
BACHELORARBEIT

Herr
Adrian Hardt

**Implementierung eines
Archivsystems für
Videoproduktionen der Fakultät
Medien der Hochschule Mittweida**

Bewertung von Technologien und
Funktionsprinzipien etablierter
Archivierungssysteme durch den
wissenschaftlichen Vergleich

2018

BACHELORARBEIT

Implementierung eines Archivsystems für Videoproduktionen der Fakultät Medien der Hochschule Mittweida

Bewertung von Technologien und
Funktionsprinzipien etablierter
Archivierungssysteme durch den
wissenschaftlichen Vergleich

Autor:
Herr Adrian Hardt

Studiengang:
Angewandte Medien

Seminargruppe:
AM16sK1-B

Erstprüfer:
Prof. Mike Winkler

Zweitprüfer:
Mag. Markus Bell

Einreichung:
Mittweida, 25.07.2018

BACHELOR THESIS

The implementation of a video archive system for the faculty of applied sciences in Mittweida

Using scientific comparison of
technologies and functions to evaluate
of established archiv systems

author:

Mr. Adrian Hardt

course of studies:

Angewandte Medien

seminar group:

AM16sK1-B

first examiner:

Prof. Mike Winkler

second examiner:

Mag. Markus Bell

submission:

Mittweida, 25.07.2018

Bibliografische Angaben

Nachname, Vorname: Hardt, Adrian

Implementierung eines Archivsystems für Videoproduktionen der Fakultät Medien der Hochschule Mittweida: Bewertung von Technologien und Funktionsprinzipien etablierter Archivierungssysteme durch den wissenschaftlichen Vergleich

The implementation of a video archive system for the faculty of applied sciences in Mittweida: Using scientific comparison of technologies and functions to evaluate of established archiv systems

78 Seiten, Hochschule Mittweida, University of Applied Sciences,
Fakultät Medien, Bachelorarbeit, 2018

Abstract

Diese Bachelorarbeit dient der Identifizierung eines geeigneten Archivsystems für die Hochschule Mittweida. Diese verfügte bis zum Zeitpunkt der Erscheinens der Arbeit über kein systematisches und sicheres Archiv. Zum inhaltlichen Verständnis beschäftigt sich die Arbeit im ersten Teil mit technischen Themen rund um digitale Systeme, Speichermedien und Langzeitarchivierung. Dabei wendet der Verfasser die Literaturrecherche an. Anschließend werden verschiedene Archivsysteme durch einen wissenschaftlichen Vergleich evaluiert. Dazu erarbeitet der Verfasser für ein Langzeitarchiv relevante Kriterien. Als besonders wichtig werden dabei die Kriterien Lebensdauer und Verwaltungsaufwand herausgearbeitet. Schlussendlich führt der Vergleich zum Vorschlag der "PetaSite Library" von Sony, welche mit 100 Jahren Lebensdauer die höchste Sicherheit und ein das beste Preis-Leistungsverhältnis bietet. Mit einer anschließenden Handlungsempfehlung soll die Implementierung des Systems inhaltlich vorbereitet werden. Diese Arbeit soll die der Hochschule Mittweida bei der Gestaltung des neuen Archives unterstützen, sowie technisch versierten, an Langzeitarchivierung interessierten Lesern wichtige Informationen zu den Systemen vermitteln.

Inhaltsverzeichnis

Inhaltsverzeichnis	I
Abkürzungsverzeichnis.....	III
Abbildungsverzeichnis	IV
Tabellenverzeichnis.....	V
1 Einleitung.....	1
1.1 Problemstellung	1
1.2 Ziel der Arbeit.....	2
1.3 Methodik	2
2 Archivierung	5
2.1 Langzeitarchivierung	6
3 Terminologien und technische Grundlagen	9
3.1 Formate.....	10
3.2 Codecs.....	10
3.3 Videokompression.....	11
3.4 Hausformat: XDCAM HD422	13
3.5 Metadaten	14
3.6 Bedarfsermittlung	15
4 Aktuelle Technologien	16
4.1 IT-Infrastruktur der HS Mittweida.....	16
4.2 Speichermedien.....	17
4.2.1 Magnetbandspeicher	17
4.2.2 HDD	20
4.2.3 SSD.....	22
4.2.4 Optische Speichermedien.....	23
4.3 RAID-Level.....	24
4.3.1 RAID 0	25
4.3.2 RAID 1	25
4.3.3 RAID 3 und RAID 4	25
4.3.4 RAID 5	26
4.3.5 RAID 6	27
4.3.6 RAID 10	27
4.3.7 RAID 60	27
4.4 Speicherarchitekturen	27
4.4.1 DAS	28
4.4.2 NAS	28
4.4.3 SAN	29
5 Archivierungssysteme	31
5.1 Archivtechnologien.....	31
5.1.1 NAS-Lösung.....	31
5.1.2 Tape Library.....	32

5.1.3	Optical Data Archive	33
5.1.4	Cloud	33
5.2	Archivsysteme	34
5.2.1	System 1: Hybrid Bandbibliothek / NAS	34
5.2.2	System 2: NAS.....	36
5.2.3	System 3: Bandbibliothek.....	36
5.2.4	System 4: Optical Disc Archive (A).....	37
5.2.5	System 5: Optical Disc Archive (B).....	38
5.2.6	System 6: Hybrid NAS / Cloud	39
6	Vergleich.....	41
6.1	Bewertungskriterien.....	41
6.2	Bewertung.....	44
6.3	Vorschlag Archivierungslösung	48
6.4	Handlungsvorschläge.....	49
7	Schlussfolgerung	50
7.1	Forschungsablauf	50
7.2	Selbstreflexion	52
	Literaturverzeichnis	XI
	Anhang I	XXI
	Anhang II	XXIV
	Anhang III	XXV
	Eigenständigkeitserklärung	XXVII

Abkürzungsverzeichnis

Bzw. – beziehungsweise
CD – Compact Disc
CD-RW – Compact Disc ReWritable
CD-R – Compact Disc Recordable
DAS – Direct Attached Storage
FC – Fibre Channel
HDD – Hard Drive Disc
HS – Hochschule
JBOD – Just a Bunch Of Discs
LTO – Linear Tape Open
LTFS – Linear Tape File System
LAN – Local Area Network
MLC – Multi Layer Cells
NAS – Network Attached Storage
ODA – Optical Disc Archive
ODS – Optical Disc Storage
ODC- Optical Disc Cartridge
RAID – Redundant Array of Independant Discs
SAN – Storage Area Network
SAS – Serial Attached SCSI
SLC – Single Layer Cells
SSD – Solid State Drive
TLC – Triple Layer Cells
z.B – zum Beispiel

Abbildungsverzeichnis

Abbildung 1: Einheiten der Speicherkapazitäten	5
Abbildung 2: Digitale Signalverarbeitung	9
Abbildung 3: Containerformat	10
Abbildung 4: Verlustbehaftete Kompression.....	12
Abbildung 6: herkömmliche Speicherung auf HDD	21
Abbildung 7: CD Querschnitt	24
Abbildung 8: Network attached storage.....	29
Abbildung 9: Storage area network	30
Abbildung 10: Migrationshäufigkeit Magnetband / ODA	47
Abbildung 11: Infrastruktur HS Mittweida	XXIV
Abbildung 12: RAID 0	XXV
Abbildung 13: RAID 1	XXV
Abbildung 14: RAID 4	XXV
Abbildung 15: RAID 6	XXVI
Abbildung 16: RAID 10	XXVI

Tabellenverzeichnis

Tabelle 1: Kapazitäten der LTO-Generationen (ohne Kompression)	19
Tabelle 2: Vor- und Nachteile NAS	32
Tabelle 3: Vor- und Nachteile Bandbibliothek	33
Tabelle 4: Vor- und Nachteile ODA	33
Tabelle 5: Vor- und Nachteile Cloud	34
Tabelle 6: PetaSite Library Module	38
Tabelle 7: Panasonic freeze-ray Module	39
Tabelle 8: Vergleich der Systeme	44

1 Einleitung

Diese Bachelorarbeit befasst sich mit der Recherche und Bewertung bewährter Speichermedien und Archivierungsmethoden. Die Erarbeitung eines Lösungsvorschlags zur Implementierung eines Langzeitarchivsystems an der Hochschule Mittweida gilt als übergeordnetes Ziel.

1.1 Problemstellung

Ansetzend an dem omnipräsenten technischen Fortschritt und der damit verbundenen Informationsflut soll diese Arbeit zum allgemeinen Verständnis über verschiedenste Speichermedien beitragen. Die rasante Entwicklung im Bereich der Speicherung in den letzten Jahren setzt ein flexibles Wissen voraus, welches dem permanenten Hinzukommen neuer Begrifflichkeiten und neuer Herangehensweisen standhalten muss. Mit der Digitalisierung kam die Wirtschaft in eine entscheidend neue Situation, welche die Anpassung vieler Unternehmen voraussetzt¹. Unternehmen, die sich der strukturellen Veränderung vieler Bereiche durch die Digitalisierung nicht stellen, laufen Gefahr, Chancen zu verpassen. Die IT-Industrie gewinnt immer mehr an Fahrt² und die Auseinandersetzung mit technischen Themen wird wichtiger. Die hohe Komplexität und geringe Zugänglichkeit technischer Informationen führt diesbezüglich allerdings bei vielen Menschen zu einem starken Gefühl der Überforderung.

Die Digitalisierung überführt Informationen, die früher auf Kassetten oder Papier aufbewahrt wurden, auf vernetzte Datenträger. Mit der Revolution der Informationstechniken und deren neugewonnenen Vorzügen wird oft von einer Abkehr vom Haptischen gesprochen. Die daraus resultierende Intangibilität setzt voraus, dass die neuen Systeme sicher sind. Nur noch in digitaler Form existierende Daten sind bei einem Kollaps des Systems nicht abrufbar und somit unbrauchbar. Das Thema Sicherheit ist somit besonders in Zeiten der Industrie 4.0 mit hoher Aufmerksamkeit zu behandeln.

Datensicherheit als wird immer relevanter, wenn auch viele Unternehmen, und Privatpersonen weiterhin die Dringlichkeit unterschätzen. Zur Erhaltung sensibler Daten über einen langen Zeitraum hinweg existieren zahlreiche Speicherlösungen, die jeweils andere Gesichtspunkte priorisieren.

Die technischen Funktionsweisen sollen verständlich erklärt werden, um einer weiten Leserschaft mit verschiedenen Wissensständen Einblicke in die Speicherung von Daten zu geben.

¹ Vgl. Wiegand (2017): Der Weg aus der Digitalisierungsfalle: Mit Lean Management erfolgreich in die Industrie 4.0. S. 159

² Bitkom (2017): Umsatz mit Informationstechnik in Deutschland von 2007 bis 2018 (in Milliarden Euro).

1.2 Ziel der Arbeit

Die Fakultät Medien der Hochschule Mittweida bildet im Studiengang „Medienmanagement“, „Media Production“ und „Media Acoustic Engineering“ Studenten aus und lehrt sie die Arbeit mit Video- und Audiomaterial zur Erzeugung audiovisueller Beiträge. Die Studenten erarbeiten dabei über mehrere Semester hinweg von der Planung, über die Ausführung bis hin zur Nachbearbeitung diverse Video- und Audioprojekte. Die dabei entstehenden Daten sollen zur Erhaltung kreativen Gutes und zur Wiederverwertung der Werke zuverlässig gespeichert werden. Wegen der zahlreichen Studentenprojekte erfordert das zukünftige System eine beachtliche Speicherkapazität und setzt damit auch eine durchdachte Verkettung verschiedener Komponenten voraus. Dabei sollen vor allem Videodateien, welche in der Regel viel Speicherplatz beanspruchen, über viele Jahre sicher aufbewahrt werden. Das Ergebnis dieser Arbeit ist die Bewertung diverser Systeme mit der Identifizierung der auf den Bedarf abgestimmten besten Lösung der langfristigen Sicherung sensibler Daten. Anlässlich der Planung eines Langzeitarchives in der Fakultät Medien der Hochschule Mittweida soll ein passgenaues System recherchiert bzw. kombiniert und projiziert werden. Durch die planmäßige Erarbeitung des nötigen Wissens und der Klärung relevanter Begriffe soll das Verständnis für das technisch-wissenschaftliche Thema „Archivierung“ gestärkt werden. Bisher fehlt der Hochschule Mittweida ein geeignetes System, das die Kriterien eines ausfallsicheren und langlebigen Archivs erfüllt. In Absprache mit der Hochschule Mittweida identifizierte der Verfasser das voraussichtlich zu sichernde Datenvolumen und erarbeitete relevante Kriterien zur Evaluation des optimalen Systems.

1.3 Methodik

Dabei nimmt der Verfasser eine Analyse bestehender Systeme vor und erklärt deren technische Funktionen bzw. Eigenschaften. Der Leser soll nach der Lektüre Technologien einordnen und deren Vor- und Nachteile selbst abwägen können. Als erstes wird die bereits existierende IT-Infrastruktur der Fakultät beschrieben, um dann eine lückenlose Eingliederung der technischen Lösungsvorschläge verständlich und nachvollziehbar zu beschreiben. Das neue Archivsystem muss kompatibel mit der bereits bestehenden Infrastruktur sein und sich ohne die Modifikation des Gesamtsystems umsetzen lassen. Das Ziel ist ein leistungsstarkes, anpassbares und sicheres Langzeitarchiv.

Methodisch geht der Verfasser mit Hilfe der Literaturrecherche vor und wendet den wissenschaftlichen Vergleich an. Zur Informationsbeschaffung verwendet der Verfasser wissenschaftliche Quellen aus Fachbüchern, aus Fachzeitschriften und aus dem Internet. Die gewonnenen Informationen werden in seinen eigenen Worten wiedergegeben. Nach der Darstellung der technischen Grundlagen werden die Systeme miteinander verglichen.

Um die Erklärung der Speicherung von Daten einzuleiten, beschäftigt sich der Verfasser zuerst mit den Eigenschaften von Dateien, genauer von Videodateien. Formate, Codecs und Komprimierungsmethoden werden dabei zum näheren Verständnis erläutert.

Der Verfasser unterscheidet weiter zwischen Speichertechnologien und Speichersystemen. Die Technologien behandelnd geht der Verfasser auf die Funktionen etablierter Speichermedien bzw. Speichertechnologien ein. Mit Hilfe eines präziseren Einblicks in die technischen Eigenschaften sollen verschiedene Medien gegeneinander abgewogen werden. Zum Nutzen eines besseren Verständnisses wird eine genauere Erläuterung der Funktionsweise von Dateien, insbesondere von Videodateien erfolgen. Somit wird der Verfasser Hardware und Software gemeinsam behandeln, um durchweg die Nachvollziehbarkeit zu erhalten. Folgend auf die Evaluation verschiedener Medien bzw. Technologien wird ein Vergleich verschiedener auf dem Markt befindlicher Systeme vorgenommen. Mit einem durchdachten Erwartungshorizont und evidenten Bewertungskriterien werden dabei Archivierungssysteme auf ihre Passgenauigkeit analysiert und schlussendlich gegeneinander abgewogen. Attribute wie Schreibgeschwindigkeit, Kapazität und Kosten spielen beim Vergleich der Systeme eine tragende Rolle. Nur so kann der Verfasser gewährleisten, dass das von ihm vorgeschlagene System den Anforderungen des Arbeitsalltages entspricht und sich von anderen Systemen durch seine praktische Funktionalität positiv abhebt.

Archivierungslösungen existieren in vielen Formen und Ausführungen. Viele Technologien stehen zur Verfügung und lassen sich unterschiedlich kombinieren. Dabei ist die Anzahl der Speichermedien und die optimale Redundanz der Daten stets ein wichtiges Merkmal eines Archivs.³ Mehrere Datenträger mit gleichen Datensätzen minimieren das Risiko des Datenverlustes, fällt ein Speicher aus, muss es ein weiteres Medium geben, auf dem die Daten weiterhin bestehen. Die Wahl des Speichermediums ist dabei von entscheidender Bedeutung. Sichert man Daten auf einer Festplatte, unterscheidet sich deren Lebensdauer im Vergleich zu einer Sicherung auf einem Magnetband erheblich.⁴ Hat man sich erst einmal für ein Medium entschieden, gibt es eine Vielzahl an Herstellern mit zahlreichen an Angeboten. Speziell auf Archivierung optimierte Systeme grenzen den Markt zwar ein, nehmen diesem dennoch nicht die Komplexität. Systeme mit vielen Komponenten steigern natürlich diese Komplexität. Je mehr Nutzer zugreifen sollen, desto sicherer muss der Zugang sein.

Das Ziel ist die Darstellung unterschiedlicher Systeme, um die vielfältigen Möglichkeiten der Archivierung aufzuzeigen. Um die Archivsysteme trotzdem vergleichbar zu machen, greift der Verfasser auf möglichst hardwareunabhängige Faktoren zurück. Dadurch lassen sich strukturell unterschiedliche Systeme aufgrund ihrer Leistung hinsichtlich der Langzeitarchivierung vergleichen.

Der Verfasser nimmt eine objektive, auf Fakten basierende Bewertung vor und nimmt Stellung dazu. Die Bewertung ist dabei eine persönliche Meinung ohne Anspruch auf universelle Gültigkeit. Weiterhin analysiert der Verfasser aktuelle Systeme und nennt dabei die Namen von Herstellern, ohne persönliche Beweggründe finanzieller Natur,

³ Vgl. Ullrich (2010): Bitstream preservation. In: Neuroth; Oßwald; Scheffel et al. (Hrsg.). nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. S. 4

⁴ Vgl. Friedewald; Leimbach (2011): Computersoftware als digitales Erbe: Probleme aus Sicht der Technikgeschichte. In: Robertson-von Trotha; Hauser (Hrsg.): Neues Erbe: Aspekte, Perspektiven und Konsequenzen der digitalen Überlieferung, S. 204f

die seine Meinung beeinflussen könnten. Zur Erschließung des nötigen Wissens trat der Verfasser im Laufe seiner Recherchen mit verschiedenen Herstellern von Speichertechnologien in Kontakt, ohne den oben erwähnten Bewertungsrahmen zu verlassen. Das Ergebnis der Arbeit ist somit eine Leitlinie und kann zur späteren Implementierung mit einem effizienzorientierten Vorschlag zur passenden Lösung beitragen.

2 Archivierung

Mit dem enormen Zuwachs an Informationen hat sich im 21. Jahrhundert die weltweite Infrastruktur einer großen Herausforderung stellen müssen. Mehr Informationen bedeuten unweigerlich die Erfordernis neuer Systeme zur Speicherung und Erhaltung von Daten. Die 1971 erstmals für den Verbraucher erhältliche Floppy Disk Drive (FDD) in der 8-Inch-Ausführung fasste ganze 80 Kilobyte⁵. Die für die damaligen Verhältnisse revolutionär großen Speicher wurden über die Jahre hinweg erweitert und so kamen 1976 die 5,2-inch⁶ Disks mit 360KB auf den Markt. Derart kleine Datenträger sind in Zeiten zunehmender Digitalisierung inzwischen längst nicht mehr aktuell. Durch diese Entwicklung bildete sich um digitale Produkte als auch um die dafür erforderlichen Speichermedien und Technologien ein milliardenschwerer Markt. Deutlich macht dies eine Statistik von Gartner, aus der hervorgeht, dass alleine 2017 der weltweite Absatz von Computern 262,54 Millionen Einheiten betrug⁷. Computer sind dabei nur ein kleiner Teil der Digitalisierung, andere elektronische Unterhaltungs- und Speichergeräte sind in der Rechnung nicht eingerechnet. Tausende Hersteller sprangen seitdem auf den Digitalisierungs-Zug auf und entwickelten neue Produkte, die zu Meilensteinen des technologischen Fortschrittes wurden.

Mit der Entwicklung unterschiedlicher Computersysteme kamen über Jahrzehnte hinweg zahlreiche neue Formate und Datei-Systeme auf den Markt. Die steigende Anzahl an Informationen erfordert dabei unweigerlich immer mehr Speicherplatz. Besonders erkennbar wird dies bei den steigenden Auflösungen im Bereich der Film- bzw. Video-Produktion.

Processor or Virtual Storage	Disk Storage
· 1 Bit = Binary Digit	· 1 Bit = Binary Digit
· 8 Bits = 1 Byte	· 8 Bits = 1 Byte
· 1024 Bytes = 1 Kilobyte	· 1000 Bytes = 1 Kilobyte
· 1024 Kilobytes = 1 Megabyte	· 1000 Kilobytes = 1 Megabyte
· 1024 Megabytes = 1 Gigabyte	· 1000 Megabytes = 1 Gigabyte
· 1024 Gigabytes = 1 Terabyte	· 1000 Gigabytes = 1 Terabyte
· 1024 Terabytes = 1 Petabyte	· 1000 Terabytes = 1 Petabyte
· 1024 Petabytes = 1 Exabyte	· 1000 Petabytes = 1 Exabyte
· 1024 Exabytes = 1 Zettabyte	· 1000 Exabytes = 1 Zettabyte
· 1024 Zettabytes = 1 Yottabyte	· 1000 Zettabytes = 1 Yottabyte
· 1024 Yottabytes = 1 Brontobyte	· 1000 Yottabytes = 1 Brontobyte
· 1024 Brontobytes = 1 Geopbyte	· 1000 Brontobytes = 1 Geopbyte

Abbildung 1: Einheiten der Speicherkapazitäten

⁵ Vgl. Vaughan-Nichols (2017): Floppy disk history: The evolution of personal computing.

⁶ Vgl. Bähring (2013): Mikrorechner-Systeme: Mikroprozessoren, Speicher, Peripherie. S. 510

⁷ Vgl. Business Wire (2017) : Absatzzahlen der PC-Hersteller weltweit von 2006 bis 2017 (in Millionen Stück).

Mit den zunehmenden Datenmengen und den steigenden Nutzerzahlen digitaler Systeme ist das Thema Archivierung zunehmend wichtiger geworden. Das wachsende Bedürfnis nach der Sicherung wichtiger Daten hat dabei einen eigenen Markt geschaffen, der Lösungen für jeden Bedarf bietet. Für große Rechenzentren mit einem Bedarf von mehreren Petabyte (siehe Abbildung 1) bis hin zum Privatverbraucher mit einigen hundert Gigabyte an Datenbedarf, entwickeln die Hersteller für jede Situation passende Produkte. Erzeugte Dateien sollen dabei möglichst sicher und schnell verfügbar gespeichert werden.⁸

2.1 Langzeitarchivierung

Das Thema der Langzeitarchivierung ist wichtiger denn je. Die digitale Revolution bringt dem Menschen in fast allen Bereichen große Vorteile. Die Möglichkeiten der Kommunikation im 21. Jahrhundert unterscheiden sich unter nahezu allen Gesichtspunkten von der Kommunikation der Gesellschaft vor 100 Jahren. Mit dem Fortschritt der Technik begeben sich die Menschen aber auch in eine Abhängigkeit von den neuen Medien. Je mehr diese das Leben eines Menschen bestimmen und steuern, desto abhängiger wird er davon.⁹ Betrachtet man Videoproduktionen, deren Qualität und damit auch das Datenvolumen deutlich zugenommen haben, kann sich eine Abhängigkeit von besseren Speichersystemen abzeichnen. Daten zu erzeugen ist ein schneller einfacher Vorgang, diese Daten aber auf lange Zeit zu erhalten eine technische Herausforderung.

Das Video als etabliertes Medium unserer Zeit gibt Aufschluss darüber, wie die Menschen am Anfang des 21. Jahrhundert leben. Wird das Material erhalten, können zukünftige Generationen umfassende Geschichtsforschung betreiben, als vergangene Generationen es je konnten. Dieser grundlegende Aspekt dient Gesellschaften, Institutionen und Privatpersonen und kann natürlich auch zur Erhaltung kultureller Schöpfungen an der Hochschule Mittweida von Nutzen sein. Die Langzeitarchivierung sichert die vergangenen, aktuellen und zukünftigen digitalen Werke der Studenten und schafft so die Möglichkeit der späteren Dokumentation.

Der Wandel vom Analogen zum Digitalen ist seit dem Zuwachs der Nutzung leistungsfähiger Rechenmaschinen im Privatbereich omnipräsent. „Neue“ Medien erfreuen sich eines wirtschaftlichen Booms und weltweiter Akzeptanz. Attraktive digitale Angebote aktueller Medienunternehmen versprechen Nutzern digitaler Endgeräte ein besseres, interessanteres Leben. Die Nutzer digitaler Technologien haben dennoch nur selten eine Vorstellung davon, was sich hinter dem Begriff „digital“ verbirgt. Das Wissen darüber, wie Daten gespeichert werden, wer darauf Zugriff hat und wie sicher die persönlichen Daten eines jeden Nutzers sind, erfordert eine Auseinandersetzung mit dem Thema. Es gilt als selbstverständlich, dass Daten jederzeit abrufbar sind und dass die

⁸ Vgl. Scheuring; Waldspühl (2012): Informatik für technische Kaufleute und HWD. S.160

⁹ Vgl. Grunwald (2010): Technikfolgenabschätzung: eine Einführung. S. 30

digitalen Systeme den Nutzer in seinem Alltag unterstützen.¹⁰ Sollen digitale Informationen auch in vielen Jahren funktionieren, sollte die Langzeitarchivierung thematisiert werden.

Langzeitarchivierung grenzt sich von der einfachen Ablage digitaler und analoger Daten ab. Sie soll Informationen für einen langen Zeitraum haltbar und jederzeit verfügbar machen.¹¹ Das von mehreren Raumfahrt-Institutionen entwickelte OAIS-Referenzmodell (Open Archival Information System) sieht die Hauptaufgaben eines Langzeitarchivs in der:

„Sicherung von Authentizität und Integrität durch dauerhafte Stabilisierung und Zugänglichmachung von authentischen unikalen Kontexten“.¹²

Obwohl anzunehmen wäre, dass das bereits 1997 entwickelte Modell aufgrund des starken technischen Wandels der letzten zwei Jahrzehnte inzwischen keine Aktualität mehr beanspruchen kann, lässt sich die Leitlinie des OAIS noch heute auf die Konzeption eines Langzeitarchivs anwenden. Das OAIS dient dabei der Herleitung wichtiger Kriterien zur Schaffung eines Systems und betrachtet das Thema Archivierung unabhängig von der technischen Umsetzung.

