Grundzüge der Wirtschaftsinformatik
Introduction to Business Information Systems

Unit 3

Prof. Dr. Martin Hepp
http://www.heppnetz.de
mhepp@computer.org

http://www.heppnetz.de/teaching/gwi/

Logistics

- **Lecture**
  - Tuesdays, 13:15 - 14:45, Auditorium Maximum (Building 33)
- **Tutorial and Exercises** **(Begin: October 17)**
  - Wednesdays, 11:30 – 13:00, Building 33 Room 2401 (in German)
  - Thursdays, 09:45 - 11:15, Building 43 Room 4/126 (in German)
  - Thursdays, 15:00 - 16:30, Building 33 Room 2216 (in German)
  - Thursdays, 16:45 - 18:15, Building 33 Room 2116 (in German)
- **Exam (tentatively, to be confirmed)**
  - ~~December 18, 2007 – 13:00 – 15:00~~

-> Waiting for feedback on available classrooms...

Structure of the Lecture

Unit 1: Introduction
Unit 2: Central Processing Units
Unit 3: Storage and Data Structures
Unit 4: Input and Output Devices
Unit 5: Software
Unit 6: Networks, Data Interchange, and the Internet
Unit 7: Design, Development, Deployment, and Operations of Information Systems
Unit 8: Office Applications
Unit 9: Enterprise Applications
Unit 10: Supply Chain Applications and E-Business
Unit 11: Management Support Systems
Unit 12: Exam Review

Link to the Previous Unit

- Last Unit: Central Processing Units – how does a computer execute a program? Main memory (RAM) – temporary storage; loses content when power supply is interrupted.
- Today: How can data be stored persistenly? How do we make sure that a fact can be found when needed?

Assignment from Last Week

- WI2: pp. 93-211, 387-515
- IBIS, pp. 55-78
- Review the slides

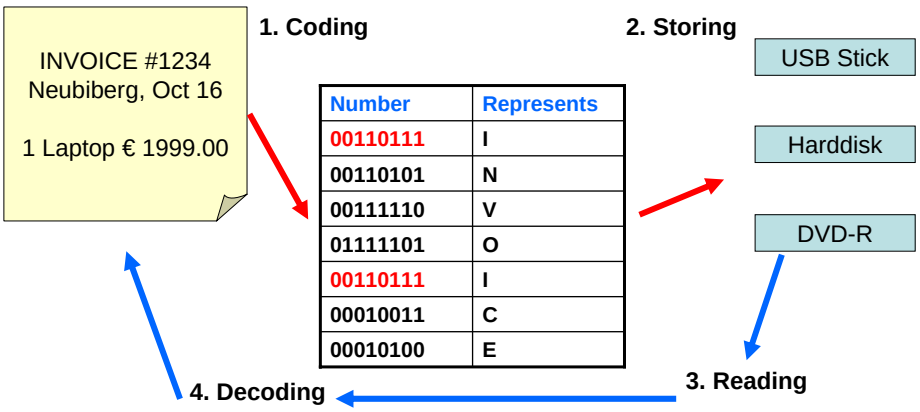
WI1 = Hansen/Neumann: Wirtschaftsinformatik 1; WI2 = Hansen/Neumann: Wirtschaftsinformatik 2; IBIS = Wigand et al: Introduction to Business Information Systems.

Storage and Data Structures

If we want to store information in a computer system, we need to

- Develop a **representation structure** and
- Have a **media** to write to and read from

Storage and Data Structures



Number	Represents
00110111	I
00110101	N
00111110	V
01111101	O
00110111	I
00010011	C
00010100	E

All information must be converted into
a sequence of numbers between 0 and 255.

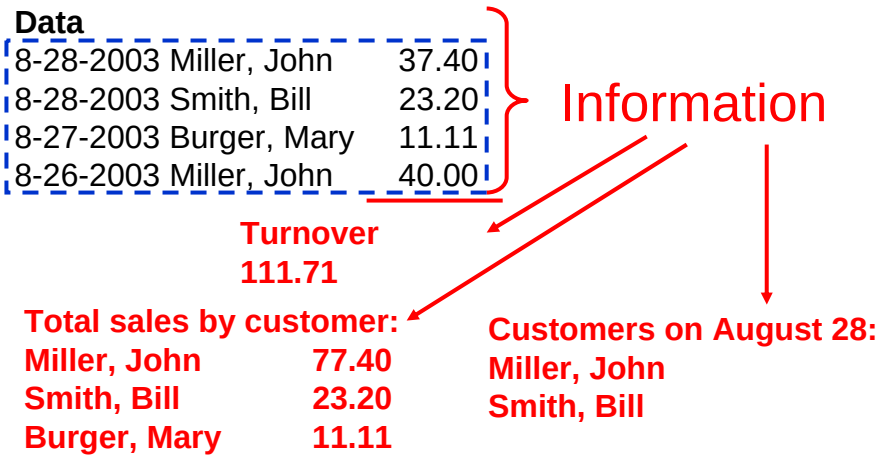
Storage

- Exploiting physical phenomena for storing and retrieving data
- Examples:
  - Iron and iron oxide (ferrous oxide) can keep magnetism
  - A laser can be used to modify the surface of a media so that it reflects (or no longer reflects) light

