreuzschiene. Dafür waren auf dem Mac Rechner DADman V5.7 und Pro Tools Ultimate 2024.10.1 installiert. Als Audio-Interface kam ein RME HDSPe MADI FX für Sequoia und ein Avid HDX für Pro Tools zum Einsatz.

Sämtliche Mikrofoneingänge waren mit einer MTRX verbunden, in der sowohl die A/D-Wandlung als auch die Mikrofon-Vorverstärkung der Signale erfolgte. Die digitalen Signale wurden anschließend an eine zweite MTRX geschickt, in der das Routing sämtlicher Ein- und Ausgangssignale stattfand. Abhängig vom jeweiligen Testszenario wurde das Signal von der ersten auf die zweite MTRX auf unterschiedlichem Weg übertragen: Entweder direkt über eine MADI-Koaxialverbindung, über Pro Tools mittels DigiLink oder über Sequoia/TotalMix durch eine optische MADI-Verbindung. In der zweiten MTRX wurden die Signale in einen „Andiamo“ von DirectOut weitergeleitet, wo die abschließende D/A-Wandlung stattfand. Von dort gelangten die analogen Signale zurück ins Studio und wurden über einen Lake People Phone-Amp als Kopfhörersignal an den Musiker ausgegeben. Diese etwas komplexe Konstellation mit mehreren MTRX-Modulen ergab sich aus der verfügbaren Studioteknik im Hörspiel-Studio des SWR, in dem alle in dieser Arbeit beschriebenen Latenztests durchgeführt wurden.

Für die Messungen kamen hochwertige Messmikrofone der Marken „Schoeps“ und „Neumann“ sowie drei Mini-Monitore des Herstellers „Klein + Hummel“ zum Einsatz. Die Monitore dienten unterschiedlichen Zwecken: Einer wurde zur Ausgabe des Musikersignals verwendet, ein zweiter zur Wiedergabe des verzögerten Kopfhörersignals aus dem Phone-Amp und ein dritter für die parallele Wiedergabe des zeitlich unverzögerten Referenzsignals in einem separaten Raum. So konnte ein direkter Vergleich zwischen verzögertem und unverzögertem Signal hergestellt werden. Die Testsignale wurden mithilfe eines Focusrite Scarlett 8i6 (3rd Gen) Audio-Interfaces auf einem separaten Laptop über Cubase Pro 12 abgespielt und aufgezeichnet. Zur präzisen Latenzmessung wurde ein perkussives Testsignal mit kurzem Impuls aus der DAW ausgegeben und über zwei Messmikrofone erneut aufgenommen. Die simultane Aufzeichnung des verzögerten und unverzögerten

Signals auf getrennte Spuren in Cubase ermöglichte eine exakte Bestimmung der Latenz, die im jeweiligen Signalweg entstanden war. Für die Stabilitätstests wurde anstelle des perkussiven Impulses ein kontinuierlicher 440 Hz Sinuston verwendet, der über einen längeren Zeitraum abgespielt und aufgezeichnet wurde. Auf diese Weise konnten potenzielle Audiofehler bei unterschiedlicher Spurenzahl oder Puffergröße identifiziert werden.

Im folgenden Abschnitt werden die vier durchgeführten Testszenarien im Detail beschrieben und deren Ergebnisse im Hinblick auf Latenzverhalten und Stabilität ausgewertet. Für jeden Test bieten Tabellen eine übersichtliche Darstellung der Testergebnisse, die in Kapitel 5.2 nochmals gegenübergestellt werden.

5.1.1 Test 1: Sequoia im Alleinbetrieb

Im ersten Test lief Sequoia im Alleinbetrieb. Es wurde die minimal mögliche Latenz ermittelt und die Grenzen von Audioaussetzern und Dropouts ausgelotet. Das Musikermonitoring wurde hierbei über die DAW gelöst.

Sequoia lief hierbei im Mixer FX Monitoring/Hybrid Engine Modus. Dabei wird das Monitoring Signal durch die DAW geschleift und der komplette Signalfuss des Sequoia Mixers übernommen. Hintergrund dieser Überlegung war es, die Kopfhörermischung und die eigentliche Mischung vollständig softwareseitig in Sequoia zu lösen, um den Einsatz zusätzlicher Software zu vermeiden. Diese Arbeitsweise hat allerdings zur Folge, dass sich sämtliche auf den Spuren eingesetzte Effekte auf die Latenz aller Monitoring-Signale in der Regie und auf dem Kopfhörer auswirken. Auch die ASIO-Puffergröße spielt dabei eine entscheidende Rolle für die korrekte Verarbeitung und Wiedergabe der Signale und wirkt sich ebenso auf die Gesamtlatenz aus (MAGIX, 2025, S. 736).

Für den Test wurde in Sequoia ein Projekt mit 32 Spuren angelegt, um ein volles Aufnahmeszenario zu simulieren (z. B. ein Ensemble mit 32 Mikrofonen). Jeder Spur wurde ein individuelles Eingangssignal der MTRX zugewiesen. Parallel wurde auf jeder Spur eine Kopfhörerausspielung eingerichtet. Dafür gibt es in Sequoia pro Spur einen AUX-Bus, auf dem eine Kopfhörermischung gemacht werden kann und das Signal als Pre-Fader auf einen getrennten Stereo-Ausgang geroutet wird. Auf jede Spur wurden drei Plugins geladen: Ein EQ (MAGIX EQ116), ein Kompressor (MAGIX AM-Track) und ein Limiter (MAGIX sMax11), um eine realistische DSP-Last zu erzeugen (insgesamt 96 Plugins aktiv). Für den Latenztest wurden alle Spuren in Sequoia parallel aufgezeichnet.

Die Messungen wurden bei verschiedenen ASIO-Puffergrößen durchgeführt: 32, 64 und 128 Samples. Für die Latenzmessung wurde zunächst das perkussive Testsignal über einen Mini-Monitor (in Abb. 2: Spk 1) aus Cubase wiedergegeben und über ein Mikrofon (in Abb. 2: Mic 1) aufgezeichnet. Das Signal des Musikers wurde zunächst in der MTRX A/D-gewandelt und anschließend per Lichtwellenleiter an die RME MADI-Karte übertragen, über welche es in Sequoia eingespeist wurde. Dort erfolgte die Verarbeitung für das Monitoring. Das bearbeitete Signal wurde anschließend in Form eines MADI-Signals an eine weitere MTRX-Einheit zurückgesendet. Nach der D/A-Wandlung wurde das Signal schließlich analog an den Kopfhörer des Musikers ausgegeben. Der Kopfhörer des Musikers wurde für die Messungen durch einen weiteren Mini-Monitor (in Abb. 2: Spk 3) ersetzt und konnte so das Signal an ein Messmikrofon ausgeben.

Um die Latenz zwischen verzögertem Signal und unverzögertem Originalsignal korrekt beurteilen zu können und Latenzen aus Cubase oder dem Focusrite Scarlett Interface auszuschließen, wurde für die Messung das perkussive Testsignal parallel auf einen weiteren Mini-Monitor (in Abb. 2: Spk 2) ausgespielt. Spk 3 als Musikerskopfhörer und Spk 2 als unverzögertes Testsignal konnten somit in einem getrennten Raum über ein Schoeps Mikrofon mit einer MK 2S Kapsel in Cubase auf eine Spur aufgezeichnet werden. Die Kugelcharakteristik der Mikrofonkapsel ermöglichte eine relativ genaue Messung der Signale, die über die Mini-Monitore in gleichem Abstand und Winkel zum Mikrofon ausgerichtet waren.

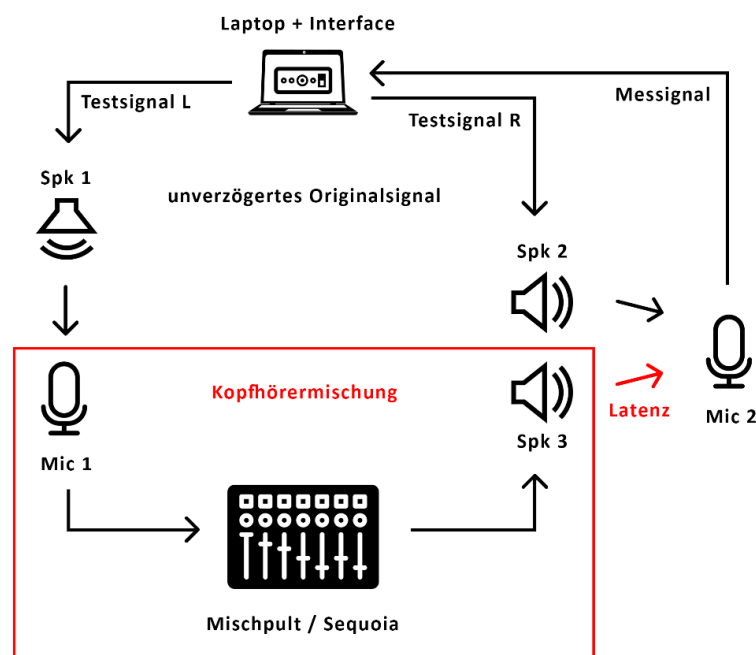


Abbildung 2: Testaufbau „Latenzmessung über ein Mikrofon“

Um neben der Latenz die Systemstabilität beurteilen zu können, wurde in einem weiteren Durchlauf ein Sinuston über einen längeren Zeitraum anstelle des perkussiven Testsignals aus Cubase wiedergegeben.

Im Test zeigte sich zunächst, dass bei geringen Puffergrößen durchaus niedrige Latenzwerte möglich waren, jedoch auf Kosten der Stabilität: Bei einem ASIO-Puffer von 32 Samples lag die Roundtrip-Latenz, also Zeit von Aufnahme bis Ausgabe des Signals, bei ca. 4,8 ms (≈ 230 Smp). Subjektiv war beim Abhören im Studio ein leichtes Phasen-Flanger hörbar, wenn man gleichzeitig das direkte Schallsignal und das Kopfhörersignal abhörte – was andeutet, dass ≈ 5 ms durchaus wahrnehmbar sind, aber für einen Sprecher oder manche Musiker womöglich noch tolerierbar. Allerdings traten bereits nach wenigen Sekunden Aufnahme Aussetzer auf, welche Sequoia selbst als "Aufnahmefehler" verzeichnete. Dies wurde mit dem Sinustest eindeutig, bei dem Audioknackser und Lücken hörbar waren. Auch bei 64 Samples ASIO-Puffer kam es zu Pufferüberläufen, die sich als Audioknackser äußerten. Die erreichte Latenz betrug ca. 6,3 ms (≈ 300 Smp).

Bei 128 Samples kam es zu einer Roundtrip-Latenz von ca. 9 ms (≈ 430 Smp). Hierbei wurden bei einem fünfminütigen Sinustests allerdings keine Aufnahmefehler festgestellt. Dies bestätigte sich auch bei der Erhöhung der Spurenzahl auf 64 Spuren, welche in einem weiteren Durchlauf gleichzeitig aufgenommen wurden.

Tabelle 1: Auswertung Latenztest 1 – Sequoia im Alleinbetrieb

H/W Puffer	Sequoia
32 Samples	4,8 ms (≈ 230 Smp)
64 Samples	6,3 ms (≈ 300 Smp)
128 Samples	9 ms (≈ 430 Smp)

Dropouts / Nicht akzeptable Latenz

In Summe lieferte der erste Test die Erkenntnis, dass im reinen Betrieb mit Sequoia ein Puffer von mindestens 128 Samples erforderlich ist, um eine stabile Aufnahme zu gewährleisten. Dabei war jedoch deutlich, dass die daraus resultierende Latenz von rund 9 Millisekunden insbesondere für anspruchsvolle Musiker oder Sprecher als zu hoch einzustufen ist. Nach Diskussionen beim SWR war ebenso klar, dass für eine durchgehend stabile Aufnahme über mehrere Stunden der ASIO-Puffer nicht an der Grenze der

Systemstabilität betrieben werden sollte. Aus diesem Grund wäre es aus technischer Sicht empfehlenswert, den Puffer auf 256 Samples zu erhöhen – was jedoch mit einer deutlich höheren Latenz einhergeht und das latenzfreie Monitoring zusätzlich erschwert.