Ein Aspekt ist, dass Daten mehr als 10 Jahre lang gespeichert werden sollen, ohne Gefahr zu laufen, durch den Defekt eines Geräts gelöscht oder beschädigt zu werden:

„Zielrichtung einer Langzeitarchivierung ist, einem Informationsverlust entgegen zu wirken, der seine Ursache im rasanten technischen Fortschritt und den damit verbundenen kurzen Lebenszyklen DV-technischer Systeme hat. Langzeitarchivierung strebt nach Langzeitverfügbarkeit, also nach einer Sicherung der Nutzbarkeit der Inhalte.“¹³

Ständig neuentwickelte Technologien führen dazu, dass deren Nutzbarkeit in den folgenden Jahren nur schwer einschätzbar ist. Können alte Datenträger nicht mehr von neuen Lesegeräten gelesen werden, muss ein hoher Aufwand betrieben werden, um die Daten auf ein neues System zu überführen.¹⁴ Bei der Langzeitarchivierung sollen also Überlegungen zur Sicherung der Daten so getroffen werden, dass sie über viele Jahrzehnte hinweg funktionieren und der Aufwand für eventuelle Umschichtungen so gering wie möglich ausfällt. Die Lebensdauer einzelner Trägermedien spielt bei der

¹⁰ Vgl. Naumann (2013): Digitalisierung als Katalysator für die Fotografieforschung. In: Ziehe; Hägel (Hrsg.): Fotografie und Film im Archiv: Sammeln, Bewahren, Erforschen. S. 204

¹¹ Vgl. Wallaszkovits (2010): Qualitätsaspekte der Langzeitarchivierung von Archiven im Web. In: Fennesz-Juhász; Fröschl; Hubert: Digitale Verfügbarkeit von audiovisuellen Archiven im Internet-Zeitalter: Beiträge zur Tagung der Medien Archive Austria und des Phonogrammarchivs der Österreichischen Akademie der Wissenschaften : Dietrich Schüller zum 70. Geburtstag. S. 85ff

¹² Brübach (2010): Das Referenzmodell OAIS. In: Neuroth; Oßwald; Scheffel et al. (Hrsg.): nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. S. 6

¹³ Schwartz (2015): Digitale Langzeitarchivierung: Definition und Problemstellung. In: Umlauf; Hobohm (Hrsg.): Erfolgreiches Management von Bibliotheken und Informationseinrichtungen. Abschn. 9.4.2.2.1

¹⁴ Vgl. Schwartz (2015): Abschn. 9.4.2.2.1

Langzeitarchivierung also eine tragende Rolle. Ist darüber hinaus keine Sicherheit der Funktionalität gewährleistet, ist das Medium als unbrauchbar einzustufen und entspricht nicht dem Ziel einer langfristigen Erhaltung von Daten. CD-ROMS, DVD und Magnetbänder haben eine Haltbarkeit von ca. 30-50 Jahren.¹⁵ Bei einem Bedarf der Speicherung über einen Zeitraum von mehr als 50 Jahren gäbe es keine Garantie auf Sicherheit der Verwahrung und ein erhebliches Verlustrisiko. Daraus lassen sich zwei wichtige Eigenschaften der Langzeitarchivierung erkennen: Langlebigkeit und Kompatibilität.

Die Digitalisierung hat das Thema der Langzeitarchivierung in ein neues Licht gerückt. Inhalte werden immer häufiger nur noch digital aufbewahrt und ältere analoge Medien müssen nachträglich digitalisiert werden. Die Vorteile der digitalen Aufbewahrung liegen dabei auf der Hand. Digitale Speicherung erspart dem Menschen die Lagerung große Mengen Papier oder Kassetten und schafft damit Platz in der Wohnung und im Arbeitsumfeld. Darüber hinaus lassen sich digitale Inhalte bei sachgemäßer Ablage gut verwalten.

Konkret auf ein Archivsystem angewendet, weist das Konzept der Langzeitarchivierung ein breites Faktorenspektrum auf. Dabei sind Faktoren unterschiedlich zu priorisieren.(siehe Kapitel 6.1)

¹⁵ Vgl. Schwartz (2015): Abschn. 9.4.2.2.1

3 Terminologien und technische Grundlagen

Zum Verständnis der Funktionsweisen digitaler Systeme ist es wichtig, zuerst das Wort „digital“ zu definieren, um eine Grundlage für weitere Ausführungen zu schaffen. Digitale Signale folgen dem Binärsystem und bestehen aus Matrizen diskreter Werte, die aus der Kombination der Zahlen Null und Eins bestehen.¹⁶ Digitale Signale sind zudem zeitlich nicht kontinuierlich. Zur Wandlung eines analogen in ein digitales Signal wird ein A/D-Wandler benötigt. Dieser misst mit einem Abtast-Verfahren in gleichen zeitlichen Abständen (Sampling) die Frequenz des Analogsignals und codiert dieses als Zahlen.¹⁷ Durch den zeitlichen Abstand einzelner Abtastpunkte weist das digitale Signal weniger Informationen auf, da nicht abgetastete Werte im codierten Digitalsignal nicht vorkommen. Die digitale Signalverarbeitung ist dadurch fehlerbehaftet und deutlich informationsärmer. Die Reduktion der Informationen bei digitalen Daten ist dabei aber nicht zwingend als Nachteil anzusehen. Digitale Daten können einfacher und platzsparender gelagert werden und unterliegen keinem Qualitätsverlust.¹⁸ Werden Datenträger gepflegt und sensible Daten redundant gespeichert, können digitale Daten viele Jahrzehnte problemlos überstehen.

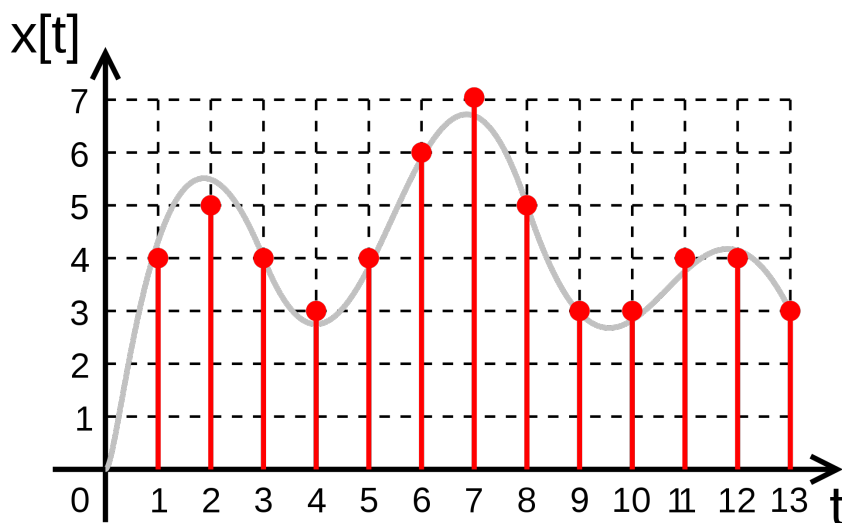


Abbildung 2: Digitale Signalverarbeitung

¹⁶ Vgl. Meinel; Sack (2009): Digitale Kommunikation: Vernetzen, Multimedia, Sicherheit. S. 11

¹⁷ Vgl. Lehner (2001): Einführung in Multimedia: Grundlagen, Technologien und Anwendungsbeispiele. S. 64f

¹⁸ Vgl. Schmitt-Egenolf (1990): Kommunikation und Computer: Trends und Perspektiven der Telematik. S. 183

3.1 Formate

Unter einem Format versteht man in der Informatik die innere Struktur einer Datei, welche festlegt, wie Informationen angeordnet sind.¹⁹ Ein Format unterliegt Konventionen, die eingehalten werden müssen. Es folgt Richtwerten und Regeln, damit gewährleistet werden kann, dass möglichst jedes Format interpretiert werden kann. Jede Datei endet mit einer zwei-, drei- oder vierbuchstabigen „File Extension“, zu Deutsch Dateinamenerweiterung²⁰. Diese teilt dem Betriebssystem mit, um welche Datei es sich handelt und wie sie interpretiert werden kann. Durch die Digitalisierung in den Anwendungsbereichen gibt es verschiedenste Arten von Formaten wie z.B. Grafikformate, Audioformate oder Containerformate. Zudem lassen sich Formate auch nach ihrem Inhalt kategorisieren. Grafikformate enthalten dementsprechend grafische Informationen; Audioformate hingegen auditive Informationen. Sogenannte Containerformate dienen der Bündelung verschiedener Codierformate zu einer einzigen Datei.²¹ Als Beispiel sei hier das AVI-Format (Audio Video Interleave) erwähnt, welches ein Audio- sowie ein Videoformat in einer Datei vereint. Containerformate können neben Video- und Audioinformationen z.B. auch Textinformationen oder Metadaten enthalten.²² Beim Interpretieren der im Containerformat enthaltenen Informationen sind verschiedenste Codecs nötig.

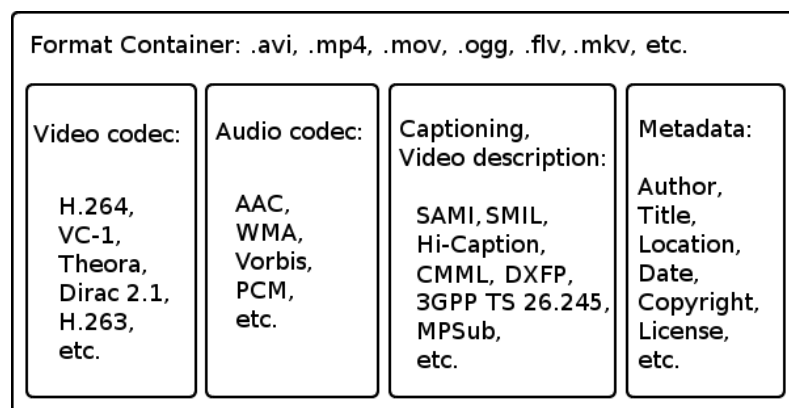


Abbildung 3 Containerformat

3.2 Codecs

Das Kunstwort Codec (gebildet aus den Wörtern „code“ und „decode“) beschreibt einen Algorithmus zum Kodieren und Dekodieren von Daten oder Signalen. Genauer handelt sich um ein Paar aus Algorithmen, wobei der eine Teil jeweils das Kodieren,

¹⁹ Vgl. IT Wissen (2014): Dateiformat.

²⁰ Vgl. ebenda

²¹ Vgl. Hansch; Rentschler (2012): Emotion@Web: Emotionale Websites durch Bewegtbild und Sound-Design. S.24

²² Vgl. Pilgrim (2010): HTML5: Up and Running: Dive into the Future of Web Development. S.82

der andere das Decodieren übernimmt.²³ Die Algorithmen sind dabei interdependent voneinander. Dies bedeutet, dass ohne installierten Codec ein ehemals kodierte Signal nicht mehr dekodiert werden kann, da das Betriebssystem die Funktionsweise des Codes nicht zu interpretieren versteht.²⁴

Daten bzw. Signale können analoger oder digitaler Natur sein. Codecs sind auch bekannt als A/D-Wandler („analoge to digital converter“), sowie als D/A-Wandler (digital to analoge converter).²⁵ Am Beispiel einer Videoaufnahme mit einer Videokamera wandelt der Codec die audiovisuellen analogen Signale, die von der Kamera aufgezeichnet werden, in eine digitale Datei um, die auf dem Kameraspeicher gespeichert werden. Dieser Vorgang ist als A/D-Wandlung anzusehen. Bei der D/A-Wandlung wird aus einer digitalen Datei ein analoges Signal.

Neben dem Wandeln von Informationen können Codecs auch digitale Daten in andere digitale Daten konvertieren.²⁶ Besonders gebräuchlich ist hierbei die Videokompression, bei der Videodaten komprimiert werden. Rohdaten einer Videoaufnahme zeichnen sich in der Regel durch ihre hohe Dateigröße aus. Ein Codec kann die Rohdaten wandeln und die Dateigröße signifikant reduzieren.

3.3 Videokompression

Videodaten unterscheiden sich von Text oder Bilddaten besonders in ihrer Dateigröße. Durch die Aufnahme mehrerer Bilder pro Sekunde können sich bei Videos in guter Qualität rapide große Datenmengen ansammeln. Die Videokompression kann helfen, Speicherprobleme zu lösen und Videos nutzbarer und übertragbarer zu machen. Videokompression bedeutet also nicht weniger als die Reduzierung von Information zur Reduzierung der Datenmenge. Datenreduktion lässt zunächst vermuten, dass das Ergebnis einer Videokompression sichtbar schlechter ist, dennoch ist die Videotechnik bereits derartig ausgereift, dass eine Kompression der Daten kaum wahrnehmbar ist. Zur Verdeutlichung der Sinnhaftigkeit von Datenkompression wird zunächst die hohe Datenmenge einer Rohdatei in einem Beispiel erläutert:

„Man möchte einen Film in Full HD-Qualität (1920 x 1080 50p) auf einem Bildschirm wiedergeben. Jede der drei Grundfarben soll dabei mit einer Auflösung von 10 Bit dargestellt werden. Ein RGB-Pixel beansprucht also 30 Bit. Bei einer Auflösung von 1920 (horizontalen) x 1080 (vertikalen) Pixeln belegt ein einziges Vollbild 2.073.600 Pixel x 30 Bit = 62.208.000 MBit (= 7,776 MByte) Speicher. Bei 50 Vollbildern pro Sekunde käme man also auf 388,8 MByte an Daten pro Sekunde! Eine DVD mit 4,7GByte (bei 1GByte=1000 MByte) könnte gerade

²³ Vgl. Jackson, Wallace (2012): Learn Android App Development. S.274

²⁴ Vgl. Maslo; Vonhoegen (2005): Audio, Video, Foto. S.707

²⁵ Vgl. Wintermeyer; Kempgen (2009): Asterisk 1.4 + 1.6: Installation, Programmierung und Betrieb. S. 173

²⁶ Vgl. Adam (2010): Micromovie: Ein kreatives Medium für mobile Endgeräte. S.104

*einmal 12 Sekunden Film in diesem Urformat aufnehmen. Diese Rechnung berücksichtigt dabei noch nicht einmal den Ton und den Overscan des Bildes!*²⁷

Datenmengen eines solchen Umfangs sind weder effizient speicher-, noch übertragbar. Um große Datenmengen zu reduzieren, muss ein Kompromiss eingegangen werden. Lediglich die Reduktion von Informationen erlaubt eine Reduktion der Datenmenge. In der Videotechnik wird von „verlustfreier“ und „verlustbehafteter“ Videokompression unterschieden:

„Kompression heißt, eine Information (verlustfrei) oder eine deren ähnliche (verlustbehaftet), in einer Form darzustellen, in der weniger Daten zu ihrer Beschreibung benötigt werden.“²⁸

Example of Lossy Compression



Abbildung 4: Verlustbehaftete Kompression

Verlustfreie (lossless) Datenkompression sei in diesem Kontext nur erwähnt und findet in den folgenden Kapiteln keine weitere Beachtung. Verlustfreie Kompression verwirft bei der Redundanzreduktion²⁹ lediglich redundante Informationen und ist dadurch reversibel.³⁰ Verlustbehaftete (lossy) Kompression hingegen entfernt irrelevante Informationen und gewährleistet somit eine deutlich kleinere Informationsmenge. Die sogenannte Irrelevanzreduktion³¹ definiert das planmäßige Entfernen von unwichtigen Daten nach der rechnerischen Vorgabe eines Kompressionsverfahrens. Unwichtige bzw. irrelevante Informationen sind solche, die die richtige Wahrnehmung des Videos nicht schmälern:

²⁷ Infotip Service GmbH (2018): Grundlagen Videokompression.

²⁸ Kaiser, Alexander (2003): Transformationscodierung. S.2

²⁹ Vgl. Schmitz; Kiefer; Maucher et al (2006): Kompendium Medieninformatik: Mediennetze. S. 4

³⁰ Vgl. ebenda, S. 4

³¹ Vgl. ebenda, S. 4

„Da Helligkeitsunterschiede stärker wahrgenommen werden als Farbunterschiede, können die Farbdifferenzsignale (U, V) mit einer geringeren Abtastfrequenz digitalisiert werden als das Luminanzsignal.“³²

Entfernen von Farbinformationen kann das menschliche Auge dementsprechend schlechter wahrnehmen. Beim Chroma Subsampling, einer gängigen Methode in der Videokompression, werden Farbinformationen stärker komprimiert als Helligkeitsinformationen. Das YCbCr-Farbmodell unterteilt das Farbbild in die drei Kanäle Y (Luminanz), Cb (Chrominanz Blau) und Cr (Chrominanz Rot).³³ Des Weiteren wird beim Chroma Subsampling versucht, die YCbCr-Informationen eines 4-Pixelblocks oder auch zweier benachbarter Pixel gleichzusetzen und mit anderen Worten Farbinformationen zu entfernen. Bei einem Block aus vier Pixeln wäre die Information somit nur noch ein Viertel so groß.³⁴ Bei einer zu starken Reduktion der Farbinformationen können wahrnehmbare Artefakte auftreten.

Videokompression ist immer ein Kompromiss. Der Verlust von Informationen ist dabei unumgänglich und je nach Kompressionsverfahren mehr oder weniger sichtbar.

3.4 Hausformat: XDCAM HD422

Bei diesem Format handelt es sich um das einheitlich genutzte Dateiformat für Videodaten in der Hochschule Mittweida. Der Gebrauch eines festen Formats dient dabei besonders der Einheitlichkeit und Kompatibilität der Daten. Das von Sony entwickelte Format verspricht hohe Qualität zu moderater Datengröße. Die Rohdaten der Kameras müssen vor der Arbeit mit dem Videomaterial erst ins Hausformat gewandelt werden. Der Nachfolger des Formats „XDCAM HD“ arbeitet dabei mit anderen Kompressionsparametern und einer höheren Datenrate. Mit einer 4:2:2-Signalverarbeitung werden Daten bei 50Mbit/s so quantisiert,³⁵ dass möglichst wenig Qualität bei der Wandlung der Rohdaten in XDCAM HD422 verloren geht. Die Auflösung von 1920x1080 erlaubt qualitativ hochwertige HD-Aufnahmen zur Verwendung im Bereich der Fernsehübertragung. Als Speichermedium wird eine Dual-Layer Professional Disk von Sony verwendet. Laut eigenen Angaben der Hochschule soll eine Stunde Filmmaterial mit dem Hausformat um die 50 Gigabyte beanspruchen.

³² Lichtscheidl (2018): Video.

³³ Vgl. Klimpke; Keimel; Diepold (2012): Visuelle Qualitätsmetrik basierend auf der multivariaten Datenanalyse von H.264/AVC Bitstream-Features. S. 13

³⁴ Vgl. Malaka; Butz; Hußmann (2009): Medieninformatik: eine Einführung. S. 99

³⁵ Vgl. Sauter (2009): Video. In: Neuroth; Oßwald; Scheffel et al. (Hrsg.): Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Kap. 17 S.26f

3.5 Metadaten

Metadaten sind Informationsdaten, welche der Beschreibung anderer Daten (Objekte) dienen.³⁶ Sie enthalten mannigfaltige Informationen über das Objekt. Die Metadaten beschreiben verschiedene Aspekte des Objekts wie z.B. den Inhalt oder die Beschaffenheit. Je nach Situation und Kontext sind einige der Metadaten relevant und andere nicht. Als Beispiel sei an dieser Stelle das folgende Zitat erwähnt:

„In einer Bibliothek z. B. würde dies dazu führen, dass das Erscheinungsjahr und der Autor eines Buches sehr wohl Metadaten darstellen, die Qualität des Papiers oder die Farbe des Einbandes jedoch nicht, obwohl es sich in beiden Fällen um Daten über die verwalteten Daten (die Bücher) handelt.“³⁷

Metadaten lassen sich also nach ihrer Relevanz für eine gegebene Situation klassifizieren. Besonders wichtig können Metadaten beim Datenmanagement sein. Bei einer konsequenten Einhaltung der Vergabe von Metadaten zu Daten lassen sich diese schnell sortieren und filtern. Metadaten lassen sich in verschiedene Kategorien unterteilen. Die nestor-Arbeitsgruppe kategorisiert Metadaten wie folgt:

„deskriptive Metadaten: Es handelt sich um klassische Erschließungsmetadaten. Diese beschreiben ein Objekt und gewährleisten, dass es wiederauffindbar ist. Erschließungsmetadaten unterteilen sich in formale Metadaten wie Titel, Urheber, Erscheinungsjahr usw. und inhaltsbeschreibende Metadaten wie Schlagworte, ein Abstract, ein Transkript, Sequenzbeschreibungen usw.

administrative Metadaten: Administrative Metadaten sind für die Verwaltung der Objekte im Langzeitarchiv und den Betrieb des Systems erforderlich, z.B. systeminterne Identifier, Nutzerdaten, Rollen- und Rechtevergabe.

rechtliche Metadaten: Rechtliche Metadaten beschreiben den Nutzerzugriff auf die Objekte. Sie legen auch fest, ob die Objekte bearbeitet und wie sie bereitgestellt werden dürfen.

technische Metadaten: Technische Metadaten beschreiben z.B. Dateiformat und -version, den Codec, Laufzeit, Dateigröße, Bitrate usw. Technische Metadaten können aus der Datei extrahiert werden.

strukturelle Metadaten: Strukturelle Metadaten beschreiben die Zusammensetzung eines Objekts und Zusammenhänge mit anderen, z.B. wenn ein Objekt aus mehreren Dateien in einer festgelegten Reihenfolge besteht oder wenn Objekte Teil einer Serie sind [sic!].

Provenienzmetadaten: Diese Metadaten beschreiben die Objektgeschichte, um zu dokumentieren, welcher Nutzer wann mit welchen Tools welche Bearbeitung vorgenommen hat. Diese Metadaten sind von großer Wichtigkeit für den Nachweis der Authentizität.“³⁸

³⁶ Vgl. Anahory; Murray (1997): Data Warehouse: Planung, Implementierung und Administration. S. 155

³⁷ Huthmann (2004): Metadaten und Datenqualität in Data Warehouses. S. 5

³⁸ Hasler (2017): Empfehlungen für größere Einrichtungen.

3.6 Bedarfsermittlung

Die Studenten der Hochschule Mittweida stellen in verschiedenen Lehrmodulen verschiedenste Bewegtbild-Produktionen her. Neben kleinen Projekten sind besonders die großen Produktionen Soundcheck Rock, Soundcheck Pop, Soundcheck Urban sowie das Campusfestival zu nennen. Dabei arbeiten oft bis zu 60 Studenten gemeinsam an der Entstehung der beliebten Formate.

Die Bedarfsermittlung für das Langzeitarchiv hat der Verfasser zusammen mit dem Studiendekan für „Media and Acoustical Engineering“ Prof. Dipl. Toningenieur Mike Winkler erarbeitet. Dabei wurde grob ein Bedarf errechnet, der einen Zeitraum von etwa 40 Jahren abdeckt. Nach dieser Berechnung blieben der Hochschule Mittweida weitere Investitionen in die Archivierung während dieses Zeitraumes weitgehend erspart.

Berechnung:

1 Std Videomaterial entspricht 50 GB³⁹

6 Std pro Woche = 300GB

24 Std pro Monat = **1,2 TB**

288 Std pro Jahr = **14,4 TB**

10 Jahre = 144 TB, **40 Jahre = 576 TB**

Es besteht bereits ein Datensatz von 30 TB, der direkt archiviert werden soll. Neben Videodaten können andere Daten, wie Fotos oder Audiodateien, unvorhergesehen dazukommen. Der Bedarf orientiert sich somit grob an der Menge von **600 Terabyte** für eine Langzeitarchivierung von 40 Jahren.

Auf das Jahr gerechnet erbringt zur Zeit jeder Student allein, und mit anderen Kommilitonen zusammen vier bis fünf bewertete Bewegtbild-Leistungen. Nicht jedes Projekt muss dabei gespeichert werden, da sich die Arbeiten zum einen im Umfang, zum anderen in der Relevanz und in der Qualität unterscheiden. Es sollen lediglich fertiggestellte Produktionen archiviert werden.

³⁹ interne Information (Hochschule Mittweida, 2018)

4 Aktuelle Technologien

Der Technikmarkt ist ein sich ständig verändernder Markt. Neues kommt und ersetzt Altes. Dabei sind Flexibilität und Wachsamkeit wichtige Eigenschaften eines jeden Anwenders.⁴⁰ Das Ziel der Industrie, sich ständig zu verbessern, zeigt sich in der omnipräsenten Weiterentwicklung bereits bestehender und der Entwicklung neuartiger Technologien. Aus diesem Grund werden Informationen vermehrt digital, und immer häufiger ausschließlich digital aufbewahrt.⁴¹ Es kann sich also für jeden Anwender, ob Konsument oder Unternehmen, lohnen, sich mit den neuen digitalen Speichersystemen zu beschäftigen. In den folgenden Kapiteln sollen deshalb Speichermedien und Technologien zur Administration und Organisation von Daten im Fokus stehen. Darüber hinaus werden Speichernetzwerke behandelt, welche bei den Erläuterungen zu Archivsystemen (Kapitel 5) näher zum Tragen kommen.

