

Betriebssysteme

Tutorium 12

Philipp Kirchhofer

`philipp.kirchhofer@student.kit.edu`

`http://www.stud.uni-karlsruhe.de/~uxbtt/`

**Lehrstuhl Systemarchitektur
Universität Karlsruhe (TH)**

3. Februar 2010

Was machen wir heute?

1 Organisatorisches

2 Tutorien Übungsblatt

- Hard Disks
- Disk Scheduling
- Swap Space Management
- RAID
- Device Drivers

Programmieraufgaben Vorstellung

Montag, den 8. Februar, 14:00 Uhr bis 15:30 Uhr

Poolraum G im Rechenzentrum (Gebäude 20.21)

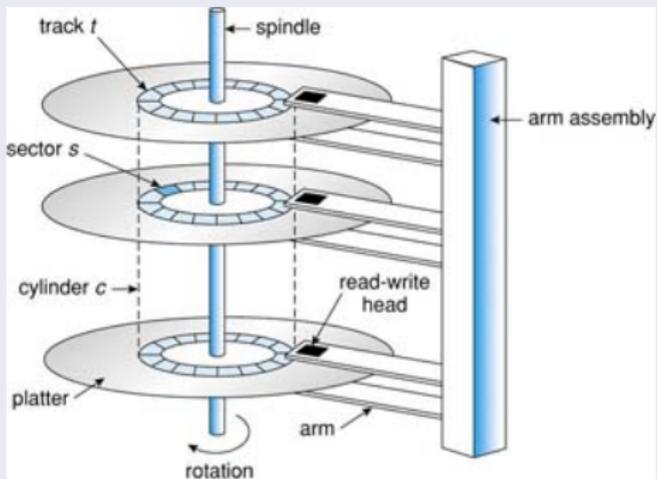
Lösung der Syscall Aufgabe sowie das Design-Dokument vorstellen.

Wichtig: Beide Gruppenteilnehmer müssen vorstellen!

Frage 12.1.a

Was sind Zylinder, Spuren und Sektoren?

Antwort



Frage 12.1.b

Eine Festplatte rotiert mit 7200 Umdrehungen pro Minute, hat eine Sektorgröße von 512 Byte und 160 Sektoren pro Spur. Wie groß ist die maximale Lesegeschwindigkeit der Festplatte?

Antwort

In einer Spur werden $512 * 160 = 81920$ Byte gespeichert.

Bei 7200 Umdrehungen pro Minute können $7200/60 = 120$ Spuren pro Sekunde gelesen werden.

Dies ergibt eine maximale Lesegeschwindigkeit von $120 * 81920 = 9600$ KB/s.

Frage 12.1.c

Was passiert bei „Sector sparing“? Welche Probleme können dadurch ausgelöst werden?

Antwort

„Sector sparing“ ist eine Methode, um defekte Sektoren ausblenden zu können. Eine Festplatte enthält Reservesektoren, die für das Betriebssystem nicht zugänglich sind. Nach der Erkennung eines defekten Sektors wird der betroffene Sektor durch einen Reservesektor ausgetauscht, d.h. alle Lese-/Schreibanforderungen an den defekten Sektor werden auf den neuen Sektor umgeleitet.

„Sector sparing“ kann zu Problemen beim Scheduling von Lese-/Schreibanforderungen führen, da aufeinanderfolgende Sektoren nicht mehr aufeinanderfolgend auf der Festplatte liegen müssen.

Frage 12.1.d

Was ist Sector slipping?

Antwort

„Sector slipping“ ist eine weitere Methode, um mit defekten Sektoren umgehen zu können. Bei „Sector slipping“ werden alle auf den defekten Sektor folgenden Sektoren und der defekte Sektor selber um einen Sektor nach hinten umgeleitet.