Damit war die Notwendigkeit einer anderen Lösung für das Monitoring klar untermauert, was uns direkt zu Test 2 führte.

5.1.2 Test 2: Hybrid-Setup mit Pro Tools (HDX)

Im zweiten Test wurde überprüft, ob ein hybrides Verfahren mit Pro Tools zu besseren Ergebnissen führt. Hierbei wurde das Monitoring über Pro Tools HDX umgesetzt und Sequoia parallel für die Aufnahme verwendet.

Das Mikrofonsignal wurde dazu gesplittet: Über ein Doppelrouting wurden 32 Inputs gleichzeitig in Sequoia und in Pro Tools eingespeist. Eine Pro Tools Session wurde mit 32 Mono-Audiospuren und einer Stereo-AUX-Spur eingerichtet, die das Stereosummen-Signal von Sequoia im Falle eines Kopfhörer-Zuspielens aufnimmt. In Pro Tools wurde auf allen Kanälen das Monitoring aktiviert und ein Post-Fader Send auf einen Kopfhörer-Ausgang gelegt. Das bedeutet: Der Kopfhörermix konnte in Pro Tools über die Hauptfader der 32 Kanäle plus dem Stereo-AUX erstellt werden. In Sequoia wurden parallel dieselben 32 Spuren aufgenommen. Somit konnten in Sequoia Plugins für die Mastermischung unabhängig vom ASIO-Puffer geladen werden, während der Musiker die Signale aus Pro Tools abhörte. Es wurden zwei Betriebsmodi in Pro Tools verglichen:

Mit „DSP Disabled“ arbeitet Pro Tools im Native Mode wie eine native DAW ohne DSP-Beschleunigung. Der Puffer wurde in Pro Tools auf 64 Samples gestellt, da mit der verwendeten HDX-Hardware kein kleinerer Puffer im Native Mode möglich war.

„DSP Enabled“ schaltet Pro Tools in den Hybrid Engine Modus: Dafür wurden alle Inputs auf "DSP-Mode" geschaltet. Das bewirkt, dass ausschließlich unterstützte Plugins (AAX DSP) berücksichtigt werden und das Signal direkt auf der HDX-Karte latenzarm geroutet wird (Rodgers, 2021).

Zusätzlich wurde im DSP-Enabled Zustand getestet, wie sich unterstützte Plugins auf die Latenz auswirken: Dazu wurde exemplarisch ein 3-Band EQ (Oxford EQ 3 Band) auf allen Spuren in Pro Tools inseriert. So konnte ermittelt werden, ob durch Signalverarbeitung eine Latenz entsteht.

Für die Tests wurde in einem ähnlichen Verfahren wie im ersten Test (Kapitel 5.1.1) die effektive Latenz ermittelt, indem ein perkussives Testsignal und ein Sinuston in mehreren Durchläufen bei unterschiedlichen Pro Tools Puffereinstellungen aufgezeichnet wurden. Für eine genauere Analyse der Latenz wurden diesmal und für alle weiteren Tests jedoch das originale Testsignal sowie das zeitverzögerte Kopfhörersignal getrennt über zwei Testmikrofone in Cubase aufgezeichnet. Diese Methode ermöglichte auch bei sehr kurzen Latenzen unterhalb von 100 Samples eine exakte Trennung der perkussiven Signale, während bei der Verwendung eines einzelnen Mikrofons wie im ersten Test die zeitliche Überlagerung der Signale eine genaue Trennung erschwert hätte. Zur gerichteten Aufnahme des verzögerten und unverzögerten Signals kamen Neumann KM 140 Mikrofone mit AK 40 Nierenkapseln zum Einsatz.

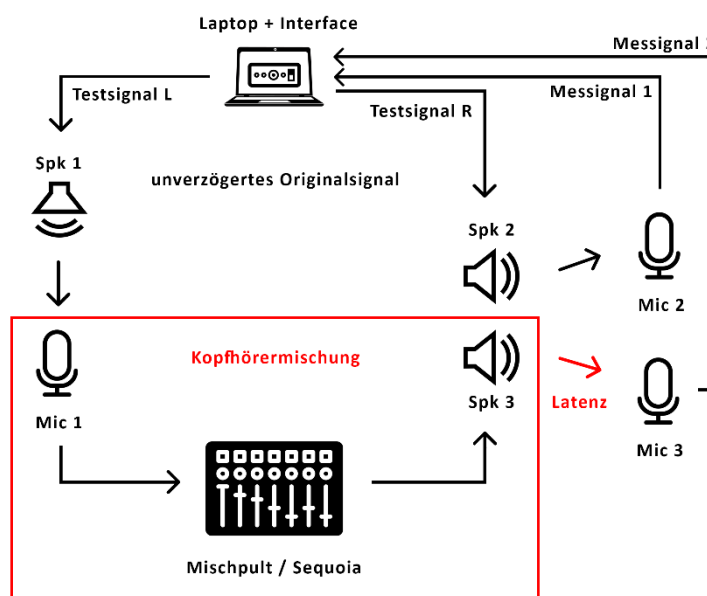


Abbildung 3: Testaufbau „Latenzmessung über zwei Mikrofone“

Test 2 zeigte sehr interessante und ermutigende Resultate bezüglich Latenz und Stabilität des hybriden Ansatzes: Im Pro Tools DSP Disabled Modus fungierte Pro Tools im Prinzip ähnlich wie Sequoia im Alleinbetrieb, nur dass 64 Samples Puffer die untere mögliche Grenze der HDX-Hardware war. Die gemessene Latenz betrug ca. 4,2 ms (≈ 200 Smp), bei 128 Samples etwa 7,1 ms (≈ 340 Smp) und bei 256 Samples rund 12,5 ms (≈ 600 Smp) und war verglichen zu den ASIO-Puffern von Sequoia minimal niedriger. Problematisch waren ebenso die Audio-Dropouts bei niedrigen Puffern: Während eines mehrspurigen Tests mit 32 Mikrofonsignalen bei 64 Samples kam es zu 2 deutlichen Aussetzern im Audiosignal. Diese manifestierten sich als kurze Stotterer im Sinustest. Dies deckt sich mit

dem Verhalten, das wir auch in Sequoia bei niedrigen Puffern gesehen hatten. Ohne DSP-Unterstützung ist auch Pro Tools nicht in der Lage, die geforderte Stabilität bei niedriger Latenz zu gewährleisten. Somit bestätigte sich, dass reine Native-Lösungen in diesem Szenario riskant sind.

Sobald der Hybrid Engine Modus auf allen Spuren aktiviert wurde, änderten sich die Ergebnisse deutlich. Die Monitoring Latenz lag bei ca. 1,7 ms (≈ 80 Smp) – unabhängig vom Hardware-Puffer (64, 128, 256, 512 oder 1024 Samples) war kein Latenzunterschied spürbar und messbar. Zudem traten keinerlei Dropouts oder Fehler bei der Aufnahme auf. Offensichtlich entlastete die HDX-Karte den Rechner so stark, dass selbst mit hohen Puffergrößen das Monitoring stabil blieb. Das Recording in Sequoia lief parallel ohne Fehler bei einem stabilen ASIO-Puffer weiter. Auch mit aktivierten DSP-Plugins (ein 3-Band Oxford EQ auf allen 32 Pro Tools-Kanälen) blieb die Latenz nahezu identisch (ca. 1,7 ms). Beim Hinzufügen eines Oxford Dynamics Kompressors auf allen Spuren stieg die Latenz auf rund 2,3 ms (≈ 110 Smp) an sowie beim weiteren Hinzufügen eines Oxford Dynamic EQ auf etwa 3,5 ms (≈ 170 Smp). Die HDX hat insgesamt also genug Reserve, um ein realistisches Processing ohne spürbare Verzögerung einzuschleusen.

Tabelle 2: Auswertung Latenztest 2 – Hybrid-Setup mit Pro Tools (HDX)

H/W Puffer	Pro Tools (native)	Pro Tools (HDX) / + EQ	+ Dynamic Comp	+ Dynamic EQ
64 Samples	4,2 ms (≈ 200 Smp)	1,7 ms (≈ 80 Smp)	2,3 ms (≈ 110 Smp)	3,5 ms (≈ 170 Smp)
128 Samples	7,1 ms (≈ 340 Smp)	1,7 ms (≈ 80 Smp)	2,3 ms (≈ 110 Smp)	3,5 ms (≈ 170 Smp)
256 Samples	12,5 ms (≈ 600 Smp)	1,7 ms (≈ 80 Smp)	2,3 ms (≈ 110 Smp)	3,5 ms (≈ 170 Smp)

Dropouts / Nicht akzeptable Latenz

Akzeptable Latenz

In diesem Test zeigte sich, dass eine Kopfhörermischung in Pro Tools latenzarm möglich ist und zu guten Ergebnissen führt, während die eigentliche Aufnahme und Mischung in Sequoia bei einem stabilen ASIO-Puffer erfolgt.

5.1.3 Test 3: Sequoia mit DADman und TotalMix

Im dritten Test wurde nun die Latenz für den Musiker/Sprecher in einem DAW-unabhängigen Setup getestet. Es wurden die Monitoring-Optionen über DADman mit dem MTRX-Interface sowie der zusätzliche Einsatz von RME TotalMix für Hardware-Monitoring inklusive möglicher Echtzeiteffekte verglichen. Ziel war es, praxisnahe Aussagen über Verzögerungen und Systemstabilität zu treffen und mögliche Vor- und Nachteile der beiden Monitoring-Methoden klar herauszuarbeiten.

Im ersten Szenario („DADman und MTRX ohne TotalMix“) wurde das Signal zunächst direkt über das Musikermikrofon in eine MTRX eingespeist, dort A/D-gewandelt und direkt auf den Ausgang geroutet. Während für das reale Studiokonzept eine große MTRX-Einheit mit mehreren analogen Ausgängen vorgesehen war, wurde in diesem Test aufgrund des vorhandenen Setups das digitale Signal via MADI-Koaxialverbindung in eine zweite MTRX geführt, von der das Signal anschließend auf einen Kopfhörerausgang geroutet wurde. Die Steuerung und Anpassung der Kopfhörerpegel erfolgten mittels DADman-Software. Sequoia lief hierbei im Mixer FX Monitoring/Hybrid Engine Modus. Das Problem hierbei war, dass DADman keine Plugins unterstützt und z. B. Trittschallfilter oder eine Equalizeranpassung für den Musiker nicht im DADman vorgenommen werden kann. Aus diesem Grund wurde eine weitere Lösung getestet.

Im zweiten Szenario („DADman und MTRX mit TotalMix“) durchlief das Mikrofonsignal ebenfalls zuerst die MTRX-Wandlung, wurde jedoch anschließend optisch an ein RME MADI FX Interface weitergeleitet, das sich über die Software TotalMix steuern lässt. Dort können zusätzliche EQ- und Kompressions-Effekte in Echtzeit eingesetzt werden, bevor das Signal schließlich zurück auf die zweite MTRX und schließlich auf den Kopfhörerausgang weitergeroutet wurde. Dieses Verfahren ermöglicht es, latenzarmes Hardware-Monitoring mit DSP-Effekten für Musiker zu realisieren, was insbesondere für anspruchsvolle Künstler von großer Bedeutung sein könnte. Sequoia lief hierbei im Hardware Monitoring/Hybrid Engine Modus, um den TotalMix über die Sequoia Fader und somit über die Avid S6 fernsteuern zu können.

Die Ergebnisse des Delaytests mit DADman und MTRX zeigten eine konsistente und äußerst geringe Verzögerung von rund 70 Samples, was etwa 1,46 ms bei 48 kHz entspricht. Dabei blieb die Latenz unabhängig von der verwendeten ASIO-Puffergröße in Sequoia stabil. Dies war zu erwarten, da das Monitoring-Signal des Musikers nicht den Signalweg durch die DAW durchläuft.

Der Einsatz von TotalMix führte zu einer minimal erhöhten Latenz. Ohne Effekte betrug die Latenz ca. 1,77 ms (≈ 85 Smp), und mit aktivierten EQ- und Kompressor-Effekten auf allen Spuren in TotalMix erhöhte sich die Latenz marginal auf etwa 1,79 ms (≈ 87 Smp). Diese Werte waren ebenfalls unabhängig von größeren ASIO-Puffern wie 4096 Samples und immer noch deutlich unterhalb kritischer Wahrnehmbarkeit.