4.1 IT-Infrastruktur der HS Mittweida

Die Implementierung eines Archivsystems erfordert das Wissen von der bereits bestehenden Infrastruktur (siehe Anhang II). Nur so kann gewährleistet werden, dass das Langzeitarchiv eingegliedert werden kann. Nicht aufeinander abgestimmte Hardware kann zu irreversiblen Fehlern führen. So können mangelhaft zusammengestellte Systeme zu Zeitpunkten, bei denen es um akute zeitliche Präzision geht, ausfallen oder fehlerhafte Vorgänge durchführen. Im folgenden Abschnitt soll der Fokus auf der Infrastruktur rund um die Video- und Audioproduktionen der Fakultät Medien an der Hochschule Mittweida liegen. Eine Erläuterung der gesamten Vernetzung aller Bereiche würde den Rahmen dieser Arbeit sprengen und nur wenig zur Analyse beitragen.

Die Hochschule Mittweida verfügt über umfangreiche, professionelle Kamera- und Aufnahmetechnik. Alle erfassten Videodaten werden direkt auf zwei verschiedene Video- bzw. Produktionsserver eingespielt. Vor der Ablage werden die Dateien zusätzlich ins Hausformat XDCAM HD422 (Siehe Kapitel 3.4) gewandelt. Die beiden Videoserver „K2“ von Grassvalley⁴² und „Venice“ von DVS⁴³ bieten jeweils vier Kanäle. Die auf den Produktionsservern abgelegten Videodaten sind dabei jederzeit abrufbar und übertragbar. Jeder Kanal (Ingest-Channel) bleibt dabei unabhängig bearbeitbar. Zusammen ergeben sich aus beiden Lösungen acht Einspielkanäle.

Nach der Ablage bzw. dem Einspielen der Daten werden diese durch die Software „Strawberry“ von FlavourSys für den Schnitt bereitgestellt. Strawberry ist ein für Video-

⁴⁰ Vgl. Urbach; Ahlemann (2016): IT-Management im Zeitalter der Digitalisierung: Auf dem Weg zur IT-Organisation der Zukunft. S. 161

⁴¹ Vgl. Anderson (2017): Die Idee einer menschlichen Digitalisierung. In: Keese; Anderson; Pousttchi et. al: Digital human: Der Mensch im Mittelpunkt der Digitalisierung. S. 13ff

⁴² Vgl. Grassvalley (2018): Produktbeschreibung K2 Summit 3G.

⁴³ Vgl. DVS (2018): Produktbeschreibung DVS Venice.

produktionen entwickeltes Content Management System.⁴⁴ Die Vorteile sollen dabei bei der einheitlich durchdachten Struktur der Softwareumgebung liegen. Als CMS (Content Management System) liegt der Fokus der Anwendung auf der Verwaltung von Inhalten. Entscheidend zum Tragen kommen hierbei die Metadaten, durch die Inhalte kategorisiert und geordnet werden können.

Bei einem erwünschten Zugriff der einzelnen Schnitträume auf Inhalte können verschiedene berechnigte Nutzer über Strawberry auf sämtliche vom System erfasste Daten zugreifen. Strawberry selbst ist kein Speicher, sondern lediglich eine Management Software. Die Schnitträume verfügen über eigene Speicherlösungen auf denen nach Bedarf Inhalte abgelegt werden können. Nach der Produktion werden die Endprodukte auf Festplatten gespeichert und somit archiviert.

Die Aufbewahrung auf einzelnen ungeordneten Festplatten ist keineswegs eine durchweg sichere Archivierungsmethode. An dieser Stelle soll das Langzeitarchivsystem ansetzen und über einen unbestimmten Zeitraum für mehr Datensicherheit sorgen (siehe Kapitel 6).

4.2 Speichermedien

Das Speichern von Daten bezeichnet die Lagerung von Informationen zur späteren Nutzung.⁴⁵ Das Speichermedium hält dabei die Informationen. Ein Speichermedium, auch Datenträger genannt, kann ein Gerät, z.B. eine Festplatte oder ein Material wie z.B. eine CD-ROM, sein. Die technischen Weiterentwicklungen haben verschiedenste Arten und Modelle von Speichermedien hervorgebracht. Dabei war stets die Sicherheit der Daten ein entscheidender Faktor für das Überleben einer Technologie. Speichert ein Datenträger die ihm übermittelten Daten nicht, erfüllt der Speicher seine Kernaufgabe – Erhalt der Daten – nicht. Bei sensiblen relevanten Daten steigt die Wichtigkeit der Datensicherheit, da beim Datenverlust Folgeschäden auftreten können. Da Defekte oft unerwartet entstehen, ist es wichtig, bei der Sicherung von Daten möglichst die höchste Sicherheitsstufe zu wählen. Sollen Daten mehrfach gespeichert werden, steigen natürlich schon wegen der zusätzlichen Speichermedien die Kosten des Systems. Im Folgenden werden einige Speichermedien beschrieben und Ihre Funktionsweise erklärt. Die Auswahl der untersuchten Techniken beschränkt sich dabei lediglich auf für das Langzeitarchivsystem relevante Technologien.

4.2.1 Magnetbandspeicher

Ein Magnetbandspeicher ist ein sequentiell arbeitendes Speichermedium.⁴⁶ Sequentiell bedeutet in dem Zusammenhang, dass es Informationen seriell, d.h. aufeinanderfol-

⁴⁴ Vgl. Flavorsys (2018): Startseite.

⁴⁵ Vgl. Freyer (2013): Medientechnik: Basiswissen Nachrichtentechnik, Begriffe, Funktionen, Anwendungen. S.67

⁴⁶ Vgl. Ottmann; Widmayer (2017): Algorithmen und Datenstrukturen. S. 142

gend abspeichert. Das Band wird dabei über Spulen in einer Hülle bewegt.⁴⁷ Informationen werden auf dem Band gespeichert und befinden sich somit an einem festen Ort auf dem Band. Sollen bestimmte Informationen gelesen werden, muss dieser Ort erst gesucht werden. Dies geschieht durch das Vor- oder Zurückspulen des Bandes:

„Das magnetische Speicherverfahren beruht auf der Änderung von Magnetisierungszuständen magnetischer Teilchen einer Schicht, die auf Folien (Diskette, Magnetband) oder Platten (Festplatte) aufgebracht ist. Das Schreiben und Lesen von Daten erfolgt auf elektromagnetischer Basis. Dazu wird ein kombinierter Schreib- und Lesekopf benutzt. Zum Schreiben der Daten erzeugt der Schreibkopf ein Magnetfeld, welches in der Nähe befindliche Bereiche der Magnetschicht magnetisiert (magnetische Remanenz). Die magnetischen Teilchen behalten auch nach dem Entfernen des Magnetfeldes ihren Zustand bei. Der Magnetisierungszustand kann durch ein Magnetfeld auch wieder verändert werden. Das Auslesen der Daten erfolgt durch Abtastung der Magnetisierung mit einem Lesekopf. Durch die Bewegung des Kopfes relativ zur magnetischen Schicht wird im Lesekopf ein Strom induziert, der ausgewertet werden kann.“⁴⁸

Das ursprüngliche Magnetband, auch Trägerband, besteht aus Kunststoff und ist ca. 37 Mikrometer dick. Auf dem Kunststoffband befindet sich zusätzlich eine 15 Mikrometer dicke Schicht aus magnetisierbarem Eisenoxyd.⁴⁹ Anfang und Ende des Bandes sind mit einer Reflektorschicht aus Aluminium markiert. Die magnetisierbare Fläche auf dem Band ist der Länge nach in 7,8 oder 9 Spuren aufgeteilt. Das Magnetband befindet sich dabei auf zwei Spulen in einer Kassette.

Magnetbänder haben einen großen Nachteil. Um Informationen von einem Band abzurufen, muss es erst von Spule zu Spule laufen und durch den Magnetkopf gelesen werden. Dies benötigt relativ viel Zeit. Wird eine bestimmte Information benötigt, muss das Band vom Anfang bis zum Zielort bewegt werden.⁵⁰ Magnetbänder eignen sich somit besser für die Speicherung von den Daten, die selten gebraucht werden.⁵¹ Zu groß wäre der Aufwand für eine einzelne kleine Datei. Durch den Gebrauch der Magnetbänder kommt es zu einem Verschleiß.⁵² Der bewusste Gebrauch der Magnetbänder sollte somit bei der Erhaltungsstrategie bedacht werden. Die regelmäßige Reinigung von Magnetbändern muss zusätzlich beachtet werden, da sie mit der Zeit verstauben können.

⁴⁷ Vgl. Goldammer (1994): Informatik für Wirtschaft und Verwaltung: Einführung in die Grundlagen. S.109

⁴⁸ Däßler, Rolf (2010): Datenträger und Speicherverfahren für die digitale Langzeitarchivierung. In: Neuroth; Oßwald; Scheffel et al. (Hrsg.): Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Kap. 10, S. 6

⁴⁹ Vgl. Fahrion (1989): Wirtschaftsinformatik: Grundlagen und Anwendungen. S. 116

⁵⁰ Vgl. Ortman (2003): Einführung in die PC-Grundlagen. S.112

⁵¹ Vgl. Fahrion (1989): S.115

⁵² Vgl. Ullrich (2008): Magnetbänder. In: Neuroth; Oßwald; Scheffel et al. (Hrsg.). nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. S. 164

LTO-Standard

Die Technologie der Magnetbandaufzeichnung (MAZ) hat sich über die Jahrzehnte deutlich gewandelt. In heutiger Zeit ist es mit einem LTO-Band (Linear Tape Open) möglich, mehrere Terabyte an Daten auf einem Band zu speichern. Der von IBM, Hewlett Packard und Seagate entwickelte Standard findet in vielen professionellen Einrichtungen Anwendung.⁵³ Als Standard wird die LTO-Technologie von vielen Speicherherstellern lizenziert genutzt. LTO-Bänder sind besonders für ihr große Speicherkapazität und das gute PreisLeistungsverhältnis bekannt. Der bereits in der achten Generation befindliche LTO-Standard (LTO-8) kann 12 Terabyte Rohdaten bzw. 30 Terabyte⁵⁴ komprimierte Daten speichern.⁵⁵ Das Kompressionsverfahren (2,5:1) beim LTO-System kann somit über die Hälfte an Dateigröße einsparen. Die Kosten eines Bandes in Kombination mit der gelieferten Kapazität macht diese Technologie sehr preiswert. Angaben zur Lebensdauer eines Bandes wird seitens der Hersteller auf ca. 30 Jahre⁵⁶ geschätzt.

Ein Linear Tape Open befindet sich, anders als bei klassischen Kassetten, in einer Cartridge, einer rechteckigen Plastikbox und nur auf einer Spule (single reel). Beim Lesevorgang wird das Magnetband eines LTOs in einem LTO-Streamer abgerollt. Die Lesegeschwindigkeit und Kapazität variieren je nach genutzter Generation des LTO-Bandes stark. Während die erste Generation noch 384 Spuren⁵⁷ zur Datenspeicherung nutzt, kann das LTO-8 mit 6,684⁵⁸ Spuren deutlich mehr Daten speichern. Eine LTO-Cartridge verfügt über einen eigenen kleinen Speicher, welcher Daten zu Cartridge-Nummer, aber auch zu den Zugriffszeiten sowie zu Error-Aufkommen enthält. LTO-Generationen sind zum Teil abwärts kompatibel und können somit über einen längeren Zeitraum als Archivierungsmethode fungieren.

LTO 1	LTO 2	LTO 3	LTO 4	LTO 5	LTO 6	LTO 7	LTO 8
100 GB	200 GB	400 GB	800 GB	1,5 TB	2,5 TB	6 TB	12 TB

Tabelle 1: Kapazitäten der LTO-Generationen (ohne Kompression)⁵⁹

⁵³ Vgl. Reiner (2005): HEAVEN. Eine hierarchische Speicher- und Archivierungsumgebung für multidimensionale Array Datenbankmanagement Systeme. S. 27

⁵⁴ Vgl. Hewlett Packard Enterprise (2018): Produktbeschreibung HPE LTO-8 Ultrium 30TB RW Data Cartridge.

⁵⁵ An dieser Stelle ist die Kapazität von 30 Terabyte irreführend. Das Band kann nur 12 Terabyte speichern, möchte der Anwender allerdings 30 Terabyte speichern, kann er diese durch das LTO-System komprimieren und die Datenmenge auf 12 Terabyte „schrumpfen“. Damit geht ein Qualitätsverlust einher.

⁵⁶ Vgl. Hewlett Packard Enterprise (2018)

⁵⁷ Vgl. Gloning (2008): Digitalisiertes Langzeitarchiv eines Fernsehsenders: Analyse der vernetzten, bandlosen Content-Produktion, -Speicherung und -Verwertung und der dazu nötigen technischen sowie infrastrukturellen Grundlagen. S. 27

⁵⁸ Vgl. Hewlett Packard Enterprise (2018)

⁵⁹ Vgl. Coyne; Engelbrecht; Browne (2018): IBM Tape Library Guide for Open Systems. S.19

LTFS

Das Linear Tape File System ergänzt die LTO-Technologie mit einer neuen Eigenschaft. Mit LTFS werden LTO-Bänder in zwei Partitionen geteilt. Die eine Partition enthält die Metadaten aller in Partition 2 gespeicherten Daten. Die Metadaten-Partition ist deutlich kleiner, sagt dennoch viel über die Gesamtheit der Daten auf dem Band aus. Zur Nutzung dieses Features benötigt das Betriebssystem die offene LTFS-Software, welche hardware- und softwareunabhängig arbeitet. Somit wird ein weltweit lesbarer Standard für LTO-Bänder ab LTO 5 gewährleistet. Im Vergleich zu herkömmlichen Lesevorgängen muss bei LTFS lediglich die kleine Metadaten-Partition gelesen werden und nicht das gesamte Band. Dem Anwender erscheint der LTO-Speicher wie eine herkömmliche Festplatte, mit der per „Drag & Drop“, Dateien hinzugefügt oder entfernt werden können.

4.2.2 HDD

Eine Hard Disk Drive, im allgemeinen Sprachgebrauch als Festplatte bezeichnet, beschreibt ein elektromagnetisches Speichermedium. Festplatten speichern, ähnlich wie Magnetbänder, Informationen durch elektromagnetische Ladungen.

„Die elektromagnetische Aufzeichnung kann im Analog- oder im Digitalverfahren vorgenommen werden. Die Aufzeichnung erfolgt, indem das verstärkte Signal dem Aufzeichnungskopf zugeführt wird. Im Aufzeichnungskopf wird ein magnetisches Feld erzeugt, das eine örtlich unterschiedliche Magnetisierung des am Aufzeichnungskopf vorbeigeführten magnetisierbaren Materials bewirkt. Bei der Wiedergabe induzieren die Magnetfelder im Wiedergabekopf elektrische Schwingungen. Diese werden elektronisch verstärkt und in Bild-, Ton- oder Binärsignale umgesetzt.“⁶⁰

Bei einer HDD werden mehrere Aluminiumplatten auf einer Spindel beidseitig mit einem ferromagnetischen Material versehen⁶¹. Auf der Oberfläche der Platte lassen sich durch einen Magnetkopf elektrische Ladungen auf dem ferromagnetischen Material erzeugen. Die sogenannten Weiß'schen Bezirke bezeichnen Felder aus kleinen elektromagnetischen Domänen (Dipolen), die nach einer Widerstandsänderung durch das Magnetfeld ihre elektromagnetische Ladung ändern können.⁶² Plus und Minus werden als „1“ für Plus und „0“ für Minus verstanden. Beim Lesevorgang erzeugen die magnetischen Ladungen im Lesekopf einen Spannungsimpuls, die die Interpretation der Ladungen „Plus“, „Minus“ bzw. „0“ und „1“ möglich machen.⁶³ Die Daten auf einer Festplatte werden in Sektoren und Spuren gespeichert. So wird ein schneller Abruf der Daten ermöglicht. Die Festplatte kann jeder geschriebenen Information einen Ort,

⁶⁰ Umlauf, Konrad (2006): Medienkunde. S. 43

⁶¹ Vgl. Stickel; Groffmann; Rau (Hrsg.) (2013): Gabler Wirtschaftsinformatik-Lexikon. S. 179

⁶² Vgl. Huebener (2014): Leiter, Halbleiter, Supraleiter - Eine Einführung in die Festkörperphysik: Für Physiker, Ingenieure und Naturwissenschaftler. S. 159

⁶³ Vgl. Manz (2016): Fehlerkorrigierende Codes: Konstruieren, Anwenden, Decodieren. S. 192

technisch ausgedrückt eine Adresse zuordnen. Die Adresse eines Sektors besteht aus dem „Zylinder“, der den Ort auf der Platte angibt, der „Kopfnummer“, die die Platte sowie die Seite der Platte angibt, und der „Sektornummer“, die letztlich den gesuchten Sektor auswählt.⁶⁴

Die Domänen befinden sich horizontal auf der Platte und benötigen einen freien Übergangsbereich nebeneinander.⁶⁵ Eine Festplatte kann also nur so viele Daten speichern, wie sie über Platz für Domänen verfügt. Der Wissenschaft gelingt im Jahre 2005 allerdings mit dem „perpendicular recording“ oder „PMR“ ein Durchbruch. So werden Domänen fortan vertikal angeordnet und benötigen einen kleineren Übergangsbereich. Die Dichte an möglichen Domänen auf Flächen aller Platten, ergo die Kapazität, ist deutlich höher. Der Hersteller Seagate entwickelte den PMR-Standard weiter und stellte das „shingled magnetic recording“ oder SMR vor. Dieses Verfahren lässt die Spuren auf der Platte überlappen. Auch so konnte die Kapazität weiterhin vergrößert werden. Mit dem SMR Verfahren sollen Datendichten von über einem Terabit/Inch² möglich sein.⁶⁶ Die bereits in der Realisierung steckenden Technologien „Microwave-Assisted Magnetic Recording“, kurz MAMR und „Heat-Assisted Magnetic Recording“, kurz HAMR sollen in der nächsten Dekade den Sprung auf bis zu 40 Terabyte pro Festplatte schaffen.⁶⁷

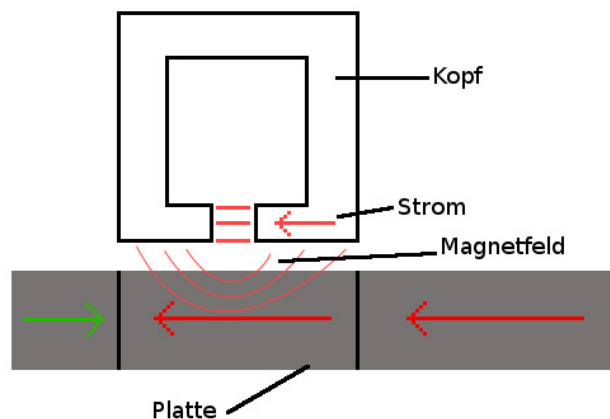


Abbildung 5: herkömmliche Speicherung auf HDD

Vergleicht man die Festplatte mit einem Magnetbandspeicher, wird schnell klar, dass die HDD deutlich schnellere Zugriffszeiten ermöglicht. Bei mehreren tausend Umdrehungen pro Minute⁶⁸ ist der Datendurchsatz sehr hoch und punktueller. Oft genutzte

⁶⁴ Vgl. Schiffmann; Bähring; Hönig (2011): Technische Informatik 3: Grundlagen der PC-Technologie. S. 169

⁶⁵ Vgl. ebenda, S. 169

⁶⁶ Vgl. Labs (2015): Dicke Brummer. In: c't Redaktion. c't TESTGUIDE: Die wichtigsten Tests des Jahres: PC & Zubehör, Grafikkarten, Mobile, Printing, Laptops, Storage, Netzwerke, Monitore. S. 237

⁶⁷ Vgl. Wachenschwanz (2018): HAMR vs MAMR: Which Technology Will Be Used For the Next Generation of Hard Disk Drives?

⁶⁸ Vgl. Giancoli (2006): Physik. S. 326

Daten können so einfacher gefunden werden. Zusätzlich kommt es zu einem nur geringen Verschleiß der Platte, da der Lesekopf diese nie berührt.

4.2.3 SSD

Die SSD, auch Solid State Drive, ist ein rein elektronischer Datenspeicher ohne bewegliche Teile. Auch oft als Halbleiterspeicher bekannt speichert eine SSD Daten auf Halbleiterbausteinen:

„Digitale Halbleiterspeicher dienen der zeitlich befristeten oder weitgehend unbefristeten Informationsspeicherung der logischen Zustände 0 (Low) oder 1 (High) in Speicherzellen, die mit elektronischen Schaltungen realisiert sind. Man unterscheidet das Abspeichern der Informationen (Write/Schreiben) und das Auslesen (Read/Lesen) der Information.“⁶⁹

Halbleiterspeicher können als flüchtiger und nicht-flüchtiger Speicher fungieren. Nicht-flüchtig bezieht sich dabei auf das Fortbestehen der Daten nach Trennung der Stromzufuhr.⁷⁰ Zu den flüchtigen Speichern zählt beispielsweise der RAM⁷¹, dem Arbeitsspeicher eines Rechensystems. Daten werden im RAM mit hoher Geschwindigkeit verarbeitet, erlöschen allerdings, wenn die Stromquelle keine Energie mehr liefert. Infolgedessen sind zum langfristigen Speichern von Informationen lediglich nicht-flüchtige Speicher sinnvoll.

Nicht-flüchtige Halbleiterspeicher, auch Flash-Speicher genannt, erhalten die Informationen also auch nach dem Abtrennen der Stromzufuhr. Ein Flash-Speicher besteht zum einen aus seinen Bausteinen, zum anderen verwaltet ihn ein Controller.⁷² Die Daten werden im Inneren des Speichers in einzelne Speicherzellen gespeichert. Jede Zelle kann, abhängig vom Modell, einige Bits speichern.

Bei SSDs wird zwischen SLC (Single Layer Cell), MLC (Multi Layer Cell) und TLC (Triple Layer Cell) unterschieden. Eine Speicherzelle einer SSD fasst somit 1-3 Bits pro Zelle. Bei einer SLC (ein Bit pro Zelle) ist die Kapazität folglich kleiner, da weniger Informationen in einer Zelle gespeichert werden als bei beispielsweise TLC (drei Bits pro Zelle). Single Layer Cells werden dafür beim Schreiben der SSD weniger beansprucht und weisen somit eine längere Lebensdauer auf.

Sogenannte „Pages“ bestehen aus Tausenden von Zellen. Soll nur ein Bit geändert werden, muss dabei der gesamte Block neu geschrieben werden. Eine SSD ist auf Verfügbarkeit ausgerichtet und reserviert mit „Over-Provisioning“ stets einige freie Blocks. Bei nicht vollständig beschriebenen Blocks verteilt die SSD bei der „Garbage

⁶⁹ Kramer (2014): Digitale Halbleiterspeicher. In: Siemers; Sikora (Hrsg.): Taschenbuch Digitaltechnik. S. 298

⁷⁰ Vgl. Leimeister (2015): Einführung in die Wirtschaftsinformatik. S.52

⁷¹ Vgl. ebenda, S.52

⁷² Eine SSD hat weitere Komponenten, diese tragen der Betrachtung der SSD-Speichertechnik in diesem Zusammenhang nur wenig bei.

Collection“ im Hintergrund die Daten so um, dass immer möglichst viele freie Blocks zur Verfügung stehen.

Aus physikalischen Gründen kann eine Zelle nur begrenzt oft beschrieben werden. Um dem entgegen zu wirken, kommt das für SSDs sehr wichtige Wear-Leveling (WL) zum Tragen. Beim WL werden Zellen nach ihrer Benutzungshäufigkeit kategorisiert. Besonders häufig genutzte Zellen werden seltener zum Schreiben von Daten genutzt. Der sogenannte Counter einer Zelle gibt dabei Auskunft über die Häufigkeit der bereits getätigten Schreibvorgänge. Zellen mit niedrigem Counter werden bevorzugt zum Speichern genutzt. Damit kann die Abnutzung der Zellen auf der SSD gleichmäßig verteilt und die Lebensdauer verlängert werden.

