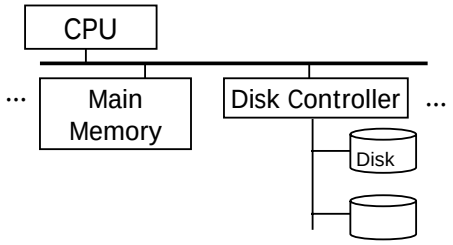
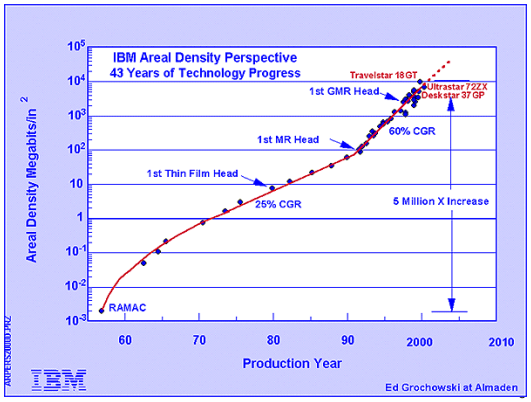
CS143: Disks and Files

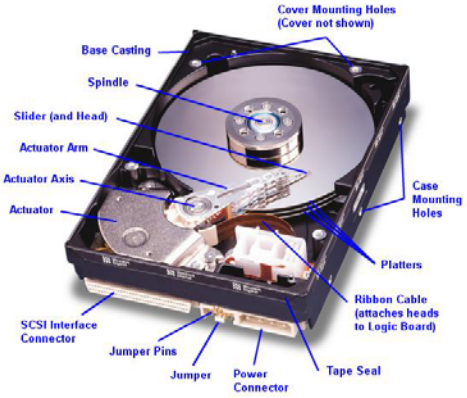
1

System Architecture

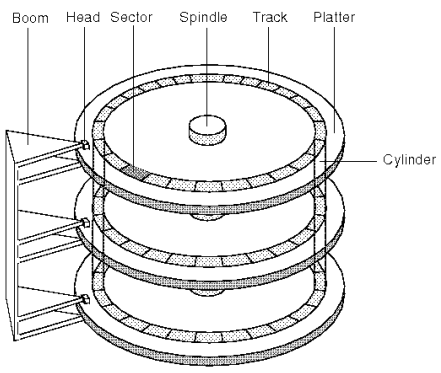


2



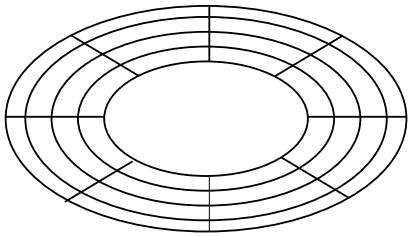


4



5

Structure of a Platter



- Track, cylinder, sector (=block)

6

A Typical Disk

- Platter diameter: 1-5 in
- Cylinders: 100 – 2000
- Platters: 1 – 20
- Sectors per track: 200 – 500
- Sector size: 512 – 50K
- Rotation speed: 1000 – 15000 rpm
- Overall capacity: 1G – 300GB

- Q: 2 platters, 500 tracks/platter, 200 sect/track, 1KB/sector. What is the overall capacity?

7

Access Time

- Q: What needs to be done to read data from a particular section?