Tabelle 3: Auswertung Latenztest 3 – Sequoia mit DADman und TotalMix

H/W Puffer	DADman	DADman + TotalMix	DADman + TotalMix (EQ & Comp)
64-4096 Samples	1,46 ms (≈ 70 Smp)	1,77 ms (≈ 85 Smp)	1,79 ms (≈ 87 Smp)

Akzeptable Latenz

Aus Sicht des Musikers ist das Monitoring über DADman/MTRX oder mit TotalMix nahezu latenzfrei und damit ideal geeignet, um ein natürliches und unverfälschtes Musikerlebnis während der Aufnahme zu gewährleisten. Gerade beim Einsatz von TotalMix bietet sich der Vorteil, in nahezu Echtzeit moderate Klangbearbeitung durchführen zu können, um den Künstlern eine individuell angepasste Klangqualität anzubieten.

Allerdings kann TotalMix nicht direkt über das von Avid unterstützte EUCON-Protokoll gesteuert werden. Um eine Fernsteuerung über die Avid S-Konsole zu ermöglichen, kann Sequoia in den Hardware Monitoring-Modus umgeschaltet werden. In diesem Modus übernimmt Sequoia das Monitoring über die verwendete Audiokarte, wodurch eine Fernsteuerung von TotalMix über die Avid S-Konsole möglich wird. Ein wesentlicher Nachteil ist jedoch, dass im Hardware Monitoring-Modus nicht die interne Mischung von Sequoia abgehört werden kann. Stattdessen wird die TotalMix-Konfiguration – und damit das Kopfhörermonitoring – zur Regie zurückgeführt, was eine korrekte Beurteilung der eigentlichen Mischung erschwert. Alternativ kann Sequoia in den Mixer FX Monitoring-Modus zurückgeschaltet werden, um die Mischung von Sequoia abzuhören. Dies geht allerdings zulasten der Steuerbarkeit des Kopfhörermonitorings, da die Fader der Avid S-Konsole in diesem Fall nicht mehr indirekt für TotalMix, sondern für die internen Sequoia-Fader genutzt werden können. Anpassungen am Monitoring müssen dann manuell per Maus über die Benutzeroberfläche von TotalMix vorgenommen werden.

Letztlich ist die Wahl des passenden Modus eine Frage der Prioritätensetzung: Minimale Latenz ohne Klangeffekte, aber dafür bessere Steuerungsmöglichkeit über die Avid S-Konsole (DADman/MTRX) oder niedrige Latenz mit zusätzlichen Möglichkeiten der

Klanggestaltung, aber weniger Flexibilität bei der Steuerung durch eine Avid S-Konsole (TotalMix).

5.1.4 Test 4: Latenzausgleich beim Overdub

Im vierten Test wurde schließlich überprüft, welche Latenzen bei der Verwendung von Zuspielern aus Sequoia in der Regie während der Aufnahme entstehen und ob ein anschließender Latenzausgleich stattfindet.

In diesem Szenario würde in einer Realsituation dem Musiker ein Zuspielsignal aus Sequoia auf den Kopfhörer geroutet werden, auf welches er dann ein präzises Overdubbing einspielt. Dabei spielt das Kopfhörer-Monitoring für den Musiker keine Rolle, da er in diesem Setup über DADman oder TotalMix abhören könnte (siehe Kapitel 5.1.3). In der Regie hingegen entstehen während der Aufnahme beim Abhören von Sequoia aufgrund von ASIO-Puffer und Insert-Effekten Latenzen zwischen dem Zuspielsignal und dem Overdub-Signal. Diese Latenz galt es im folgenden Test herauszufinden und in einem weiteren Durchlauf zu messen, ob nach Stoppen der Aufnahme ein Latenzausgleich stattfindet.

Hierfür wurde das bereits in den vorherigen Versuchen verwendete perkussive Testsignal in Sequoia auf den Kopfhörerausgang geroutet und gleichzeitig in der Regie wiedergegeben. Das Kopfhörersignal wurde über einen Mini-Monitor im Aufnahmerraum wiedergegeben und über ein Musiker-Mikrofon erneut auf eine weitere Spur aufgenommen. Der Gedanke war, dass ein Musiker im Idealfall sein Overdubbing exakt im Timing zum Kopfhörersignal einspielt. In der Regie wurden dann das aufgenommene Overdub-Signal und das zeitunverzögerte Zuspielsignal gemeinsam wiedergegeben. Zur Messung der Zeitdifferenz zwischen Wiedergabe und Aufnahme wurde in Sequoia das Zuspielsignal hart auf den rechten Abhörkanal und das Overdubsignal auf den linken Kanal gepannt. Wie bereits im dritten Test kamen zwei gerichtete Neumann-Testmikrofone zum Einsatz, die jeweils direkt auf den linken bzw. rechten Monitor ausgerichtet waren. Die resultierende zweispurige Aufnahme in Cubase ermöglichte eine präzise Auswertung des Versatzes zwischen beiden Signalen – und somit eine exakte Bestimmung der Latenz während der Aufnahme sowie beim anschließenden Abspielen in Sequoia.

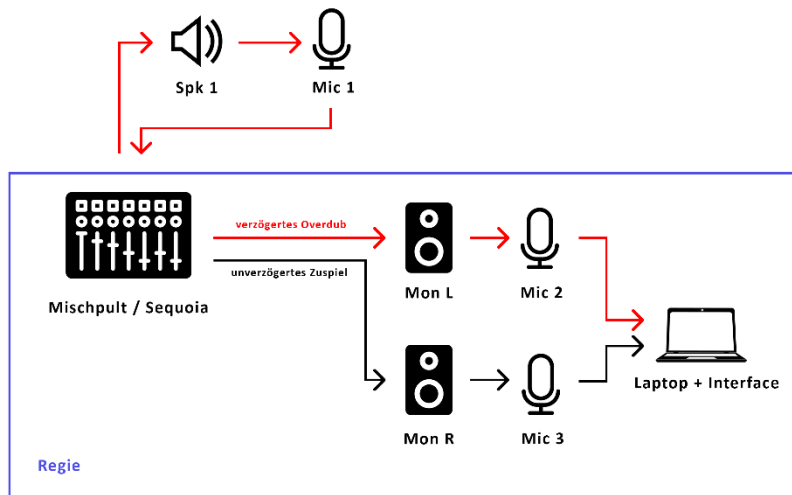


Abbildung 4: Testaufbau „Latenzmessung Overdub“

Außerdem wurde mit verschiedenen Puffern und Spurenzahlen experimentiert sowie Plugins auf der Overdub-Spur eingefügt. Dadurch konnte beobachtet werden, wie sich die Latenz in unterschiedlichen Konstellationen verhält und ob es Auswirkungen auf den Latenzausgleich nach der Aufnahme gibt. Zu bemerken war, dass während der Aufnahme und dem Abhören von Sequoia ein Versatz zwischen Zupspiel- und Aufnahmesignal hörbar war. Dieser Versatz wurde nach dem Stoppen und dem erneuten Abspielen unabhängig der verwendeten Einstellungen nahezu ausgeglichen. Die Latenz-Messung lieferte konkrete Zahlen:

Beim Abhören über Sequoia während der Aufnahme von einer Spur ohne Plugins zeigte sich bei einem ASIO-Puffer von 64 Samples eine gemessene Latenz von etwa 6 ms (≈ 290 Smp). Mit zunehmender Puffergröße stieg die Latenz erwartungsgemäß an: Bei 128 Samples lag sie bei ca. 8,7 ms (≈ 420 Smp), bei 256 Samples bei rund 14 ms (≈ 680 Smp) und bei 512 Samples bei ca. 24 ms (≈ 1160 Smp). Bei größeren Pufferwerten – 1024 Samples bzw. 2048 Samples – lagen die ermittelten Latenzen bei rund 46 ms (≈ 2210 Smp) bzw. ca. 88ms (≈ 4250 Smp).

Darüber hinaus wurde untersucht, inwieweit sich die Anzahl der Spuren sowie das Einfügen von Insert-Effekten auf die Latenz auswirkt. Selbst bei einer Erweiterung auf 32 Spuren, gleichzeitiger Aufnahme aller Spuren sowie dem Einsatz von drei ressourcenschonenden Effekten – EQ112, AM-Track und FabFilter Pro-Q3 – blieben die Messergebnisse quasi identisch. Erst durch das Hinzufügen eines vierten, deutlich rechenintensiveren Effekts – iZotope Ozone 11 mit dem Preset „Ozone Vintage Mix“ – kam es zu signifikanten Latenzsteigerungen. Bei einem ASIO-Puffer von 64 Samples betrug die gemessene Latenz

ca. 72 ms (≈ 3.450 Smp), bei 128 Samples etwa 75 ms (≈ 3.580 Smp) und bei 256 Samples rund 81 ms (≈ 3.850 Smp). Diese Ergebnisse bestätigen, dass sich Plugins grundsätzlich auf die Latenz auswirken, wobei diese in praxisnahen Szenarien mit typischen ressourcenschonenden Insert-Effekten wie den obigen vernachlässigbar ist.

Tabelle 4: Auswertung Latenztest 4 – Latenzausgleich beim Overdub

H/W Puffer	Ohne Effekte	EQ, Comp, Pro-Q3	EQ, Comp, Pro-Q3, iZotope Ozone 11
64 Samples	6 ms (≈ 290 Smp)	6 ms (≈ 290 Smp)	72 ms (≈ 3.450 Smp)
128 Samples	8,7 ms (≈ 420 Smp)	8,7 ms (≈ 420 Smp)	75 ms (≈ 3.580 Smp)
256 Samples	14 ms (≈ 680 Smp)	14 ms (≈ 680 Smp)	81 ms (≈ 3.850 Smp)
512 Samples	24 ms (≈ 1160 Smp)	24 ms (≈ 1160 Smp)	96 ms (≈ 4.600 Smp)

Im Rahmen der Überprüfung des Latenzausgleichs in Sequoia wurde festgestellt, dass unabhängig von Puffer, Spurenzahl und Anzahl der Plugins nach der Aufnahme eine Restlatenz von etwa 1,94 Millisekunden (≈ 93 Smp) bestehen blieb. Unsere Vermutung war, dass diese Restlatenz durch die D/A-A/D-Wandlung und Kabelwege entsteht, also physikalisch bedingt ist. Das System selbst gleicht alle restlichen Latenzen offenbar aus – Sequoia rückt die aufgenommenen Spuren beim Abschluss der Aufnahme so, dass sie im Projekt nahezu synchron liegen. Auf die Frage, ob es eine Möglichkeit gibt, diese Hardwarelatenz per negativem Delay auszugleichen, bestätigte uns der Hersteller MAGIX, dass diese Restlatenz vom System kommt und unabhängig der internen Puffereinstellungen nur per negativem Spurdelay oder händischem Synchronisieren eliminiert werden kann. Die 93 Samples Differenz sind allerdings am Ende so gering, dass sie beim Podcast oder der Kammermusik praktisch vernachlässigbar sind, da ein so kleiner Versatz z. B. durch Mikrofonabstand leicht entstehen kann und weit unterhalb der kritischen Schwelle liegt

Allerdings offenbarte der Test auch eine praktische Herausforderung im Produktionsalltag: Beim Abhören über Sequoia während der Aufnahme ist es für die Tonregie schwierig, die Synchronität und Präzision eines Takes zuverlässig zu beurteilen, da das Monitoring über den latenzkompensierten Signalpfad erfolgt. Im Versuch zeigte sich, dass selbst bei der

Verwendung von ressourcenschonenden Plugins eine hörbare Latenz in der Regie entsteht, obwohl die Aufnahme exakt im Timing war.