4.2.4 Optische Speichermedien

Optische Speichermedien unterscheiden sich von den bereits genannten Speichermedien in der Art des Lesens und Schreibens. Am Beispiel von Compact Discs (CD), die in den achtziger Jahren⁷³ an Bedeutung gewannen, funktioniert die Datenspeicherung mit Hilfe eines Lasers. Die Compact Disk misst 120mm Durchmesser und besteht aus Polycarbonat. Auf dem Polycarbonat ist eine Aluminiumschicht aufgebracht, die zusätzlich von einer Schutzschicht geschützt wird.⁷⁴

„Das Konzept des Verfahrens besteht darin, die Bitfolge des zu speichernden Signals auf einer Platte mit verspiegelter Oberfläche mechanisch nachzubilden. Die Wiedergabe erfolgt durch Abtastung der Bitfolge mit einem Lichtstrahl bei konstanter Geschwindigkeit der optischen Platte.“⁷⁵

Da auf der CD nur digitale Signale gespeichert werden können, muss es auch nur zwei interpretierbare Zustände geben, die folglich für „0“ oder „1“ stehen.⁷⁶ Dies geschieht bei einer CD mit Hilfe von sogenannten „Pits“. Pits sind Vertiefungen auf der Polycarbonat-Schicht einer Disk. Bei einer Compact Disc werden bei der Produktion durch das Spritzguss-Verfahren Pits auf eine spiralförmige Spur gepresst.⁷⁷ Ein Laser liest beim Lesevorgang diese Vertiefungen aus. Dabei reflektiert der Lichtstrahl unterschiedlich, je nachdem, ob dieser auf ein Pit oder Land trifft.⁷⁸ Diese Zustandsänderungen können interpretiert werden. Der Übergang von einem Pit zu einem Land und anders herum fasst der Laser als logische Zahl 1 auf. Keine Änderungen der Reflektion werden hingegen als Zahl 0 interpretiert.⁷⁹ Die Compact Disc ist dabei lediglich ein lesbarer, aber kein schreibbarer Speicher. Die CD-R (Compact Disc Recordable) und die CD-RW

⁷³ Vgl. Dotzler; Roesler-Keilholz (2017): Mediengeschichte als Historische Techno-Logie. S. 91

⁷⁴ Vgl. Zollner; Zwicker (2013): Elektroakustik. S. 333

⁷⁵ Freyer (2013), S.63

⁷⁶ Vgl. ebenda, S. 63

⁷⁷ Vgl. Roth (2010): Chemische Köstlichkeiten. S.159

⁷⁸ Vgl. Flik; Liebig (2013): Mikroprozessortechnik: CISC, RISC Systemaufbau Assembler und C. S.513

⁷⁹ Vgl. Messmer; Dembowski (2003): PC-Hardwarebuch: Aufbau, Funktionsweise, Programmierung ; ein Handbuch nicht nur für Profis. S.26

(Compact Disc Rewriterable) können im Gegensatz zur Compact Disc beschrieben werden. Zu diesem Zweck wurde die Plattenarchitektur geändert. Die Platten der CD-Rs oder CD-RWs enthalten Materialien, welche durch Hitze ihre Struktur verändern und Pits formen können.^{80 81}

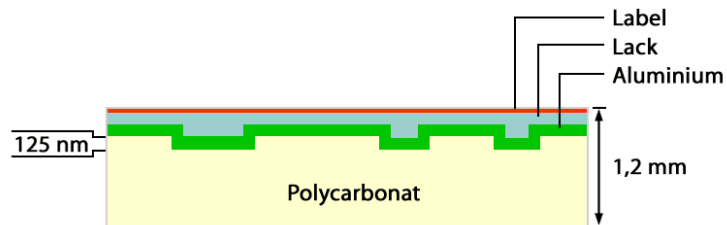


Abbildung 6: CD Querschnitt

Der zunehmende Bedarf an Speicherkapazität resultierte in der Weiterentwicklung der Technologie. So können optische Datenträger, wie die Blu Ray bis zu 100 Gigabyte Daten speichern.⁸² Die Steigerung der Datendichte auf einer Disc wurde durch die Verkleinerung der Pits sowie durch den geringeren Abstand zueinander erreicht.⁸³ Mit der Entwicklung der HVD (Holographic Versatile Disc) sollen in Zukunft bis zu 10 Terabyte pro Platte möglich sein.⁸⁴

4.3 RAID-Level

Ein RAID (Redundant Array of Independent Disks) sorgt für den Verbund mehrerer einzelner Festplatten.⁸⁵ Damit soll ein großer Speicher simuliert werden. Ein Controller regelt das Datenmanagement innerhalb des RAID. Der Zugriff auf ein RAID-System geschieht zentral und wird vom Betriebssystem als eine Einheit interpretiert.⁸⁶ Der Festplattenverbund, auch Array genannt, weist ein großes koordiniertes System aus Speichermedien auf, das aus allen Kapazitäten eine große Gesamtkapazität zur Verfügung stellt. Ein RAID gibt es in vielen verschiedenen Ausführungen, auch Levels genannt. Das Konzept des Verbunds mehrerer Platten bleibt dabei durchweg gleich. Je nach Level ändert sich die Anzahl der Platten und die Art, wie Daten gespeichert werden. Ziel eines RAID ist das Fortbestehen der Datenintegrität im Falle eines Defekts.

⁸⁰ Das Prinzip der Pits und Lands ist dabei bei allen optischen Speichermedien gleich, weswegen die Funktionsweisen aller Compact-Disc-Varianten an dieser Stelle den Rahmen der Arbeit sprengen würden.

⁸¹ Vgl. Steinmetz (2013): Multimedia-Technologie: Grundlagen, Komponenten und Systeme. S. 219ff

⁸² Vgl. Gautzsch (2010): Entwicklung Und Markteinführung Der Blu-Ray Disc: Die Blu-Ray Disc Als Home-entertainment-standard Und Dvd-nachfolger. S.33

⁸³ Vgl. Lutzke (2013): Surfen in die digitale Zukunft. S. 53

⁸⁴ Vgl. Royon; Bourhis; Bellec et. al (2010): Silver Clusters Embedded in Glass as a Perennial High Capacity Optical Recording Medium. In: Advanced Materials. 22. Ausgabe, S. 5282

⁸⁵ Vgl. Koslowski; Hensel (2013): Betriebswirtschaftliche Formelsammlung. S. 98

⁸⁶ Vgl. Schneider (2012): Digitale Amnesie: Langzeitarchivierung digitaler Dokumente im betrieblichen Umfeld. S. 62

Auf diese Weise können Institutionen, aber auch Verbraucher das Risiko des Datenverlustes minimieren.

Es wird zwischen RAID 0, RAID 1, RAID 2, RAID 3, RAID 4, RAID 5, RAID 6 und RAID 10 unterschieden⁸⁷. Dabei stellt dies keine Rangordnung der Sicherheitsstufen dar.⁸⁸ RAID 2 wird bei der Betrachtung bewusst ausgelassen, da es sich um ein wenig gebräuchliches Modell handelt.

Allgemein ist bei der Wahl des passenden RAIDs zu beachten, dass auch die Faktoren Sicherheit und Kosten relevant sind. Mehr Sicherheit wird erreicht indem Daten mehrfach gespeichert und besser verteilt werden. Daraus resultiert ein höherer Bedarf an Festplatten, was die Kosten steigert. Je nach Datenvolumen, Art des Gebrauchs und Relevanz der Daten kann ein passendes Modell ausgewählt werden. Besonders bei großen Datenmengen, kann ein Ausfall des Datenträgers Datenverluste von mehreren Gigabyte oder gar Terabyte bedeuten.

4.3.1 RAID 0

Das RAID 0-Array⁸⁹ (siehe Anhang III) gewährleistet keine Steigerung der Datensicherheit. Mehrere Festplatten befinden sich in einem Festplattenverbund und weisen somit eine gesteigerte Gesamtkapazität auf. Daten werden mittels des „Data Striping“ block-weise bzw. bit-weise aufgespalten und auf die Festplatten verteilt. Fällt ein Datenträger aus, sind die Daten geschädigt, irreparabel und verloren.

4.3.2 RAID 1

RAID 1⁹⁰ (siehe Anhang III) sieht, hierarchisch gesehen, als erstes Level die redundante Sicherung vor. Daten werden dafür dupliziert und auf mehreren Platten gespeichert. Dieses Verfahren sorgt dafür, dass die Gesamtkapazität der Platten in der Nutzbarkeit zwar eingeschränkt wird, beim Ausfall eines Datenträgers besteht aber das Duplikat weiter auf den anderen RAID-Elementen.

4.3.3 RAID 3 und RAID 4

Diese beiden RAID-Systeme ähneln sich stark und können zusammengefasst werden. RAID 3 bzw. RAID 4⁹¹ (siehe Anhang III) wendet das bereits erwähnte Data Striping an. Die Daten werden blockweise oder bitweise auf die Festplatten verteilt. Hinzu

⁸⁷ Es existieren deutlich mehr RAID-Lösungen, auf welche in diesem Kontext bewusst verzichtet wurde.

⁸⁸ Vgl. Broy; Spaniol (Hrsg.) (1999): VDI-Lexikon Informatik und Kommunikationstechnik. S.572

⁸⁹ Vgl. Koslowski; Hensel (2013): S. 98

⁹⁰ Vgl. ebenda, S. 98

⁹¹ Vgl. Kopacek; Zauner (2013): Leitfaden der technischen Informatik und Kommunikationstechnik. S.55

kommt die Erstellung von Paritätsdaten, welche der Rekonstruktion von beschädigten oder verlorenen Daten übernehmen. Eine Parität kann wie folgt definiert werden:

„Eine Technik zur Fehlererkennung, die zur Kontrolle bei Datenübertragungen und der Prüfung des Arbeitsspeichers eingesetzt wird. Dazu wird jedem Byte mit Daten ein zusätzliches Paritätsbit hinzugefügt, das bei der Einstellung gerade (even) auf 1 gesetzt wird, wenn die Anzahl der 1er-Bits im Byte gerade ist. Es ist aber auch die umgekehrte Einstellung ungerade (odd) möglich, bei der das Paritätsbit bei einer ungeraden Anzahl 1er-Bits gesetzt wird.“⁹²

Bei der Identifikation eines Fehlers reagiert ein Mechanismus und wiederholt die Übertragung der fehlerhaften Bit-Folge.⁹³ Die Paritätsdaten werden bei RAID 3 bzw. 4 auf einer ausschließlich für Paritätsdaten reservierten Festplatte gespeichert. Beim Ausfall der Paritätsplatte können diese neu geschrieben werden. Beim Defekt einer anderen Platte können die Paritätsdaten die verlorenen Daten rekonstruieren.

Der Unterschied zwischen RAID 3 und RAID 4 liegt in der Aufteilung der Daten in Blöcke (Data Striping). RAID 3 spaltet die Daten bitweise, wohingegen RAID 4 größere Blöcke aus Bits erstellt. Die Wahl des RAID hängt vom Anwendungsbereich ab. Bei bedeutend mehr anfallenden Lese- und Schreiboperationen, beispielweise rechenintensiven Anwendungen, ist RAID 4 die geeignetere Lösung.⁹⁴

4.3.4 RAID 5

Das Prinzip hinter RAID 5 macht sich die bereits erwähnten Paritätsdaten zu Nutze. Bei RAID 5 gibt es dafür keine gesonderte Paritätsplatte. Stattdessen fungiert jede Festplatte im Verbund zum Teil als Paritätsspeicher. Am Beispiel von drei Platten findet für jede Information auf zwei der Datenträger Data Striping statt, während der dritte Datenträger die dazugehörige Parität erhält. Nach diesem Vorgang wird eine andere Platte als Paritätsplatte gewählt. Die Festplatten werden nach einer Reihenfolge entweder mit Datenblöcken oder Paritäten bespielt. Fällt eine Platte aus, gibt es keinen Datenverlust. Grund dafür sind die einzigen zwei möglichen Zustände einer Datei beim Ausfall einer Festplatte im RAID 5 (am Beispiel dreier Festplatten im Verbund):

Die Parität zu einer Datei ist erloschen.

Lösung: Die beiden Blöcke der Datei bestehen weiterhin auf den laufenden Platten

Ein Teil der Datei ist erloschen.

⁹² Beisecker (2005): Das Lexikon der PC-Fachbegriffe. S.263

⁹³ Vgl. Beisecker (2005): S. 263

⁹⁴ Vgl. Kopacek; Zauner (2013): S.55

Lösung: Ein Block sowie die Parität der Datei bestehen weiterhin auf den laufenden Platten. Mit Hilfe der Parität kann der fehlende Block rekonstruiert werden

Wichtig dabei ist, dass nur eine Festplatte ausfallen darf. Fallen mehrere aus, kommt es zum irreparablen Datenverlust.

4.3.5 RAID 6

Um das Risiko bei dem bereits erwähnten RAID 5 zu senken, erlaubt RAID 6 (siehe Anhang III) den Ausfall von zwei Platten im Verbund.⁹⁵ Damit aber zwei Platten ausfallen können, sind mindestens vier Datenträger nötig. Das Prinzip von RAID 6 ähnelt dabei dem des RAID 5, nur dass Paritätsdaten nun auf zwei Festplatten dupliziert gespeichert werden. Die Verteilung der Paritäten bleibt weiterhin verteilt über alle Elemente im Verbund. Bei dem Ausfall von zwei Festplatten wären alle Informationen rekonstruierbar.

4.3.6 RAID 10

Die Nummer dieses RAID-Modells setzt sich - anders als es der Name vermuten lässt - aus mehreren RAID 1-Systemen und einem RAID 0 zusammen und lautet korrekt RAID 1+0. Das RAID 0 gilt dabei als Deckel aller RAID 1-Systeme.⁹⁶ Die Mindestanzahl der Festplatten beträgt 4. Es werden die Verfahren des Data Stripings und des Data Mirroring in einem RAID 10 (siehe Anhang III) kombiniert. Eine Datei wird hierbei im ersten Schritt gespalten und anschließend dupliziert. Jedes Array hat nun doppelt gesicherte „Stripes“. RAID 10 übersteht in jedem RAID 1 Array einen Ausfall.

4.3.7 RAID 60

Ähnlich wie bei RAID 10 befinden sich hier zwei RAID 6-Systeme unter einem RAID 0.⁹⁷ Pro System dürfen zwei Platten ausfallen.

4.4 Speicherarchitekturen

Der Umgang mit Daten jeglicher Art verlangt dem Anwender eine selbstgewählte Art der Ordnung ab. Unstrukturierte Dateiverzeichnisse verlängern den Workflow und hindern den Benutzer eines Rechners daran, zeiteffizient zu arbeiten. An Relevanz gewinnt die Struktur der Verzeichnisse besonders, wenn mehrere Anwender auf den selben zugreifen wollen. Durchdachte Strukturen eines Anwenders können bei den

⁹⁵ Vgl. Kopacek; Zauner (2013): S.56

⁹⁶ Vgl. Vetter; Narendra Babu; Balakrishnan (2015): IBM Power Systems RAID Solutions Introduction and Technical Overview. S. 9

⁹⁷ Vgl. Vetter; Narendra Babu; Balakrishnan (2015): S. 10

restlichen Anwendern möglicherweise fehlinterpretiert werden. Um den gemeinsamen, geregelten und sicheren Datenverkehr innerhalb einer Benutzergruppe zu gewährleisten, empfiehlt es sich, eine durchdachte Speicherarchitektur zu nutzen. Im Folgenden werden einige Architekturen behandelt und in ihrer Funktionsweise erklärt.

4.4.1 DAS

Allen folgenden Architekturen vorangehend gilt ein DAS in seiner Struktur als besonders unkompliziert. Direct Attached Storage (DAS) bezeichnet einen Speicher, der direkt an einen Computer oder Server angeschlossen wird.⁹⁸ Zwischen beiden Elementen besteht somit eine Punkt-zu-Punkt⁹⁹ Verbindung. Im Fall eines an einem Server angeschlossenem DAS¹⁰⁰ muss der Anwender bei Informationsbedarf beim Server eine Anfrage stellen. Der Server kommuniziert daraufhin mit dem Massenspeicher und leitet die Informationen zum Anwender:

„Damit steht dieser lokale Festplattenverbund in der Regel exklusiv nur jeweils dem physikalischen Server (Host) und deren VMs (Gäste) zur Verfügung“¹⁰¹

Solche exklusive Verbindungen, auch dedizierte Verbindungen genannt, sind nur über das Servermedium möglich. Ein Umstand, dem Architekturen wie NAS (Kapitel 4.4.1.2) und SAN (Kapitel 4.4.1.3) entgegenwirken sollen.

4.4.2 NAS

Ein NAS (Network-Attached-Storage) ist ein Massenspeicher, der, verbunden mit einem Netzwerk, Daten für verschiedene Benutzer im Netzwerk (Clients) bereitstellt. Das NAS ist dabei in ein LAN-Netzwerk (Local Area Network) integriert.¹⁰² Als „File-Server“ hat ein NAS die Kernaufgabe der Bereitstellung von Daten für unterschiedliche Clients im Netzwerk. Das NAS zeigt sich dem Nutzer als ein zentrales Speichersystem, d.h. die im NAS enthaltenen Festplatten werden nicht gesondert angezeigt. Zur Sicherung beinhalten neue NAS-Modelle meist RAID. In puncto Datensicherheit sind NAS weitaus weniger störanfällig sein.

Bei der Installation eines Network Attached Storage kann dieses per Ethernet-Port an das bestehende LAN angeschlossen werden. Somit kann das NAS schnell genutzt werden und bedarf keiner aufwendigen Konfiguration. Darüberhinaus verfügt NAS über

⁹⁸ Vgl. Müller (2003): IT-Sicherheit mit System: Strategie - Vorgehensmodell - Prozessorientierung – Sicherheitspyramide. S. 101

⁹⁹ Vgl. Eisentraut; Helmle (2013): PostgreSQL-Administration. S. 368

¹⁰⁰ Anders als bei einer Verbindung mit einem Computer, sind hierbei servertypisch mehrere Anwender möglich.

¹⁰¹ Hengstberger (2016): Storage Systeme erfolgreich managen. In: Tiemeyer (Hrsg.). Handbuch IT-Systemmanagement: Handlungsfelder, Prozesse, Managementinstrumente, Good-Practices. S. 255

¹⁰² Vgl. Badertscher; Scheuring (2007): Wirtschaftsinformatik: Wartung und Betrieb eines Informations- und Kommunikationssystems. S.178

ein eigenes Betriebssystem, welches die Datenströme koordiniert und als Kommunikationseinheit dient. Die einfache Struktur und Kompatibilität des Betriebssystems sorgt dafür, dass auch heterogene^{103,104} Clients auf die Daten zugreifen können. Im System können darüberhinaus Back-Ups geplant werden.

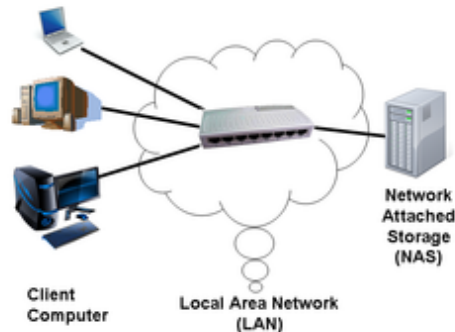


Abbildung 7: Network attached storage

Ein NAS transportiert Daten stets dateibasiert (file-based). Der Zugriff auf die Daten erfolgt über das LAN-Netzwerk. Der Datenfluss ist deshalb maximal so schnell wie das Netzwerk selbst. Durch den Transfer der Daten durch das LAN-Netzwerk wird dieses zusätzlich belastet. Bei einem häufigen Gebrauch der Daten durch mehrere Benutzer kann die Leistung des LANs deshalb deutlich sinken.

4.4.3 SAN

Das Storage Area Network (SAN) verbindet mehrere Massenspeicher zusammen in einem einzigen Datenpool, dem SAN, und stellt diese einem Netzwerk aus Servern zur Verfügung. Ein Massenspeicher kann dabei auch als Speichersubsystem bezeichnet werden. Vergleichend mit der klassischen Kommunikation zwischen Speicher und Server, kann ein SAN wie folgt definiert werden (Robbe, 2004):

„Ein SAN ist ein Hochgeschwindigkeitsnetzwerk zwischen Servern (Hosts) und Speichersubsystemen. Dabei ermöglicht ein SAN eine „any-to-any“-Verbindung durch das gesamte Netzwerk. Es eliminiert die traditionell dedizierten Verbindungen zwischen Servern und Subsystemen und den bisherigen Anspruch, dass Server Speichersubsysteme „besitzen“ und „handhaben“.“¹⁰⁵

¹⁰³ Vgl. Riggert (2009): ECM - Enterprise Content Management. S.171

¹⁰⁴ Heterogen bezieht sich hierbei auf das mögliche Vorhandensein verschiedener Betriebssysteme einzelner Clients. Die Kompatibilität des NAS zu den meisten Systemen gewährleistet hierbei die Kommunikation zwischen NAS und Client.

¹⁰⁵ Robbe (2004): SAN - Storage Area Network: Technologie, Konzepte und Einsatz komplexer Speicherumgebungen. S.33

Ein SAN stellt sich den Servern dabei als Gesamtsystem mit mehreren logischen Einheiten dar.¹⁰⁶ So kann bewusst auf einen Bereich zugegriffen werden, z.B. auf das Langzeitarchiv in Form einer Tape Library:

„Ein SAN benötigt eine eigene Infrastruktur, bestehend aus mindestens einem SAN-Server, einer SAN-Infrastruktur (Switches, Verkabelung) sowie SAN-Speicher-Array. Jeder Server bzw. jedes Storage-Device in einer SAN-Struktur benötigt einen Host-Bus-Adapter (HBA).“¹⁰⁷

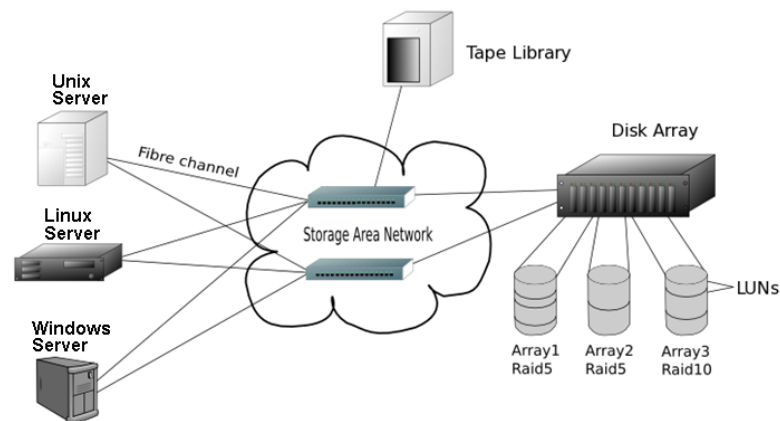


Abbildung 8: Storage area network

Um eine möglichst schnelle Datenrate zu ermöglichen, sind die Server mittels Fibre-Channel-Technologie mit den Subsystemen verbunden.¹⁰⁸ Dazu wird ein Fibre-Channel-Switch zwischen ihnen angelegt. Die erstellte Fibre-Channel-Verbindung ermöglicht nun einen Austausch von mehreren Gbits pro Sekunde zwischen Server und Subsystem.¹⁰⁹ Die bereits im Zitat von Robbe erwähnte „any-to-any“-Verbindung wird dadurch möglich, dass alle Server und Subsysteme mit dem Switch (Fibre-Channel) verbunden sind und damit auch untereinander. Durch die Bildung eines Netzwerkes neben dem LAN kann das LAN entlastet werden¹¹⁰, weil, anders als beim NAS, die Massenspeicher direkt mit den Servern statt über das LAN kommunizieren. Rechenaufwendige Datenspeicherung wird also optimalerweise über die Fibre-Channel-Verbindung realisiert.

¹⁰⁶ Vgl. Mandl (2010): Grundkurs Betriebssysteme: Architekturen, Betriebsmittelverwaltung, Synchronisation, Prozesskommunikation. S. 279

¹⁰⁷ Cattini; Kammermann; Gut (2016): CompTIA Server+: Serversysteme einrichten und betreiben. Vorbereitung auf die Prüfung SKO-004. S.106

¹⁰⁸ Vgl. Scheuring; Waldspühl, 2012, S.89

¹⁰⁹ Vgl. ebenda

¹¹⁰ Vgl. Lassmann (Hrsg.) (2006): Wirtschaftsinformatik: Nachschlagewerk für Studium und Praxis. S.98

5 Archivierungssysteme

Anschließend an die Speichertechnologien wird im Folgenden ein Überblick über einige aktuell verwendete Archivsysteme bzw. Archivtechnologien vorgestellt. Diese können möglicherweise zu einem späteren Zeitpunkt obsolet sein.

Die Anzahl der Speichertechnologien ist relativ groß. Festplatten, Flash-Drives, Optical Discs oder Magnetbänder – alle dienen der Sicherung von Daten. Dabei gibt der Begriff „Sicherung“ noch keine Angabe zum Zeitrahmen der Speicherung. Nicht jeder Datenträger kann physikalisch über mehrere Dekaden Daten sicher speichern. Zur Langzeiterhaltung von Daten werden speziell Magnetbänder, Festplatten (HDDs) und optische Disks präferiert.¹¹¹ SSD fällt dabei zur Zeit noch aus Gründen der Wirtschaftlichkeit und der quantitativ begrenzten, Beschreibbarkeit heraus. Langzeitarchivierung soll in erster Linie sicher sein. Dennoch ist auch der Faktor „Kosten pro Gigabyte“ als Kriterium relevant. Die wichtigen Parameter und Faktoren für die Auswahl der passenden Technologie und weiter eines Archivierungssystems werden im folgenden Abschnitt behandelt. Sie dienen dem Verständnis für den wissenschaftlichen Teil (Kapitel 6).

5.1 Archivtechnologien

Die nachfolgenden ausgewählten Archivierungstechnologien leitet der Verfasser von aktuellen auf dem Markt als relevant eingestuften Produkten aus dem professionellen Bereich ab.

5.1.1 NAS-Lösung

Die bereits in Kapitel 4.4.1.2 erwähnte Technologie des Network Attached Storage wird in der Archivierung häufig genutzt. Besonders im Bereich der Back-Up-Sicherung ist ein NAS oft vertreten. Ein Verbund aus vielen Festplatten hat neben der RAID-Sicherung den Vorteil eines hohen Datendurchsatzes. Das eingebaute Betriebssystem bietet dem Anwender vielfältige Möglichkeiten des Datenmanagements und durch das „Scheduling“ (Ablauf-Planung) lassen sich Sicherungen automatisch in geregelten zeitlichen Abständen planen.

NAS-Systeme werden in unterschiedlichen Kapazitäten angeboten. Für den Privatanwender gibt es bereits Modelle ab 1 Terabyte. Für professionelle Anwender z.B. Unternehmen, sind auch Speicher mit der Kapazität von mehreren hundert Terabytes erhältlich.¹¹² Dies wird meist durch Erweiterungsmodule erreicht. Zur effizienteren Nutzung des Potenzials werden NAS-Systeme durch das installierte Betriebssystem virtualisiert. Der hohe Datendurchsatz ermöglicht einen schnellen Zugriff auf Daten,

¹¹¹ Dickreiter; Dittel; Hoeg; et. al (2008): Digitale Tonsignalspeicherung. In: ARD ZDF medienakademie (Hrsg.): Handbuch der Tonstudioteknik. S. 791f

¹¹² QNAP (2018): Produktbeschreibung TS-EC2480U.

wodurch der Workflow verbessert werden kann. Durch zusätzliches Storage Tiering¹¹³ können relevante Daten schneller verfügbarer gemacht werden.