8

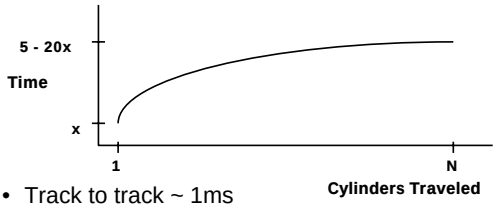
Access Time

- Access time =
  (seek time) + (rotational delay) + (transfer time)

9

Seek Time

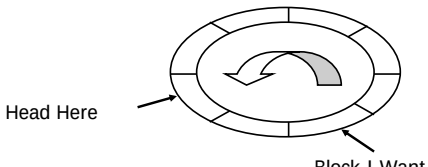
- Time to move a disk head between tracks



- Track to track ~ 1ms
- Average ~ 10 ms
- Full stroke ~ 20 ms

10

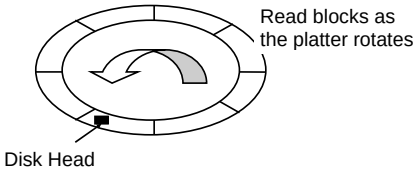
Rotational Delay



- Typical disk:
  - 1000 rpm – 15000 rpm
- Q: For 6000 RPM, average rotational delay?

11

Transfer Rate



6000 RPM, 200 sectors/track

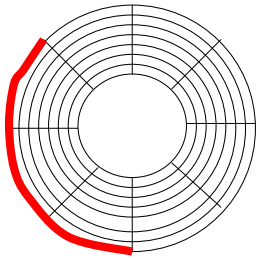
- Q: How long to read one block?

- Q: What is the transfer rate (bytes/sec)?

12

Sequential vs. Random I/O

- Q: How long to read 3 sequential blocks?

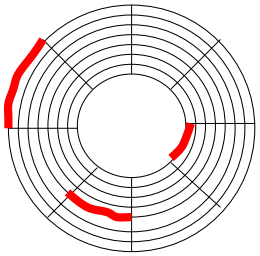


- ❑ 6000 RPM
- ❑ 200 sectors/track
- ❑ Assume the head is above the first block

13

Sequential vs. Random I/O

- Q: How long to read 3 random blocks?



- ❑ 6000 RPM
- ❑ 200 sectors/track
- ❑ 10ms seek time
- ❑ Assume the head is above the first block

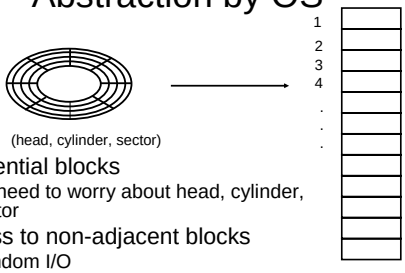
14

Data Modification

- Byte-level modification not allowed
  - Can be modified by blocks
- Q: How can we modify only a part of a block?

15

Abstraction by OS



(head, cylinder, sector)

- Sequential blocks
  - No need to worry about head, cylinder, sector
- Access to non-adjacent blocks
  - Random I/O
- Access to adjacent blocks
  - Sequential I/O

16

Buffers, Buffer pool

- Temporary main-memory “cache” for disk blocks
  - Avoid future read
  - Hide disk latency
  - Most DBMS let users change buffer pool size

17

Reference

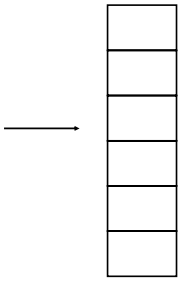
- Storage review disk guide
  - http://www.storagereview.com/guide2000/ref/hdd/index.html

18

Files: Main Problem

- How to store tables into disks?

Jane	CS	3.7
Susan	ME	1.8
June	EE	2.6
Tony	CS	3.1



19

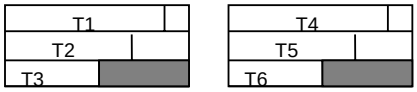
Spanned vs Unspanned

- Q: 512Byte block. 80Byte tuple. How to store?

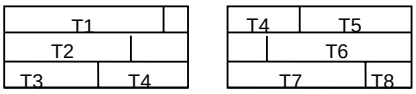
20

Spanned vs Unspanned

- Unspanned



- Spanned

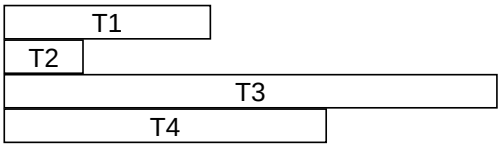


- Q: Maximum space waste for unspanned?