Frage 12.2

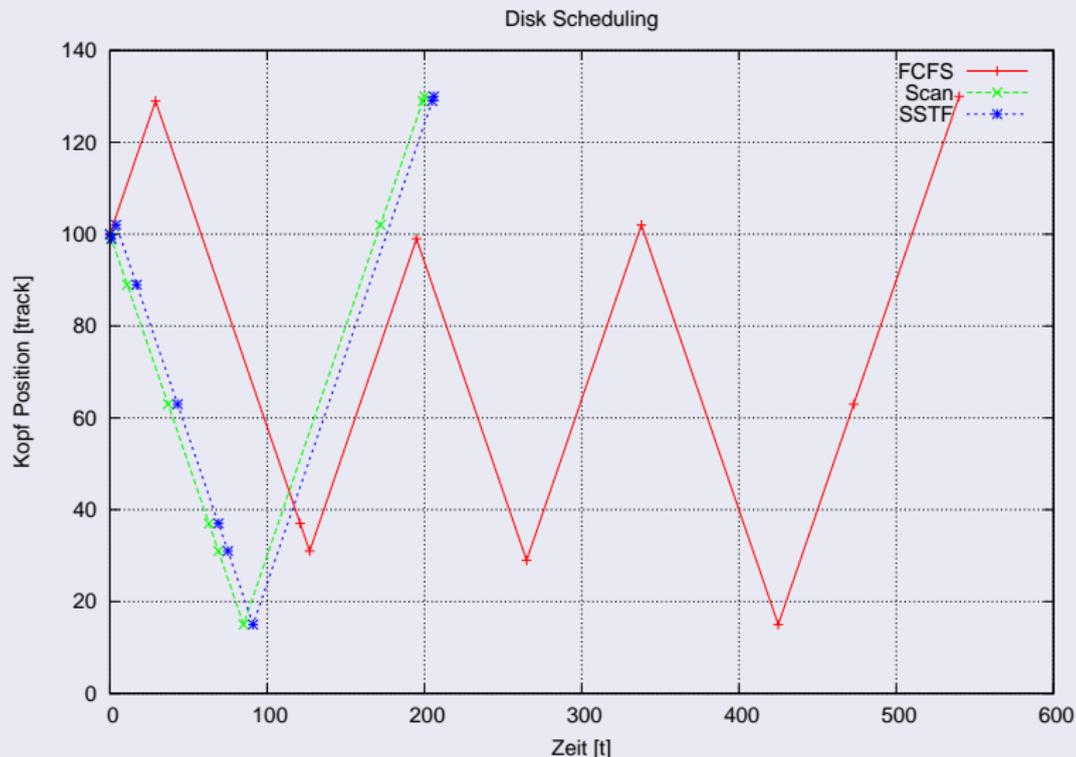
Ein Betriebssystem erhält eine Leseanfrage für die Spuren 129, 37, 31, 99, 89, 102, 15, 63, 130. Der Festplattenkopf steht am Anfang auf Spur 100 und bewegt sich in Richtung Spur 0.

Berechne die durchschnittliche Kopfbewegung für die Verfahren „FCFS“, „Scan“, „SSTF“. Welches Verfahren benötigt am wenigsten Zeit für die Bearbeitung der Anfrage?

Tutorien Übungsblatt - Disk Scheduling

FCFS		Scan		SSTF	
Spur	Delta	Spur	Delta	Spur	Delta
100		100		100	
129	29	99	1	99	1
37	92	89	10	102	3
31	6	63	26	89	13
99	68	37	26	63	26
89	10	31	6	37	26
102	13	15	16	31	6
15	87	102	87	15	16
63	48	129	27	129	114
130	67	130	1	130	1
Durchschnitt:	46.67		22.22		22.89

Visualisierung der Verfahren



Frage 12.3.a

Bereiche für das Swapping können sowohl in einer Datei als auch auf einer eigenen Partition liegen. Welche Vor- und Nachteile haben beide Vorgehensweisen?

Antwort

Swap Datei

- ⊕ Einfache Implementierung
- ⊕ Kann bei Bedarf größer und kleiner werden
- ⊖ Höherer Overhead durch VFS und Inode Lookup
- ⊖ Fragmentierung möglich

Swap Partition

- ⊕ Kein Dateisystem nötig
- ⊕ Optimierung der Datenspeicherung möglich
- ⊖ Feste Größe

Frage 12.3.b

Was ist anonymer Speicher?

