

ENERGIESPARMECHANISMEN DER FESTPLATTE

SEMINARARBEIT

VON

GERRIT GLASER

31. MÄRZ 2012

BETREUER:

PROF. DR. THOMAS LUDWIG

TIMO MINARTZ

UNIVERSITÄT HAMBURG

SEMINAR „GREEN IT“

Inhaltsverzeichnis

1	Einleitung	2
2	Hard-Disk-Drive	3
2.1	Aufbau	3
2.2	Speicherplatzaufteilung	4
2.3	Schreib-/Lesekopf	4
2.4	Cache und Interface	5
3	Energieeinsparungsmöglichkeiten	6
3.1	bei der Herstellung	6
3.1.1	Spindelgeschwindigkeit	6
3.1.2	Datenscheibengröße	6
3.1.3	Datendichte	6
3.1.4	Solid-State-Drive	7
3.1.5	Hybridspeicher	8
3.2	im Betrieb	8
3.2.1	Advanced Configuration and Power Interface	8
3.2.2	Schreibcache	8
3.2.3	Automatic Acoustic Management	8
3.2.4	Advanced Power Management	9
3.2.5	Spindown	9
3.2.6	Steuern der Energieoptionen	9
4	Probleme und Fazit	10
4.1	Probleme	10
4.2	Fazit	10
5	Quellen	11
5.1	Literatur	11
5.2	Bilder	11

Kapitel 1

Einleitung

Laut einer Studie der Gartner Inc. stiegen die Ausgaben für Speicherlösungen fast Drei mal so schnell wie die IT-Budgets an sich. Der Prozentuale Anteil am Serverstromverbraucht beläuft sich auf ca. 16%, das ergibt ein großes Einsparpotential. Zudem lohnen sich Reduzierungen des Stromverbrauchs gleich doppelt, da mit der geringeren Wärmeabstrahlung weniger Energie für die Kühlung benötigt wird. Außerdem haben Festplatten nur eine stark begrenzte Haltbarkeit, so das man auch an eine umweltfreundliche Entsorgung denken muss. Aufgrund dessen und der immer weiter steigenden Energiepreise sind die Hersteller zu einem Umdenken gezwungen und Investitionen in Green IT werden immer wichtiger.



Abbildung 1.1: Innensicht auf eine Festplatte

Kapitel 2

Hard-Disk-Drive

2.1 Aufbau

Vereinfacht gesagt besteht eine Festplatte aus Magnetplatten, Schreib- und Leseköpfen, sowie einem Interface zur Kommunikation mit dem Computer. An die Magnetplatten werden besondere mechanische Anforderungen gestellt, einerseits müssen sie bei hohen Temperaturschwankungen und andererseits bei hohen Drehgeschwindigkeiten ihre Form behalten. Zusätzlich dazu muss ihr Material noch sehr wenig Strom leiten und es darf nicht magnetisch sein. Wenn das Plattenmaterial zu leitfähig ist können Wirbelströme entstehen und die Rotation sehr stark abbremmen. Je höher die Drehzahl ist, desto stärker ist die Bremswirkung durch die Wirbelströme, dies ist aber bei Festplatten unerwünscht. Die Magnetplatten werden anschließend mit einer $1\ \mu\text{m}$ dicken magnetischen Oberfläche beschichtet und darauf wird noch eine Kunststoffschicht aufgetragen. Moderne Festplatten haben, je nach Kapazität, zwischen 1 und 4 Datenplatten.