21

Variable-Length Tuples

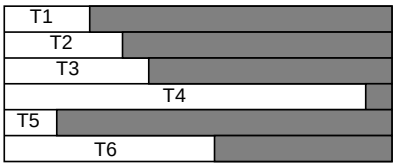
- How do we store them?



22

Reserved Space

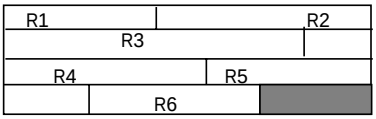
- Reserve the maximum space for each tuple



- Q: Any problem?

23

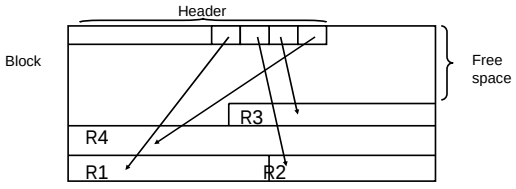
Variable-Length Space



- For each tuple
  - Tuple length in the beginning
  - End-of-record symbol (⊥)
- Pack tuples tightly
- Q: What to do for delete/update?

24

Slotted Page



25

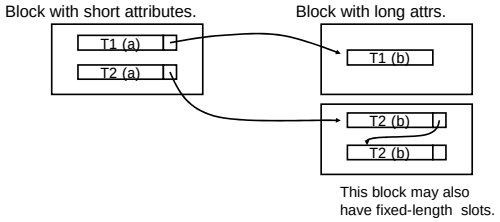
Long Tuples

- ProductReview(
  pid INT,
  reviewer VARCHAR(50),
  date DATE,
  rating INT,
  comments VARCHAR(1000))
- Block size 512B
- How should we store it?

26

Long Tuples

- Spanning
- Splitting tuples



27

Sequential File

- Tuples are ordered by some attributes (search key)

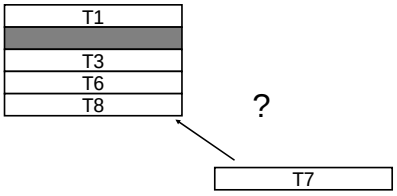
Elaine	CS	3.7
James	ME	2.8
John	EE	1.8
Peter	EE	3.9
Susan	CS	1.0
Tony	EE	2.4

– Search key: Name

28

Sequencing Tuples

- Inserting a new tuple
  – Easy case

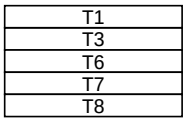


29

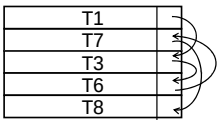
Sequencing Tuples

Two options

1) Rearrange



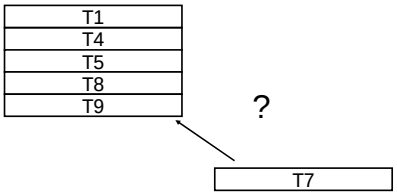
2) Linked list



30

Sequencing Tuples

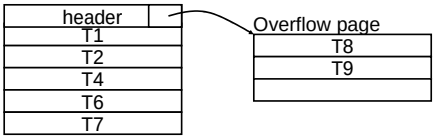
- Inserting a new tuple
  - Difficult case



31

Sequencing Tuples

- Overflow page



- Reserving free space to avoid overflow
  - PCTFREE in DBMS

    CREATE TABLE R(a int) PCTFREE 40

32

Things to Remember

- Disk
  - Platter, track, cylinder, sector, block
  - Seek time, rotational delay, transfer time
  - Random I/O vs Sequential I/O
- Files
  - Spanned/unspanned tuples
  - Variable-length tuples (slotted page)
  - Long tuples
  - Sequential file and search key
    - Problems with insertion (overflow page)
    - PCTFREE

33