Antwort

Anonymer Speicher ist Speicher im Heap, Stack oder nicht initialisierter Speicher eines Prozesses.

Frage 12.3.c

Weshalb werden anonymer und nicht anonymer Speicher beim Swapping üblicherweise unterschiedlich behandelt?

Antwort

Nicht anonymer Speicher ist Speicher, der mit einer bestimmten Datei verknüpft ist. Beim Auslagern einer Seite mit nicht anonymen Speicher kann die Speicherseite verworfen werden (z.B. Programmcode) und bei Bedarf in die richtige Datei zurückgeschrieben werden (memory mapped file). Wenn die Seite neu angefordert wird kann sie einfach aus der zugehörigen Datei neu geladen werden.

Anonymer Speicher dagegen muss ausgelagert werden, da er nicht mit einer Datei verknüpft ist. Beim Auslagern einer Seite mit anonymen Speicher wird also die Seite in den Swap Bereich ausgelagert und bei einem späteren Zugriff wieder aus diesem eingelagert.

SLED/RAID

- SLED: Single Large Expensive Disk
- RAID: Redundant array of Independent/Inexpensive Disks

RAID0: Block Striping

Jeder Datenblock wird auf eine von n Festplatten abgebildet (Striping).
Gesamtkapazität: Summe der Festplattengrößen

- ⊕ Erhöhter Schreib- und Lesedurchsatz durch Parallelität
- ⊖ Erhöhte Datenverlustwahrscheinlichkeit

RAID1: Mirroring

Jeder Datenblock wird auf 2 Festplatten geschrieben.
Zum Lesen reicht eine der beteiligten Festplatten aus.
Gesamtkapazität: Größe der kleinsten Festplatte

- ⊕ Erhöhter Lesedurchsatz
- ⊕ Erhöhte Datensicherheit bei Festplattenausfall
- ⊖ Geringe Kapazität

RAID2: Bit Striping + ECC

Striping auf Bit Ebene, Einsatz eines ECC (Hamming Code).
Gesamtkapazität abhängig von ECC Wortlänge.
Nicht mehr eingesetzt.

- ⊕ Erhöhte Datensicherheit
- ⊖ Geringerer Durchsatz

RAID3: Bit Striping + Parity

Jedes Datenbit wird auf eine von $n - 1$ Festplatten abgebildet.

Die letzte Festplatte speichert ein dazugehöriges Paritätsbit.

Gesamtkapazität: $n - 1 * \text{Festplattengröße}$

Nicht mehr verwendet.

- ⊕ Erhöhter Lesedurchsatz
- ⊖ Festplatten müssen synchron laufen
- ⊖ Paritätsfestplatte wird bei jedem Schreibzugriff beschrieben

RAID4: Block Striping + Parity

Jeder Datenblock wird auf eine von $n - 1$ Festplatten abgebildet.

Die letzte Festplatte speichert einen dazugehörigen Paritätsblock.

Gesamtkapazität: $n - 1 * \text{Festplattengröße}$

Nicht mehr verwendet.

- ⊕ Erhöhter Lesedurchsatz
- ⊖ Paritätsfestplatte wird bei jedem Schreibzugriff beschrieben

RAID5: Block Striping + Distributed Parity

Ähnlich wie RAID4, der Paritätsblock zu einem Datenblock wechselt allerdings zwischen den Festplatten.

Gesamtkapazität: $n - 1 * \text{Festplattengröße}$

- ⊕ Erhöhter Lesedurchsatz
- ⊕ Gute Gesamtkapazität

Es sollen die RAID Level 0 bis 5 miteinander verglichen werden.
Die nutzbare Datenkapazität soll $4 * \text{Festplattengröße}$ betragen.
RAID 2 soll 3 Paritätsbits verwenden.

Frage 12.4.a

Wieviele Festplatten werden benötigt?

Frage 12.4.b

Wieviele Blöcke müssen für das Schreiben eines Bytes gelesen und geschrieben werden?

Frage 12.4.c

Wie können Daten bei dem Ausfall einer Festplatte rekonstruiert werden?