Daraus ergibt sich ein Zielkonflikt: Für den Tonmeister ist es ungünstig, bei der Aufnahme ein Timing zu beurteilen, das durch Latenz verschoben ist – selbst wenn es durch die DAW nachträglich korrekt kompensiert wird. Als praktikable Alternative zeigte sich in der Testumgebung, dass beim Monitoring über DADman oder TotalMix – also außerhalb von Sequoia – keine hörbare Latenz auftrat, wodurch der Tonmeister zwar nur den „rohen“ Mix ohne Echtzeit-Postprocessing hörte, jedoch mit einem zeitlich korrekten Klangbild.

Es ergibt sich folgende Abwägung für die Praxis: Möchte man während der Aufnahme bereits die eigentliche Mischung hören (z. B. mit EQ, Kompression oder Raumsimulation aus Sequoia), muss man eine gewisse Monitoring-Latenz in Kauf nehmen. Steht hingegen die zeitliche Präzision der Performance im Vordergrund, ist ein latenzfreies Monitoring – z. B. über DADman oder TotalMix – zu bevorzugen.

Mit diesen Messwerten im Gepäck konnte in der Diskussion im nächsten Unterkapitel nun sehr gezielt argumentiert werden, welche Konfiguration die Anforderungen erfüllt und wie man mit den identifizierten Hürden umgehen kann.

5.2 Diskussion der Ergebnisse

Im Lichte der empirischen Befunde sollen in diesem Kapitel die Ergebnisse kritisch diskutiert und in den Gesamtkontext eingeordnet werden.

Die nachfolgende Tabelle liefert eine Übersicht über die erzielten Testergebnisse des Kopfhörermonitorings im praktischen Einsatz von Sequoia zusammen mit verschiedenen Signalführungen. Untersucht wurden Konfigurationen mit Sequoia im Alleinbetrieb, Pro Tools als vorgeschaltetem Mischpult – sowohl ohne als auch mit aktiviertem DSP-Modus über HDX – sowie Varianten mit Hardware-Monitoring direkt über die MTRX oder durchgeschleift über TotalMix. Die jeweils verwendeten Puffereinstellungen sind in der linken Spalte aufgeführt. Die Messergebnisse werden in Millisekunden und Samples dargestellt. Farblich hervorgehoben sind dabei rot die Ergebnisse mit Audioaussetzern und Aufnahmefehlern, blau die Werte, die aufgrund ihrer Latenz (über 5 ms) außerhalb des akzeptablen Bereichs liegen, und grün die stabilen Ergebnisse ohne Dropouts und mit akzeptabler Latenz.

Tabelle 5: Gegenüberstellung der Testergebnisse aus Kapitel 5.1

H/W Puffer	Sequoia	Pro Tools (native)	Pro Tools (HDX)	DADman	DADman + TotalMix
32 Samples	4,8 ms (≈ 230 Smp)			1,5 ms (≈ 70 Smp)	1,8 ms (≈ 85 Smp)
64 Samples	6,3 ms (≈ 300 Smp)	4,2 ms (≈ 200 Smp)	1,7 ms (≈ 80 Smp)	1,5 ms (≈ 70 Smp)	1,8 ms (≈ 85 Smp)
128 Samples	9 ms (≈ 430 Smp)	7,1 ms (≈ 340 Smp)	1,7 ms (≈ 80 Smp)	1,5 ms (≈ 70 Smp)	1,8 ms (≈ 85 Smp)
256 Samples	11,9 ms (≈ 570 Smp)	12,5 ms (≈ 600 Smp)	1,7 ms (≈ 80 Smp)	1,5 ms (≈ 70 Smp)	1,8 ms (≈ 85 Smp)

Dropouts / Nicht akzeptable Latenz

Akzeptable Latenz

Die vier durchgeführten Tests erlauben einen direkten Vergleich verschiedener Monitoring-Ansätze hinsichtlich Latenz, Systemstabilität und Praxistauglichkeit. Insgesamt zeigt sich dabei deutlich, warum rein softwarebasierte DAW-Lösungen an Grenzen stoßen und welche Vorteile hybride oder DSP-gestützte Monitoring-Ansätze bieten.

Test 1 (reines Sequoia-Setup) hat verdeutlicht, dass eine native DAW zwar theoretisch sehr geringe Monitoring-Latenzen erreichen kann, allerdings nur mit kleinen Puffern und entsprechend instabiler Leistung. Bei einem minimalen ASIO-Puffer von 32 Samples (≈ 4,8 ms Latenz) kam es zu hörbaren Effekten (leichter Phasing/Flanger-Effekt) und Aufnahmefehlern. Dieser Wert überschreitet bereits die in Kapitel 3.4 ermittelte unkritische Latenzschwelle von etwa 3,75 ms. Erst ab 128 Samples (≈ 9 ms Latenz) lief das System zuverlässig, diese Verzögerung liegt jedoch deutlich oberhalb dessen, was anspruchsvolle Künstler als unproblematisch empfinden. Die Befunde machen klar, dass im reinen DAW-Betrieb ein Zielkonflikt besteht: Entweder wählt man einen kleinen Puffer für geringe Latenz und riskiert Systeminstabilität, oder man erhöht den Puffer für stabile Aufnahmen, was aber eine höhere Latenz zur Folge hat. In der Praxis einer langen Aufnahme-Session (z. B. mehrstündige Produktionen) wäre es gemäß den Ergebnissen geboten, sogar einen noch größeren Puffer (256 Samples oder mehr) einzusetzen, um genügend Reserve zu haben – was die Latenz weiter erhöhen würde. Damit zeigt Test 1 die Grenze des rein softwarebasierten Monitorings deutlich auf: Unter realistischen Bedingungen ist eine latenzarme und gleichzeitig stabile Abhörsituation allein mit Sequoia nicht zu erreichen.

Im Gegensatz dazu zeigte Test 2 (hybrides Setup mit Pro Tools HDX), wie ein DSP-gestützter Ansatz die Latenzproblematik spürbar reduziert. Durch die Auslagerung des Kopfhörermonitorings auf die HDX-Hardware bei paralleler Aufnahme in Sequoia und Pro Tools sank die effektive Monitoring-Latenz auf etwa 1,7 ms (≈ 80 Smp) – ein Wert, der

deutlich unterhalb der kritischen Latenzschwelle von 3,75 ms liegt. Entscheidend ist, dass diese Latenz unabhängig von der eingestellten Hardware-Puffergröße in Sequoia war. Auch das Hinzufügen einfacher DSP-Effekte im HDX-Mixer (hier ein 3-Band-EQ auf allen Kanälen) veränderte die Latenz nicht merklich. Erst deutlich rechenintensivere DSP-Plugins (wie zusätzliche Dynamics-Module) ließen sie geringfügig auf $\approx 2,3$ ms bzw. maximal $\approx 3,5$ ms ansteigen, was immer noch innerhalb des unkritischen Bereichs liegt. Diese Ergebnisse sind ermutigend, da sie belegen, dass ein hybrides System die Vorteile beider Welten vereint: Der Musiker erhält ein praktisch latenzfreies Monitoring, während die eigentliche Aufnahme weiterhin in Sequoia stattfinden kann – dort mit einem großzügigen Puffer, der für Systemstabilität sorgt und die DAW auch bei vielen Spuren oder Plug-ins nicht überlastet. Zudem bietet ein weiteres Pro Tools System eine vollwertige Backup-Möglichkeit durch die Verwendung von zwei unterschiedlichen DAWs auf getrennten Rechnern. Allerdings bringt dieses Setup auch praktische Implikationen mit sich: Die parallele Nutzung zweier DAWs führt zu einem erhöhten Administrations- und Schulungsaufwand, da sowohl das doppelte Routing aller Eingangssignale als auch die Bedienung beider Systeme zusätzliche Kompetenzen erfordert. Darüber hinaus steigen die Gesamtkosten durch die Anschaffung und den Betrieb eines weiteren Systems deutlich an, sodass die Wirtschaftlichkeit dieser Variante kritisch hinterfragt werden muss. Insgesamt zeigte Test 2 aber, dass der Nutzen der hybriden Variante mit Pro Tools diesen Ansatz zu einer aussichtsreichen Option für die Studioerneuerung macht.

Test 3 (Monitoring über DADman/MTRX und RME TotalMix) führte die Idee des hardwarebasierten Monitorings noch einen Schritt weiter und untersuchte ein vollständig DAW-unabhängiges Abhörsystem. Hier wurden die eingehenden Mikrofondsignale direkt in der Audio-Hardware (Avid MTRX) geroutet und auf den Kopfhörer gegeben – zunächst ohne jegliche zusätzliche Signalbearbeitung, anschließend mit Einbindung von RME TotalMix, um DSP-Effekte bereitzustellen. Die Messergebnisse untermauern erneut die Vorteile eines solchen Ansatzes: Ohne TotalMix lag die Latenz bei konstant ≈ 70 Samples ($\approx 1,46$ ms bei 48 kHz) und war quasi nicht wahrnehmbar. Wie zu erwarten, blieb diese Verzögerung vom Sequoia ASIO-Puffer vollkommen unberührt, da das Monitoring-Signal den Signalweg über die DAW umgeht. Mit dem Einsatz von TotalMix erhöhte sich die Latenz geringfügig auf ca. 1,77 ms (ohne Effekte) bzw. 1,79 ms (mit zugeschaltetem EQ und Kompressor auf allen Kanälen). Dieser Unterschied ist praktisch vernachlässigbar – das Monitoring bleibt auch mit leichter Klangbearbeitung unter 2 ms Latenz und damit deutlich unterhalb der kritisch wahrnehmbaren Schwelle.

Allerdings brachte Test 3 auch operative Herausforderungen zutage, die für die Praxistauglichkeit relevant sind. Beim reinem DADman/MTRX-Monitoring ohne TotalMix

fehlte die Möglichkeit, dem Musiker Effekte im Monitorsignal bereitzustellen (die MTRX-Routingmatrix verarbeitet nur das direkte Mikrofonsignal). TotalMix schließt diese Lücke, führt aber ein Steuerungsproblem ein: Die RME-Software lässt sich nicht ohne weiteres in eine Regie-Bedienoberfläche mit einer Avid S-Konsole integrieren, da TotalMix das EUCON-Protokoll nicht unterstützt. In der Testumgebung wurde dies durch Sequoias Hardware-Monitoring-Modus umgangen, sodass die Sequoia-Fader indirekt die TotalMix-Mischung steuern konnten. Dieses Konstrukt funktioniert technisch, hat aber den Nachteil, dass die Regie in diesem Modus nicht die Summe aus Sequoia abhören kann, sondern die vom Hardware-Mischer zurückgeführte Kopfhörermischung. Um die DAW-Summe abzuhören, müsste man Sequoia wieder in den normalen Mixer-Modus (Mixer-FX Monitoring) schalten, was jedoch die Fernsteuerbarkeit der Kopfhörermischung aufhebt. In diesem Fall bliebe lediglich die manuelle Anpassung von TotalMix per Maus, was sich im laufenden Aufnahmebetrieb als unpraktisch erweist. In einer zukünftigen Studio-Installation muss für dieses Problem eine praktikable Lösung gefunden werden – etwa durch klare Arbeitsabläufe, in welchen Phasen der Produktion welcher Modus genutzt wird, technische Erweiterungen wie separate Controller, die eine Steuerung von TotalMix ermöglichen, oder in einfachen Fällen das Steuern von TotalMix per Maus. Trotz dieser Hürden bleibt festzuhalten, dass das hardwarebasierte Monitoring in Bezug auf Latenz und Stabilität überzeugt hat: Es bietet die gleichen Latenz-Vorteile wie das HDX-System aus Test 2 mit eingeschränkten, für ein Kopfhörermonitoring aber wahrscheinlich ausreichenden DSP-Einsatzmöglichkeiten, und umgeht vollständig die Schwächen eines rein nativen Workflows. Somit stellt diese Konfiguration insgesamt eine überzeugende Option dar.

