

nestor Handbuch:
**Eine kleine Enzyklopädie
der digitalen Langzeitarchivierung**

11 Hardware

Herausgeber

Heike Neuroth
Hans Liegmann †
Achim Oßwald
Regine Scheffel
Mathias Jehn
Stefan Strathmann

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Im Auftrag von

nestor – Kompetenznetzwerk Langzeitarchivierung und Langzeitverfügbarkeit
digitaler Ressourcen für Deutschland
nestor – Network of Expertise in Long-Term Storage of Digital Resources
<http://www.langzeitarchivierung.de>

Kontakt

editors@langzeitarchivierung.de

c/o

Niedersächsische Staats- und Universitätsbibliothek Göttingen

Dr. Heike Neuroth

Forschung und Entwicklung

Papendiek 14

37073 Göttingen

Tel. +49 (0) 55 1 39 38 66

Der Inhalt steht unter folgender Creative Commons Lizenz:
<http://creativecommons.org/licenses/by-nc-sa/2.0/de/>



11 Hardware

11.1 Hardware-Environment

Dagmar Ulrich

Abstract

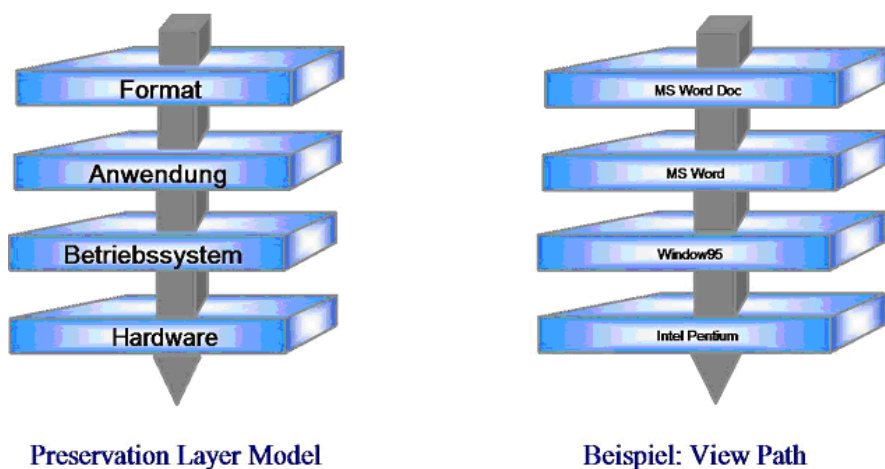
Digitale Datenobjekte benötigen eine Interpretationsumgebung, um ihren Inhalt für Menschen zugänglich zu machen. Diese Umgebung kann in unterschiedliche Schichten gegliedert werden, deren unterste die Hardware-Umgebung bildet. Diese Einteilung wird anhand eines Schichtenmodells, dem „Preservation Layer Model“ veranschaulicht. Die Hardware-Umgebung umfasst nicht nur eine geeignete Rechnerarchitektur zur Darstellung der Inhalte, sondern auch eine funktionsfähige Speicherumgebung für den physischen Erhalt und die Bereitstellung des digitalen Datenobjektes.

Gliederung

Interpretationsumgebung digitaler Objekte und „Preservation Layer Model“
Speicherung und Bereitstellung des digitalen Objekts

Interpretationsumgebung digitaler Objekte und „Preservation Layer Model“

Um ein digitales Datenobjekt lesbar zu halten, muss eine entsprechende Interpretationsumgebung verfügbar sein. Diese umfasst Hardware, Betriebssystem und Anwendungssoftware. Um z.B. eine Word-Datei anzuzeigen wird eine passende Version von MS-Word benötigt. Für die Installation der Anwendungssoftware muss ein geeignetes Betriebssystem verfügbar sein, das seinerseits auf eine entsprechende Rechnerarchitektur angewiesen ist. In der Regel gibt es mehrere mögliche Kombinationen. Die Lesbarkeit digitaler Daten ist nur so lange sichergestellt, wie mindestens eine solche gültige Kombination einsatzfähig ist. Dieser Zusammenhang wird im Konzept des „Preservation Layer Models“ herausgearbeitet. Die nachstehende Grafik veranschaulicht dieses Konzept.¹



Eine funktionsfähige Kombination der verschiedenen Ebenen wird als gültiger „View Path“ eines digitalen Datenobjektes bezeichnet und kann dem entsprechenden Objekt zugeordnet werden. Das Preservation Layer Model wurde an der Nationalbibliothek der Niederlande gemeinsam mit IBM entwickelt, um rechtzeitig zu erkennen, wann ein Datenobjekt Gefahr läuft, ohne gültigen View Path und damit nicht mehr lesbar zu sein. Zeichnet sich der Wegfall einer Komponente ab, lässt sich automatisch feststellen, welche View Paths und