Vorteile:	Nachteile:
- Schneller Datenzugriff	- Geringe Datensicherheit
- Kein Mounting	- Migration nach wenigen Jahren
- Hohe Skalierbarkeit	- Anfällig gegen Hacker
	- Energie- und Wartungsintensiv
	- Hoher Preis

Tabelle 2: Vor- und Nachteile NAS¹¹⁴

5.1.2 Tape Library

Eine Tape Library (Bandbibliothek) bezeichnet ein System zur Aufbewahrung und Nutzung von Magnetbändern. Dieses wird von einem Roboter bedient. Dieser hat die Funktion, Cartridges zu fassen und innerhalb des Geräts zu bewegen (Mounting). Zur Nutzung der Daten muss das gewählte Band gelesen werden. Dieser Vorgang findet im eingebauten Laufwerk statt. Der Roboter erkennt Bänder anhand des auf jeder Cartridge angebrachten Barcodes. Alle Cartridges werden in Fächern (Slots) gelagert und für einen Zugriff bereitgehalten.

Abhängig vom Modell variiert die Zahl an Laufwerken, Slots und damit der Kapazität. Die gebräuchlichste Bandtechnologie im professionellen Bereich ist die LTO-Technologie. Mit der aktuellsten Generation LTO-8 lassen sich 12 Terabyte¹¹⁵ in einer Cartridge speichern (siehe Kapitel 4.2.1). Große Bandbibliotheken lassen sich mit Erweiterungsmodulen auf mehr als drei Petabyte erweitern. Der Datendurchsatz einer Tape Library ist von der verwendeten LTO-Generation und der Anzahl an Laufwerken abhängig.

Vorteile:	Nachteile:
- Lange Lebensdauer	- Mounting-Zeit
- Geringe Kosten / TB	- Mögliche Inkompatibilität
- Schnelle Lesegeschwindigkeit (nach Mounting)	- Verschleiß

¹¹³ Beim Storage Tiering, werden Daten nach ihrem Gebrauch klassifiziert. Oft genutzte Daten können bei passendem NAS-Modell z.B. auf SSDs transferiert werden und bieten somit eine kürzere Zugriffszeit. (Quelle: Hengstberger (2016): Storage Systeme erfolgreich managen. In: Tiemeyer (Hrsg.). Handbuch IT-Systemmanagement: Handlungsfelder, Prozesse, Managementinstrumente, Good-Practices. S. 256)

¹¹⁴ Vgl. Noppenberger (2009): Kostenminimierung in Speichernetzwerken. S. 26

¹¹⁵ Vgl. IBM (2018): Produktbeschreibung IBM LTO Ultrium 8 Data Cartridge.

- Offline-Lagerung	- Lagerungsvoraussetzungen ¹¹⁶
--------------------	---

Tabelle 3: Vor- und Nachteile Bandbibliothek

5.1.3 Optical Data Archive

Sogenannte ODA-Systeme nutzen die Technologie der optischen Speicherung. Da ein einziger optischer Datenträger nur eine Kapazität bis maximal 300GB (Archival Disc) bietet, arbeiten ODAs mit einer großen Anzahl an Disks. Als besonderes Produktvorteil heben der Hersteller von ODA-Systemen die Haltbarkeit der optischen Datenträger hervor. Diese liegt nach Angaben der Hersteller bei über 50 Jahren.¹¹⁷ Die gesteuerte Archivierung mit optischen Datenträgern ist noch in der Entwicklung, weshalb ODAs zur Zeit noch deutlich teurer als NAS und Bandbibliotheken sind.

Vorteile:	Nachteile:
- Sehr lange Lebensdauer	- Hohe Kosten
- Geringer Energieverbrauch	- Mounting-Zeit
- Unempfindlichkeit	- Ungeeignet für Hot Data ¹¹⁸
- Keine Migration nötig	
- Kein Verschleiß ¹¹⁹	

Tabelle 4: Vor- und Nachteile ODA

5.1.4 Cloud

Eine Cloud (Wolke), ist eine Dienstleistung. Cloud-Anbieter stellen ihren internen Speicher für Geld zur Verfügung. Der Nutzer benötigt lediglich einen Internetzugang, um auf die Daten zugreifen zu können.

Vorteile:	Nachteile:
Platzersparnis	Abhängigkeit vom Internet
Flexibel skalierbar	Anbieterabhängig
Investitionsarm	Datenzugriff langsam
Pay-as-you-go ¹²⁰	Fremdzugriff ungewiss

¹¹⁶ von Löher, Franz (1976): Archivalische Zeitschrift. S. 92

¹¹⁷ Vgl. Sony; Panasonic (2015): White Paper Archival Disc Technology

¹¹⁸ Hot Data bezeichnet häufig verwendete Dateien.

¹¹⁹ Die Datenträger erleiden keinen Verschleiß. Die Robotik kann dennoch durch Verschleiß versagen.

¹²⁰ Durch das Pay-as-you-go-Modell bezahlt der Nutzer lediglich den Cloud-Speicher, den er auch wirklich verwendet.

Tabelle 5: Vor- und Nachteile Cloud

5.2 Archivsysteme

Ein Archivsystem ist eine Verkettung von Hardware- und Softwarelösungen. Die Software spielt dabei eine relevante Rolle, da die Hardware zwar die Daten sichert, diese aber nicht ordnet und verwaltungsfähig macht. Zum Datenmanagement ist ein der Hardware übergeordnetes System nötig, das den Speicher virtualisiert:

*Speichervirtualisierung, oft auch als Storagevirtualisierung bezeichnet, ermöglicht eine effizientere Nutzung von vorhandenen Datenspeichern, zum Beispiel Serverfestplatten. Die Grundidee bei der Speichervirtualisierung liegt darin den Datenspeicher von den klassischen Fileservern zu trennen und die physischen Speicher in Pools zusammenzufassen. Durch Speichervirtualisierung erscheint Nutzern Speicherplatz virtuell.*¹²¹

Durch die virtuelle Bündelung aller Speicherressourcen greift der Anwender zentral auf den Pool zu. Daten sind so leichter zu finden, da die Lokation der Daten für den Anwender irrelevant ist.¹²² Virtuelle Speicher haben weiter den Vorteil, dass sich der Speicher partitionieren lässt. Bestimmte Nutzergruppen können so auf softwaredefinierte Speicher zugreifen, die hardware-unabhängig auf mehreren Datenträgern liegen können.¹²³ Dadurch steigt die Effizienz der Speichernutzung.¹²⁴ Im Bezug auf das Archivsystem kann eine Virtualisierung des Speichers die Nutzung der Ressourcen sowie das Datenmanagement verbessern. Wird eine bestimmte Datei benötigt, findet die Archivsoftware durch Metadaten das gewünschte Datenobjekt in der Hardwareumgebung. Einheitliche Software ermöglicht einen kontinuierlich planbaren, effizienten und übersichtlichen Datenbestand.

Im Folgenden werden verschiedene Archivsysteme und ihre Komponenten beschrieben. Dabei sind die erwähnten Produkte lediglich eine Anregung auf Basis einer Recherche. Die Kapazität der Systeme soll bei ca. 600 TB liegen, kann jedoch durch feste Konfigurationen seitens der Hersteller abweichen.

Die Systeme werden zudem als „Hybrid“ oder „Standalone“ klassifiziert. Hybride Systeme nutzen verschiedene Speichermedien (z.B. Festplatte und Magnetband), während Standalones lediglich einen Medium-Typ zur Archivierung benötigen.

5.2.1 System 1: Hybrid Bandbibliothek / NAS

Dieses Archivsystem setzt sich aus den Komponenten actiLib Kodiak 6807 (Tape Library), dem actiNAS WIN 424 (NAS-System) und der Archivsoftware „P5 Archive“ der

¹²¹ Bedner (2013): Cloud Computing: Technik, Sicherheit und rechtliche Gestaltung. S.43

¹²² Vgl. ebenda

¹²³ Vgl. Xu (2009): Virtualisierung als Möglichkeit der Optimierung des IT- Managements. S. 48ff

¹²⁴ Vgl. ebenda

Firma Archiware zusammen. Das NAS dient mit über 600 Terabyte Kapazität dem schnellen Zugriff auf die Daten. Die Bandbibliothek hingegen soll lediglich als Offline-Speicher und zur Redundanz dienen. Die Archivsoftware P5 steuert die doppelte Speicherung auf NAS und Bibliothek. Alle Produkte kommen von deutschen Unternehmen.

actiNAS WIN 424¹²⁵

Das erste Glied im Verbund ist das NAS WIN 424 als Hauptsteuerungseinheit. Als Server fungierend verfügt der WIN 424 über einen Prozessor (8-Kerner), 16 GB RAM und 24 Festplattenslots. Eine JBOD-Erweiterungseinheit erweitert den Server auf 48 Festplatteneinschübe und eine Gesamtkapazität von 672 Terabyte. Zusätzlich befinden sich zwei Slots für SSDs im hinteren Teil des NAS. Zur Steigerung der Sicherheit des Systems wird das Betriebssystem durch RAID 1 redundant auf den SSDs gespeichert. Das Betriebssystem nutzt die Archivierungssoftware und sichert somit die konforme Archivierung der Daten. Eine redundante Stromversorgung schützt das NAS beim Ausfall eines Netzteils vor dem Absturz. Das WIN 424 verfügt über Ethernet-Ports zur Anbindung an das LAN als Schnittstelle für den Anwender. Zusätzlich gibt es SAS-Schnittstellen zur Verbindung mit dem Kodiak 6807.

actiLib Kodiak 6807¹²⁶

Die Kodiak 6807 Tape Library vom Hersteller actidata ist eine Bandbibliothek, welche die LTO-Technologie nutzt. Es wird dabei die neuste Generation LTO-8 genutzt, bei Bedarf aber auch ältere Generationen wie LTO-7 und LTO-6. Die Abwärtskompatibilität ermöglicht so einen Mixbetrieb. In einen bestehenden Server-Rack¹²⁷ lässt sich das 19-Zoll Rackmount-Gehäuse¹²⁸ integrieren.

Passend zum Bedarf wird die Bandbibliothek neben dem Basismodul (80 Slots) mit einer Erweiterungseinheit (24 Slots) ausgestattet und erreicht eine Kapazität von 624 Terabyte bei LTO-7. Voll ausgestattet kann das Kodiak 6807 bei sechs zusätzlichen Modulen eine Kapazität von 560 Slots und somit 6,7 Petabyte (LTO-8) fassen. Durch sechs LTO-8-Laufwerke können sechs LTO-Bänder gleichzeitig gelesen werden. Den schnellen Transport der Daten auf einen Server ermöglichen mehrere SAS-Ports mit der Übertragungsrate von 6000 Gigabit pro Sekunde (750 MB/s).

Archiware P5¹²⁹

Die browserbasierte Archivierungssoftware P5 nimmt im „NAS-zu-Tape“-Verbund eine Kontrollfunktion ein. Schon beim Eingang der Daten auf das NAS werden die Daten

¹²⁵ Vgl. actidata (2018): Produktbeschreibung actiNAS WIN424.

¹²⁶ Vgl. actidata (2018): Produktbeschreibung Kodiak 6807.

¹²⁷ Ein Rack bezeichnet in der Informationstechnik einen Schrank in dem mehrere Server oder Speichereinheiten integriert werden können. So lassen sich Hardware-Lösungen zentral und geordnet an einem Platz halten.

¹²⁸ Rackmount bezeichnet die Eigenschaft eines Gehäuses in handelsübliche Racks einbau- bzw. integrierbar zu sein.

¹²⁹ Vgl. Archiware (2018): Produktbeschreibung Archiware P5 Archive.

von P5 erfasst. Durch den softwareeigenen Browser¹³⁰ können Verzeichnisse sortiert und verwaltet werden. Zudem lassen sich Metadaten einsehen und bei Bedarf bearbeiten. Eine HTML-5-Funktion bietet die Möglichkeit eines visuellen Previews, wodurch der Inhalt eines Videos schon vor dem Lesen der Datei einzusehen ist. P5 ist kompatibel zu Windows, Mac und Linux und kann dadurch interdisziplinär mit mehreren Betriebssystemen simultan arbeiten.

Durch das Aufrufen eines Kontext-Menüs lässt sich eine Datei auf dem Langzeitspeicher (Tape Library) archivieren. Die Software wird so konfiguriert, dass Daten jeweils auf das NAS und auf die Bandbibliothek geschrieben werden.

5.2.2 System 2: NAS

Festplatten zählen zu den Speichermedien mit schnellen Zugriffszeiten, sie sind jedoch weniger für Langzeitarchive geeignet. Dem begegnet die Firma FastLTA mit der Archivierungslösung Silent Brick. Das aus einem Controller, mehreren optionalen Erweiterungsmodulen und Software bestehende NAS speichert die Daten auf einem Verbund aus Festplatten. Der Hersteller verspricht 10 Jahre Lebensdauer für den Silent Brick.¹³¹

Silent Brick Library¹³²

Das Konzept hinter Silent Brick beruht auf einem Festplatten-Magazin (Brick) mit eingesteckten HDDs oder SSDs. Für den Schutz der Daten kann ein SecureNAS-Verfahren verwendet werden, welches einem RAID ähnelt. Beim Erasure Coding erstellt der Controller aus einer Datei zwölf Fragmente und teilt diese auf die zwölf Festplatten eines Bricks auf. Dabei sind 4 Fragmente redundant, wodurch beim Ausfall von vier Platten kein Datenverlust entsteht. Im Vergleich zu RAID 6 ist diese Methode doppelt so sicher. Dieses Verfahren kann durch den Controller bei jedem Brick einzeln initiiert werden.

Der Controller führt neben der Server-Hardware fünf Einschübe für Bricks an. Ein Erweiterungsmodul bietet mit 14 Einschüben flexible Skalierbarkeit. Pro Controllermodul sind 10 Erweiterungseinheiten möglich, wodurch die potenzielle Brutto-Kapazität eines Racks auf 2,32 Petabyte steigt. Ein Brick lässt sich mit zwölf 2,5 Zoll HDDs oder SSDs bestücken. Ein einzelner Brick kann maximal 24 Terabyte an Daten fassen. Bei einer Konfiguration aus Controller und vier Erweiterungseinheiten lassen sich 652 Terabyte¹³³ Netto-Kapazität speichern. Des Weiteren ermöglicht das einfache Entfernen eines Bricks aus dem Verbund durch Auslagerung einen Offline-Speicher.

5.2.3 System 3: Bandbibliothek

¹³⁰ Als Browser wird die Anwendung P5 verstanden, welche eine Benutzeroberfläche zur Verwaltung von Daten darstellt. Sie ähnelt dabei dem klassischen Windows-Explorer.

¹³¹ Vgl. Fast LTA (2018): Backup optimieren.

¹³² Vgl. FastLTA (2018): Produktbeschreibung Silent Brick.

¹³³ Bei der Rechnung wurde angenommen, dass alle Daten per SecureNAS gesichert sind.

Dieses System kombiniert einen Server mit zwei Tape Libraries in einer Archivierungslösung. Das Management der Daten wird durch eine Archivierungssoftware auf dem Server kontrolliert. Produkte der Firmen QStar und HP stellen die Komponenten der Lösung dar.

HP MSA 2040¹³⁴

Die herstellerinterne Kompatibilität zu HP-Bandbibliotheken macht den MSA 2040 zu einem passenden Server für die Archivierung. Mit 24 Festplatteneinschüben lassen sich bei RAID 50 insgesamt 44 Terabyte interner Speicher nutzen. Dadurch können Inhalte für das Archiv zwischengelagert werden. Als leistungsstarker Server dient er neben dem temporären Erhalt von Daten auch dem Betrieb der Archivsoftware. Ein SAS-Port sichert die Verbindung zu beiden MSL-4048-Bandbibliotheken.

Qstar Archive Manager¹³⁵

Als bekannte Archivierungssoftware bietet der Archive Manager die Möglichkeit, Tape Libraries mittels des LTFS-Standards in ein Windows-Dateisystem, genauer ins NFS/CIFS, zu integrieren. Der Nutzer kann folglich auf das Archiv wie auf ein normales Laufwerk im Netzwerk zugreifen. Zusätzlich, neben der Datensicherung, lässt der Archive Manager eine Replikationsmöglichkeit zu. Per „Volume Spanning“ kann der Speicherplatz effizienter genutzt werden, da Datensätze auf unterschiedliche Datenträger verteilt werden, wenn ein Band keine ausreichende Kapazität bietet. Im Archive Manager lassen sich Archivierungszeitpläne einstellen. So können Daten im Cache des Servers vorlagern und zu einem späteren Zeitpunkt archiviert werden. Der QStar Archive Manager repliziert in diesem Verbund jede Datei und speichert diese doppelt auf beiden Bandbibliotheken. Eine auf HP optimierte Version der Software ist erhältlich.

HP Storeever MSL 4048¹³⁶

Die mit 48 Fächern bestückten Bandbibliotheken bieten bei LTO 8 eine Kapazität von jeweils 576 Terabyte. Durch die Replikation aller Daten auf beide Bibliotheken stehen dementsprechend nur 576 Terabyte redundante Kapazität zur Verfügung. Mit 6 Laufwerken lässt eine MSL 4048 z.B. einen Datendurchsatz von ca. 6 Terabyte pro Stunde zu.

5.2.4 System 4: Optical Disc Archive (A)

Die folgende Lösung unterscheidet sich von den bereits beschriebenen Systemen in der Technologie. Durch optische Speicherung in Sony ODC (Optical Disc Cartridge) lassen sich 3,3 Terabyte pro Cartridge speichern. Das Archivsystem besteht aus dem Optical Disc Archive und dem zugehörigen File-Manager.

¹³⁴ HPE (2018): Produktbeschreibung MSA 2040.

¹³⁵ QStar (2018): Produktseite Archive Manager.

¹³⁶ HPE (2018): Produktbeschreibung MSL 4048.

Sony PetaSite Library¹³⁷

Als einer der Vorreiter in der optischen Speicherung großer Datensätze bietet Sony mit der PetaSite Library eine Bibliothek für optische Speichermedien. Das Produkt kann je nach Skalierung einige Terabytes bis mehrere Petabytes an Daten erfassen. Die bereits in der zweiten Generation befindlichen ODC (Modell ODC3300R) können auf 12 Blu-Rays pro Magazin 3,3 Terabyte Daten speichern. Als WORM-Disk lassen sich die Cartridges, auch als Magazin betitelt, lediglich einmal beschreiben.

Kombiniert in einer Library können durch einen Robotermechanismus, ähnlich wie bei einer Bandbibliothek, Magazine aus einem Fach ins Laufwerk transportiert werden. Je nach Ausstattung lässt eine PetaSite Library Platz für 30-535 Fächer in einem Rack zu. Pro Rack lassen sich also maximal 1,765 Petabyte speichern. Die Master-Einheit kann mit fünf Erweiterungsmodule vergrößert werden.

Modellname	Einheit	Fächer	Anzahl Laufwerke
ODS-L30M	Master-Einheit	30	2
ODS-L60E	Erweiterungseinheit	60	4
ODS-L100E	Erweiterungseinheit	100	0

Tabelle 6: PetaSite Library Module

Neben den Ablagefächern für Magazine lässt sich die PetaSite auf 18 Laufwerke erweitern, um den Datentransfer zu maximieren. Ein Laufwerk (ODS-D280F) kann mit 125 MB/s schreiben und 250 MB/s lesen. Magazine lassen sich durch Schubladen schnell entfernen und bieten die Möglichkeit eines nahezu unendlichen Offline-Speichers.

Die Hardware unterstützend arbeitet der File Manager mit den Inhalten der PetaSite Library. Installiert auf einem Host-Computer können Clients über eine web-basierte grafische Oberfläche auf das Archiv zugreifen. Dateien lassen sich per Drag & Drop transferieren. Durch eine erweiterte Metadaten-Funktion kann der Wiederauffindbarkeit von Informationen zugearbeitet werden.

5.2.5 System 5: Optical Disc Archive (B)

Konkurrierend zur PetaSite von Sony bietet die Firma Panasonic ein Archiv mit optischen Datenträgern an. Das in Zusammenarbeit mit Facebook entwickelte Produkt „freeze-ray“ soll Daten bis zu 100 Jahre speichern können. Zudem wird ein Server für die dazugehörige Software benötigt, um den freeze-ray nutzen zu können.

Panasonic freeze-ray¹³⁸

¹³⁷ Sony (2018): Optical Disc Archive.

Ähnlich wie die PetaSite Library ist das „freeze-ray“ in einem großen Rack eingebaut. In der Grundkonfiguration¹³⁹ stellt „freeze-ray“ mit 228 Magazinen 736 Terabyte (RAID 60) an Speicher zur Verfügung. Durch einen Robotermechanismus werden Magazine innerhalb des „freeze-ray“ bewegt. Optional kann der Speicher eines „freeze-ray“, bei sechs zusätzlichen Erweiterungsmodulen auf 1,9 Petabyte aufgestockt werden. Durch eine maximale Anzahl von 18 Laufwerken verspricht Panasonic eine kumulierte Lese- und Schreibgeschwindigkeit von 1080 MB/s.

Modellname	Einheit	Fächer	Anzahl Laufwerke
LB-DH70A0G	Basis-Einheit	76	3
LB-DF81Z1G	"Bottom"-Einheit	0	0
LB-DH82Z1G	Erweiterungseinheit	76	0
LB-DF72A0G	Erweiterungseinheit	76	3

Tabelle 7: Panasonic freeze-ray Module

Die Erweiterungseinheit LB-DH82Z1G bietet eine Skalierung des Systems ohne zusätzliche Laufwerke. Um die „freeze-ray“ Bibliothek verwalten zu können, bedarf es eines Servers (Linux oder Windows) mitsamt der „freeze-ray“-Software „DA Manager“. Dieser virtualisiert die Bibliothek und macht sie durch entsprechende Verbindungen über das LAN zugänglich und verwaltungsfähig.

5.2.6 System 6: Hybrid NAS / Cloud

On-Premise und Off-Premise kombinierend, verbindet diese Lösung ein NAS mit einem Cloud-Speicher.

Storinator XL60¹⁴⁰

Der als NAS konfigurierte Server Storinator XL60 des Herstellers fasst bei 60 HDD-Steckplätzen eine Brutto-Kapazität von 720 Terabyte, wird jedoch durch eine RAID 6+0-Konfiguration auf 611 Terabyte nutzbare Netto-Kapazität beschränkt. Durch den Einbau vor Ort dient dieser zur schnellen Bereitstellung der Archivdaten.

Oracle Cloud¹⁴¹

¹³⁸ Panasonic (2018): Produktseite freeze-ray.

¹³⁹ Bei einem „bottom module“, einem „base module“ und einem „expansion module“

¹⁴⁰ 45 Drives (2018): Produktseite Storinator XL60.

¹⁴¹ Oracle (2018): Produktseite Oracle Cloud.

Der Gebrauch einer Cloud unterscheidet dieses Archivsystem von allen bisher genannten. Hierbei handelt es sich nicht um eine Hardware, sondern um eine Dienstleistung. Dabei stellt der Anbieter Oracle gegen Zahlung eigenen Speicher zur Verfügung. Die in Oracle-Datenzentren gespeicherten Daten werden mehrfach redundant gesichert, und Oracle verspricht Schutz durch höchste Sicherheitsanforderungen. Ein Cloud-Speicher bzw. dessen Dienstleistung lässt sich nahezu unendlich skalieren. Mit einer Zahlung nach dem „Pay-as-you-go“-Modell muss der Anwender lediglich den Archivspeicher bezahlen, den er tatsächlich belegt.

Der Zugriff auf Daten findet über ein Web-Interface statt, kann jedoch auch als Netzwerk-Laufwerk virtualisiert werden. Der Zugriff auf die Daten in der Cloud ist abhängig von der Übertragungsgeschwindigkeit des Internets eines Clients. Kunden einer Cloud erhalten nach Erstellung eines Vertrages die Zugangsdaten zu ihrer Cloud, wodurch ermöglicht wird, dass nahezu jedes internetfähige Gerät auf die Daten zugreifen kann.

6 Vergleich

Anschließend an die Dokumentation der Systeme nimmt der Verfasser einen Vergleich der Archivierungslösungen vor. Die Bewertungskriterien sollen auf Hybrid- sowie Standalone-Lösungen anwendbar sein. Dabei wird stets das System als Ganzes dargestellt. Für eine Hybridlösung bedeutet dies, dass die Kapazität aus beiden Geräten kumuliert angegeben wird.

6.1 Bewertungskriterien

Gesamtkapazität

Die kumulierte Gesamtkapazität aller Komponenten im System, die benötigte Kapazität, geht aus der Bedarfsermittlung hervor. Sie kann auch Brutto-Kapazität genannt werden.¹⁴²

Kapazität (redundant)

Die Menge an Daten, die das Archivierungssystem redundant fasst, kann auch Nettokapazität genannt werden.¹⁴³

Beispiel 1: (mit RAID)

zwei Festplatten à 50 GB im RAID 1 (Spiegelung)

Gesamtspeicher: 100 GB

Redundante Kapazität: 50 GB

Beispiel 2: (ohne RAID)

NAS-System 1 = 150 TB, NAS-System 2 = 250 TB

Gesamtspeicher: 400 TB

Redundante Kapazität: 150 TB¹⁴⁴

Redundante Kapazität ist essenziell, wenn es um Datensicherheit geht. Datenredundanz bedeutet:

„[...]in der Datenorganisation das mehrfache Führen der gleichen Daten, also das mehrfache Vorkommen desselben Sachverhaltes, [...]“¹⁴⁵

¹⁴² Herminghaus; Scriba (2006): Veritas Storage Foundation. S. 46

¹⁴³ Vgl. ebenda, S.46

¹⁴⁴ Zwar hat das NAS-System 2 noch 100 TB zur Verfügung, dennoch wären die 100 TB dann nicht mehr redundant auf NAS-System 1 speicherbar.

Die Anzahl an Redundanzen ist nicht begrenzt, mit der Anzahl steigt jedoch der Hardwareaufwand.

Lebensdauer des Datenträgers

Die Lebensdauer hängt von der Speichertechnologie ab.