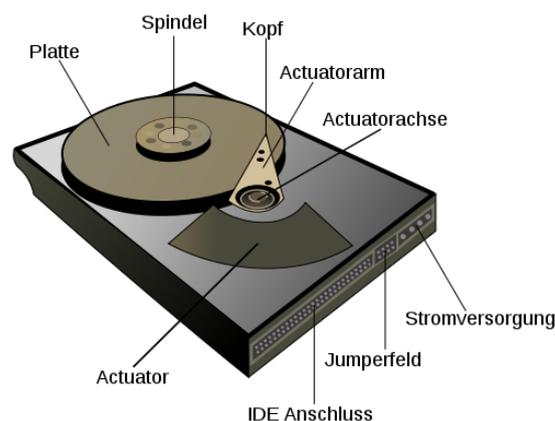


Abbildung 2.1: Schematische Darstellung einer Festplatte

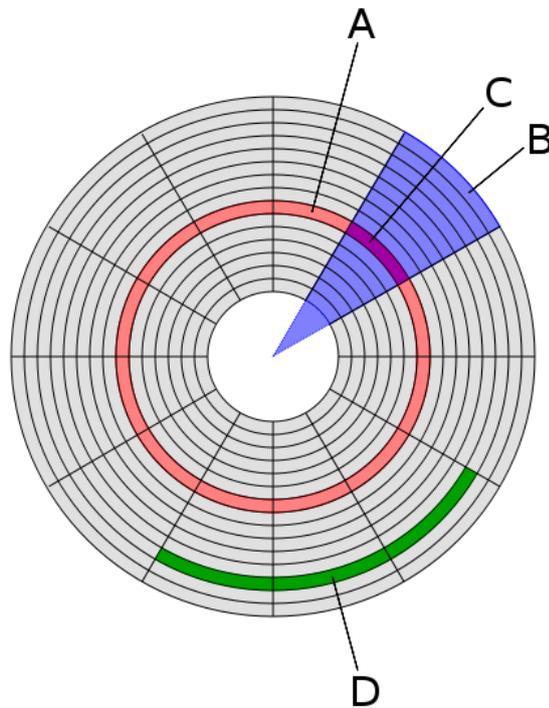


Abbildung 2.2: Schematische Darstellung einer Magnetplatte

2.2 Speicherplatzaufteilung

Auf der Abbildung 2.2 kann man die verschiedenen Bereiche einer Magnetplatte sehen. A ist eine komplette Spur, so wie sie der Magnetkopf abtastet. Moderne Festplatten haben ca. $5,3 \text{ Spuren}/\mu\text{m}$. B ist ein Sektor. Früher wurden die Daten in Sektoren aufgeteilt, dies hat aber einen gravierenden Nachteil. Bei gleich bleibender Drehzahl steigt, je weiter man nach außen kommt, die Umfangsgeschwindigkeit. Deswegen muss man Außen, bei gleicher Abtastrate, eine deutlich niedrigere Datendichte verwenden, genau so wie es bei Schallplatten der Fall ist. Jetzt ist es üblich die Datenplatte in Zonen aufzuteilen, jede Zone enthält gleich viele Blöcke, innen gibt es weniger Zonen als außen, dadurch hat man eine konstante Datendichte. Bei C handelt es sich um einen Block. Heutzutage enthält jeder Block 4096 kB Daten. D ist ein Cluster, ein Verbund mehrerer Blöcke. Cluster sind nicht unbedingt vorhanden, es kommt auf das verwendete Dateisystem an.

2.3 Schreib-/Lesekopf

Die Schreib- und Leseköpfe sind kleine, bewegliche Arme die über den kompletten Radius der Datenscheibe gefahren werden können und so zusammen mit der Rotation der Magnetplatte die komplette Oberfläche abtasten können. Jede Platte hat ihre eigenen Schreib- und Leseköpfe. Der Abstand des Kopfes zur Platte ergibt sich durch ein Luftpolster. Die drehende Platte wirbelt Luft

mit im Kreis, diese Luft staut sich unter dem Kopf auf und drückt ihn nach oben. Dieses Phänomen, das auch z.B. Bei Flugzeugen auftritt, nennt man Bodeneffekt. Aufgrund dieser Tatsache benötigt man mindestens eine bestimmte Spindeldrehzahl, da ansonsten der Kopf die Platte berühren würde. Die Daten werden durch Remanenz, also den Restmagnetismus gespeichert. Wenn man eine Stelle auf der Datenscheibe magnetisiert wird sie hinterher selbst magnetisch, beim entmagnetisieren ist es im Prinzip das selbe, und so kann man Informationen speichern.