¹ Eine ausführliche Beschreibung des Preservation Layer Models findet sich in: Van Diessen, Raymond J. (2002): *preservation requirements in a deposit system*. Amsterdam: IBM Netherlands. S. 7-15. <http://www-05.ibm.com/nl/dias/resource/preservation.pdf> [2007, 20. August]

somit welche Datenobjekte betroffen sind. Auf dieser Grundlage kann dann entweder eine Emulationsstrategie entwickelt oder eine Migration betroffener Datenobjekte durchgeführt werden. Im Falle einer Formatmigration werden alle darunter liegenden Ebenen automatisch mit aktualisiert. Die Hard- und Softwareumgebung des alten Formats wird nicht mehr benötigt. Will man jedoch das Originalformat erhalten, müssen auch Betriebssystem und Rechnerarchitektur als Laufzeitumgebung der Interpretationssoftware vorhanden sein. Nicht immer hat man die Wahl zwischen diesen beiden Möglichkeiten. Es gibt eine Reihe digitaler Objekte, die sich nicht oder nur mit unverhältnismäßig hohem Aufwand in ein aktuelles Format migrieren lassen. Hierzu gehören vor allem solche Objekte, die selbst ausführbare Software enthalten, z.B. Informationsdatenbanken oder Computerspiele. Hier ist die Verfügbarkeit eines geeigneten Betriebssystems und einer Hardwareplattform (nahezu) unumgänglich. Um eine Laufzeitumgebung verfügbar zu halten, gibt es zwei Möglichkeiten. Zum einen kann die Originalhardware aufbewahrt werden (vgl. hierzu Kapitel 12.4 Computermuseum). Zum anderen kann die ursprüngliche Laufzeitumgebung emuliert werden (vgl. hierzu Kapitel 12.3 Emulation). Es existieren bereits unterschiedliche Emulatoren für Hardwareplattformen² und Betriebssysteme.

Speicherung und Bereitstellung des digitalen Objekts

Aber nicht nur die Interpretierbarkeit der Informationsobjekte erfordert eine passende Umgebung. Bereits auf der Ebene des Bitstream-Erhalts wird neben dem Speichermedium auch eine Umgebung vorausgesetzt, die das Medium ausliest und die Datenströme an die Darstellungsschicht weitergibt. So brauchen Magnetbänder, CD-ROMs oder DVDs entsprechende Laufwerke und zugehörige Treiber- und Verwaltungssoftware. Bei einer Festplatte sind passende Speicherbusse und ein Betriebssystem, das die Formatierung des eingesetzten Dateisystems verwalten kann, erforderlich.

Literatur

Van Diessen, Raymond J. (2002): preservation requirements in a deposit system. Amsterdam: IBM Netherlands. S. 7-15. <http://www-05.ibm.com/nl/dias/resource/preservation.pdf> [2007, 20. August]

2 Als Beispiel für die Emulation einer Rechnerarchitektur kann „Dioscuri“ genannt werden. Dioscuri ist eine Java-basierte Emulationssoftware für x86-Systeme. <http://dioscuri.sourceforge.net/> [2007, 20. August]

11.2 Digitale Speichermedien

Dagmar Ulbrich

Abstract

Datenträger, egal ob analog oder digital, sind nur begrenzt haltbar und müssen früher oder später ausgewechselt werden, um Informationsverlust zu verhindern. Digitale Datenträger veralten in der Regel wesentlich schneller als übliche analoge Medien. Zudem hängt ihre Lesbarkeit von der Verfügbarkeit funktionsstüchtiger Lesegeräte ab. Zu den gängigen digitalen Speichermedien zählen Festplatten, Magnetbänder und optische Medien wie CD-ROM oder DVD. Die Unterschiede in Haltbarkeit und Speichereigenschaften entscheiden darüber, in wie weit und in welcher Kombination sie für die Langzeitarchivierung eingesetzt werden können.

Gliederung

Lebensdauer von Trägermedien

Die wichtigsten digitalen Speichermedien

Speichermedien in der Langzeitarchivierung

Lebensdauer von Trägermedien

Um Informationen über die Zeit verfügbar zu halten, müssen sie auf einem zuverlässigen Trägermedium vorliegen. Die Haltbarkeit des Trägermediums ist von wesentlicher Bedeutung für die Verfügbarkeit der Information. Seine begrenzte Lebensdauer erfordert ein rechtzeitiges Übertragen auf ein neues Medium. Mündlich tradierte Gedächtnisinhalte werden durch Auswendiglernen von einer Generation an die nächste weitergereicht. Schriftstücke wie Urkunden, Bücher oder Verträge werden bei Bedarf durch Kopieren vor dem Verfall des Trägermediums geschützt. Auch digitale Daten benötigen Trägermedien, die erhalten und ggf. erneuert werden müssen.³ Im Vergleich zu herkömmlichen analogen Datenträgern sind digitale Datenträger jedoch in der Regel deutlich kurzlebiger. Neben ihrer Kurzlebigkeit spielt für digitale Datenträger noch ein

3 Der Nachweis der Authentizität ist bei analogem Material wesentlich stärker als bei digitalen Daten an das Trägermedium gebunden. Bei Kopiervorgängen muss dies berücksichtigt werden. Vgl. hierzu Kapitel 8.1.

weiterer Aspekt eine Rolle: Es wird eine Nutzungsumgebung benötigt, um die Datenobjekte zugänglich zu machen. Um ein digitales Trägermedium, z.B. ein Magnetband oder eine CD-ROM lesen zu können, ist ein entsprechendes Laufwerk und die zugehörige Treibersoftware nötig. Wenn man von der Lebensdauer eines digitalen Datenträgers spricht, muss dabei stets auch die Verfügbarkeit der entsprechenden Nutzungsumgebung (Lesegerät und Betriebssystem mit Treibersoftware) im Auge behalten werden. Eine CD-ROM ohne Laufwerk enthält verlorene Daten, selbst wenn die CD-ROM völlig intakt ist.