Ein spezieller Aspekt der Praxistauglichkeit wurde in Test 4 (Latenzausgleich beim Overdub) untersucht: Hier stand weniger die Latenz für den Musiker im Fokus (diese ließ sich durch Hardware-Monitoring minimieren), sondern die Frage, wie sich Latenzen während einer Aufnahme auf die Regie-Beurteilung auswirken und wie gut Sequoia nachträglich die Aufnahmen synchronisiert. Der Versuchsaufbau simulierte ein typisches Overdub-Szenario, bei dem ein Musiker zu einem bereits aus der DAW kommenden Zuspieldesignal (Playback) etwas einspielt. Während der Musiker das Kopfhörersignal nahezu latenzfrei hört, kann die Tonregie die Mischung aus Playback und dem gerade eingehenden Overdub entweder direkt über die DAW abhören oder ebenfalls über einen externen Mischer. Im Test 4 wurde das Szenario betrachtet, bei dem die Regie das Live-Signal über Sequoia abhört, also mit aktivem DAW-Pfad und dessen Latenz. Die Ergebnisse zeigten deutlich, dass während der Aufnahme ein hörbarer Versatz zwischen dem Playback und dem neu eingespielten Signal auftreten kann, wenn die Regie rein softwarebasiert abhört: Je nach Puffergröße betrug dieser Versatz etwa 6 ms (bei 64 Samples Puffer), ≈ 9 ms (128 Samples) bis hin zu zweistelligen Millisekundenwerten bei größeren Puffern (z. B. ≈ 24 ms).

bei 512 Samples). Selbst bei optimierten Einstellungen (64–128 Samples ohne Effekte) war also in der Regie ein kleiner Zeitunterschied hörbar, der die Beurteilung der Timing-Präzision des Overdubs erschwert. Wichtig ist hierbei die Feststellung, dass Sequoia nach Beenden der Aufnahme praktisch alle Verzögerungen automatisch ausgleicht und die aufgenommene Spur zeitsynchron zum Playback verschiebt. Lediglich eine minimale Restlatenz von ca. 1,9 ms (≈ 93 Smp) blieb bestehen – vermutlich bedingt durch die unvermeidliche Wandlungsverzögerung (D/A am Kopfhörerausgang und A/D am Mikrofoneingang) und Kabelwege. Diese Abweichung ist so gering, dass sie in musikalischen Anwendungen kaum ins Gewicht fällt und notfalls manuell durch ein negativ eingestelltes Spur-Delay korrigiert werden könnte.

Dennoch offenbart Test 4 das praktische Problem, dass der Tonmeister während laufender Aufnahme und Abhören über Sequoia nicht sicher das Timing des Musikers beurteilen kann, da er einen latenzbehafteten Mix hört. So entstand im Versuch der Eindruck eines leicht verzögerten Spiels – obwohl die Aufnahme hinterher exakt im Takt lag, was im Produktionsalltag kritisch ist. Abhilfe schafft auch hier der hybride/DSP-Ansatz: Würde man das Overdub-Szenario komplett latenzfrei abhören – z. B. indem man das Playback-Signal ebenfalls über das Hardware-Monitoringsystem an den Regieraum schickt und dort mit dem direkt abgehörten Livesignal mischt – ließe sich dieser Versatz eliminieren. In der Praxis könnte das bedeuten, dass die Regie für kritische Einspielpassagen auf den gleichen Kopfhörermix (oder einen entsprechenden Abzweig daraus) umschaltet, den auch der Musiker erhält. Dieses Vorgehen müsste in einem Studiokonzept mit eingeplant werden, bestätigt aber erneut den Vorteil flexibler Hardware-Monitoringlösungen: Sie ermöglichen nicht nur dem Künstler, sondern auch dem Tonmeister eine fast latenzfreie Hörkontrolle.

Zusammenfassend liefern die empirischen Ergebnisse aus Kapitel 5 eine solide Entscheidungsgrundlage für das weitere Vorgehen. Sie unterstützen die bereits im SWR diskutierte Vorentscheidung, nicht ausschließlich auf eine softwarebasierte DAW-Lösung zu setzen, sondern ein hybrides System zu konzipieren. Konkret legen die Resultate nahe, das Beste aus beiden Welten zu kombinieren: Sequoia kann weiterhin als Haupt-DAW für Aufnahme, Schnitt und Mischung dienen, während für das Monitoring ein dedizierter, latenzarmer Signalpfad vorgesehen wird – sei es über eine HDX/DSP-Lösung oder über die direkte Nutzung der Audio-Hardware mit entsprechender Steuerungsmöglichkeit.

Die in den Tests aufgezeigten Vor- und Nachteile der jeweiligen Variante werden in die Konzepterstellung in Kapitel 8 einfließen. Dort gilt es, einen Weg zu finden, die Vorteile (Latenzfreiheit, Stabilität, Klangflexibilität) maximal zu nutzen und die Nachteile (Bedienungsaufwand, Integrationshürden, Kostenfaktor) zu minimieren.

6 Monitoring-Konzept

Ein sorgfältig geplantes und flexibles Monitoring-Konzept ist elementar für effiziente Produktionsabläufe. Dabei sind die Anforderungen in den DSZs und dem Podcast-Studio bzw. der Kammermusik geringfügig unterschiedlich. Die folgende Tabelle gibt einen Überblick über die benötigten Abhörquellen sowie deren mögliche Ansteuerung für die Verbraucher. Im Anschluss werden die unterschiedlichen Anforderungen der DSZs im Vergleich zu Kammermusik und Podcast näher erläutert.

Tabelle 6: Monitoring-Konzept

Verbraucher / Nutzer	Notwendige Abhörquellen	Ansteuerung über
Regie-Lautsprecher 1	- Mischung (Stereo, Mono, RR, LL, Mehrkanalformate bis 7.1.4) - AUX-Wege - Kopfhörer-Wege - PFL - Dante - SessionLink (o. Ä.) - Office-PC - SWR1-4 (evtl. Kammer) - Medienzuspiel wie CD, DVD, DigAS Spion (Kammer/Podcast)	Monitorcontroller (z. B. DADman in Kombination mit DAD MOM)
Regie-Lautsprecher 2	identisch zu Regie-Lautsprecher 1 (Mixkontrolle auf anderer Abhöre)	Avid S6 oder direkt in DADman
Vorhör-Lautsprecher (PFL)	- Einzelkanäle PFL - SessionLink - Medienzuspiel - Office-PC - Spion o. Ä.	Separater Lautsprecherweg (z. B. über DADman)
Kopfhörer Toningenieur	- Mischung - AUX-Wege - Quellenvorhörung - Dante - SessionLink (o. Ä.) - Medienzuspiel - Office-PC - Follow Main-Speaker (Regie-Lsp)	Kopfhörerverstärker
Kopfhörer Tonmeister	- Mischung - AUX-Wege - Quellenvorhörung - Dante - SessionLink (o. Ä.) - Medienzuspiel - Office-PC - Follow Main-Speaker (Regie-Lsp)	Kopfhörerverstärker, eigenes DAD MOM oder Stream Deck

Kopfhörer Musiker/Sprecher (Studio)	• Individuelle Mixe (AUX-Wege) • Talkback (min. 3 separate, z. B. mit Riedel & Hausnetzanbindung) • Mikrofon-Direktsignale (latenzarm) • Office-PC	Kopfhörerverstärker im Studio
Einspielung in Aufnahmerraum	• Mischung • Talkback • Medienzuspiel wie CD, DVD, DigAS	Einspiel-Lautsprecher
Office-PC	• Mischung • Talkback • Dante • Follow Main-Speaker	SessionLink o. Ä.

Im Zentrum steht die Regieabhöre, welche eine Vielzahl unterschiedlicher Quellen flexibel abhören können muss, wie die Mischung in Stereo, Mono, Summierung der rechten oder linken Kanäle und Mehrkanalformate bis 7.1.4, AUX-Wege, Kopfhörerwege, Dante, SessionLink oder ähnliche Tools, der Office-PC sowie Medienzuspiele wie CD, DVD oder DigAS. Für die Kammermusik könnte zusätzlich eine Anbindung an das Schaltraum-Signal (SWR1-4) und ein Spion für schnelle, direkte Abhörmöglichkeiten notwendig sein. Diese Quellen werden idealerweise über eine zentrale Kreuzschiene wie die MTRX verwaltet und mittels eines Monitorcontrollers wie DADman in Kombination mit DAD MOM – ein tastengesteuerter Controller für DADman – flexibel angesteuert.

Zusätzlich zur primären Regieabhöre sollte eine sekundäre Regie-Abhöre identisch ausgestattet werden, um eine kritische Mixkontrolle über separate Abhörlautsprecher zu gewährleisten. Diese zweite Abhöre kann direkt über die Avid S6-Konsole oder das DADman-System gesteuert werden.

Vorhörlautsprecher sind für einzelne Kanäle (PFL), SessionLink-Zuspierungen, Medienzuspierungen und Office-PC vorgesehen. Ein eigener, separater Lautsprecherweg, etwa über DADman, ermöglicht dabei ein unabhängiges Abhören, das den Mischvorgang nicht unterbricht.

Kopfhörermonitoring spielt insbesondere im Podcast-Studio eine wesentliche Rolle. Hier werden fast 100% der Aufnahmen mit Kopfhörermonitoring durchgeführt, anders bei der Kammermusik: Rund 90% der Aufnahmen werden ohne Kopfhörermonitoring durchgeführt, jedoch erfordern etwa 9% der Produktionen, insbesondere bei Overdubs oder Einzelaufnahmen, individuelle Kopfhörermixe. Diese Mixe müssen flexibel und latenzarm gestaltet sein und sollten direkt über die MTRX geroutet werden können. Außerdem sind

Punch-In/Out-Szenarien, wenn auch selten, bei ca. 1% der Produktionen relevant und benötigen schnelle Umschaltmöglichkeiten. Die angegebenen Werte basieren auf Schätzungen und dienen lediglich der groben Orientierung. In welchem Umfang Overdubs und Punch-In/Out-Szenarien künftig tatsächlich auftreten werden, kann in einem dynamischen Produktionsumfeld nicht vorausschauend exakt beurteilt werden.

Toningenieure und Tonmeister erhalten jeweils eigene Kopfhörerverstärker. Dabei ist es wichtig, dass der Tonmeister zusätzlich einen eigenen Controller wie ein DAD MOM oder Stream Deck erhält, um schnell zwischen der Gesamtmischung und einzelnen Musikerwegen wechseln zu können, was die Kontrolle und Bewertung einzelner Performances vereinfacht.

In den DSZs liegt der Fokus auf lokalen Quellen. Netzwerkverteilungen oder komplexe Talkbacks sind hier nicht notwendig. Ein einfaches Kommando für Remote-Kommunikation via SessionLink reicht hier aus. Dies könnte z. B. über einen Button am DAD MOM gelöst werden.

Für den Podcast-Bereich ist hingegen eine flexible Kopfhörerkommunikation entscheidend. Jeder Sprecher benötigt ggf. individuelle Mixe und Talkback-Wege, zudem müssen Einspieler und externe Remote-Teilnehmer gezielt und schnell zu- und abgeschaltet sowie individuell abgehört werden können. Daher ist hier eine externe Kommando-Lösung von Riedel empfehlenswert, die eine hohe Flexibilität und Skalierbarkeit für komplexere Produktionssituationen gewährleistet.

Für das Kommando-System steht somit eine grundsätzliche Entscheidung zwischen einer integrierten DADman-Lösung (kostengünstig, aber weniger flexibel) und einem externen Riedel-System (höhere Flexibilität, jedoch höhere Kosten und mehr Geräte) an. Aufgrund der unterschiedlichen Anforderungen könnte für Podcast die Riedel-Lösung bevorzugt werden, während in der Kammermusik eine interne DADman-Lösung ausreichend sein könnte, sofern keine umfangreiche Kommunikation zwischen verschiedenen Studios erforderlich wird.

In der Praxis treten spezifische Situationen auf, in denen die Flexibilität des Monitoring-Konzepts entscheidend ist. Beispielsweise bei einem einseitigen Kopfhörerausfall müssen Techniker den physischen Ausgang schnell überprüfen können, um Ursachen wie defekte Kabel oder Routingfehler auszuschließen. Auch bei Rückmeldungen von Remote-Teilnehmern über Verzögerungen oder Echos ist es essenziell, die einzelnen AUX-Wege separat abhören zu können, um zügig die Ursachen wie hohe Latenzen oder falsches Routing zu identifizieren und zu beheben. Ebenso ist eine schnelle und unkomplizierte Prüfung verschiedener Mischungen wie Mono-Kompatibilität, Downmixes oder diverse

Lautsprecher setups ein häufiger Anwendungsfall, der einen raschen Wechsel der Abhörquellen erfordert und im hier vorgestellten Konzept unterstützt wird.