2.4 Cache und Interface

Weitere wichtige Komponenten sind das Interface und der Cache. Das Interface ermöglicht die Kommunikation der Festplatte, über eine standardisierte Schnittstelle, mit dem Computer. Der frühere Standard war IDE (Integrated Device Electronics) im Desktopbereich und im Serverbereich SCSI (Small Computer System Interface), mittlerweile sind diese Interfaces veraltet und man verwendet SATA (Serial Advanced Technology Attachment) und im Serverbereich SAS (Serial Attached SCSI). Der Cache wird als Puffer verwendet. Wenn die Festplatte Daten ausliest werden vorsorglich nicht angeforderte Daten zwischengespeichert. Viele Festplatten lesen die Spuren komplett ein und speichern die Daten im Cache, damit diese, falls sie doch benötigt werden, schnell ausgegeben werden können. Dadurch kann man im Idealfall viel Zeit sparen, da bei einer Festplatte das Ausrichten des Kopfes auf die richtige Spur einen großen Teil der Zeit in Anspruch nimmt. Die zuschreibenden Daten werden auch zwischengespeichert und dem System wird gemeldet, dass die Daten geschrieben sind, dadurch muss das Betriebssystem nicht auf die langsame Festplatte warten und kann direkt weitermachen. Anschließend sucht die Festplatte eine freie Stelle und schreibt die Daten hintereinander auf die Platte. Ein großer Cache ist nur bedingt sinnvoll, ab einer gewissen Größe nimmt es mehr Zeit in Anspruch den Cache zu verwalten als man an Zeitersparnis gewinnt.

Kapitel 3

Energieeinsparungsmöglichkeiten

3.1 bei der Herstellung

3.1.1 Spindelgeschwindigkeit

Die Spindelgeschwindigkeit spielt eine entscheidende Rolle beim Stromverbrauch. Das schnellere Drehen der Datenscheiben ermöglicht es Spuren schneller einzulesen, es ergibt sich eine geringere Zugriffszeit und ein deutlich höherer Stromverbrauch. Im Umkehrschluss heißt das, je langsamer die Festplatte dreht, desto weniger Strom verbraucht sie. In der nachfolgenden Tabelle findet man ungefähre Angaben zum Zusammenhang zwischen Spindeldrehzahl, Zugriffszeit und Verbrauch. Wie man sehen kann steigt bei doppelter Drehzahl der Stromverbrauch stark an. Bei der 2,5" Festplatte verdreifacht sich der Verbrauch, wohingegen die Zugriffszeit sich nicht mal halbiert.

Server	2,5"	10.000 1/min	3.0 ms	8.32 W
(Seagate 600gb)	3,5"	15.000 1/min	2.0 ms	16.35 W
Desktop	2,5"	5.400 1/min	5.5 ms	2,5 W
(Western Digital 640gb)	3,5"	7.200 1/min	8.9 ms	8,3 W

3.1.2 Datenscheibengröße

Ein weiter entscheidender Faktor der schon bei der Herstellung berücksichtigt wird ist die Größe der Datenscheiben. Grob gesagt braucht man bei größeren Datenscheiben, durch die höhere Masse, mehr Energie für die Rotation, außerdem müssen vom Kopf zwischen den Spuren längere Wege überwunden werden, dadurch wird die Zugriffszeit erhöht. Im Serverbereich werden oft 2,5" Festplatten verwendet, da diese ein gutes Leistungsaufnahme/Speicherplatz-Verhältnis haben und somit weniger Abwärme erzeugen als der gleiche Speicherplatz mit 3,5" Festplatten.

3.1.3 Datendichte

Eine höhere Datendichte bedeutet bessere Ausnutzung des vorhandenen Platzes, dadurch kann man seine Festplatten kleiner bauen, oder den selben Speicherplatz mit weniger Datenscheiben realisieren. Moderne Festplatten haben

sehr hohe Datendichten mit $34,3 \text{ Bit}/\mu\text{m}$ und $5,3 \text{ Spuren}/\mu\text{m}$ ergibt sich eine Datendichte von $182 \text{ Bit}/\mu\text{m}^2$. Der größte Fortschritt bei der Verbesserung der Datendichte in den letzten Jahren war der Advanced-Format-Technology zu verdanken. Hierbei wurden die Blockgrößen von 512 kB auf 4096 kB angehoben. Jeder Block auf der Festplatte braucht etwas Abstand zum vorherigen Block, eine Synchronisationsmarke und anschließend ein Feld in dem die Adresse steht, erst dann kommt der Datenblock und anschließend noch ein Block mit Fehlerkorrekturdaten. Durch das steigern der Blockgröße auf 4096 kB kann man Sieben Lücken, Synchronisationsmarken und Ad