Die wichtigsten digitalen Speichermedien

In den folgenden Kapiteln werden die drei wichtigsten digitalen Speichermedien, nämlich Festplatte, Magnetbänder und optische Medien vorgestellt. Die genannten Trägermedien lassen sich in zwei Gruppen einteilen: magnetische Medien wie Festplatten und Magnetbänder und optische Medien wie CD-ROM oder DVD. Eine andere mögliche Gruppierung unterscheidet nach Online- und Offline-Speicher. Festplatten werden als Online-Speicher bezeichnet, da sie in der Regel konstant eingeschaltet und für den Zugriff verfügbar sind, Offline-Speichermedien (Magnetbänder, CD-ROM, DVD) dagegen werden nur im Bedarfsfall in ein Laufwerk eingelegt und ausgelesen. Eine dritte mögliche Einteilung der drei Medientypen trennt Medien mit Direktzugriff von so genannten sequentiellen Medien. Beim Direktzugriff kann ein Schreib-/Lesekopf direkt über der gesuchten Stelle positioniert werden. Beim sequentiellen Zugriff muss einer Schreib-/Lesespur gefolgt werden, bis der relevante Abschnitt erreicht wurde. Festplatten arbeiten mit Direktzugriff. Magnetbänder sind dagegen sequentielle Medien. Durch die Online-Verfügbarkeit und den Direktzugriff ist die Festplatte nach wie vor das schnellste der drei gängigen Speichermedien. Dafür ist sie derzeit noch das verschleißanfälligste und teuerste Speichermedium.⁴ Die genannten Medientypen werden oft in Kombination eingesetzt. Dabei werden die Medien so angeordnet, dass teure und performante Medien, zumeist Festplatten, Daten mit hoher Zugriffshäufigkeit vorhalten, weniger oft angeforderte Daten dagegen auf preiswerte Offline-Medien ausgelagert werden. Eine solche Anordnung von Speichermedien wird auch als „Hierarchisches Speichermanagement“ (HSM) bezeichnet. Eine entsprechende Empfehlung findet

4 Ob Festplatten immer noch teurer sind als Bandspeicher ist eine derzeit viel diskutierte Frage. Eine interessante Untersuchung findet sich in: McAdam, Dianne (2005): Is Tape Really Cheaper Than Disk?. White Paper. Nashua: Data Mobility Group. http://www-03.ibm.com/industries/media/doc/content/bin/DMG_tape_disk.pdf?g_type=pspot [2007, 20. August]

sich in Calimera Guidelines for Digital Preservation:

Strategies for both online and offline storage will be needed. Delivery files in continual use will need to be stored online, on servers. Master files are best stored offline since they are less frequently accessed.⁵

Bei größeren Unternehmen und Rechenzentren werden die unterschiedlichen Speichermedien zu umfangreichen Speichernetzwerken zusammengeschlossen. Die verschiedenen Arten von Speichernetzwerken ermöglichen eine gut skalierbare, redundante Speicherung auf unterschiedlichen Medien. In den meisten Fällen kommen hierfür gängige Backup- oder Spiegelungsmechanismen in lokalen Speichernetzwerken zum Einsatz. Andere Konzepte sehen das Zusammenwirken räumlich weit voneinander entfernter Speicherkomponenten vor. Hierzu gehören auch Peer-to-Peer-Netzwerke, wie sie z.B. von der Open Source Software „Lots of Copies Keep Stuff Safe“ (LOCKSS)⁶ eingesetzt werden. Speichermedien in der Langzeitarchivierung

Die nachstehende Tabelle vergleicht Festplatte, Bandspeicher und Optische Medien hinsichtlich ihrer Eignung für unterschiedliche Archivierungszeiträume.⁷

Anforderung	Disk	Bandspeicher	Optische Medien
Häufiger Zugriff	Y	N	N
Schnelle Zugriffszeit	Y	N	Vielleicht
Kurze Archivierung (< 1 Jahr)	Y	Y	Y
Mittlere Archivierung (< 10 Jahre)	N	Y	Y
Lange Archivierung (< 20 Jahre)	N	Y	Y
Auslagerung	N	Y	Y
Unveränderbar	N	mit WORM Tape	mit WORM Disc

Die Eignung eines Speichermediums hängt von den Nutzungsanforderungen und ggf. seiner Kombination mit anderen Speichermedien ab. In diesem Sinne gibt es kein für die Langzeitarchivierung in besonderer Weise geeignetes Speichermedium. Vielmehr empfiehlt es sich, eine Speicherstrategie aufzustellen, die

5 o.V. (o.J.) *Digital preservation*. Calimera Guidelines. S.6. http://www.calimera.org/Lists/Guidelines%20PDF/Digital_preservation.pdf [2007, 20.August]

6 <http://www.lockss.org/lockss/Home> [2007, 20.August]

7 Arbeitsgemeinschaft für wirtschaftliche Verwaltung e.V. (AWV) (2003): *Speichern, Sichern und Archivieren auf Bandtechnologien. Eine aktuelle Übersicht zu Sicherheit, Haltbarkeit und Beschaffenheit*. Eschborn: AWV-Eigenverlag, S. 45.

den unterschiedlichen Anforderungen der Archivdaten und der durchschnittlichen Lebensdauer der eingesetzten Speichertechniken gerecht werden kann.