Zusammenfassend stellt das vorgeschlagene Monitoring-Konzept eine ausgewogene Mischung aus Flexibilität, Effizienz und Wirtschaftlichkeit dar, wobei die jeweiligen Besonderheiten der Produktionsräume DSZs, Kammermusik und Podcast individuell berücksichtigt werden.

7 Ergonomie

Die ergonomische Gestaltung der Arbeitsplätze in den DSZs, im Podcast-Studio und in der Kammermusik ist entscheidend für effizientes und gesundes Arbeiten. Im Rahmen von Mitarbeiterbefragungen wurden hierzu konkrete Anforderungen, Wünsche und Vorschläge gesammelt und bewertet.

In den DSZs dominiert die computergestützte Editierarbeit. Dies erfordert eine zentrale Arbeitsfläche, auf der Maus, Tastatur, Partituren und weitere Materialien übersichtlich angeordnet werden können. Daraus ergibt sich, dass DAW-Controller, beispielsweise der Avid S1, nicht zentral, sondern seitlich oder flexibel verschiebbar positioniert sein sollten. Diese Gestaltung bietet maximale Flexibilität und ist kosteneffizient.

In der Kammermusik und dem Podcast-Studio findet die Hauptarbeit am Mischpult statt, was eine gesonderte Anordnung der Fader erforderlich macht. Im Rahmen der Planung wurden hierzu mehrere Varianten diskutiert:

Variante 1 setzt auf verschiebbare Ablageflächen, die oberhalb des im Tisch versenkten Controllers angeordnet sind. So kann je nach Bedarf entweder eine freie Arbeitsfläche genutzt oder direkt auf die Steuerungselemente zugegriffen werden. Eine vergleichbare Konstruktion wird u. a. bereits in manchen Regionen des Bayerischen Rundfunks eingesetzt. Denkbar wären außerdem verschiebbare Controller auf Schienen, die je nach Arbeitssituation flexibel positionierbar sind. Diese Variante bietet ergonomische Vorteile, birgt jedoch eine höhere mechanische Komplexität und potenziell erhöhten Wartungsbedarf.

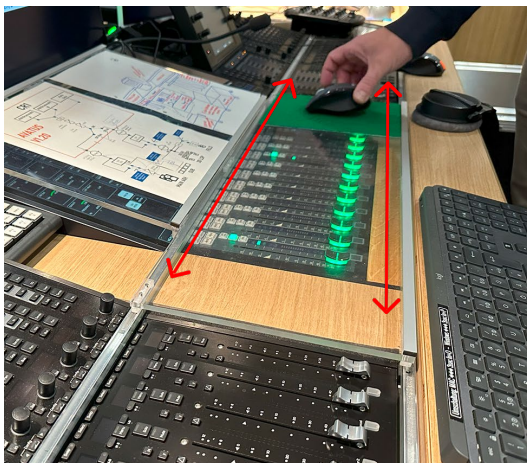
Variante 2 orientiert sich am Aufbau der PR04-Regie des SWR in Baden-Baden. Dabei sind die Fadermodule seitlich in einer leichten V-Form angeordnet, während der mittlere Bereich für Maus und Tastatur freigehalten wird. Diese Anordnung bietet eine gute Trennung der Bedieneinheiten, erfordert jedoch eine sorgfältige Abstimmung der Abstände – insbesondere, um eine komfortable Bedienung für alle Mitarbeitenden sicherzustellen. Zu große Distanzen könnten gerade für ältere Kolleginnen und Kollegen eine ergonomische Herausforderung darstellen.

Variante 3 sieht die Platzierung des Controllers auf einer zweiten Ebene oberhalb der Tischplatte vor. Die Fadermodule könnten dabei flexibel vor- und zurückgeschoben werden, was sowohl die Nutzung der unteren Fläche für Maus, Tastatur oder Partituren als auch eine komfortable Bedienung während des Mischvorgangs ermöglicht. Welche Ebene dabei als Hauptebene gestaltet wird, sollte sich nach der jeweiligen Nutzungshäufigkeit richten:

Häufig benötigte Bedienelemente sollten auf der ergonomisch günstigeren Ebene platziert werden, während seltener genutzte Elemente auf der sekundären Ebene Platz finden können. Variante 3 erlaubt ein besonders flexibles Arbeiten, erfordert jedoch eine stabile und technisch ausgereifte Konstruktion. Ein Beispiel für den Aufbau eines Controllers auf einer zweiten Ebene findet sich derzeit im Bimedialen Studio auf Ebene 2 des SWR Stuttgart.

Alternativ wäre auch als 4. Variante eine fest installierte Lösung mit einer zentralen Aussparung für Maus, Tastatur und Notenmaterial denkbar. Ein solches Setup wurde bereits im Hörspielstudio des SWR erfolgreich umgesetzt. Es bietet den Vorteil, dass keine beweglichen Teile erforderlich sind, was potenzielle Fehlerquellen reduziert und die Wartung vereinfacht.

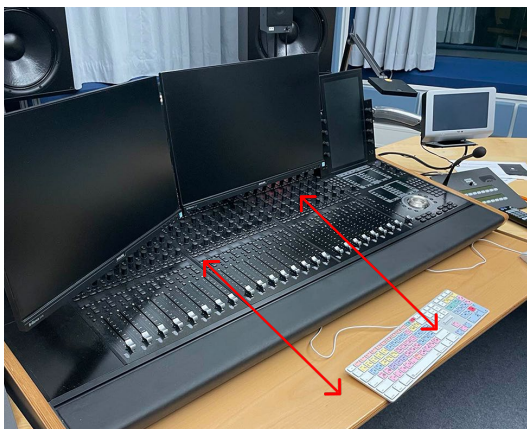
In der folgenden Abbildung werden die vier aufgeführten Varianten am Beispiel der genannten Regien gezeigt.



Variante 1: Verschiebbare Ablagefläche über Controller



Variante 2: V-förmige Anordnung der Module



Variante 3: Controller auf zweiter verschiebbarer Ebene



Variante 4: Feste Installation mit zentraler Aussparung

Abbildung 5: Mögliche Varianten der Faderanordnung (Ergonomie)

Ein weiteres wesentliches ergonomisches Kriterium ist ein höhenverstellbarer Hauptarbeitstisch, der sowohl sitzendes als auch stehendes Arbeiten ermöglicht und somit zur Förderung der Rückengesundheit beiträgt. Die Monitore sollten direkt mit dem Tisch verstellbar sein, um stets optimale Blickwinkel und ergonomische Sitz- oder Stehpositionen, insbesondere für Brillenträger mit Gleitsichtbrillen, zu gewährleisten. Zusätzliche Seitentische wurden als sinnvoll, jedoch nicht zwingend höhenverstellbar bewertet.

Für die Aufbewahrung von Arbeitsmaterialien wie Partituren, Festplatten oder persönlichen Gegenständen wurden abschließbare, akustisch optimierte Schränke empfohlen, die zusätzlich positive akustische Eigenschaften besitzen. Ergänzend wurde vorgeschlagen, einen leichten Notenständer sowie einen flexibel positionierbaren Tischarm als Notenhalter direkt am Arbeitsplatz zu integrieren, um kleinere Partituren bequem zu handhaben.

Weiterhin wurden Anforderungen an die Integration privater Laptops in das Studioumfeld genannt, insbesondere aufgrund individuell genutzter Plugins und Konfigurationen, die nicht auf den SWR-eigenen Rechnern verfügbar sind. Hierzu wurde eine Einbindung per Dante-Schnittstelle über lokale Netzwerke (ohne direkte Verbindung zum Hausnetz) als praktikable Lösung vorgeschlagen. Alternativ wurde die Nutzung von USB-Dockingstationen diskutiert, welche über eine KVM-Lösung externe Laptops komfortabel in die bestehende Infrastruktur einbinden können. Diese Lösung reduziert den Kabelaufwand erheblich, ermöglicht die Verbindung des Dante-USB-Geräts und damit die Konfiguration eines festen, statischen Routings in Dante. Somit lässt sich eine besonders einfache und schnelle Steuerung realisieren, was die Übersichtlichkeit des Setups deutlich verbessert.

Aus technischer Sicht wurde auf gut erreichbare Anschlüsse Wert gelegt. Steckdosen und USB-3.0-Anschlüsse mit hoher Datenübertragungsrate sollten unmittelbar am Arbeitsplatz verfügbar sein, um externe Geräte wie Festplatten oder Mobilgeräte schnell und ohne Kabelgewirr anschließen zu können.

Zur Reduzierung der Geräuschbelastung wurde vorgeschlagen, den Hauptrechner in einem separaten Raum oder in einem schalldämmenden Silent-Rack zu platzieren. Dies würde die akustische Belastung verringern, könnte jedoch den direkten physischen Zugriff auf den Rechner erschweren. Ebenfalls ist auf eine korrekte Wärmeabfuhr zu achten.

In Bezug auf die Bildschirmanordnung äußerten die Mitarbeitenden unterschiedliche Präferenzen. Einige bevorzugten das klassische Zwei-Monitor-Setup, während andere einen großen Monitor (ggf. Curved Monitor) zur besseren ergonomischen Zentrierung bevorzugten. Während ein großer zentraler Monitor für Einzelpersonen Vorteile bieten kann, erweist sich der Einsatz von zwei nebeneinander angeordneten Bildschirmen als deutlich praktikabler, wenn zwei Personen gemeinsam an einem Arbeitsplatz sitzen und

eine gleichberechtigte Sicht haben möchten. Ergänzend wurde ein dritter Monitor für Funktionen wie Kreuzschienensteuerung, Monitoring und Inputpegelkontrolle gewünscht, dessen Positionierung über flexible Schwenkarme individuell angepasst werden kann.

Ein zusätzlicher ergonomischer Faktor war der Einsatz eines Touchdisplays in etwa DIN A4-Größe, um digitale Partituren intuitiv zu blättern und den Papierverbrauch zu reduzieren. Alternativ wurden Bluetooth-Fußpedale oder -Knöpfe vorgeschlagen, die eine freihändige Steuerung ermöglichen und so eine einfache und ablenkungsfreie Bedienung der DAW unterstützen. Allerdings sind Fußpedale weniger intuitiv nutzbar, erfordern eine Einarbeitungszeit und bedeuten zusätzlichen technischen Aufwand wie Batteriewechsel oder potenzielle Verbindungsprobleme.

Auch die Positionierung und Zugänglichkeit von Kopfhöreranschlüssen für Toningenieur und Tonmeister wurde berücksichtigt, um störende Kabelquerungen auf dem Arbeitsplatz zu vermeiden. Ebenso wurde empfohlen, ergonomische Monitoring-Steuerungen wie Stream Decks oder programmierbare Tastenfelder zu integrieren, um schnell zwischen unterschiedlichen Abhörkonfigurationen wechseln zu können und den Arbeitsfluss intuitiver zu gestalten.

Zusammenfassend zeigt die Mitarbeiterbefragung klar, dass neben technischen Aspekten wie Latenz und Routing insbesondere ergonomische Faktoren maßgeblich sind, um langfristig komfortables und effizientes Arbeiten in den DSZs, im Podcast-Studio und in der Kammermusik zu gewährleisten.

8 Schlusskonzept

Basierend auf den umfangreichen Analysen, praktischen Tests und der strategischen Einschätzung innerhalb des Projektrahmens ergibt sich folgendes Schlusskonzept für die Erneuerung der Tonregien und digitalen Schnittplätze beim SWR.

Ziel ist es, ein flexibles, stabiles und zukunftssicheres Studiokonzept umzusetzen, das insbesondere die Herausforderungen im Bereich Monitoring-Latenz sowie Workflow-Effizienz adressiert. Zentrales Element bleibt die DAW „Sequoia“ von MAGIX als Hauptschnitt- und Bearbeitungsplattform, da sie in der SWR-Struktur fest verankert ist, umfangreiche Editierfunktionen bietet und mit den bestehenden Archiven vollständig kompatibel ist. Die benötigten Sequoia-PCs können aus vorhandenen Ressourcen im SWR beschafft werden, entsprechende Lizenzen sind ebenfalls bereits verfügbar.