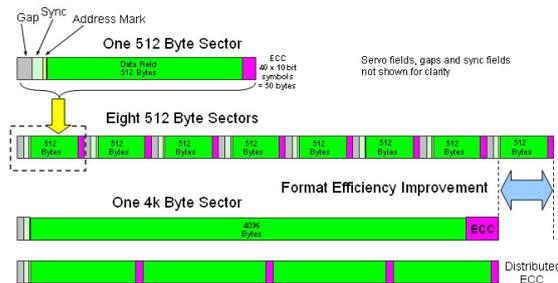


Abbildung 3.1: Advanced Format Technology

3.1.4 Solid-State-Drive

SSD steht für Solid State Drive. Es gibt hauptsächlich zwei Arten von SSDs, zum einem die Flash-Speicher basierten SSDs und zum anderen die SDRAM basierten SSDs. Jede Bauart hat ihre eigenen Vor- und Nachteile. Der größte Vorteil von SSDs ist die sehr kurze Zugriffszeit. Da keine mechanischen Bauteile angesprochen werden müssen ist eine SSD um ein vielfaches schneller als eine HDD, außerdem ist sie erschütterungsresistent, wodurch sie sich perfekt für den mobilen Einsatz eignet. SSDs mit Flashspeicher haben eine begrenzte Anzahl von Schreibzyklen (bis zu 100.000 Schreibzyklen), danach ist der entsprechende Sektor defekt und kann nicht neu beschrieben werden. Deshalb versuchen die SSDs, sofern möglich, alle Zellen relativ gleichmäßig zu belasten. Außerdem werden, wie bei Festplatten auch, Reservespeicherzellen integriert. Dadurch ist es möglich sobald ein Fehler erkannt wurde die Daten umzulagern und weiterhin die volle Kapazität zur Verfügung zu stellen. SSDs mit SDRAM haben den gravierenden Nachteil, dass sie ihre Daten ohne Strom „vergessen“, außerdem haben sie, verglichen mit Flash-SSDs, einen hohen Stromverbrauch. Die Vorteile sind, dass es keine Schreibzyklenbegrenzung gibt und sie zudem noch 80 mal schneller sind. Bei Schreib- und Lesevorgängen ist der Stromverbrauch mit dem von 2,5" HDDs vergleichbar, jedoch gibt es aufgrund der höheren Geschwindigkeit viel längere Idle-Phasen.

3.1.5 Hybridspeicher

Hybridspeicher sind eine Kombination aus HDD und SSD. Sie sammeln die Daten im Flashspeicher und lagern diese, falls der Speicher voll ist auf die HDD aus. Es werden die Vorteile von SSDs und HDDs kombiniert, dadurch ergeben

sich längere Idle-Zeiten, in denen die Datenplatten nicht rotieren müssen, zusammen mit dem großen Speicherplatz den nur HDDs bieten können. Häufig verwendete Daten werden nicht ausgelagert, sondern bleiben im Flash-Speicher. Obwohl dieses Modell gut kling erreicht es nur einen geringen Geschwindigkeitsgewinn von gerade einmal 20%. Jedoch ist der Stromverbrauch deutlich geringer als bei einer konventionellen HDD.

3.2 im Betrieb

3.2.1 Advanced Configuration and Power Interface

ACPI ist ein Standard zur Energieverwaltung durch das Betriebssystem, er wurde von Intel, Microsoft und Toshiba entwickelt und 1996 veröffentlicht und seitdem ständig erweitert. ACPI arbeitet mit Power States. Jedes Gerät kann einen individuellen State haben. Für Festplatten sind die Device States wichtig. Es gibt insgesamt Vier verschiedene Device States: D0 bedeutet, dass das Gerät normal läuft, D1 und D2 sind hardware-spezifische States, das heißt der Festplattenhersteller kann festlegen, welcher State welche Funktion hat und der letzte State ist D3. D3 bedeutet das Gerät ist ausgeschaltet.