Literatur

McAdam, Dianne (2005): Is Tape Really Cheaper Than Disk?. White Paper. Nashua: Data Mobility Group.

http://www-03.ibm.com/industries/media/doc/content/bin/DMG_tape_disk.pdf?g_type=pspot [2007, 20.August]

o.V. (o.J.) Digital preservation. Calimera Guidelines. http://www.calimera.org/Lists/Guidelines%20PDF/Digital_preservation.pdf [2007, 20.August]

Arbeitsgemeinschaft für wirtschaftliche Verwaltung e.V. (AWV) (2003): Speichern, Sichern und Archivieren auf Bandtechnologien. Eine aktuelle Übersicht zu Sicherheit, Haltbarkeit und Beschaffenheit. Eschborn: AWV-Eigenverlag.

11.2.1 Magnetbänder

Dagmar Ulbrich

Abstract

Magnetbänder speichern Daten auf einem entsprechend beschichteten Kunststoffband. Dabei können zwei unterschiedliche Verfahren eingesetzt werden, das Linear-Verfahren oder das Schrägspur-Verfahren. Gängige Bandtechnologien verfügen über Funktionen zur Datenkompression und Kontrollverfahren zur Sicherung der Datenintegrität. Die wichtigsten aktuellen Bandtechnologien werden im Überblick vorgestellt. Als Lesegeräte können Einzellaufwerke, automatische Bandwechsler oder umfangreiche Magnetband-Bibliotheken dienen. Verschleiß der Magnetbänder und damit ihre Lebensdauer hängen von der Nutzungsweise und Laufwerksbeschaffenheit ab und fallen daher unterschiedlich aus. Die Haltbarkeit hängt darüber hinaus von der sachgerechten Lagerung ab. Regelmäßige Fehlerkontrollen und -korrekturen sind für einen zuverlässigen Betrieb erforderlich. Magnetbänder eignen sich für die langfristige Speicherung von Datenobjekten, auf die kein schneller oder häufiger Zugriff erfolgt, oder für zusätzliche Sicherungskopien.

Gliederung

Funktionsweise von Magnetbändern

Übersicht der wichtigsten Bandtechnologien

Einzellaufwerke und Bandbibliotheken

Verschleiß und Lebensdauer von Magnetbändern und Laufwerken

Magnetbänder in der Langzeitarchivierung

Funktionsweise von Magnetbändern

Die Datenspeicherung erfolgt durch Magnetisierung eines entsprechend beschichteten Kunststoffbandes. Dabei können zwei unterschiedliche Verfahren eingesetzt werden: das Linear-Verfahren und das Schrägspur-Verfahren. Beim Linear-Verfahren wird auf parallel über die gesamte Bandlänge verlaufende Spuren nacheinander geschrieben. Dabei wird das Band bis zum Ende einer Spur in eine Richtung unter dem Magnetkopf vorbeibewegt. Ist das Ende des Bandes erreicht, ändert sich die Richtung, und die nächste Spur wird bearbeitet. Dieses Verfahren wird auch lineare Serpentinenaufzeichnung genannt. Beim Schrägspur-Verfahren (Helical Scan) dagegen verlaufen die Spuren nicht parallel zum

Band, sondern schräg von einer Kante zur anderen. Der rotierende Magnetkopf steht bei diesem Verfahren schräg zum Band. Die wichtigsten Bandtechnologien, die auf dem Linear-Verfahren beruhen, sind „**L**inear **T**ape **O**pen“ (LTO), „**D**igital **L**inear **T**ape (DLT), die Nachfolgetechnologie Super-DLT und „**A**dvanced **D**igital **R**ecording“ (ADR). Für das Schrägspurverfahren können als wichtigste Vertreter „**A**dvanced Intelligent Tape“ (AIT), Mammoth-Tapes, „**D**igital **A**udio **T**apes“ (DAT) und „**D**igital **T**ape **F**ormat“ (DTF) genannt werden. Die jeweiligen Technologien nutzen verschiedene Bandbreiten. Gängige Bandformate sind 4 mm, 8 mm, 1/4 Zoll (6,2 mm) und 1/2 Zoll (12,5 mm). Die Kapazitäten liegen im Gigabyte-Bereich mit aktuellen Maximalwerten bei bis zu 1,6 Terabyte (LTO4, mit Datenkompression). Ebenso wie die Bandkapazität hat sich auch die erreichbare Transferrate in den letzten Jahren stark erhöht. Die meisten Bandtechnologien nutzen Datenkompressionsverfahren, um die Kapazität und die Geschwindigkeit zusätzlich zu steigern. Diese Entwicklung wird durch den Konkurrenzdruck immer preiswerteren Festplattenspeichers gefördert. Zur Sicherung der Datenintegrität verfügen die meisten Bandtechnologien über Kontrollverfahren, die sowohl beim Schreiben als auch bei jedem Lesezugriff eingesetzt werden.

Übersicht der wichtigsten Bandtechnologien

Die nachstehende Tabelle listet die oben genannten Technologien im Überblick.⁸ Es wurden bewusst auch auslaufende Technologien in die Tabelle aufgenommen (ADR, DTF). Das hat drei Gründe: Erstens werden diese Technologien noch vielerorts eingesetzt, zweitens erlauben die älteren Angaben eine anschauliche Darstellung des Kapazitäts- und Performance-Wachstums in den letzten Jahren und drittens zeigt sich hier, wie schnell Bandtechnologien veralten und vom Markt verschwinden, auch wenn die Medien selbst eine wesentlich längere Lebensdauer haben.