Datendurchsatz

Die Menge an Daten die das System lesen und schreiben kann. Dies ist von der Leistung des Trägermediums (z.B. HDD) abhängig. Im Vergleich wird stets die maximale Geschwindigkeit eines einzigen Datenträgers in einem Laufwerk angegeben.

Skalierbarkeit

Kann das System zu einem Zeitpunkt dem Speicherbedarf nicht mehr genügen, muss es erweiterbar sein. Dies geschieht in den meisten Fällen über Erweiterungsmodule.

Migration

Die Migration von digitalen Informationen ist der Zeitpunkt, zu dem Datenträger ein sogenanntes „refreshment“¹⁴⁶ benötigen. Beim „refreshment“ werden die Informationen eines Trägermediums auf ein anderes gleichen Typs übertragen. Damit sollen die gespeicherten Daten aufgefrischt werden. Die Frequenz der Migration eines Datenträgers hängt von der Speichertechnologie ab. Der Vorgang der Datenmigration muss mit äußerster Präzision durchgeführt werden, da die Bits identisch übertragen werden müssen. Durch sogenanntes Hashing wird einem Datenobjekt ein Identifikationswert zugeteilt, welcher die Datenintegrität sichert.¹⁴⁷

Verwaltungsaufwand

Die Zeit, die der Anwender aufwenden muss, um die Betriebsbereitschaft des Systems zu gewährleisten. Dazu zählt z.B. die Wartung oder Reinigung.

Schnittstellen

Ein Speicher braucht eine oder mehrere Schnittstellen, um mit anderen Geräten kommunizieren zu können. Bei Archivsystemen sind Schnittstellen wie Fibre-Channel, SATA, SAS und Ethernet relevant.

Kompatibilität

¹⁴⁵ Lackes; Siepermann (2018): Datenredundanz.

¹⁴⁶ Vgl. Brübach (2010) .S. 12

¹⁴⁷ Vgl. Hackel; Schäfer; Zimmer (2010): Vertrauenswürdigkeit von digitalen Langzeitarchiven. In: Neuroth; Oßwald; Scheffel et al. (Hrsg.). nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. S.9

Die Implementierung eines Systems wird normalerweise entsprechend des gegenwärtigen Bedarfs vorgenommen. Durch den technischen Wandel ist es allerdings nötig, die Anpassbarkeit eines Systems an eventuell veränderte Voraussetzungen einzuplanen. Kompatibilität sichert, dass das System zum einen mit der bestehenden Infrastruktur kommuniziert, zum anderen die einzelnen Datenträger auch bei zukünftigen Erweiterungen mit dem Archiv arbeiten.

Service

Alle technischen Geräte haben eine begrenzte Lebenszeit. Beim Defekt eines Datenträgers sowie eines ganzen Systems wird zur zeitnahen Problembekämpfung meist der Service des Herstellers benötigt. Auf welchen zeitlichen Rahmen ein Hersteller die Garantie für sein Produkt festsetzt und in welcher Form Service angeboten wird, muss bei der Projektierung beachtet werden.

Garantie

Gibt an, wie lange ein Hersteller die Gewährleistung einer Reparatur auf eigene Kosten festlegt.

Energieeffizienz

Der Stromverbrauch ist bei einem Archiv zwar aus funktioneller Sicht nicht relevant, wird dennoch in Bezug auf die Wirtschaftlichkeit eines Systems wichtig.

Kosten

Die Beachtung der Kosten eines Systems ist evident. Dennoch sollten bei der langfristigen Speicherung von Daten möglichst keine Kompromisse eingegangen werden, die zu Lasten der Datensicherheit gehen. Natürlich dessen müssen Kosten trotzdem dokumentiert werden. Anfallende Zusatzapplikationen oder Zubehör müssen dabei selbstverständlich auch beachtet werden. Datenträger sind oft nicht im Preis inbegriffen und steigern, abhängig von der Technologie, zusätzlich die Gesamtkosten.

6.2 Bewertung

Bewertungs-kategorie	Archivsystem					
	1	2	3	4	5	6
Systemnummer:						
Name	Hybrid /NAS Bandbibliothek	NAS ¹⁴⁸	Bandbibliothek	Optical Disc Archive 1	Optical Disc Archive 2	Hybrid NAS / Cloud ¹⁴⁹
Gesamtkapazität	1,632 PB	978 TB	1,200 PB	811 TB	806 TB	1,320 PB
Redundante Kapazität	672 TB	652 TB	600 TB	738 TB	736 TB	611 TB
Lebensdauer des Datenträgers	5 Jahre / 30 Jahre	10 Jahre	30 Jahre	100 Jahre	100 Jahre	5 Jahre / x
Datendurchsatz	200 MB/s / 300 MB/s	300 MB/s	300 MB/s	250 MB/s	360 MB/s	300 MB/s
Datenverfügbarkeit	Sofort	Sofort	Mind. 30 Sek.	Mind. 25 Sek.	Mind. 25 Sek.	Sofort
Skalierbarkeit	> 5 PB	2,32 PB	2,4 PB	1,7 PB	1,9 PB	unendlich
Verwaltungsaufwand / Wartung	Sehr hoch	mittel	hoch	leicht	leicht	mittel
Schnittstellen	SAS/FC/Ethernet	FC / Ethernet	SAS/FC/Ethernet	FC / Ethernet	SAS/FC/Ethernet	Ethernet
Kompatibilität	Windows/Mac/ Linux	Windows/Mac/ Linux	Windows/Mac/ Linux	Windows/Mac/ Linux	Windows/Mac/ Linux	Windows/Mac/ Linux
Service	3 Jahre	5-10 Jahre	5 Jahre	2 Jahre	unklar	3 Jahre
Garantie	3 Jahre	5-10 Jahre	1 Jahr	2 Jahre	3 Jahre	3 Jahre
Energieeffizienz	1000 W	400 W	851W	200W-700W	160W-480W	1200 W
Kosten (mit Service)	65.600 €	288.015,00 €	172.520 ,00 €	150.694,92 €	167.270,00 €	119.900,00 €

Tabelle 8: Vergleich der Systeme

Bewertungskriterium Kapazität

Betrachtet man die Systeme im Vergleich, fällt als erstes der große Unterschied zwischen der Gesamtkapazität und der redundanten Kapazität auf. Die Systeme 1, 3 und 6 speichern die Daten jeweils identisch auf mehreren Komponenten (1:1-Kopie), wohingegen die Standalone-Systeme (System 2, 4 und 5) die Daten nur auf einer Komponente lagern. Grundsätzlich sind Speicherungen auf verschiedenen Komponenten sicherer. Ortsverschiedene Standorte der einzelnen Komponenten verbessern zudem die Sicherheit, da ein Brand oder eine andersartige Zerstörung meist nur einen Stand-

¹⁴⁸ Der Anbieter FastLTA berechnet pro Jahr im Zuge der Serviceleistung eine Pauschale auf die Produkte (6%) und Lizenzen (18%). In der Rechnung wurden 5 Jahre Service (92.485 €) eingerechnet.

¹⁴⁹ Die Dienstleistungskosten sind hierbei für 5 Jahre gerechnet (1,405 € * 60)

ort zur gleichen Zeit gefährden. Am ähnlichsten sind die Kapazitäten der Optical Disc Archives. Die durch RAID 60 gesicherten Standalone-Lösungen weisen nur innerhalb des Geräts eine Redundanz durch Paritätsdaten auf. Ungeachtet dessen zählen optische Datenträger mit 100 Jahren Lebensdauer zu den zuverlässigsten Speichermedien. Dadurch sinkt die Dringlichkeit einer 1:1-Kopie, da der Datenträger ein geringeres Ausfallrisiko aufweist. Aus den oben genannten Gründen ist eine Kopie trotzdem zu empfehlen.

Bewertungskriterium Datendurchsatz / Datenverfügbarkeit

Die Identifikation des passenden Systems verlangt das Wissen darüber, was eigentlich archiviert und wie das archivierte Material verwendet werden soll. Ein Archiv soll nicht dem ständigen Abruf von kürzlich erstellten Produktionen dienen, vielmehr soll es Daten langfristig sicher erhalten. Damit grenzt es sich vom klassischen Produktionsserver ab. Besteht dennoch der Bedarf des erneuten Zugriffs auf archiviertes Material, z.B. für eine eventuelle Wiederverwertung, kommt der Datendurchsatz zum Tragen. Eine hohe Lesegeschwindigkeit reduziert die Dauer dieses Prozesses.

Die Systeme leisten zwischen 200 MB/s und 360 MB/s. Es gibt also keines, dessen Geschwindigkeit als ungenügend einzustufen wäre. Die bei Bandbibliotheken und Optical Disc Archives benötigte Mounting-Zeit eines Datenträgers in das Laufwerk verlängert allerdings den Zeitraum zwischen dem Lesebefehl und dem tatsächlichen Lesevorgang. Für einen schnellen Zugriff auf eine Datei eignet sich ein NAS also besser.

Bewertungskriterium Lebensdauer

Das wohl relevanteste Kriterium für Langzeitarchivierung ist die Lebensdauer eines Systems. Die Optical Disc Archives (System 4 und 5) liegen weit vor den restlichen Systemen und haben mit einer Lebensdauer von ca. 100 Jahren die besten Werte. Dies kommt vor allem dadurch zustande, dass die Cartridges (Sony) bzw. Magazine (Panasonic) nahezu keiner Pflege bedürfen. Die Widerstandsfähigkeit der Blu-rays trägt maßgeblich zur Lebensdauer bei, da die Disks Temperaturen zwischen -10 bis 55° Celsius¹⁵⁰ keinen Datenverlust erleiden. Dadurch sind die Systeme weniger ortsgebunden und bedürfen weniger Kontrolle.

Bandbibliotheken reihen sich hinter den Optical Disc Archives ein, können dennoch mit 30 Jahren eine gute Lebensdauer aufweisen. Um die Lebensdauer von 30 Jahren tatsächlich zu erreichen, ist eine regelmäßige Pflege der Cartridges nötig. Die meisten modernen Bandbibliotheken erkennen selbst, wann ein Band gereinigt werden muss. Nichtsdestotrotz sind die Schreib- und Lesezyklen, die ein Band verträgt, zu gering, um der Technologie allein den kompletten Datenbestand anzuvertrauen. Manche Hersteller sprechen dabei von ca. 250 Zyklen, bis das Band nicht mehr vertrauenswürdig ist. Faktoren wie Temperatur, Luftfeuchtigkeit und Abschottung von magnetischen Quellen müssen bei der Erhaltungsstrategie berücksichtigt werden. Eine Offline-Lagerung wäre bei LTO-Bändern der verlässlichste Gebrauch der Technologie – allerdings auf Kosten

¹⁵⁰ Sony (2018): Optical Disc Archive.

der Datenverfügbarkeit. Mit der Datenmigration, die Magnetbänder nach einigen Jahren benötigen, steigt der Aufwand der LTO-Nutzung weiter.

Priorisiert der Anwender die größtmögliche Datenverfügbarkeit, kann der Silent Brick die Archivierung über 10 Jahre lang gewährleisten. Bei der einfachen Archivierung mit wenigen Datenzugriffen verschenkt diese Lösung allerdings ihr Potenzial.

Bewertungskriterium Verwaltungsaufwand

Die Wartung einer Archivlösung ist unumgänglich. Durch den unregelmäßigen Zugriff nimmt das Archiv neben einem Produktionsserver meist eine passivere Rolle ein, muss aber trotzdem gepflegt werden. Zwar wäre eine „Set-it-forget-it“-Lösung attraktiver, aber unrealistisch. Jede Archiv-Hardware sollte in gewissen zeitlichen Abständen kontrolliert werden.

Die verschiedenen Datenträger benötigen unterschiedlichen Wartungsaufwand. Ein Server mit Festplatten muss zur Sicherung der Datenintegrität und Datenverfügbarkeit oft überprüft werden. Unabhängig ob ein Server der Steuerung eines anderen Systems dient oder selbst ein Massenspeicher ist, funktioniert dieses auf Basis eines Betriebssystems. Betriebssysteme sind grundsätzlich anfällig für moderne Schädlingsoftware. Kommt RAID zum Einsatz, gilt es zudem als empfehlenswert, gelegentlich den RAID-Controller auf volle Funktionstüchtigkeit zu überprüfen. Durch die kurze Lebensdauer einer Festplatte sind Migrationszyklen häufiger und erfordern das Umschichten des Datenspeichers nach weniger als 5 Jahren. Das Auslassen von Wartung gefährdet maßgeblich die Datenintegrität.

Bandbibliotheken bedürfen neben der bereits erwähnten Reinigung der Magnetbänder auch eines Mechanik-Checks. Der Transport-Roboter einer Bandbibliothek unterliegt während des Betriebes mechanischen Verschleißerscheinungen. Die erforderliche Migration wird bei Magnetbändern auf einen Zeitraum zwischen 6-10 Jahren¹⁵¹ festgelegt.

Optical Disc Archives benötigen neben der Wartung der inneren Mechanik wenig Zeitinvestment. Dadurch, dass das Auslesen keinen physischen Kontakt zwischen Lesekopf und Datenträger erfordert, entsteht kein Verschleiß. Die Disks sind zudem unabhängig von Umwelteinflüssen und erfordern keine Migration. Der Verwaltungsaufwand beschränkt sich also lediglich auf die Mechanik. Wird die Wartung vernachlässigt, leidet lediglich das Gerät, aber nicht der Datenträger. Die Sicherung der Datenintegrität erfordert bei Optical Disc Archives somit deutlich weniger Zeit.

¹⁵¹ Pavuza (2012): Detailprobleme und Lösungen aus dem Digitalisierungsalltag eines kleinen Videoarchivs. In: Österreichische Mediathek. Reale Probleme und virtuelle Lösungen: eine Bestandsaufnahme anlässlich 50 Jahre Österreichische Mediathek und des UNESCO-World-Day for Audiovisual Heritage 2010 : Beiträge zur Tagung der Medien Archive Austria und der Österreichischen Mediathek : Gerhard Jagschitz zum 70. Geburtstag. S. 51

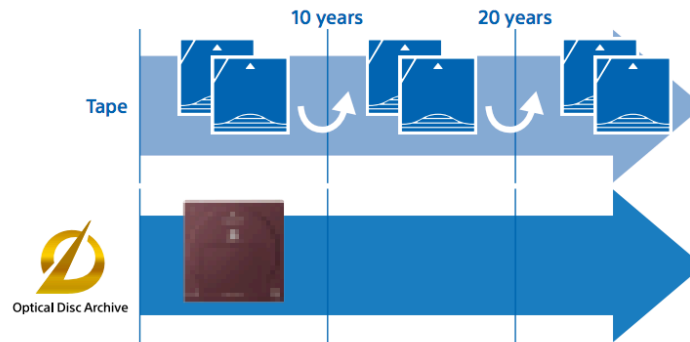


Abbildung 9: Migrationshäufigkeit Magnetband / ODA

Bewertungskriterium Garantie / Service

Die Garantieleistungen sowie das Serviceangebot beeinflussen die Wahl eines Gerätes stark. Serviceangebote sind oft recht teuer, erhalten dafür aber meist einen schnellen Vor-Ort-Austausch, Beratung und Remote-Wartungen. Je nach Hersteller sind Servicepakete unterschiedlich im Umfang und Preis.

Der Anbieter actidata (System 1) bietet von Haus aus 3 Jahre Garantie und Service ohne Aufpreis. Durch das Erweitern der Leistungen kann das „Sorglos“-Paket erworben werden. Hersteller 45 Drive (System 6) bietet zwar die gleichen Leistungen, lässt sich jedoch nicht erweitern und hat durch den Sitz in den USA eine schlechtere Anbindung. Die kostenfreie Garantie und Serviceleistung von actidata grenzt dieses System positiv gegen die anderen Anbieter ab, da diese entweder kurze Garantiezeiträume oder lediglich kostenpflichtige Services anbieten.

FastLTA bietet mit 5-10 Jahren den langfristigsten Service an. Dieser Hersteller berechnet bei einer Laufzeit von 5 Jahren pro Jahr 18% Gebühren auf die Kosten der Softwarelizenzen und 6% auf die Hardwarekosten. Der Preis der Serviceleistungen treibt die Kosten somit signifikant in die Höhe. Als Pendant dazu bietet der Hersteller HP bei einer sehr kurzen Garantie eine zeitliche Serviceerweiterung zu einem festen Preis an. Das System 3 wurde mit einer Erweiterung auf 5 Jahre ausgestattet.

Sony bietet für einige seiner Produkte einen Supportservice namens Primesupport. Aus verschiedenen Service-Paketen lässt sich die Dauer auf bis zu fünf Jahre verlängern. Dabei können technischer Support und schnelle Reparaturen per Vertrag erworben werden.¹⁵² Die Garantie von zwei Jahren ließe sich damit dementsprechend erweitern. Die genauen Preise ließen sich dafür nicht recherchieren. Anders verhält es sich bei Panasonic, aus dessen Angebot kein Serviceangebot ersichtlich wurde. Die Garantie des freeze-ray gilt allerdings, anders als bei Sony, für drei Jahre¹⁵³.

¹⁵² Sony (2018): Produktinformationen Primesupport.

¹⁵³ Panasonic (2015): Because we believe in our products.

Bewertungskriterium Kosten¹⁵⁴

Bei der Betrachtung der Kosten fällt auf, dass es zwischen einigen Lösungen gravierende Unterschiede gibt. So kann System 6 mit gerade einmal 118.105,75 € die zweitgünstigste Archivlösung bieten, wobei der Cloud-Service für 5 Jahre und ganze 600 Terabyte eingerechnet sind. Durch das „Pay-as-you-go“-Modell können die Kosten sogar deutlich unter 100.000 Euro liegen. Im Gegensatz dazu veranschlagt System 2 (FastLTA) mit 288.015,00 € über das doppelte an Kosten. Das günstigste System mit einem Gesamtpreis von 65.600 € ist System 1 von actidata. Werden nur die Kosten priorisiert, bietet dieses System die beste Lösung.

Beim Preis kann sich Sony positiv von seinem Konkurrenten Panasonic abheben. Mit nahezu identischer Kapazität, aber 25 % weniger Kosten ist die PetaSite Library nach System 6 die drittgünstigste Lösung

6.3 Vorschlag Archivierungslösung

Der Schutz der Daten - das oberste Ziel eines Archivs - stellt den Hauptgesichtspunkt für die Qualität einer Lösung dar. Er korreliert mit der zeitlichen Perspektive. Ein System, das die Daten fünf Jahre lang sicher speichert, fungiert im genannten Zeitraum als sicheres Archiv. Nach fünf Jahren muss es allerdings als potentiell unsicher eingestuft werden. Die Erneuerung eines Archivs ist prinzipiell nach einigen Jahren möglich, dennoch geht damit ein deutlicher Mehraufwand einher. Die Migration der Daten, Beschaffung neuer Hardware und Neukonfiguration erfordern einen Zeit- und Finanzaufwand, der sich kaum rechtfertigen lässt.

Wenn man die Datensicherheit und die Kosten priorisiert, ist das System 4 die beste Lösung für ein Langzeitarchiv. Die Sony PetaSite Library ist eine gute Alternative zum Konkurrenten Panasonic, der mit seinem System (System 5) preislich deutlich über Sony liegt. Weiter kann die PetaSite Library durch eine stärkere Präsenz am europäischen Markt mehr Zugänglichkeit zum Hersteller, aber auch zur Community gewähren. Die bereits vor über zehn Jahren eingeführte Technologie der PetaSite lässt darauf schließen, dass das Produkt ausgereift ist. Mit 739 € einmalige Kosten für den Sony File Manager bietet Sony eine Softwarelösung zur Verwaltung des Materials. Funktionen wie Thumbnails, Metadatenverwaltung und Proxys machen den File Manager zu einem passenden Werkzeug für die Anwendung im Videobereich. Die Software zum Betreiben der PetaSite ist zudem komplett kostenfrei. Im Vergleich dazu kostet allein die Software zum freeze-ray umgerechnet 21.856 €. Zudem ist freeze-ray nur im englischsprachigen Raum und in Asien bekannt.

Die Verbindung zwischen einfacher Handhabung und Strapazierfähigkeit des Datenträgers machen die optische Speicherung nahezu unschlagbar. Zwar bieten Magnet-

¹⁵⁴ Die Kosten betreffend wurden die verfügbaren Preise aus dem Internet als Referenz genommen. Sie beanspruchen keinerlei Aktualität und sind nicht alle durch den Hersteller verifiziert. Sie dienen somit nur der Einschätzung eines groben Kostenrahmens.

bänder auch eine adäquate Lebensdauer, können aber durch den deutlichen Mehraufwand in der Verwaltung kaum mit dem geringeren Preis konkurrieren. Dennoch soll der LTO-Technologie an dieser Stelle nicht die Daseinsberechtigung abgesprochen werden, da sie deutlich länger haltbar ist als Festplatten. Gerade bei Unternehmen ist wegen der rechtlichen Compliance-Vorschriften das Aufbewahren von Geschäftsdaten für zehn Jahre Pflicht. Dabei können Bandbibliotheken bei richtigem Umgang ohne Probleme diese Anforderungen erfüllen. Für die Sicherung von kreativem Gut gibt es keine Richtlinien, dennoch sollten den kreativen Arbeiten der Hochschule Mittweida ein hoher Wert zugesprochen werden. Mit der Einführung der optischen Speicherung kann das Thema Langzeitarchivierung wirtschaftlich und effizient umgesetzt werden.

Würde die Hochschule Mittweida neben der optischen Speicherung eine zusätzliche Sicherung anstreben, böte sich das Prinzip der Hybrid Cloud mit der PetaSite an. Durch das attraktive Pay-as-you-go-Bezahlmodell sind die Kosten der redundanten Auslagerung sehr niedrig. Ob die Sicherheit der Daten einer anderen Organisation anvertraut werden kann, liegt hierbei im Ermessen der Hochschule.

6.4 Handlungsvorschläge

Anknüpfend an die Empfehlung der Sony PetaSite gibt der Verfasser einen Handlungsvorschlag zur schnelleren Implementierung des Systems. Sony hat durch sein umfangreiches Produktportfolio Niederlassungen auf allen Kontinenten. Für den besten Zugang zum Produkt PetaSite empfiehlt es sich, Sony zu kontaktieren. Zwar bieten diverse Onlineshops einige PetaSite-Produkte, der direkte Kontakt zum Hersteller vereinfacht aber die Beschaffung und Installation des Systems. So kann auch eine perfekt abgestimmte Serviceleistung nachträglich vereinbart werden.

Auf Software-Ebene bleibt vorab mit Sony zu klären, ob das Hausformat XDCAM HD422 mit dem File Manager kompatibel ist. Die Ursache für die Unklarheit ist die Produktseite des File Manager, welche das XDCAM Format nicht unter den Kompatibilitäten aufzählt. Aus anderen Onlinequellen wird jedoch ersichtlich, dass einige Sendeanstalten in der Vergangenheit PetaSite und XDCAM simultan verwendeten. Durch ein großes Software-Portfolio von „Sony Creative Software Inc.“ speziell für den Broadcasting-Sektor kann eine Softwarelösung mit dem Hersteller zusammen entwickelt werden.

Generell empfiehlt sich die Beobachtung des Archivierungsmarktes, da neue Ideen und Ansätze stets das eigene Archiv verbessern können. Als Beispiel sei hier der „BlueGlacier“ von Panasonic und QNAP genannt. Dieser vereint einen Panasonic freeze-ray mit Hochleistungs-NAS in einer Standalone-Lösung. Der BlueGlacier verspricht gerade für Sendeanstalten mit der Symbiose aus den Vorteilen eines NAS und eines Optical Disc Archives eine leistungsstarke Hybridlösung. Das Produkt ist bis dato dennoch nur in Taiwan und Japan erhältlich und wurde deshalb nicht in den Vergleich aufgenommen. Vor dem Hintergrund des am 17.07.2018 unterschriebenen Freihandelsabkommens zwischen Japan und der EU wäre eine derartige Lösung vielleicht in Zukunft realisierbar. Inwieweit sich das Produkt in den europäischen Raum bringen lässt und wie ein Service gewährleistet werden kann, ist aktuell aber noch ungewiss.

7 Schlussfolgerung

Der Verfasser arbeitete im Vergleich der Archivierungssysteme die Sony PetaSite als geeignetste Lösung heraus. Die miteinander verglichenen Systeme orientieren sich an den am Markt gebräuchlichen Archivierungslösungen. Der Verfasser wählte die analysierten Systeme mit Bedacht. Keines der Systeme ließ sich aufgrund des erlangten Wissens als grundsätzlich unpassend identifizieren. Bei jedem System ließen sich also Stärken und Schwächen erkennen. Durch die Beschreibung der Funktionsweisen einzelner Technologien konnte klargemacht werden, warum sich das vorgeschlagene System am besten mit dem Bedarf deckt.

7.1 Forschungsablauf

Der Verfasser recherchierte im Vorfeld des Vergleichs ihm zunächst fremde Funktionsweisen technischer und digitaler Systeme. Dabei konnte er kein Vorwissen vorweisen und eignete sich alle Informationen in dieser Zeitspanne selbst an. Als er die Systeme nach eigener Einschätzung verstand, unternahm er den Versuch, diese in einer Sprache zu formulieren, die von semi-technikaffinen Menschen verstanden werden kann. Dennoch sah der Verfasser vor dem Hintergrund der Anforderung wissenschaftlichen Arbeitens davon ab, die tatsächliche Komplexität der Systeme bis zu einem Grad zu simplifizieren, bei dem die Beschreibungen ungenau wirken könnten. Der Fokus auf die Zusammensetzung einer digitalen Information und die Möglichkeit der Datenreduktion halfen dem Verfasser, den Forschungsgegenstand besser zu verstehen. Die ausgedehnte theoretische Vorarbeit verschaffte dem Verfasser eine Einsicht in die Verknüpfung vieler Aspekte mit dem Ergebnis, Archivsysteme beurteilen zu können. Der anfangs noch unscharfe Begriff eines Archivsystems klärte sich, als der Verfasser die Funktionsweisen verschiedener Komponenten verstand.