Antwort

	a.	b.	c.
SLED	1	1 lesen, 1 schreiben	Wiederherstellung nicht möglich
RAID 0	4	1 lesen, 1 schreiben	Wiederherstellung nicht möglich
RAID 1	8	1 lesen, 2 schreiben	Daten von Spiegelfestplatte lesen
RAID 2	7	4 lesen, 7 schreiben	Mit Hamming rekonstruieren
RAID 3	5	4 lesen, 5 schreiben	Mit Parity rekonstruieren
RAID 4	5	2 lesen, 2 schreiben #	Mit Parity rekonstruieren
RAID 5	5	2 lesen, 2 schreiben #	Mit Parity rekonstruieren

jeweils Daten + Parity

Frage 12.5.a

Viele Betriebssysteme deaktivieren Interrupts nur bei der Ausführung von wenigen kritischen Funktionen.

Angenommen das Betriebssystem führt gerade eine Methode f aus und es wird ein Interrupt ausgelöst. Kann während der Interruptbehandlung in jedem Fall die Methode f erneut aufgerufen werden?

Antwort

Um eine erneute Ausführung der Methode zu gewährleisten muss die Methode reentrant sein. Eine Funktion, die globale oder statische Variablen verwendet ist nicht reentrant. In diesem Fall kann ein zweiter Aufruf den Programmzustand verändern, so dass die Änderungen auch für den ersten Funktionsaufruf sichtbar sind.

Frage 12.5.b

Viele moderne Betriebssysteme teilen die Behandlung von Interrupts in zwei Phasen auf: eine Phase mit hoher Priorität und eine weitere Phase mit niedriger Priorität, die auch später ausgeführt werden kann. Was soll durch dieses Design erreicht werden?

Linux: „Top halves“, „Bottom halves“

Windows: „Deferred Procedure Call“ (DPC)

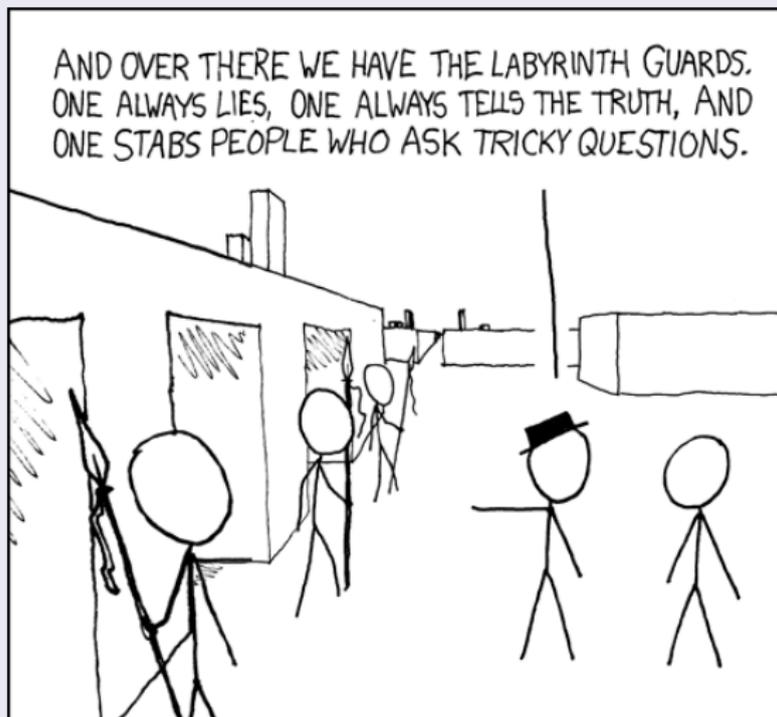
Antwort

Interrupts sollten möglichst schnell behandelt werden:

- Zeit, die ein Programm unterbrochen ist, soll möglichst kurz sein
- Das Wiedereinschalten von Interrupts soll möglichst schnell geschehen
- Das Gerät, das den Interrupt ausgelöst hat, soll möglichst schnell weiterarbeiten können

Fragen & Kommentare?

Nächste Woche Wiederholungsstunde
Themen bitte vorher per Mail an mich schicken



xkcd: Labyrinth Puzzle