Als zentrale Kreuzschiene für alle Ein- und Ausgänge dient eine MTRX-II mit flexiblen Routingmöglichkeiten. Ebenso ist die Verwendung von RME HDSPe MADI FX Interfaces in den Sequoia PCs vorgesehen.

Die technischen Tests haben ergeben, dass ein hybrider Ansatz – mit Sequoia für die Bearbeitung und Avid Pro Tools HDX für das Monitoring – eine leistungsfähige Kombination aus Stabilität, niedriger Latenz und maximaler Flexibilität bietet. Während rein native DAW-Lösungen im Test bei Puffergrößen unter 128 Samples instabil wurden, konnte Pro Tools im DSP-Modus konstant Latenzen von unter 2 ms liefern – unabhängig von der verwendeten Puffergröße. Selbst die Verwendung von DSP-Plugins wie Equalizer und Kompressor lieferten bei den Latenztests ähnliche Ergebnisse. Dies erlaubt es, auch komplexe Aufnahmesituationen mit gleichzeitiger Echtzeitwiedergabe zuverlässig zu bewältigen. Ein weiterer positiver Aspekt ist, dass Pro Tools als vorgeschaltetes Mischpult eine vollwertige Backup-Möglichkeit bietet. Durch die parallele Aufnahme auf zwei getrennten Systemen können mögliche Fehler, die bei der Aufnahme entstehen, zusätzlich abgesichert werden. Außerdem ist die Vorgehensweise mit Pro Tools bereits in den Hörspielstudios des SWR etabliert und ermöglicht es, den Bearbeitungsworkflow weiterhin vollständig in Sequoia abzubilden. Die Avid S-Serie kann dabei beide DAWs effizient ansteuern und stellt damit einen erheblichen Vorteil gegenüber herkömmlichen Hardware-Mischpulten dar. Auch im Hinblick auf eine mögliche Standardisierung in verschiedenen Produktionsbereichen ist die Option vorteilhaft, da Pro Tools im TV-Bereich bereits flächendeckend im Einsatz ist. Es wären somit auch in den neuen Tonstudios beide Systeme vorhanden und die Mitarbeiter mit beiden DAWs zwangsläufig vertraut.

Gleichzeitig muss jedoch berücksichtigt werden, dass diese Variante mit höheren Kosten und größerem Administrationsaufwand einhergeht. Ein zusätzlicher Mac mit Pro Tools und HDX-Hardware bedeutet eine erhebliche Investition. Darüber hinaus erfordert die parallele Nutzung zweier DAWs zusätzlichen Wartungsaufwand aufgrund von Updates für beide Systeme sowie zusätzlichen Schulungsaufwand, insbesondere für weniger erfahrene Nutzer. Längere Abstände zwischen Produktionen machen das Arbeiten schwieriger, da beide DAWs gelernt werden müssen und das Wissen im Idealfall aufrechterhalten werden muss. Auch stellt sich die Frage der Notwendigkeit: Während im Podcast Bereich regelmäßig Kopfhörermonitoring erforderlich ist, ist dies in der Kammermusik nur in rund geschätzt 5 % der Produktionen der Fall.

Unabhängig von der Verwendung von Pro Tools als Mischpult wurde ein Backupsystem mit einem zweiten Rechner zunächst verworfen. In seltenen Fällen eines Ausfalls kann dieser durch eine einfache Wiederholung der Aufnahme kompensiert werden – was die Investition in ein vollständiges Redundanzsystem aus wirtschaftlicher Sicht nicht rechtfertigt. Außerdem wäre eine spätere Aufrüstung von Sequoia oder Pro Tools als zweites Backupsystem und die Einbindung in die KVM-Umschaltung durchaus möglich und in absehbarer Zeit umsetzbar, falls diese Variante zu einem späteren Zeitpunkt aufgrund der Vorteile oder einem SWR-weiten Umstieg auf Pro Tools angedacht wird.

Daher wird vorgeschlagen, in der ersten Ausbaustufe auf Pro Tools als Monitoringlösung zu verzichten. Technisch überzeugend, aber wirtschaftlich unverhältnismäßig, erscheint der parallele Betrieb beider DAWs zunächst nicht sinnvoll. Stattdessen kann das Monitoring im Kammermusik-Studio effizient über die MTRX in Kombination mit der Software DADman erfolgen. Diese bietet eine direkt in der Hardware umgesetzte, latenzarme Monitoringlösung mit hoher Flexibilität durch frei definierbare Abhörpunkte. Gerade im klassischen Bereich, wo Musiker häufig ohne Kopfhörer aufnehmen, ist diese Lösung in den meisten Fällen ausreichend. Sollten sich im Laufe des Betriebs erweiterte Anforderungen ergeben – etwa der Bedarf an DSP-basierten Effekten oder komplexeren Mischszenarien – lässt sich TotalMix als ergänzende Monitoring-Ebene problemlos nachrüsten. Hierzu wären lediglich Anpassungen der Signalpfade und eine Umkonfiguration der Monitoring-Instanzen in der MTRX notwendig, da die grundlegenden Verbindungen bereits bestehen.

TotalMix wäre vor allem in Produktionen mit aufwendigerem DSP-basierten Monitoring gefragt, wenn komplexere Kopfhörermischungen mit Effekten und Klangbearbeitung erforderlich sind – beispielsweise in Podcasts oder modernen Musikproduktionen. In diesen Fällen müsste Sequoia vom Mixer-FX- auf den Hardware-Monitoring-Modus umgeschaltet werden, wodurch allerdings das Abhören der eigentlichen Mischung in Sequoia entfällt.

Da bei Kammermusik-Produktionen Anpassungen des Kopfhörermonitorings eher im laufenden Betrieb schnell vorgenommen werden müssen und die Verwendung von Plugins meistens nicht notwendig ist, erscheint die DADman-Lösung als ideale Einstiegslösung für die Kammermusik: Sie ist intuitiv, stabil, latenzarm und in der Lage, durch vorbereitete Abhörpunkte flexibel auf unterschiedlichste Produktionsszenarien zu reagieren.

Für das Podcast-Studio wird dagegen empfohlen, Sequoia in Verbindung mit DADman und TotalMix zu verwenden, da häufiger vor der Aufnahme ein Kopfhörermonitoring erstellt wird und Sprecher oder Musiker eine Klangbearbeitung wünschen. Zwar bedeutet der Einsatz von TotalMix eine gewisse Einschränkung in der Steuerbarkeit über die Avid S-Konsole, da das EUCON-Protokoll nicht von TotalMix unterstützt wird, dennoch überwiegen die Vorteile durch die Möglichkeit, DSP-Plugins flexibel in den Signalweg einzubinden. Ggf. kann die Bedienung auch per Maus über das TotalMix-Interface erfolgen, ohne den Monitoring-Modus in Sequoia umzuschalten.

Die Entscheidung für eine Standardisierung auf Avid-Controller (S1/S6) bietet ebenfalls viele Vorteile: Die Geräte sind im SWR bereits an mehreren Standorten im Einsatz, der Support durch die interne IT ist erprobt und auch die Nutzer profitieren von einer einheitlichen Bedienstruktur. LAWO als Hersteller für digitale Mischpulte wurde als Alternativlösung zwar technisch evaluiert, jedoch aufgrund fehlender Unterstützung von EUCON sowie unklarer Zukunftsaussichten in Bezug auf die Integration in Sequoia ausgeschlossen.

Im Hinblick auf Audio-over-IP folgt das Konzept dem Prinzip der Zukunftsoffenheit bei gleichzeitigem Fokus auf die Wirtschaftlichkeit: Eine direkte Anbindung an das hausweite Netzwerk ist derzeit nicht vorgesehen. Jedoch könnten Leitungen (z. B. Lichtwellenleiter) bereits in den Neubau integriert werden, sodass eine spätere Anbindung an das Media-over-IP Netzwerk des SWR ohne bauliche Maßnahmen möglich ist. Die Studios könnten zudem so vorbereitet werden, dass externe Geräte – z. B. mitgebrachte Laptops – über Dante unkompliziert eingebunden werden können. So entsteht eine praxisnahe, aber erweiterbare Insellösung.

Ein wesentlicher Aspekt des vorliegenden Konzepts betrifft auch die menschliche Komponente: Die vorgeschlagene Lösung muss sich nahtlos in die bestehenden Strukturen und Arbeitsabläufe des SWR einfügen. In der Praxis sind Mitarbeitende mitunter über längere Zeiträume nicht kontinuierlich in den Studios tätig. Umso wichtiger ist es, dass die Bedienung der Technik intuitiv und unkompliziert erfolgt, sodass eine schnelle und reibungslose Einarbeitung möglich ist. Die im Rahmen dieser Arbeit durchgeführten Tests

und die darauf aufbauenden konzeptionellen Entscheidungen wurden mit Blick auf eine einfache und praxisgerechte Bedienbarkeit getroffen.

Das hier vorgestellte Konzept stellt eine Vorplanung dar, auf deren Basis in den kommenden Projektphasen konkrete Ausschreibungsunterlagen und Kostenkalkulationen entwickelt werden können. Die wesentlichen Komponenten und der Aufbau gelten dabei als stabil – kleinere Änderungen, etwa in der Ausstattung einzelner Regien, sind im Rahmen der Detailplanung durchaus möglich. Dabei könnten auch Synergien durch Rückbau bestehender Studios einfließen, etwa durch die Wiederverwendung funktionstüchtiger Hardware oder die Umnutzung bestehender Ressourcen.

Zusammenfassend lässt sich festhalten: Das vorgeschlagene Studiokonzept ist technisch ausgereift, wirtschaftlich begründet und strategisch in die bestehende Infrastruktur des SWR integrierbar. Es bietet eine hohe Flexibilität bei gleichzeitig klarer Skalierbarkeit. Die Entscheidung für DADman als initiale Monitoringlösung schafft eine zuverlässige und kosteneffiziente Grundlage, die bei Bedarf schnell erweitert bzw. für andere Monitoring-Lösungen umkonfiguriert werden kann. Die Standardisierung auf Avid-Komponenten sichert langfristige Wartbarkeit, Erweiterbarkeit und Nutzerfreundlichkeit – und legt damit ein solides Fundament für die kommenden Jahre.

9 Projektplanung und Umsetzung

Die Umsetzung des erarbeiteten Konzepts erfolgt im Rahmen des Modernisierungsprojekts des SWR voraussichtlich bis 2027. In diesem Kapitel werden die weiteren Phasen der Projektplanung und Umsetzung beschrieben.

Nach Abschluss der Konzeptionsphase steht als nächstes die formale Projektbeantragung im Rahmen des unternehmensinternen Projektportfoliomanagements (PPM) an. Der sogenannte PPM-Antrag bildet die Grundlage für die Freigabe des Budgets und die offizielle Beauftragung der weiteren Planungs- und Umsetzungsphasen. Er wird von einem Gremium – dem PPM-Board – geprüft und anhand definierter Kriterien wie Nutzen, Aufwand, Ressourceneinsatz und Dringlichkeit bewertet. Erst mit dessen Genehmigung kann das Projekt in die nächste Phase überführt werden. Dieser Schritt ist entscheidend, da ohne diese Freigabe keine weiteren Planungsressourcen aktiviert oder externe Partner beauftragt werden dürfen.

Bis zur erwarteten Entscheidung im Mai 2026 ist die weitere Zeitplanung der DSZs und Kammermusik noch offen. Eine Ausnahme bildet die bereits bewilligte Erneuerung des Podcast-Studios, das auf demselben technischen Konzept basiert. Erste Umsetzungsschritte dieser Regie können bereits zeitnah erfolgen und liefern wichtige Erkenntnisse für die spätere Umsetzung des Gesamtprojekts.