3.2.2 Schreibcache

Der Schreibcache wird verwendet, um die Abarbeitungszeit von Schreibbefehlen zu verkürzen. Die zuschreibenden Daten werden im Cache zwischengespeichert und irgendwann auf der Festplatte abgelegt. Dem System wird gesagt, dass die Daten gespeichert wurden, jedoch findet die eigentliche Speicherung verzögert statt, so dass bei einem Stromausfall Daten verloren gehen können. Der Schreibcache wird oft mit einem weiteren Cache im RAM kombiniert, da dieser eine viel größere Kapazität aufweist. Um Strom zu sparen kann man den Hardware-Cache deaktivieren und nur den RAM verwenden.

3.2.3 Automatic Acoustic Management

Das Automatic Acoustic Management (AAM) wurde ursprünglich entwickelt, um die Geräuschentwicklung beim Positionieren des Schreib- und Lesekopfes zu vermindern. Dafür wird die Geschwindigkeit verringert, mit der der Kopf positioniert wird. Dadurch kann man Strom einsparen, jedoch kommt es unter Umständen zu einer verringerten Leistung. Die Genaue Implementation des AAM ist herstellerabhängig. Western Digital nennt ihr System IntelliSeek, hierbei wird die Kopfpositionierungsgeschwindigkeit nur soweit verringert, wie es möglich ist ohne den Anfang des gewünschten Blocks in der nächsten Umdrehung zu verpassen, dadurch entsteht kein Leistungsverlust und es wird trotzdem Strom gespart.

3.2.4 Advanced Power Management

Das Advanced Power Management (APM) regelt die festplatteninternen Energiesparmechanismen, wie z.B. das automatische Verringern der Leistung nach einer bestimmten Idle-Zeit. Dazu gehören unter anderem die Verringerung der Spindeldrehzahl und das Parken des Schreib- und Lesekopfs. Auch hier ist

die tatsächliche Implementation herstellerabhängig. Western Digital nennt ihr APM-System IntelliPower und bei Seagate heißt es PowerChoice. APM am Beispiel von PowerChoice: Beim durchschnittlichem Betrieb einer 1TB HDD mit einer Spindeldrehzahl von 7.200 1/min hat man einen Stromverbrauch von 5,6W. Nach 1s Idle wechselt die HDD in den Idle_A Status, hierbei werden die Schreib- und Lesköpfe angehalten und der Verbrauch sinkt auf 2,82W. Die Festplatte ist sofort wieder einsatzbereit, es wird keine Wiederherstellungszeit benötigt. Nach insgesamt 10min Idle wird in den Status Idle_B gewechselt, der Verbrauch sinkt auf 2,18W, dafür gibt es aber eine Wiederherstellungszeit von 0,5s. Nach 30min Idle wechselt die Festplatte in den Status Idle_C dabei wird die Spindeldrehzahl verringert. Der Verbrauch beträgt nun 1,82W und das Wiederherstellen des vollen Funktionsumfangs dauert 1s. Nach 60min Idle-Zeit schaltet die Platte in den Standby_Z Status, er mach das selbe wie Idle_C, jedoch mit einer noch geringeren Spindeldrehzahl. Der Verbrauch beträgt nur noch 1,29W, dafür muss man aber eine Wiederherstellungszeit von ganzen 8s in Kauf nehmen.

3.2.5 Spindown

Bei Festplatte entspricht der Spindown dem Device State D3. Wenn der Kopf noch nicht in seiner Parkposition ist wird er dort hingefahren und anschließend wird die Festplatte komplett abgeschaltet. Eine Festplatte die in D3 ist benötigt sehr lange um wieder einsatzbereit zu sein, da erst die Datenplatten wieder auf ihre Drehzahl gebracht werden müssen und anschließend kann der Kopf ausgeparkt werden.

3.2.6 Steuern der Energieoptionen

Manuell beeinflussen kann man die Energieoptionen z.B. mit der Anwendung `hdparm`. `hdparm` gibt es für Linux und für Windows, das Programm wurde entwickelt um das Setzen und Auslesen von Laufwerks Parametern zu ermöglichen. Die wichtigsten Befehle sind:

```
hdparm -M für das AAM
hdparm -B für APM
hdparm -S ermöglicht es den Standby-Timeout zu setzen
hdparm -y erzwingt einen Standby
hdparm -Y erzwingt einen Spindown
```

Die tatsächlichen Möglichkeiten die dieses Programm hat sind Hardwareabhängig. Die verwendeten Schnittstellen zum Steuern der Energieoptionen sind bei SATA der `Set-Features`-Befehl, mit ihm kann man die Geräteeinstellungen direkt über das SATA Interface verändern. Bei Laufwerken mit dem SAS Interface verwendet man den `SCSI Mode Select` und das `Start/Stop Unit (SSU) Command`.