8 Die Tabelle wurde entnommen und modifiziert aus: Arbeitsgemeinschaft für wirtschaftliche Verwaltung e.V. (AWV) (2003): *Speichern, Sichern und Archivieren auf Bandtechnologien. Eine aktuelle Übersicht zu Sicherheit, Haltbarkeit und Beschaffenheit*. Eschborn: AWV-Eigenverlag, S. 71. Wo erforderlich, sind die Angaben über die Webseiten der Hersteller aktualisiert worden.

Tech- nologie	Aktuelle Version	Kapazi- tät ohne Kom- pressi- on	Trans- ferrate (MB/ sec)	Ver- fahren	Band- for- mat	Weiterführen- de Informationen [18.08.2007]
ADR	ADR 2	60 GB	4	Linear	8 mm	www.speichergui- de.de
AIT	AIT-4	200 GB	24	Helical Scan	8 mm	www.aittape.com
DAT	DAT-72	36 GB	6	Helical Scan	4 mm	www.datmgm. com
DLT	DLT-V4	160 GB	10	Linear	½ Zoll	www.dlttape.com
DTF	DTF-2	200 GB	24	Helical Scan	½ Zoll	www.speichergui- de.de
LTO-UI- trium	LTO-4	8400 GB	160	Linear	½ Zoll	www.lto.org
Mam- moth	M2	40 GB	12	Helical Scan	8 mm	www.speichergui- de.de
S-DLT	SDLT 600A	300 GB	36	Linear	½ Zoll	www.dlttape.com

Zu ADR siehe 9. Zu DTF siehe 10. Zu Mammoth siehe 11.

Einzellaufwerke und Bandbibliotheken

Magnetbänder werden für Schreib- und Lesevorgänge in ihre zugehörigen Bandlaufwerke eingelegt. Bei kleineren Unternehmen werden in der Regel Einzellaufwerke eingesetzt. Sie werden im Bedarfsfall direkt an einen Rechner angeschlossen und das Einlegen des Bandes erfolgt manuell. Bei steigender Datenmenge und Rechnerzahl kommen automatische Bandwechsler zum Einsatz. Diese Erweiterungen können beliebig skalierbar zu umfangreichen Bandroboter-Systemen (Bandbibliotheken) ausgebaut werden, die über eine Vielzahl von Laufwerken und Bandstellplätzen verfügen. Solche Bandbibliotheken erreichen

9 Die Herstellerfirma OnStream hat 2003 Konkurs anmelden müssen, sodass die Fortführung dieser Technologie unklar ist.

10 Die DTF-Technologie wird seit 2004 nicht fortgeführt.

11 Die Herstellerfirma Exabyte wurde 2006 von Tandberg Data übernommen. Seitdem wird das Mammoth-Format nicht weiterentwickelt.

Ausbaustufen im Petabyte-Bereich.

Verschleiß und Lebensdauer von Magnetbändern und Laufwerken

Die Lebensdauer von Magnetbändern wird üblicherweise mit 2 - 30 Jahre angegeben. Die Autoren von „Speichern, Sichern und Archivieren auf Bandtechnologie“ geben sogar eine geschätzte Lebensdauer von mindestens 30 Jahren an: Für die magnetische Datenspeicherung mit einer 50-jährigen Erfahrung im Einsatz als Massenspeicher kann man sicherlich heute mit Rückblick auf die Vergangenheit unter kontrollierten Bedingungen eine Lebensdauerschätzung von mindestens 30 Jahren gewährleisten.¹²

Die große Spannweite der Schätzungen erklärt sich durch die unterschiedlichen Bandtechnologien. Auch äußere Faktoren wie Lagerbedingungen und Nutzungszyklen spielen eine wesentliche Rolle für die Haltbarkeit. Da Magnetbänder stets ein passendes Laufwerk benötigen, hängt ihre Lebensdauer auch von der Verfügbarkeit eines funktionstüchtigen Laufwerks ab. Ein schadhafes Laufwerk kann ein völlig intaktes Band komplett zerstören und somit zu einem Totalverlust der gespeicherten Daten führen. Magnetbänder sollten kühl, trocken und staubfrei gelagert werden. Nach einem Transport oder anderweitiger Zwischenlagerung sollten sie vor Einsatz mind. 24 Stunden akklimatisiert werden. Neben der Lagerung spielt der Einsatzbereich eines Magnetbandes mit der daraus resultierenden Anzahl an Schreib- und Lesevorgängen eine Rolle. Je nach Bandtechnologie und Materialqualität ist der Verschleiß beim Lesen oder Beschreiben eines Tapes unterschiedlich hoch. Auch der Verlauf von Lese- oder Schreibvorgängen beeinflusst die Haltbarkeit der Bänder und Laufwerke. Werden kleine Dateneinheiten im Start-Stopp-Verfahren auf das Magnetband geschrieben, mindert das nicht nur Speicherkapazität und Geschwindigkeit, sondern stellt auch eine wesentlich höhere mechanische Beanspruchung von Bändern und Laufwerken dar. Aus diesem Grund bieten neuere Technologien eine anpassbare Bandgeschwindigkeit (ADR) oder den Einsatz von Zwischenpuffern. Laufwerke die einen ununterbrochenen Datenfluss ermöglichen, werden auch Streamer, die Zugriffsart als Streaming Mode bezeichnet.