Nach positiver Entscheidung beginnt die eigentliche Projektplanung, die sich in mehrere Schritte gliedert. Zunächst wird gemeinsam mit dem Einkauf ein Terminplan erstellt. Der Einkauf ist deshalb von Anfang an beteiligt, weil das Projekt – aufgrund seines finanziellen Umfangs – im Rahmen einer EU-weiten Ausschreibung vergeben werden muss. Die Einhaltung der entsprechenden Fristen und Formalitäten ist hierbei zwingend erforderlich. Für die rechtskonforme Durchführung wird in der Regel eine externe Beraterfirma eingebunden, die über Erfahrung mit komplexen Vergabeverfahren im öffentlich-rechtlichen Rundfunk verfügt. Diese unterstützt nicht nur bei der strukturellen und juristischen Ausarbeitung der Ausschreibungsunterlagen, sondern auch bei der inhaltlichen Aufbereitung technischer Anforderungen und Kriterienkataloge.

Parallel dazu erfolgt die Detailplanung des Projekts. Ziel ist es, auf Basis des zuvor erarbeiteten Konzepts konkrete technische Spezifikationen, funktionale Anforderungen und Systemkomponenten so aufzubereiten, dass sie in eine öffentliche Ausschreibung überführt werden können. Dazu gehören u. a. die Definition von Ein- und Ausspielwegen, die Gestaltung der Arbeitsplätze, Anforderungen an die Steuerungssysteme (z. B. DAW-Integration), sowie Festlegungen zur Raumakustik, Infrastruktur, Verkabelung und

Klimatisierung. Auch Erfahrungswerte aus vergleichbaren Projekten fließen in diesen Schritt ein. Im Idealfall wird ein vollständiges Lasten- bzw. Pflichtenheft erstellt, das nicht nur die gewünschten Funktionen beschreibt, sondern auch spätere Bewertungsmaßstäbe für die eingehenden Angebote liefert. Diese Phase ist sehr arbeitsintensiv, legt aber den Grundstein für eine erfolgreiche Umsetzung.

Nach Abschluss der Planung beginnt die EU-Ausschreibung. Dieses Verfahren stellt sicher, dass der Auftrag europaweit transparent vergeben wird. Die Ausschreibung wird öffentlich bekannt gemacht, und interessierte Unternehmen können innerhalb einer gesetzlich festgelegten Frist Angebote einreichen. In dieser Zeit besteht für die Bieter die Möglichkeit, Rückfragen zu stellen – eine Phase, die inhaltlich und formal gut vorbereitet sein muss. Nach Ablauf der Frist prüft der Einkauf die Angebote auf formale Richtigkeit, während die Projektverantwortlichen und ggf. die Beraterfirma die technische Eignung und Wirtschaftlichkeit bewerten. Die Angebote werden nach einem vorher definierten Bewertungssystem beurteilt. Kriterien können z. B. Preis, Konzeptqualität, technische Umsetzbarkeit oder Serviceangebote sein.

Erst nach Abschluss dieser Phase erfolgt der Zuschlag an das bestbewertete Unternehmen. Damit beginnt die Realisierungsphase, in der die konkrete Umsetzung geplant und durchgeführt wird. Diese umfasst u. a. die Montageplanung, die Terminierung der Lieferungen, die Koordination der technischen Installation sowie die Berücksichtigung bestehender Abläufe im Haus (z. B. Sendebetrieb). Auch die Schulung des künftigen Personals und die Erstellung der notwendigen Dokumentation sind Teil dieser Phase.

Die eigentliche Inbetriebnahme erfolgt in mehreren Schritten: Nach Aufbau und Verkabelung der Systeme werden diese getestet, kalibriert und – wo notwendig – an hausinterne Infrastrukturen wie Netzwerke, Audioserver oder Redaktionssysteme angebunden. Erst wenn sämtliche Funktionstests erfolgreich abgeschlossen sind und eventuelle Nachbesserungen erfolgt sind, wird das System offiziell abgenommen und in den Produktionsbetrieb überführt.

Da das Projekt aktuell noch in der Antrags- und Genehmigungsphase steckt, lässt sich der weitere Zeitplan noch nicht exakt bestimmen. Es ist jedoch bereits jetzt absehbar, dass eine enge Abstimmung zwischen Einkauf, Technik, externen Beratern und späteren Nutzenden essenziell sein wird, um das Vorhaben reibungslos umzusetzen. Auch wenn einige organisatorische Hürden und potenzielle Verzögerungen (z. B. durch Ausschreibungsfristen oder Materialengpässe) einkalkuliert werden müssen, ist mit dem abgeschlossenen Konzept eine solide Basis geschaffen worden, auf der sich das weitere Vorgehen zielgerichtet aufbauen lässt.

10 Fazit und Ausblick

Die vorliegende Arbeit hat gezeigt, wie eine moderne, multifunktionale Studioinfrastruktur beim SWR konzipiert und umgesetzt werden kann, um den steigenden Anforderungen an Flexibilität und Qualität gerecht zu werden. Ausgangspunkt war die Herausforderung, klassische Tonbearbeitungsräume so zu erneuern, dass sie sowohl für die Postproduktion als auch für Aufnahmen geeignet sind – und dies unter Beibehaltung eines weitgehend softwarebasierten Workflows. Durch eine Kombination aus theoretischer Analyse und praktischen Tests wurde ein Konzept entwickelt, das diese Anforderungen erfüllt.

Der Verzicht auf ein herkömmliches Mischpult zugunsten eines Avid S-Controllers vereint die Vorteile von Recall und Flexibilität mit einer weiterhin haptischen Bedienung. Diese Entscheidung wurde auch vor dem Hintergrund getroffen, dass Veränderung und Innovation Teil der SWR-DNA sind und der Sender sich in Richtung digitaler Workflows entwickelt.

Durch den Einbezug von Audio-over-IP in diesem Konzept wird die Anlage flexibel und zukunftssicher. Externe Geräte lassen sich z. B. mittels Dante nahtlos integrieren, und das Routing von Signalen wird softwaregestützt und damit sehr anpassbar.

In der Diskussion wurde deutlich, dass das gewählte Systemdesign gegenüber alternativen Ansätzen wie einem analogen Pult oder anderen DAW-Systemen zahlreiche Vorteile bietet. Insbesondere die Integrationstiefe, dass alles auf einer Plattform zusammenläuft, ermöglicht effiziente Workflows. So kann ein Tonmeister heute eine Sprachaufnahme machen und direkt im Anschluss in derselben Session das Material schneiden und mischen, ohne Studio- oder Systemwechsel.

Ein weiterer wichtiger Aspekt ist die Steigerung der Produktionskapazität und Qualität. Mit den neuen Räumen kann der SWR zukünftig komplexere Produktionen intern realisieren, z. B. Podcast-Serien mit mehreren Sprechern, die parallel aufgenommen werden, oder Musiksessions, die vorher extern gemischt werden mussten.

Die erfolgreiche Umsetzung in Stuttgart könnte als Blueprint für weitere Regionen dienen. Es ist denkbar, andere Studios nach dem gleichen Prinzip aufzurüsten, ggf. unter Berücksichtigung der lokalen Gegebenheiten.

Aus technischer Sicht gibt es immer Möglichkeiten zur Weiterentwicklung. Einige potenzielle Weiterführungen sind z. B. die Automatisierung und Workflow-Optimierung: So könnte in Zukunft ein Fokus auf den Einsatz von Software-Scripts oder Makros gelegt werden, um Routinehandlungen (wie das Einrichten eines Projekts mit allen nötigen

Routings für das Monitoring) per Knopfdruck zu erledigen. Dies könnte die Einstiegshürde für neue Nutzer weiter senken.

Ein weiterer Punkt ist der Einsatz von KI-Assistenz im Schnitt. Da die Hardware nun leistungsfähig ist, könnte man perspektivisch KI-Tools integrieren (z. B. automatische Transkription und Schnittvorschläge für Interviews), was die Arbeit der Cutter erleichtert. Solche Software ließe sich auf den Workstations nachrüsten.

Abschließend lässt sich festhalten, dass das Projekt "Erneuerung von hochwertigen und multifunktionalen Tonbearbeitungsräumen und Tonregionen des Südwestrundfunks" ein Beispiel dafür ist, wie durch vorausschauende Planung, den Mut zu neuen Technologien und die Einbindung der Nutzer ein großer Schritt nach vorn gemacht werden kann. Der SWR erneuert nicht nur seine Studios, sondern auch zu einem gewissen Grad seine Arbeitskultur – hin zu mehr Digitalität, Flexibilität und Kollaboration. Damit bleibt er seinem Anspruch treu, sich stetig weiterzuentwickeln und Maßstäbe zu setzen. Die Bachelorarbeit konnte einen substanziellen Beitrag leisten, indem sie das Konzept fundiert ausgearbeitet, getestet und dokumentiert hat. Die wissenschaftliche Vorgehensweise – vom Problem über Analyse, Experiment bis zur Lösung – gewährleistet, dass die Entscheidungen nachvollziehbar und gut begründet sind. Dies wird es den Entscheidungsträgern erleichtern, die vorgeschlagenen Maßnahmen umzusetzen.

Literaturverzeichnis

- Audinate's Dante Now Supported in More Than 3,000 Devices* | Audinate. (2021, März 24). Audinate. Abgerufen am 16.04.2025. <https://www.audinate.com/press/audinate-s-dante-now-supported-in-more-than-3000-device/>
- Audio over IP: The Growing Trend Amongst High End Integrators* - StormAudio. (2025, April 8). StormAudio. Abgerufen am 02.04.2025. <https://www.stormaudio.com/audio-over-ip-the-growing-trend-amongst-high-end-integrators>
- Avid. (2023). *PRO TOOLS | MTRX FAMILY COMPARISON*. Avid Technology, Inc. PRO TOOLS | MTRX FAMILY COMPARISON
- Avid. (2024). *Avid S4/S6 Guide - Version 2024.10*. Avid Technology, Inc.
- Beacham, F. (2017, Juni 13). *In the 1990s, Audio Recording Changed Forever*. The Broadcast Bridge - Connecting IT to Broadcast. Abgerufen am 15.03.2025. <https://www.thebroadcastbridge.com/content/entry/8815/in-the-1990s-audio-recording-changed-forever>
- Camou, E. (2013). *Pro Tools HD: Advanced Techniques and Workflows*. Packt Publishing.
- Dante Advanced Configuration*. (2015). Audinate. Abgerufen am 28.02.2025. <https://www.getdante.com/resources/>
- Lebowski, A. (2021, Juni 24). *Boost Your Audio Performance with HDX Hybrid Engine* - Avid. Avid. Abgerufen am 18.03.2025. <https://www.avid.com/de/resource-center/hdx-hybrid-engine>
- Lester, M., & Boley, J. (2007). The effects of latency on live sound monitoring. *Audio Engineering Society Convention Paper 7237, 123rd AES Convention, New York, NY, USA*.
- MAGIX. (2025). *Sequoia 17 | Handbuch*. MAGIX Software GmbH.
- Rodgers, J. (2021, Juli). *Pro Tools: Hybrid Engine Explained*. Sound On Sound. Abgerufen am 10.03.2025. <https://www.soundonsound.com/techniques/pro-tools-hybrid-engine-explained>
- Sequoia Pro 17 – The engineers' choice*. (o. D.). MAGIX. Abgerufen am 07.02.2025. <https://www.magix.com/de/musik-bearbeiten/sequoia/>

- Smyrek, V. (2009). *Tontechnik für Veranstaltungstechniker*, 3. Aufl. Stuttgart: S. Hirzel Verlag.
- Stange-Elbe, J. (2015). *Computer und Musik: Grundlagen, Technologien und Produktionsumgebungen der digitalen Musik*. Berlin, München, Boston: De Gruyter Oldenbourg. <https://doi.org/10.1515/9783486720198>
- SWR. (2023, August 29). *Fusion: Vor 25 Jahren wurde aus SDR und SWF der SWR - SWR Aktuell*. SWR. Abgerufen am 18.03.2025. <https://www.swr.de/swraktuell/jubilaem-25-jahre-swr-fusion-aus-sdr-und-swf-100.html>
- SWR. (o. D.). *Wir in der ARD*. SWR. Abgerufen am 18.03.2025. <https://www.swr.de/unternehmen/organisation/swr-und-ard-100.html>
- Wang, Y. (2017). *Low Latency Audio Processing*. Dissertation, School of Electronic Engineering and Computer Science, Queen Mary University of London.