Kapitel 4

Probleme und Fazit

4.1 Probleme

Die größten Probleme bei den Festplatten verursachen die mechanischen Bauteile, aufgrund der besonders hohen Belastung der Lager beim Starten der Rotation haben Festplatten eine maximale Anzahl an Start-/Stopzyklen. Viele Spindowns verkürzen die Lebenszeit. In einer Studie von Google die über mehrere Jahre hinweg über 100.000 Festplatten überwachten wurde festgestellt, dass viele Start-/Stopzyklen bei Festplatten mit einem Alter von 2 Jahren keine erkennbaren Auswirkungen haben, jedoch steigt bei Festplatten die 3 Jahre oder älter sind die Ausfallrate um mehr als 2% an. Weitere Probleme sind zum einen, dass Systemplatten oft Schreib- und Lesezugriffe haben und somit nie ein hoher Idle-State erreicht wird und zum anderen die Unberechenbarkeit von Energiesparfunktionen. Gegen die häufigen Schreib- und Lesezugriffe kann man zwei Dinge unternehmen, zum einen kann man oft gelesene Systemdaten in den RAM auslagern und zum anderen kann man die Daten, bevor man sie auf die Festplatte speichert, in einem Schreibpuffer sammeln. Gegen die Unberechenbarkeit kann man momentan leider noch nicht viel unternehmen, so dass in kritischen Bereichen oft auf die Energiesparfunktionen verzichtet werden muss.

4.2 Fazit

Energie sparen bei Festplatten beginnt schon bei der Herstellung und geht weiter mit der Auswahl der richtigen Festplatte. Eine Serverfestplatte mit 15.000 1/min ist im Desktop-Einsatz nur bedingt sinnvoll, da ihr Vorteil mit den kurzen Zugriffszeiten nur sehr selten zum Tragen kommt, aber immer viel Strom verbraucht wird. Außerdem sollte man grundsätzlich Energiesparmechanismen einsetzen, wenn es die Gegebenheiten zulassen. Natürlich gibt es auch Ausnahmen in denen eine Verzögerung nicht tragbar oder nicht sinnvoll ist. Wenn man beim Aufruf einer Website eine Verzögerung von mehr als 8s hätte, nur weil der Server die Festplatte wiederherstellen muss, dann sollte man überlegen, auf diesen Energiesparmechanismus zu verzichten.

Kapitel 5

Quellen

5.1 Literatur

- http://www.seagate.com/docs/pdf/whitepaper/tp608_powerchoice_tech_provides.pdf
- (03.12.2011)
- <http://www.wdc.com/global/products/demo.aspx?typeID=1&nameid=intelliseek>
- (25.11.2011)
- <http://de.wikipedia.org/wiki/Festplatte>
- (25.11.2011)
- <http://www.wdc.com/wdproducts/library/SpecSheet/ENG/2879-771379.pdf>
- (29.11.2011)
- <http://de.wikipedia.org/wiki/Festplattengeometrie>
- (25.11.2011)
- <http://de.wikipedia.org/wiki/ACPI>
- (02.12.2011)
- http://de.wikipedia.org/wiki/Automatic_acoustic_management
- (27.11.2011)
- http://research.google.com/en//archive/disk_failures.pdf
- (05.12.2011)

5.2 Bilder

- 1.1 http://de.wikipedia.org/w/index.php?title=Datei:Samsung_HD753LJ_03-Opened.jpg
- (25.11.2011)
- 2.1 http://de.wikipedia.org/w/index.php?title=Datei:Hard_drive-de.svg
- (25.11.2011)
- 2.2 <http://de.wikipedia.org/w/index.php?title=Datei:Disk-structure2.svg>
- (27.11.2011)
- 3.1 <http://en.wikipedia.org/wiki/File:Afdia1.jpg>
- (25.11.2011)