Da den Lebensdauerangaben von Herstellern bestimmte Lagerungs- und Nutzungsvoraussetzungen zugrunde liegen, sollte man sich auf diese Angaben nicht ohne weiteres verlassen. Eine regelmäßige Überprüfung der Funktions-tüchtigkeit von Bändern und Laufwerken ist in jedem Fall ratsam. Einige Band-

12 Arbeitsgemeinschaft für wirtschaftliche Verwaltung e.V. (AWV) (2003): *Speichern, Sichern und Archivieren auf Bandtechnologien. Eine aktuelle Übersicht zu Sicherheit, Haltbarkeit und Beschaffenheit.* Eschborn: AWV-Eigenverlag, S.85

technologien bringen Funktionen zur Ermittlung von Fehlerraten bei Lesevorgängen und interne Korrekturmechanismen mit. Aus diesen Angaben können Fehlerstatistiken erstellt werden, die ein rechtzeitiges Auswechseln von Medien und Hardware ermöglichen.

Trotz der verhältnismäßig langen Lebensdauer von Magnetbändern und deren Laufwerken sollte nicht übersehen werden, dass die eingesetzten Technologien oft wesentlich kürzere Lebenszyklen haben. Wie bereits oben aus der Tabelle hervorgeht, verschwinden Hersteller vom Markt oder die Weiterentwicklung einer Produktfamilie wird aus anderen Gründen eingestellt. Zwar wird üblicherweise die Wartung vorhandener Systeme angeboten, oft aber mit zeitlicher Begrenzung. Aber auch bei der Weiterentwicklung einer Produktfamilie ist die Kompatibilität von einer Generation zur nächsten nicht selbstverständlich. Nicht selten können z.B. Laufwerke einer neuen Generation ältere Bänder zwar lesen, aber nicht mehr beschreiben. Das technische Konzept für die Datenarchivierung des Bundesarchivs sieht daher folgendes vor:

Es sollen nur Datenträger verwendet werden, für die internationale Standards gelten, die am Markt eine ausgesprochen weite Verbreitung haben, als haltbar gelten und daher auch in anderen Nationalarchiven und Forschungseinrichtungen eingesetzt werden. Mit diesen Grundsätzen soll das Risiko minimiert werden, dass der gewählte Archiv-Datenträger vom Markt verschwindet bzw. überraschend von einem Hersteller nicht mehr produziert wird und nicht mehr gelesen werden kann, weil die Laufwerke nicht mehr verfügbar sind.¹³

Magnetbänder in der Langzeitarchivierung

Magnetbänder sind durch ihre vergleichsweise lange Haltbarkeit für die Langzeitarchivierung digitaler Datenbestände gut geeignet. Dies gilt allerdings nur dann, wenn die Daten in dem gespeicherten Format lange unverändert aufbewahrt werden sollen und die Zugriffszahlen eher gering ausfallen. Sind hohe Zugriffszahlen zu erwarten oder ein kurzer Formatmigrationszyklus sollten Bänder in Kombination mit schnellen Medien wie Festplatten zum Speichern von Sicherungskopien eingesetzt werden.

13 Rathje, Ulf (2002): *Technisches Konzept für die Datenarchivierung im Bundesarchiv*. In: Der Archivar, H. 2, Jahrgang 55, S.117-120. (Zitat S. 119).

Literatur

- Arbeitsgemeinschaft für wirtschaftliche Verwaltung e.V. (AWV) (2003): Speichern, Sichern und Archivieren auf Bandtechnologien. Eine aktuelle Übersicht zu Sicherheit, Haltbarkeit und Beschaffenheit. Eschborn: AWV-Eigenverlag.
- Rathje, Ulf (2002): Technisches Konzept für die Datenarchivierung im Bundesarchiv. In: Der Archivar, H. 2, Jahrgang 55, S.117-120.

11.2.2 Festplatten

Dagmar Ulbrich

Abstract

Festplatten sind magnetische Speichermedien. Sie speichern Daten mittels eines Schreib-/Lesekopfes, der über drehenden Platten direkt positioniert wird. Die wichtigsten Speicherbusse (S)-ATA, SCSI, SAS und Fibre Channel werden vorgestellt. Festplatten können einzeln oder im Verbund als Speichersubsysteme genutzt werden. Unterschiedliche Speicherkomponenten können komplexe Speichernetzwerke bilden. Die Lebensdauer von Festplatten wird üblicherweise zwischen 3 und 10 Jahren geschätzt. Umgebungseinflüsse wie magnetische Felder, Stöße oder Vibrationen, aber auch Betriebstemperatur und Nutzungszyklen beeinflussen die Haltbarkeit von Festplatten. Festplatten eignen sich für Kurzzeitarchivierung bzw. in Kombination mit anderen Medien zur Verbesserung von Zugriffszeiten. Für eine revisions sichere Archivierung kommen sie in „Content Addressed Storage-Systemen“ zum Einsatz, die über Inhalts-Hashes die Datenauthentizität sicherstellen.