Zu Beginn der Arbeit recherchierte der Verfasser die Rolle der Datenarchivierung in der Arbeitswelt und wandte das erlangte Wissen auf Planung der nächsten Schritte an. Den Wert der Erhaltung von Informationen konnte der Verfasser einordnen und nahm alle wichtigen Faktoren sicherer Systeme auf, um diese im späteren Vergleich erneut aufzugreifen. Zahlreiche Quellen halfen beim Erstellen von wenigen, dafür aber relevanten Kriterien für die spätere Bewertung. Die Aufgabenstellung erforderte eine Analyse der erforderlichen Eckdaten des gewünschten Archivierungssystems. Die allgemeine Information, dass die Hochschule Mittweida Videos über einen langen Zeitraum archivieren möchte, reichte nicht aus. Zusätzlich mussten die Besonderheiten audiovisueller Inhalte berücksichtigt werden.

Eine erste Struktur nahm die Arbeit schon während der Bedarfsermittlung an. Dabei musste die Kapazität bestimmt werden, die sich aus der Erfahrung mit Videodateien und Informationen der Hochschule ergab. Zu diesem Zweck musste ein Zeitrahmen festgelegt werden, der sich auf den restlichen Verlauf der Arbeit auswirkte. Die Festlegung für 40 Jahre Archivierung mit einem neuen System bei sechs Videoproduktionen pro Woche traf der Verfasser eigenverantwortlich. Bewusst wurde der Bedarf hoch

angesetzt, da das anfallende Datenvolumen durch zukünftige Produktionsgeräte voraussichtlich noch steigen wird.

Mit dieser Kapazität als Ziel recherchierte der Verfasser zunächst Informationen zu Speichermedien, welche nach dem aktuellen Stand der Technologie verwendet werden. Dabei ließen sich vier besonders häufig verwendete Speichertypen finden, die aktuell verfügbar sind. Um das Verständnis für die Datensicherung zu stärken, wurde jedes Speichermedium detailliert in seiner Funktion beschrieben. Aus den Erläuterungen zu den Speichermedien ließen sich Stärken und Eignungstendenzen herausarbeiten. Für die Langzeitarchivierung konnten vor allem optische Datenträger und Magnetbänder als besonders effizient identifiziert werden. Mit dem neuen Wissenstand zu den Speichermedien galt es herauszufinden, wie diese Technologien in der Praxis eingesetzt werden.

Aufgrund der Wertigkeit des Themas „Redundanz“ beschäftigte sich der Verfasser mit der RAID-Technik. Die Beschreibung der RAID-Level trug dazu bei, dass das vorgestellte Konzept der redundanten Speicherung allgemein verständlich dargestellt werden konnte. Diese Informationen halfen bei den späteren Recherchen, da in Fachbüchern, aber auch im Internet das Wissen um RAID zum Systemverständnis oft vorausgesetzt wird. Folglich konnte der Verfasser nun das Wissen um die Speichermedien mit der redundanten Sicherung zusammenführen. Die tatsächliche Konzeption eines Archivsystems für die Hochschule Mittweida blieb bis dahin dennoch unklar. Im nächsten Schritt wurden Speicherarchitekturen in den Fokus genommen, um die Verkettung vieler Datenträger und deren Verbindung zum Nutzer besser bewerten zu können. Aufgrund der Komplexität von Netzwerken analysierte der Verfasser nur die für die Aufgabe relevanten Informationen zur Architektur der Systeme.

Allgemein bildete die Informationsphase eine gute Grundlage, um passende Systeme auszuwählen. Die Aufgabe der Recherche über Archivsysteme erwies sich als zeitintensiv. Die Funktionen der Systeme mussten unabhängig von deren Darstellung auf ihre Kernfunktion untersucht werden. Der Verfasser entschied sich, auf dem Markt erhältliche Systeme, unter Nennung von Herstellernamen zu benennen. Der Grund hierfür liegt im Ziel dieser Arbeit, welche neben der Darstellung der technischen Themen auch einen konkreten Vorschlag unterbreiten soll, um der Hochschule Mittweida unnötige Recherchearbeit zu ersparen.

Die Vielzahl an Angeboten und Produkten auf dem Archivmarkt machten es nötig, eine Auswahl zu treffen. Die Heterogenität der Systeme wurde gewählt, um möglichst viele verschiedene Typen der Archivierung aufzuzeigen. Die Archivsysteme unterschieden sich stark voneinander, da unterschiedlich viele Komponenten benötigt werden. Um die Vergleichbarkeit trotzdem zu gewährleisten, wurden Systeme mit ähnlichen Kapazitäten verglichen. Die Auswahl war recht aufwendig, da die verschiedenen Speichermedien große Kapazitätsunterschiede aufweisen. Es ließen sich leider nicht alle Archivsysteme auf das Terabyte genau konzeptionieren.

Ziel der Arbeit war die Auswahl eines präferierten Systems. Dazu erläuterte der Verfasser zunächst die Kriterien zur Auswahl eines Systems, auf deren Grundlage er jedes System analysierte. Die Differenzen wurden in Form von Tabellen dargestellt.

Einige Bewertungskriterien wurden in der späteren Betrachtung ausgelassen, da sie weniger relevant sind beanspruchten als zunächst vermutet. Der Verfasser konnte, die Kriterien aufgrund der Erkenntnisse aus seiner Arbeit priorisieren. Die Lebensdauer des Systems wurde so als wichtigstes Merkmal festgelegt. Aufgrund der priorisierten Kriterien und der entsprechenden Werte konnte das Produkt von Sony als optimale Lösung identifiziert werden.

7.2 Selbstreflexion

Der Verfasser schaut mit einem beachtlichen Erkenntnisgewinn auf seine Arbeit zurück. Die Herausforderung, sich in ein komplett neues Thema ohne Vorwissen einzuarbeiten, zahlte sich nach anfänglichem Zweifel aus. Die Recherche der Informationen zur Datenspeicherung gestaltete sich unkompliziert. Die Identifikation des optimalen Archivsystems hingegen erforderte längere Vorarbeiten und Recherchen. Rückblickend ist der Verfasser der Ansicht, dass der direkte Kontakt zu den Herstellern zielführender und zeitsparender gewesen wäre. Stattdessen wählte er die ihm vertraute Internetrecherche, die nicht alle Fragen klären konnte. Besonders in der Ermittlung der Systemkosten war diese Methode nicht effektiv genug.

Die Literaturrecherche empfindet der Verfasser als sinnvoll und zielführend. Die vielen studierten Fachbücher konnten ihm die Unsicherheit hinsichtlich der neuen technisch komplexen Zusammenhänge nehmen. Den Umfang der Arbeit hält der Verfasser für angemessen, auch wenn manche Themen nur oberflächlich behandelt werden konnten. Einige Themen, besonders die der Speicherarchitekturen, wiesen eine derartige Komplexität auf, dass es einer ausführlicheren Beschreibung in einer speziellen wissenschaftlichen Arbeit bedurft hätte.

Den Vergleich der Systeme stuft der Verfasser als zielführend ein. Der Vorschlag wurde durch präzise Recherche erarbeitet, und das Ergebnis kann mit gutem Gewissen vorgestellt werden. Fest steht, dass Sony mit der PetaSite zum gegenwärtigen Zeitpunkt die de facto sicherste Langzeitarchivierung zu einem sehr guten Preis-Leistungsverhältnis anbietet. Keine andere frei zugängliche Technologie kann Daten sicherer speichern.

Abschließend lässt sich sagen, dass für Videoarchive die optische Speicherung, die zur Zeit vertrauenswürdigste Technologie ist. Jede Speichertechnologie hat ihre Vor- und Nachteile, so dass es in jedem Einzelfall Sinn macht, die Auswahlkriterien individuell festzulegen. Die Sicherung von Daten ist essenziell, um die Kultur des jungen 21. Jahrhunderts zu erhalten. Dafür ist ein Medium nötig, dem man genug Vertrauen entgegen bringen kann.

Literaturverzeichnis

Monografien:

- Adam, Jessika (2010): Micromovie: Ein kreatives Medium für mobile Endgeräte. Diplomica Verlag, Hamburg, S.104
- Anahory, Sam; Murray, Dennis (1997): Data Warehouse: Planung, Implementierung und Administration. Addison Wesley, o. O., S. 155
- Bähring, Helmut (2013): Mikrorechner-Systeme: Mikroprozessoren, Speicher, Peripherie. 2. Auflage, Springer Verlag, Berlin; Heidelberg; New York et al., S. 510
- Badertscher, Kurt; Scheuring, Johannes (2007): Wirtschaftsinformatik: Wartung und Betrieb eines Informations- und Kommunikationssystems. 1. Auflage, Compendio Bildungsmedien AG, Zürich, S.178
- Bedner, Mark (2013): Cloud Computing: Technik, Sicherheit und rechtliche Gestaltung. kassel university press GmbH, Kassel, S.43
- Beisecker, Michael-Alexander (2005): Das Lexikon der PC-Fachbegriffe. Verlag für die Deutsche Wirtschaft AG, Bonn, S.263
- Broy, Manfred; Spaniol, Otto (Hrsg.) (1999): VDI-Lexikon Informatik und Kommunikationstechnik. 2. Auflage, Springer Verlag, Berlin; Heidelberg; New York et. al, S.572
- Cattini, Roland; Kammermann, Markus; Gut, Mathias (2016): CompTIA Server+: Serversysteme einrichten und betreiben. Vorbereitung auf die Prüfung SKO-004. 3. Auflage, MITP-Verlags GmbH & Co. KG, Sigloch, S.106
- Coers, Alexander (2005): Verfahren zur redundanten und distributiven Datenverarbeitung in drahtlosen Sensornetzen. Doktorarbeit, Universität Duisburg, S. 1
- Coyne, Larry; Engelbrecht, Michael; Browne, Simon (2018): IBM Tape Library Guide for Open Systems. 15. Auflage, IBM Redbooks, New York, S.19
- Dotzler, Bernhard J.; Roesler-Keilholz, Silke (2017): Mediengeschichte als Historische Techno-Logie. Nomos Verlag, Baden-Baden, S. 91
- Eisentraut, Peter; Helmle, Bernd (2013): PostgreSQL-Administration. 3. Auflage, O'Reilly Verlag, Köln, S. 368
- Fahrion, Roland (1989): Wirtschaftsinformatik: Grundlagen und Anwendungen. Physica Verlag, Heidelberg, S. 116
- Flik, Thomas; Liebig, Hans (2013): Mikroprozessortechnik: CISC, RISC Systemaufbau Assembler und C. 5., vollständig neu überarbeitete Auflage, Springer Verlag, Berlin; Heidelberg, S.513

- Freyer, Ulrich (2013): Medientechnik: Basiswissen Nachrichtentechnik, Begriffe, Funktionen, Anwendungen. Carl Hanser Verlag, München, S.67
- Gautzsch, Steffen (2010): Entwicklung Und Markteinführung Der Blu-Ray Disc: Die Blu-Ray Disc Als Home-entertainment-standard Und Dvd-nachfolger. Diplomica Verlag, Hamburg, S.33
- Giancoli, Douglas C. (2006): Physik. 3. Aktualisierte Fassung, Pearson Deutschland GmbH, München, S. 326
- Gloning, Christian Fabian (2008): Digitalisiertes Langzeitarchiv eines Fernsehsenders: Analyse der vernetzten, bandlosen Content-Produktion, -Speicherung und -Verwertung und der dazu nötigen technischen sowie infrastrukturellen Grundlagen. Diplomica Verlag, Hamburg, S. 27
- Goldammer, Gerd (1994): Informatik für Wirtschaft und Verwaltung: Einführung in die Grundlagen. Gabler, Wiesbaden, S.109
- Greiling, Michael; Dudek, Matthias (2009): Schnittstellenmanagement in Der Integrierten Versorgung: Eine Analyse Der Informations- Und Dokumentationsabläufe, 1. Auflage, Verlag W. Kohlhammer, Stuttgart, S. 68
- Grunwald, Armin (2010): Technikfolgenabschätzung: eine Einführung. zweite, grundlegend überarbeitete und wesentlich erweiterte Auflage, edition sigma, Berlin, S. 30
- Hansch, Pierre; Rentschler, Christian (2012): Emotion@Web: Emotionale Websites durch Bewegtbild und Sound-Design. Springer-Verlag, Berlin; Heidelberg, S.24
- Herminghaus, Volker; Scriba, Albrecht (2006): Veritas Storage Foundation. Springer Verlag, Berlin; Heidelberg; New York, S. 46
- Huebener, Rudolf (2014): Leiter, Halbleiter, Supraleiter - Eine Einführung in die Festkörperphysik: Für Physiker, Ingenieure und Naturwissenschaftler. Springer Verlag, Heidelberg, S. 159
- Huthmann, Andreas(2004): Metadaten und Datenqualität in Data Warehouses. GRIN Verlag, München, S. 5
- Jackson, Wallace (2012): Learn Android App Development. Springer Science + Business Media, New York, S.274
- Kaiser, Alexander (2003): Transformationscodierung. Seminararbeit, Eberhard Karls Universität, Tübingen, S.2
- Klimpke, Manuel; Keimel, Christian; Diepold, Klaus (2012): Visuelle Qualitätsmetrik basierend auf der multivariaten Datenanalyse von H.264/AVC Bitstream-Features. Technical Report, Technische Universität München, München, S. 13

- Kopacek, Peter; Zauner, Martin (2013): Leitfaden der technischen Informatik und Kommunikationstechnik. Springer Verlag, Wien, S.55
- Koslowski, Linda; Hensel, Christian (2013): Betriebswirtschaftliche Formelsammlung. 3. Auflage, Books on Demand, Norderstedt, S. 98ff
- Lassmann, Wolfgang (Hrsg) (2006): Wirtschaftsinformatik: Nachschlagewerk für Studium und Praxis. 1. Auflage, Gabler Verlag, Wiesbaden, S.98
- Lehner, Franz (2001): Einführung in Multimedia: Grundlagen, Technologien und Anwendungsbeispiele. Gabler, Wiesbaden, S. 64f
- Leimeister, Jan Marco (2015): Einführung in die Wirtschaftsinformatik. 12., vollst. Neu überarb. u. ak. Auflage, Springer Verlag, Berlin; Heidelberg, S.52
- Lutzke, Dieter (2013): Surfen in die digitale Zukunft. John Wiley & Sons, Weinheim, S. 53
- Malaka, Rainer; Butz, Andreas; Hußmann, Heinrich (2009): Medieninformatik: eine Einführung. Pearson Deutschland GmbH, Hallbergmoos, S. 99
- Mandl, Peter (2010): Grundkurs Betriebssysteme: Architekturen, Betriebsmittelverwaltung, Synchronisation, Prozesskommunikation. 2. Auflage, Vieweg + Teubner / GWV Fachverlage GmbH, Wiesbaden, S. 279
- Manz, Olaf (2016): Fehlerkorrigierende Codes: Konstruieren, Anwenden, Decodieren. Springer Verlag, Worms, S. 192
- Maslo, Andreas; Vonhoegen, Helmut (2005): Audio, Video, Foto. Markt+Technik Verlag, Burgthann, S.707
- Meinel, Christoph; Sack, Harald (2009): Digitale Kommunikation: Vernetzen, Multimedia, Sicherheit. Springer, Dordrecht; Heidelberg; London et al., S. 11
- Messmer, Hans-Peter; Dembowski, Klaus (2003): PC-Hardwarebuch: Aufbau, Funktionsweise, Programmierung ; ein Handbuch nicht nur für Profis. 7. Auflage, Addison-Wesley, München, S.26
- Müller, Klaus-Rainer (2003): IT-Sicherheit mit System: Strategie - Vorgehensmodell - Prozessorientierung – Sicherheitspyramide. 1. Auflage, Friedr. Vieweg & Sohn Verlag / GWV Fachverlag, Wiesbaden, S. 101
- Noppenberger, Martin (2009): Kostenminimierung in Speichernetzwerken. Dissertation, Eul Verlag, Bergische Universität Wuppertal, Wuppertal, S. 26
- Ortmann, Jürgen (2003): Einführung in die PC-Grundlagen. 8, aktualisierte Auflage, Addison-Wesley, München; Boston, San Francisco et al., S.112
- Ottmann, Thomas; Widmayer, Peter (2017): Algorithmen und Datenstrukturen. 6., durchgesehene Auflage, Springer Verlag, Berlin, S. 142

- Pilgrim, Mark (2010): HTML5: Up and Running: Dive into the Future of Web Development. O'Reilly Media, Sebastopol, S.82
- Reiner, Bernd (2005): HEAVEN. Eine hierarchische Speicher- und Archivierungsumgebung für multidimensionale Array Datenbankmanagement Systeme. Dissertation, Technische Universität München, S. 27
- Riggert, Wolfgang (2009): ECM - Enterprise Content Management. 1. Auflage, Vieweg + Teubner / GWV Fachverlage GmbH, Wiesbaden, S.171
- Robbe, Björn (2004): SAN - Storage Area Network: Technologie, Konzepte und Einsatz komplexer Speicherumgebungen. 2., aktualisierte und erweiterte Auflage, Hanser Verlag, München, S.33
- Roth, Klaus (2010): Chemische Köstlichkeiten. John Wiley & Sons, Weinheim, S.159
- Scheuring, Johannes; Waldispühl, Candidus (2012): Informatik für technische Kaufleute und HWD. 3. Auflage, Compendio Bildungsmedien AG, Zürich, S.160
- Schiffmann, Wolfram, Bähring, Helmut, Hönig, Udo (2011): Technische Informatik 3: Grundlagen der PC-Technologie. Springer Verlag, Heidelberg, S. 169
- Schmitt-Egenolf, Andreas (1990): Kommunikation und Computer: Trends und Perspektiven der Telematik. Gabler, Wiesbaden, S. 183
- Schmitz, Roland; Kiefer, Roland, Maucher, Johannes et al (2006): Kompendium Medieninformatik: Mediennetze, Springer-Verlag, Berlin; Heidelberg, S. 4
- Schneider, Holger (2012): Digitale Amnesie: Langzeitarchivierung digitaler Dokumente im betrieblichen Umfeld. Books on Demand, Norderstedt, S. 62
- Steinmetz, Ralf (2013): Multimedia-Technologie: Grundlagen, Komponenten und Systeme. 3., überarbeitete Auflage, Springer Verlag, Berlin; Heidelberg; New York et.al, S. 219ff
- Stickel, Eberhard; Groffmann, Hans-Dieter; Rau, Karl-Heinz. (Hrsg.) (2013): Gabler Wirtschaftsinformatik-Lexikon. Springer Verlag, Wiesbaden, S. 179
- Umlauf, Konrad (2006): Medienkunde. 2., aktualisierte und neu gefaste Auflage, Otto Harrassowitz Verlag, Wiesbaden, S. 43
- Urbach, Nils; Ahlemann, Frederik (2016): IT-Management im Zeitalter der Digitalisierung: Auf dem Weg zur IT-Organisation der Zukunft, Springer Verlag, Berlin; Heidelberg, S. 161
- Vetter, Scott; Narendra Babu ,Swarna; Balakrishnan Harihara (2015): IBM Power Systems RAID Solutions Introduction and Technical Overview. IBM Redbooks, o.O, S. 9

Wiegand, Bodo (2017): Der Weg aus der Digitalisierungsfalle: Mit Lean Management erfolgreich in die Industrie 4.0, Springer Verlag, Wiesbaden, S. 159

Wintermeyer, Stefan; Kempgen, Philipp (2009): Asterisk 1.4 + 1.6: Installation, Programmierung und Betrieb. Addison-Wesley, München; Boston; San Francisco et al., S. 173

Xu, Jingli (2009): Virtualisierung als Möglichkeit der Optimierung des IT- Managements. 1. Auflage, IGEL Verlag, Hamburg, S. 48ff

Zollner, Manfred; Zwicker, Eberhard (2013): Elektroakustik. 3. Auflage, Springer Verlag, Berlin; Heidelberg; New York et. al, S. 333

Sammelbände:

Anderson, Kai (2017): Die Idee einer menschlichen Digitalisierung. In: Keese, Christoph; Anderson, Kai; Pousttchi, Key et. al: Digital human: Der Mensch im Mittelpunkt der Digitalisierung. Campus Verlag, Frankfurt; New York, S. 13ff

Brübach, Nils (2010): Das Referenzmodell OAIS. In: Neuroth, Heike; Oßwald, Achim; Scheffel, Regine et al. (Hrsg.): nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.3. Werner Hülsbusch Verlag, Boizenburg, Kap. 4 S. 6

Däßler, Rolf (2010): Datenträger und Speicherverfahren für die digitale Langzeitarchivierung. In: Neuroth, Heike; Oßwald, Achim; Scheffel, Regine et al. (Hrsg.): Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.3. Werner Hülsbusch Verlag Boizenburg, Kap. 10, S. 6

Dickreiter, Michael; Dittel, Volker; Hoeg, Wolfgang; et. al (2008): Digitale Tonsignalspeicherung. In: ARD ZDF medienakademie (Hrsg.): Handbuch der Tonstudio-technik. 7., völlig neu bearbeitete und erweiterte Auflage, K. G. Saur Verlag, München, S. 791f

Friedewald, Michael; Leimbach, Timo (2011): Computersoftware als digitales Erbe: Probleme aus Sicht der Technikgeschichte. In: Robertson-von Trotha, Caroline; Hauser, Robert (Hrsg.): Neues Erbe: Aspekte, Perspektiven und Konsequenzen der digitalen Überlieferung, KIT Scientific Publishing, Karlsruhe, S. 204f

Hackel, Siegfried; Schäfer, Tobias; Zimmer, Wolf (2010): Vertrauenswürdigkeit von digitalen Langzeitarchiven. In: Neuroth, Heike; Oßwald, Achim; Scheffel, Regine et al. (Hrsg.): Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.3. Werner Hülsbusch Verlag Boizenburg, Kap. 5, S.9

Hengstberger, Wolf (2016): Storage Systeme erfolgreich managen. In: Tiemeyer, Ernst (Hrsg.). Handbuch IT-Systemmanagement: Handlungsfelder, Prozesse, Managementinstrumente, Good-Practices. Carl Hanser Verlag, München, S. 255

- Kramer, Klaus-Dietrich (2014): Digitale Halbleiterspeicher. In: Siemers, Christian, Sikora Axel (Hrsg.). Taschenbuch Digitaltechnik. 3. Auflage, Carl Hanser Verlag, München, S. 298
- Naumann, Christoph (2013): Digitalisierung als Katalysator für die Fotografieforschung. In: Ziehe, Irene; Hägel, Ulrich (Hrsg.): Fotografie und Film im Archiv: Sammeln, Bewahren, Erforschen. Band 6, Waxmann Verlag, Münster, S. 204
- Pavuzza, Franz (2012): Detailprobleme und Lösungen aus dem Digitalisierungsalltag eines kleinen Videoarchivs. In: Österreichische Mediathek. Reale Probleme und virtuelle Lösungen: eine Bestandsaufnahme anlässlich 50 Jahre Österreichische Mediathek und des UNESCO-World-Day for Audiovisual Heritage 2010 : Beiträge zur Tagung der Medien Archive Austria und der Österreichischen Mediathek : Gerhard Jagschitz zum 70. Geburtstag. LIT Verlag Münster, Wien, S. 51
- Sauter, Dietrich: Video (2009). In: Neuroth, Heike; Oßwald, Achim; Scheffel, Regine et al. (Hrsg.): Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.3. Werner Hülsbusch Verlag Boizenburg, Kap. 17 S.26f
- Schwartz, Dieter (2015): Digitale Langzeitarchivierung: Definition und Problemstellung. In: Umlauf, Konrad; Hobohm, Hans-Christoph (Hrsg.): Erfolgreiches Management von Bibliotheken und Informationseinrichtungen. 20. Lieferung, Verlag Dashöfer, Hamburg, Abschn. 9.4.2.2.1
- Ullrich, Dagmar (2010): Bitstream preservation. In: Neuroth, Heike; Oßwald, Achim; Scheffel, Regine et al. (Hrsg.). nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung, Version 2.3, Werner Hülsbusch Verlag, Boizenburg, Kap. 8, S. 4
- Ullrich, Dagmar (2008): Magnetbänder. In: Neuroth, Heike; Oßwald, Achim; Scheffel, Regine et al. (Hrsg.). nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 1.2., Werner Hülsbusch Verlag, Boizenburg, S. 164
- Wallaszkovits, Nadja (2010): Qualitätsaspekte der Langzeitarchivierung von Archiven im Web. In: Fennesz-Juhász, Christiane; Fröschl, Gabriele; Hubert, Rainer: Digitale Verfügbarkeit von audiovisuellen Archiven im Internet-Zeitalter: Beiträge zur Tagung der Medien Archive Austria und des Phonogrammarchivs der Österreichischen Akademie der Wissenschaften : Dietrich Schüller zum 70. Geburtstag. LIT Verlag, Münster, S. 85ff

Fachzeitschriften:

- Labs, Lutz (2015): Dicke Brummer. In: c't Redaktion: c't TESTGUIDE: Die wichtigsten Tests des Jahres: PC & Zubehör, Grafikkarten, Mobile, Printing, Laptops, Storage, Netzwerke, Monitore. Heise Medien GmbH & Co. KG, Hannover, S. 237
- von Löher, Franz (1976): Archivalische Zeitschrift. Band 72, Generaldirektion der Staatlichen Archive Bayerns, S. 92

Royon, Arnaud ; Bourhis, Kevin ; Bellec, Matthieu et. al. (2010): Silver Clusters Embedded in Glass as a Perennial High Capacity Optical Recording Medium. In: Wiley-VCH (Hrsg.): Advanced Materials, 22. Ausgabe, John Wiley & Sons, Weinheim, S. 5282

Internet-Quellen:

45 Drives (2018): Produktseite Storinator XL60.

<https://www.45drives.com/products/storinator-xl60-configurations.php>, eingesehen am 25.07.2018 18:44

actidata (2018): Produktbeschreibung actiNAS WIN424.

