

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

11.2.2 Festplatte

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/>



11.2.2 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.