Gliederung

Funktionsweise und Speicherbusse

Einzelfestplatten und Festplattensubsysteme

Ausfallursachen und Lebensdauer von Festplatten

Festplatten in der Langzeitarchivierung

Revisions sichere Archivierung mit Content Addressed Storage-Systemen (CAS)

Funktionsweise und Speicherbusse

Festplatten speichern Daten durch ein magnetisches Aufzeichnungsverfahren. Die Daten werden im direkten Zugriff (random access) von einem positionierbaren Schreib-/Lesekopf auf die rotierenden Plattenoberflächen geschrieben bzw. von dort gelesen. Festplatten können beliebig oft beschrieben und gelesen werden. Die aktuelle Maximalkapazität einer einzelnen Festplatte liegt bei einem Terabyte. Festplatten zeichnen sich gegenüber sequentiellen Medien wie Magnetbändern durch schnellen Zugriff auf die benötigten Informationsblöcke aus. Die Zugriffsgeschwindigkeit einer Festplatte hängt vor allem von der Positionierzeit des Schreib-/Lesekopfes, der Umdrehungsgeschwindigkeit

der Platten und der Übertragungsrate, mit der die Daten von/zur Platte übertragen werden, ab. Die Übertragungsrate wird wesentlich von der Wahl des Speicherbusses, der Anbindung der Festplatte an den Systembus, bestimmt. Die Speicherbusse lassen sich in parallele und serielle Busse unterscheiden. Die Entwicklung paralleler Busse ist rückläufig, da bei zunehmender Übertragungsrate die Synchronisation der Datenflüsse immer schwieriger wird. Die wichtigsten Standards für Speicherbusse sind: „Advanced Technology-Attachment“ (ATA). Dieser ursprünglich parallele Bus wird heute fast ausschließlich seriell als S-ATA eingesetzt. „Small Computer Systems Interface“ (SCSI) wurde ebenfalls ursprünglich als paralleler Bus entwickelt und wird heute vorwiegend seriell als Serial-Attached-SCSI (SAS) betrieben. Dieses Bussystem zeichnet sich durch hohe Übertragungsraten und einfache Konfiguration aus. Fibre Channel¹⁴ (FC) ist ein originär serieller Bus. Er ermöglicht die Hochgeschwindigkeitsübertragung großer Datenmengen und die Verbindung von Speicherkomponenten mit unterschiedlichen Schnittstellen. Er kommt daher hauptsächlich bei größeren Speichersubsystemen oder komplexen Speichernetzwerken zum Einsatz. Festplatten werden häufig nach ihren Schnittstellen als (S-)ATA-, SCSI- oder SAS-Platten bezeichnet. SCSI- oder SAS-Platten bieten schnelle Zugriffszeiten, sind jedoch im Vergleich zu S-ATA-Platten teuer. S-ATA-Platten dienen vorwiegend dem Speichern großer Datenmengen mit weniger hohen Zugriffsanforderungen. Die ursprünglich aus dem Notebook-Umfeld stammende, heute zunehmend aber auch als mobiles Speichermedium z.B. für Backup-Zwecke eingesetzte USB-Platte basiert derzeit intern meist auf einer Platte mit (S-)ATA-Schnittstelle.

Einzelfestplatten und Festplattensubsysteme

Festplatten können intern in PCs oder Servern eingebaut oder auch als extern angeschlossener Datenspeicher eingesetzt werden. Die Kapazität einzelner Platten kann durch ihren Zusammenschluss zu Speichersubsystemen (Disk-Arrays) bis in den Petabyte-Bereich¹⁵ erweitert werden. Solche Speichersubsysteme werden meist als RAID-Systeme bezeichnet. RAID steht für „Redundant Array of Independent“¹⁶ Disks. „Redundant“ weist hier auf den wichtigsten Einsatz-

14 Die Bezeichnung Fibre Channel kann insofern irreführend sein, als dass dieser serielle Speicherbus sowohl mit Glasfaser als auch mittels herkömmlicher Kupferkabel umgesetzt werden kann.

15 Die Bezeichnung Fibre Channel kann insofern irreführend sein, als dass dieser serielle Speicherbus sowohl mit Glasfaser als auch mittels herkömmlicher Kupferkabel umgesetzt werden kann.

16 Da RAID-Systeme die Möglichkeit bieten, auch preiswerte Festplatten mit hoher Ausfallsi-

zweck dieser Systeme hin: Der Zusammenschluss von Einzelplatten dient nicht nur der Kapazitätserweiterung, sondern vorwiegend der verbesserten Ausfallsicherheit und Verfügbarkeit. Die Platten in RAID-Systemen können so konfiguriert werden, dass bei Ausfall einzelner Platten die betroffenen Daten über die verbliebenen Platten im laufenden Betrieb rekonstruiert werden können. In RAID-Systemen kommen üblicherweise SCSI-Platten zum Einsatz. Zunehmend werden aus Kostengründen auch (S-)ATA-Platten eingesetzt, wobei das Subsystem selbst über SCSI oder FC mit dem Speichernetzwerk verbunden wird. Interessant mit Blick auf ihre Langlebigkeit sind die verhältnismäßig neuen MAID-Systeme. MAID steht für „Massive Array of Idle Disks“. Im Unterschied zu herkömmlichen Festplatten-RAIDs sind die Platten dieser Speicher-Arrays nicht konstant drehend, sondern werden nur im Bedarfsfall aktiviert. Dies mindert den Verschleiß ebenso wie Stromverbrauch und Wärmeentwicklung, kann aber zu Einbußen in der Zugriffsgeschwindigkeit führen.