<https://www.actidata.com/index.php/de-de/actinas-win/actinas-win-424>, eingesehen am 24.07.2018, 18:31

actidata (2018): Produktbeschreibung Kodiak 6807.

<https://www.actidata.com/index.php/de-de/actilib-lto/actilib-kodiak-6807>, eingesehen am 24.07.2018, 18:31

Archivare (2018): Produktbeschreibung Archivare P5 Archive.

<http://p5.archivare.com/products/p5-archive>, eingesehen am 25.07.2018, 22:31

Bitkom (2017): Umsatz mit Informationstechnik in Deutschland von 2007 bis 2018 (in Milliarden Euro).

<https://de.statista.com/statistik/daten/studie/189877/umfrage/marktvolumen-im-bereich-informationstechnik-in-deutschland-seit-2007/>, eingesehen am 19. Juli 2018, 13:27

Business Wire (2017): *Absatzzahlen der PC-Hersteller weltweit von 2006 bis 2017 (in Millionen Stück)*.

<https://de.statista.com/statistik/daten/studie/151070/umfrage/absatz-der-pc-hersteller-am-weltmarkt/>, eingesehen am am 13.06.18 07:23)

DVS (2018): DVS Venice.

http://www.dvs.de/fileadmin/downloads/support/downloads/brochures/DVS_VENICE_Brochure.pdf, eingesehen am 23.07.2018, 19:01)

Fast LTA (2018): Backup optimieren. <https://www.fast-lta.de/de/anwendungen/backup/>, eingesehen am 15.07.2018 12:51

FastLTA (2018): Produktbeschreibung Silent Brick. <https://www.fast-lta.de/de/produkte/silent-bricks-uebersicht/>, eingesehen 25.07.2018 22:45

Flavorsys (2018): Startseite. <http://www.flavorsys.com/>, eingesehen am 01.07.2018, 14:41

- Grassvalley (2018): Produktbeschreibung K2 Summit 3G.
https://www.grassvalley.com/products/k2_summit_3g/ , eingesehen am 23.07.2018, 18:40
- Hasler, Tim (2017): Empfehlungen für größere Einrichtungen
<https://wiki.dnb.de/pages/viewpage.action?pageId=125432819> , eingesehen am 21.06.2018. 13:14
- Hewlett Packard Enterprise (2018): Produktbeschreibung HPE LTO-8 Ultrium 30TB RW Data Cartridge.
https://marketplace.hpe.com/pdp?prodNum=Q2078A&country=us&locale=en&catId=12169&catlevelmulti=12169_64242_342326_34648 , eingesehen am 25.06.2018, 11:48
- Hewlett Packard Enterprise (2018): Produktbeschreibung MSA 2040.
<https://h20195.www2.hpe.com/v2/GetPDF.aspx/c04123144>, eingesehen am 25.07.2018 15:06
- HPE (2018): Produktbeschreibung MSL 4048.
https://support.hpe.com/hpsc/doc/public/display?docId=emr_na-c00667627 , eingesehen am 25.07.2018 16:01
- IBM (2018): Produktbeschreibung IBM LTO Ultrium 8 Data Cartridge.
<https://www.ibm.com/de-de/marketplace/lto-8> , eingesehen am 06.07.2018 11:09
- Infotip Service GmbH (2018): Grundlagen Videokompression.
https://kompodium.infotip.de/grundlagen_videokompression.html , eingesehen am 19.06.2018, 11:52
- IT Wissen (2014): Dateiformat. <https://www.itwissen.info/Dateiformat-file-format.html> , eingesehen am 18.06.2018 21:15
- Lackes, Richard; Siepermann, Markus (2018): Datenredundanz.
<http://wirtschaftslexikon.gabler.de/definition/datenredundanz-31756/version-255307> , eingesehen 13.07.2018 11:32
- Lichtscheidl, Irene (2018): Video.
<https://www.univie.ac.at/video/grundlagen/chroma.htm> , eingesehen am 20.06.2018, 11:25
- Oracle (2018): Produktseite Oracle Cloud, https://cloud.oracle.com/en_US/home, eingesehen am 25.07.2018, 18:47
- Panasonic (2015): Because we believe in our products. <https://www.panasonic-electric-works.com/eu/3-years-warranty-because-we-believe-in-our-products.htm> , eingesehen am 17.07.2018 14:26
- QNAP (2018): Produktbeschreibung TS-EC2480U. <https://www.qnap.com/de-de/product/ts-ec2480u> , eingesehen am 05.07.2018 13:50

- Panasonic (2018): Produktseite freeze-ray.
<https://panasonic.net/cns/archiver/product/#Overview>, eingesehen am 25.07.2018 18:42
- QStar (2018): Produktseite Archive Manager. <http://www.qstar.com/index.php/archive-manager/>, eingesehen am 25.07.2018, 22:51
- Sony (2018): Optical Disc Archive. <https://pro.sony/s3/cms-static-content/file/85/1237494973885.pdf>, eingesehen am 25.07.2018 10:00
- Sony; Panasonic (2015): White Paper Archival Disc Technology. 1. Version,
https://panasonic.net/cns/archiver/pdf/E_WhitePaper_ArchivalDisc_Ver100.pdf,
eingesehen am 15.07.2018 12:28
- Sony (2018): Produktinformationen Primesupport. https://pro.sony/de_CH/support-content/warranty?q=FW-65BZ35F, eingesehen am 17.07.2018 14:19
- Vaughan-Nichols, Steven (2017): Floppy disk history: The evolution of personal computing. <https://www.hpe.com/us/en/insights/articles/the-history-of-the-floppy-disk-1703.html>, eingesehen am 12.06.2018 12:31
- Wachenschwanz, David (2018): HAMR vs MAMR: Which Technology Will Be Used For the Next Generation of Hard Disk Drives?. <https://www.linkedin.com/pulse/hamr-vs-mamr-which-technology-used-next-generation-wachenschwanz>, eingesehen am 26.06.2018, 14:25

Bildquellen:

Abbildung 1:

<https://tse1.mm.bing.net/th?id=OIP.VS7enfQY34wqPSDLATKy4gAAAA&pid=Api>

Abbildung 2:

<https://upload.wikimedia.org/wikipedia/commons/thumb/0/04/Digital.signal.discret.svg/2000px-Digital.signal.discret.svg.png>

Abbildung 3: <https://i.stack.imgur.com/tAdMw.png>

Abbildung 4: [https://encrypted-](https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcSggS1pdA_PU9X_3b70FbAN6J0-tXZ_2VPE_HdiRTvM3kn6ZawD)

[tbn0.gstatic.com/images?q=tbn:ANd9GcSggS1pdA_PU9X_3b70FbAN6J0-tXZ_2VPE_HdiRTvM3kn6ZawD](https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcSggS1pdA_PU9X_3b70FbAN6J0-tXZ_2VPE_HdiRTvM3kn6ZawD)

Abbildung 6: <https://wiki.zum.de/images/c/c6/Schreibvorgang2.jpeg>

Abbildung 7: http://deacademic.com/pictures/dewiki/67/CD_Querschnitt.png

Abbildung 8:

<https://upload.wikimedia.org/wikipedia/commons/thumb/2/29/NAS.png/260px-NAS.png>

Abbildung 9:

https://upload.wikimedia.org/wikipedia/commons/7/7f/Schema_SAN_german_V2.png

Abbildung 10: <https://pro.sony/s3/cms-static-content/file/85/1237494973885.pdf>

Abbildung 11: interne Quelle HS Mittweida

Abbildung 12: <https://www.stellardatenrettung.de/wp-content/uploads/2016/08/RAID-O.jpg>

Abbildung 13: <https://www.stellardatenrettung.de/wp-content/uploads/2016/08/RAID-1.jpg>

Abbildung 14:

https://upload.wikimedia.org/wikipedia/commons/thumb/a/ad/RAID_4.svg/640px-RAID_4.svg.png

Abbildung 15: <http://www.technologyjoint.com/wp-content/uploads/2018/03/Raid-6-Data-Recovery.png>

Abbildung 16: <https://community.spiceworks.com/topic/1809064-how-safe-is-raid-10-really>

Anhang I

Kostenrechnung System 1: (actidata)

Produkt:	Quantität:	Preis / pro Stück:
Actidata Kodiak 6807	1	15.000,00 €
actidata NAS WIN424 (inkl. Festplatten)	1	28.000,00 €
Archivare P5	1	9.000,00 €
LTO 8 Band	80	170,00 €
Ergebnis		65.600,00 €

Quelle:

Actidata- und Archivare-Preise gehen aus Kostenvoranschlag hervor.

LTO 8-Cartridge: https://shop.jacob.de/produkte/e-lto-8-ultrium-q2078a-artnr-3880758.html?gclid=CjwKCAjwp7baBRBIEiwAPTjwxFL_lmyZyllYic_qqmURDwbvxcF8vDe4BwsGldfaZmpZn0nafayjtxoCiMkQAvD_BwE

Kostenrechnung System 2: (Fast LTA)

Produkt:	Quantität	Preis pro Stück
Controller	1	5.990,00 €
Extension Shelf	4	5.390,00 €
Silent Brick 24TB	62	1.800,00 €
Service Hardware 6% / Jahr	5	8.349,00 €
Einzellizenz 500	1	37.990,00 €
Einzellizenz 152	1	18.390,00 €
Service Lizenz 18% / Jahr	5	10.148,40 €
Ergebnis		288.015,00 €

Quelle:

Fast LTA Preisliste: <https://shop.bechtle.at/medias/37PSb7S3ftqHWsQcCBYV58-30.pdf>

Kostenrechnung System 3:

Produkt:	Quantität:	Preis / pro Stück:
Serviceerweiterung (5Jahre) MSA 2040	1	3100
Serviceerweiterung (5 Jahre) MSL 4048	2	10100
Laufwerk (Quantität x6)	6	4100
LTO 8 Band (Quantität x96)	96	170
HPE MSL 4048	2	4000
Festplatte Seagate SAS SFF	24	250
HPE MSA 2040	1	4000
Qstar Archive Manager (1000TB)	1	64500

Qstar Archive Manager (250 TB)	1	25800
Ergebnis:		172520 €

Quellen (nach Auflistung):

<https://www.tiefenbach-it.com/hersteller/hp-enterprise/carepacks/18625/hp-ecare-pack-5y/4h-24x7-foundation-care-service-u2mr9e>

<https://www.redcorp.com/en/product/storage-warranty/hewlett-packard-enterprise/hp-5-years-24x7-msl4048-tape-library-fc-svc-u3bd3e--u3bd3e/1171t060>

https://www.bitinfo.shop/server-storage-pc/storage/tape-laufwerke/236677/hpe-storeever-lto-8-ultrium-30750-bandlaufwerk-lto-ultrium-12-tb/30-tb-ultrium-8-sas-2?gclid=CjwKCAjwp7baBRBIEiwAPtjwxGYPXOBQ0oXH7BdemLhI-Bubh9adYxHdSRjh2k- UJQKm2-VW2SZABoCAUoQAvD_BwE

https://shop.jacob.de/produkte/e-lto-8-ultrium-q2078a-artnr-3880758.html?gclid=CjwKCAjwp7baBRBIEiwAPtjwxFL_lmyZyIIYic_qqmURDwbvxcF8vDe4BwsGldfaZmpZn0nafajtxoCiMkQAvD_BwE

https://www.it-market.com/de/hewlett-packard/hp-storage/hp-storage-library-systeme/hp-ak381a1?gclid=CjwKCAjwp7baBRBIEiwAPtjwxC3H0ybpB2DLAWGTvV18YVR30EmOdWKbt-m98K8Wr0N0SvYEZDi-ZRoCYxUQAvD_BwE

https://www.servershop24.de/komponenten/festplatten/sas/seagate-exos-7e2000-2tb-2000-gb-12g-7-2k-sas-2-5-festplatte-hard-disk-st2000nx0273/a-118013/?ReferrerID=7&utm_source=googleproducts&utm_medium=cpc&utm_campaign=de&gclid=CjwKCAjwp7baBRBIEiwAPtjwxBHYYi1S3IBhgAp6aOma3jEqZ_RFcfGzcehDMA1OT5fCsUhdybV4XoRoCwBAQAvD_BwE

https://www.scalcom.de/hpe-msa-2040-es-san-dc-sff-storage.html?gclid=CjwKCAjwp7baBRBIEiwAPtjwxB9cAB00U01UF8FY215cAnvZZJ_aNgAWxuKOzCzA0rHLEL6DownzxoCZ5QQAvD_BwE

<http://itprice.com/hp-price-list/Qstar.html>

Kostenrechnung System 4: (Sony PetaSite)

Produkt:	Quantität:	Preis / pro Stück:
File Manager	1	739,95 €
ODA Laufwerk	6	11800 €
Cartridge ODC-3300R	250	145 €
Erweiterungseinheit 100E	1	12.192,15 €
Erweiterungseinheit 60E	2	12.192,15 €
Master-Einheit 30E	1	13.900,00 €
Ergebnis		150.694,92 €

Quellen (nach Auflistung):

<https://www.sonycreativesoftware.com/de/filemanager>

https://www.bhphotovideo.com/c/product/1273096-REG/sony_odsd280f_optical_disc_archive_fiber.html

<https://www.videodata.de/shop/products/de/Optical-Disc-Archive/Sony-ODC3300R.html>

<https://www.cnet.com/products/sony-petasite-optical-disc-archive-library-expansion-unit-ods-l100e-expansion-module/prices/>

<https://www.cnet.com/products/sony-petasite-optical-disc-archive-library-expansion-unit-ods-l60e-expansion-module/>

https://cvp.com/product/sony_broadcast_ods-l30m5

Kostenrechnung System 5: (Panasonic freeze-ray)

Produkt:	Quantität:	Preis / pro Stück:
Basis-Einheit LB-DH70A0G	1	73.625,00 €
"Bottom"-Einheit LB-DF81Z1G	1	7.500,00 €
Erweiterungseinheit LB-DH82Z1G	1	8.750,00 €
Erweiterungseinheit LB-DF72A0G	1	47.625,00 €
Cartridge	228	250,00 €
Ergebnis		167.270,00 €

Quelle:

Freeze-Ray-Preise aus PDF: https://s3-us-west-2.amazonaws.com/naspoaluepoint/1498745693_Panasonic%20FreezeRay%20SRP%20Cost%20Catalog.pdf

Kostenrechnung System 6: (45 Drives/Oracle)

Produkt:	Quantität:	Preis / pro Stück:
Storinator XL60	1	13000 €
Storinator 5 Jahre Garantie	1	4000 €
Festplatte: Western Digital Gold HD	60	335 €
Cloud /pro GB	600.000	0,0023 €
Cloud /Monat (600 TB)	60	1.380 €
Ergebnis		119.900,00 €

Quelle:

<https://www.45drives.com/products/storage/order.php?id=XL60&config=03&model=XL60-03&code=XL&software=Default&type=storage>

Cloud-Preise gingen aus einer Hersteller-Email hervor

Anhang II

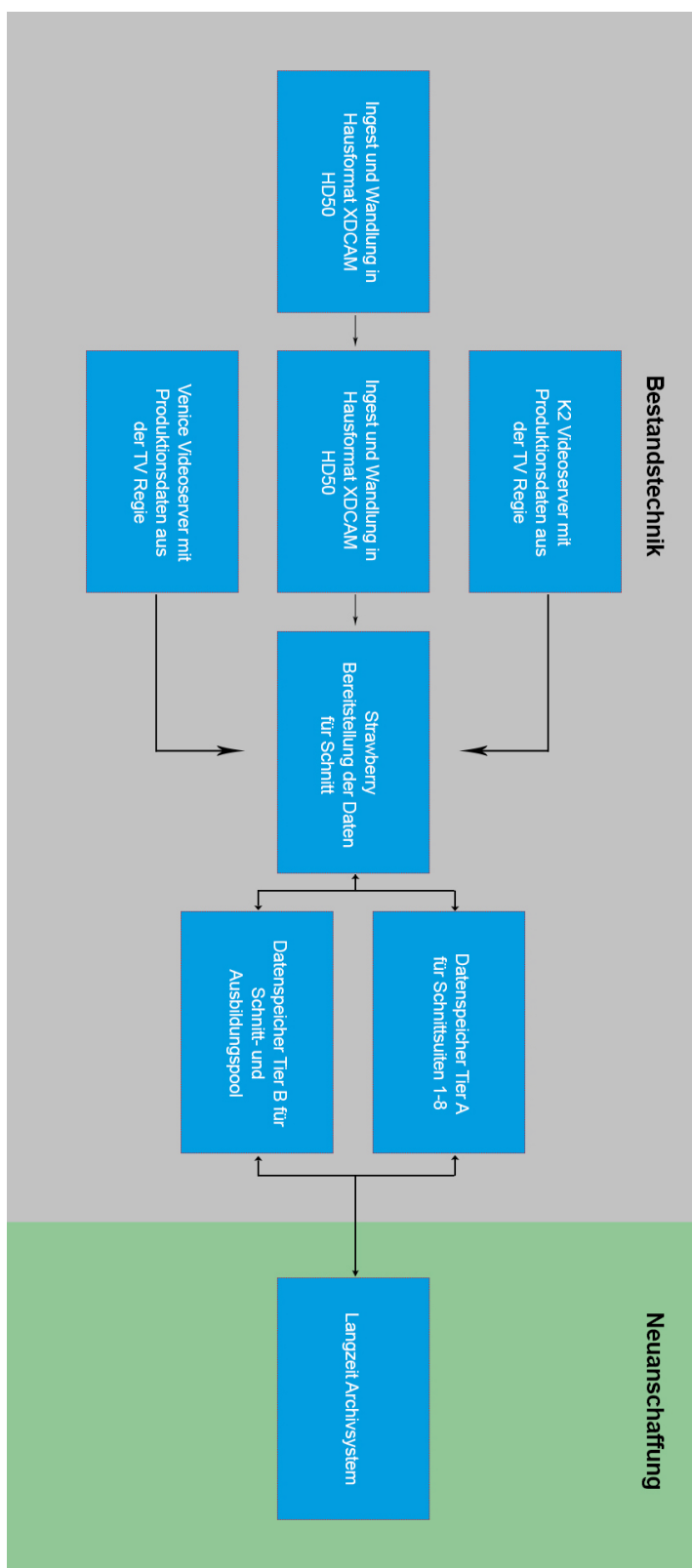


Abbildung 10: Infrastruktur HS Mittweida

Anhang III

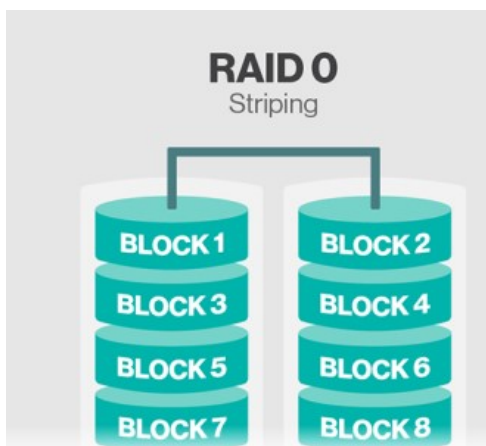


Abbildung 11: RAID 0

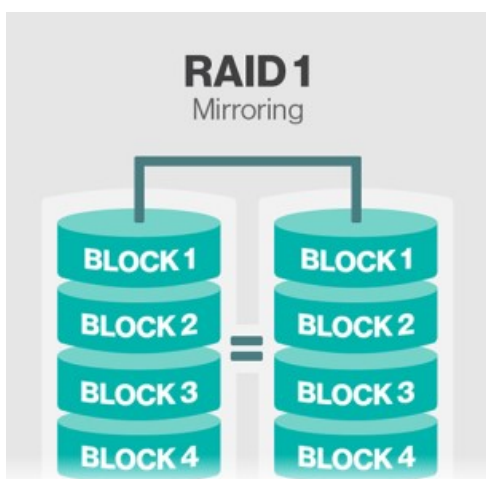


Abbildung 12: RAID 1

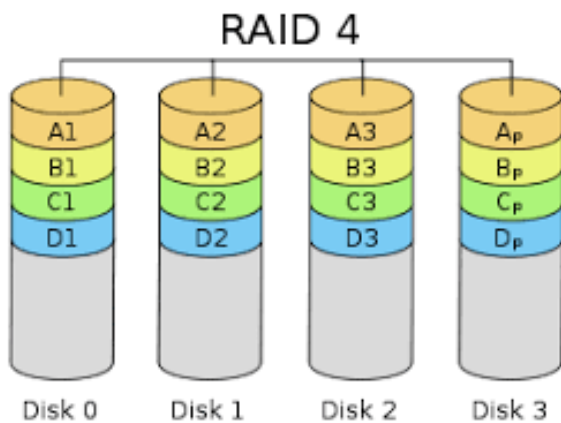


Abbildung 13: RAID 4

RAID 6

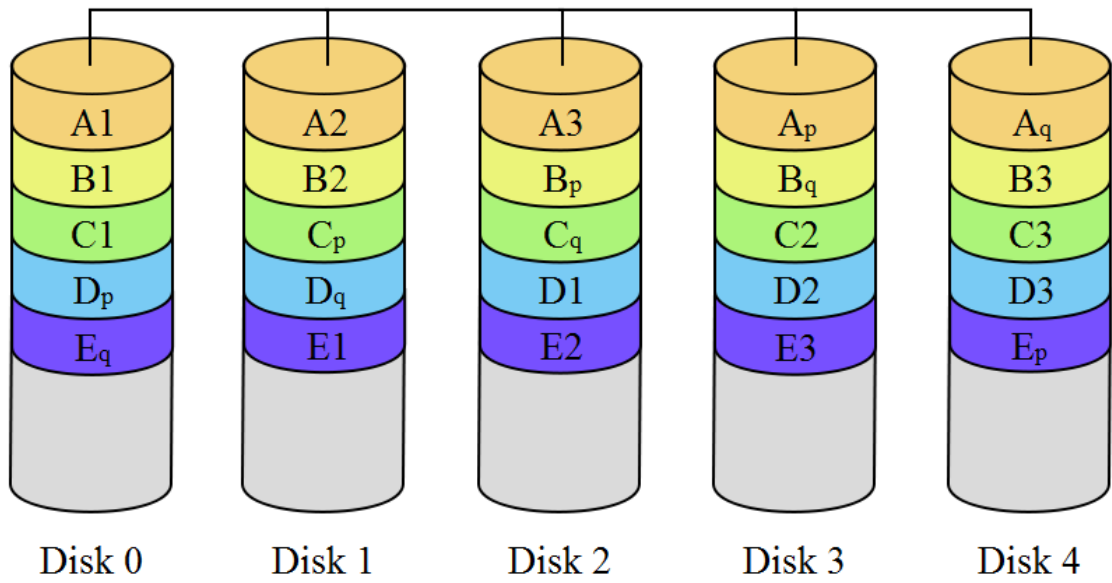


Abbildung 14: RAID 6

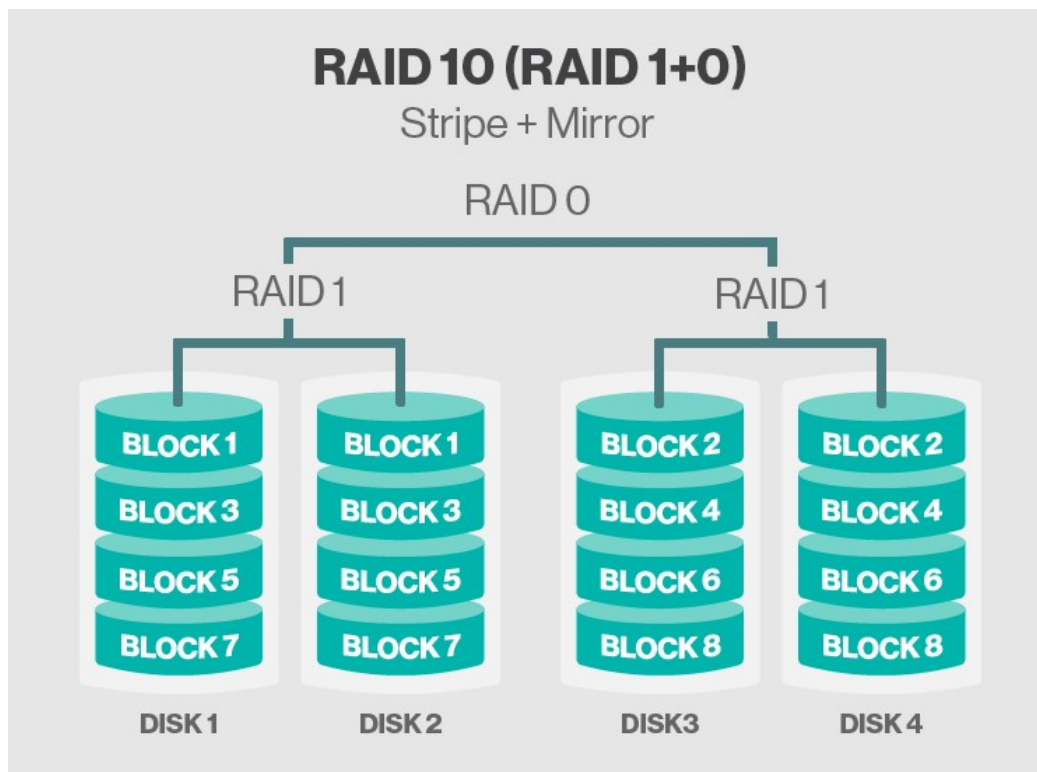


Abbildung 15: RAID 10

Eigenständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe. Stellen, die wörtlich oder sinngemäß aus Quellen entnommen wurden, sind als solche kenntlich gemacht. Diese Arbeit wurde in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegt.

Hamburg, 25.07.2018

Ort, Datum

Vorname Nachname