Ausfallursachen und Lebensdauer von Festplatten

Die Lebensdauer von Festplatten wird sehr unterschiedlich eingeschätzt. Zumeist wird eine Lebensdauer zwischen 3 und 10 Jahren angenommen. Es finden sich jedoch auch wesentlich höhere Angaben von bis zu 30 Jahren. In der Regel werden als Haupteinflüsse die Betriebstemperatur und der mechanische Verschleiß angesehen. Die übliche Betriebstemperatur sollte bei 30°-45°C liegen, zu hohe, aber auch sehr niedrige Temperaturen können der Festplatte schaden. Ein mechanischer Verschleiß ist bei allen beweglichen Teilen möglich. So sind die Lager der drehenden Platten und der bewegliche Schreib-/Lesekopf bei hohen Zugriffszahlen verschleißgefährdet. Die Gefahr, dass Platten durch lange Ruhezeiten beschädigt werden („sticky disk“), ist bei modernen Platten deutlich verringert worden. Zwei Risiken sind bei Festplatten besonders ernst zu nehmen, da sie einen Totalverlust der Daten bedeuten können: zum einen der so genannte Head-Crash. Ein Head-Crash bedeutet, dass der Schreib-/Lesekopf die drehenden Platten berührt und dabei die Plattenbeschichtung zerstört. Zum anderen können umgebende Magnetfelder die magnetischen Aufzeichnungen schädigen. Festplatten sollten daher in einer Umgebung aufbewahrt werden, die keine magnetischen Felder aufweist, gleichmäßig temperiert ist und die Platte keinen unnötigen Stößen oder sonstigen physischen Beeinträchtigungen aussetzt. In welchem Maße die unterschiedlichen Einflüsse die Lebensdauer von Festplatten beeinträchtigen, wird üblicherweise durch Extrapolation von Labortests festgelegt. Hieraus resultieren die Herstellerangaben zu Lebensdauer

cherheit zu betreiben, wird das „I“ in RAID auch mit „inexpensive“ übersetzt.

und Garanzzeiten. Die Lebensdauer einer Festplatte wird üblicherweise mit „mean time before failure“ (MTBF) angegeben. Diese Angabe legt die Stunden fest, die eine Platte betrieben werden kann, bevor Fehler zu erwarten sind. Die Betriebsdauer sollte sich jedoch nicht nur an der MTBF ausrichten, da im Produktivbetrieb oft deutliche Abweichungen von diesen Werten feststellbar sind. Es empfiehlt sich stets auch der Einsatz und die Weiterentwicklung von Überwachungssoftware.

Festplatten in der Langzeitarchivierung

Welche Rolle kann ein Medium, dem eine durchschnittliche Lebensdauer von 5 Jahren zugesprochen wird, für die Langzeitarchivierung von digitalen Datenbeständen spielen? Als Trägermedium zur langfristigen Speicherung von Daten sind langlebigere Medien wie Magnetbänder nicht nur aufgrund ihrer Lebensdauer, sondern auch aus Kostengründen in der Regel besser geeignet. Festplatten können aber in zwei möglichen Szenarien auch für Langzeitarchivierungszwecke sinnvoll sein. Zum einen können sie die Zugriffszeiten auf Archivinhalte deutlich verbessern, wenn sie in Kombination mit anderen Medien in einem hierarchischen Speichermanagement eingesetzt werden. Zum anderen können beispielsweise Formatmigrationen schon nach kurzer Zeit für einen Teil der Archivobjekte erforderlich werden. In diesem Fall ist eine langfristige Speicherung der Dateien gar nicht erforderlich, sondern viel eher deren zeitnahes Auslesen und Wiedereinstellen nach erfolgter Formataktualisierung. Die veralteten Originalversionen können dann auf ein langlebiges Medium ausgelagert werden. Für die jeweils aktuellen Versionen jedoch, die möglicherweise einen kurzen Formatmigrationszyklus haben, kann eine Festplatte ein durchaus geeignetes Trägermedium sein.

Revisionssichere Archivierung mit Content Addressed Storage-Systemen (CAS)

In Wirtschaftsunternehmen und im Gesundheitswesen sind die Anforderungen an Archivierungsverfahren oft an die Erfüllung gesetzlicher Auflagen gebunden. Zu diesen Auflagen gehört oft der Nachweis der Datenauthenzität. Eine Möglichkeit, diese geforderte Revisionssicherheit herzustellen, liegt in der Verwendung von Speichermedien, die nicht überschrieben werden können. Hierfür wurde in der Vergangenheit auf WORM-Medien (Write Once Read Many) zurückgegriffen. Heute werden CD-ROM oder DVD bevorzugt. Eine Alternative hierzu stellen so genannte CAS-Systeme auf Festplattenbasis dar. CAS-Systeme nutzen gut skalierbare Festplattenspeicher in Kombination mit internen Servern und einer eigenen Verwaltungssoftware. Das Grundprinzip beruht auf der

Erstellung von Checksummen bzw. Hashes zu jedem eingestellten Inhalt. Über diese Inhalts-Hashes werden die Objekte adressiert. Der Hash-Wert sichert dabei die Authentizität des über ihn adressierten Inhalts. Dieses Verfahren ist an die Verfügbarkeit des CAS-Systems und der Funktionstüchtigkeit der eingesetzten Hardware gebunden. In der Regel können einzelne Komponenten im laufenden Betrieb ausgetauscht und aktualisiert werden.