Data vs. Information

- **Information:** Facts that they have value beyond their representation
  - Peter was born on July 4, 1980
  - This PC costs 1000 euros.
  - The temperature in Munich is 11 degrees Celsius.
- **Data:** Representation of facts

Data vs. Information



Atomic Data

- Characters and Symbols
  - A-Z, 0-9
- Numbers
  - Integer
    - 1, 2, 3
  - Floating-point
    - 3.1415
    - 6.7

Encoding Characters and Strings

- Define the set of relevant characters (an alphabet)
  - Examples
    - Just A-Z and 0-9
    - A-Z, a-z, ÄÖÜ, äöü, ß
    - Mathematical symbols
- Create a mapping between characters and binary values

Binary	Oct	Dec	Hex	Glyph
100 0000	100	64	40	@
100 0001	101	65	41	A
100 0010	102	66	42	B
100 0011	103	67	43	C
100 0100	104	68	44	D
100 0101	105	69	45	E
100 0110	106	70	46	F
100 0111	107	71	47	G
100 1000	110	72	48	H
100 1001	111	73	49	I
100 1010	112	74	4A	J
100 1011	113	75	4B	K
100 1100	114	76	4C	L
100 1101	115	77	4D	M
100 1110	116	78	4E	N
100 1111	117	79	4F	O
101 0000	120	80	50	P
101 0001	121	81	51	Q
101 0010	122	82	52	R
101 0011	123	83	53	S
101 0100	124	84	54	T
101 0101	125	85	55	U
101 0110	126	86	56	V
101 0111	127	87	57	W
101 1000	130	88	58	X
101 1001	131	89	59	Y
101 1010	132	90	5A	Z

Source: Wikipedia

Common Codesets

- ASCII: *„American Standard Code for Information Interchange (ASCII)“*
- Extended ASCII (8 Bit): ISO 8859
  - 256 characters → 0-255
- Unicode: 1,114,112 (= $2^{20} + 2^{16}$) possible code points
  - Currently about 100,000 of those used
  - Including Chinese, Korean, and Japanese characters
- UTF-8 and UTF-16
  - Varying number of bytes per character, depending on its frequency of usage

On the Importance of Codeset Standardization

> I have also approved recommendations of the Secretary of Commerce regarding standards for recording the Standard Code for Information Interchange on magnetic tapes and paper tapes when they are used in computer operations.
>
> All computers and related equipment configurations brought into the Federal Government inventory on and after July 1, 1969, must have the capability to use the Standard Code for Information Interchange and the formats prescribed by the magnetic tape and paper tape standards when these media are used.

—Lyndon B. Johnson, Memorandum Approving the Adoption by the Federal Government of a Standard Code for Information Interchange, March 11th, 1968 at Gerhard Peters (ed.), The American Presidency Project, [3]

Source: Wikipedia

Encoding Integer Values

- Non-negative Integers
  - Simple binary numbers
  - Combine multiple bytes to store larger numbers
    - Example: 16 Bit Number for values 0 – 65535
    - Represented value: lower byte + 256* higher byte
- Positive and Negative Values
  - Simple: Use highest bit for storing the sign (+/-)
  - Two-complement

Encoding Real Values

- Fixed-point representation
  - E.g., one byte or half a byte per digit
- Floating-point representation
  - Mantissa or significant: string of digits
    - Base 2 or base 10
  - Exponent: The power of the base by which the significant is multiplied
- Example
  - 2.753E3 => 2.753 * 10^3

Precision Problems

- What is 1/3 * 3?
  - 1?
  - 0.99999?
- Significant rounding errors can occur in computer systems, because such precision problems sum up

Complex Data

- Strings
  - „Peter Miller“, „Universität der Bundeswehr“
- Structured Data
  - Qtty / Description / Price

Complex Data: Fixed Length vs. Explicit Delimiters

- **Fixed Length:** All fields have the same length
  - 255 characters for each street address
- **Explicit Delimiters:** A special character that is not included in the regular alphabet indicates the end of one data field
  - Example: Comma-separated Values (CSV)
    „Peter Miller**,** Hauptstrasse 8**,** Neubiberg“

Dataset and Files

- Dataset: Set of data elements (atomic or complex) that belong together
  - Address: Name, Street, ZIP, City
  - Invoice: Customer, Items, Total
- File: Collection of datasets of the same type
  - All addresses
  - All invoices

Locating a Dataset in a File

- Fixed length: Dataset Number * length of Dataset
  - Example:
  - Each address be 250 characters long
  - 1st address starts at first byte (0), 2nd at byte 250, 3rd at byte 500, ...
- Variable length
  - Explicit dataset delimiter (different from field delimiter)
  - Table with pointers to beginning of datasets

Machine-readable Content vs. Unstructured Data

ORDER
QTTY=3
ITEMNO=1234

Further Examples

Good	Bad
Structured text message	Fax Image
Vector drawing of a floor plan	Photo

Storage

- Exploiting physical phenomena for storing and retrieving data
- Examples:
  - Iron and iron oxide (ferrous oxide) can keep magnetism
  - A laser can be used to change the surface of a media so that it reflects (or no longer reflects) light

Access Methods

- **Sequential**: records must be retrieved in order
  - Devices used are called sequential access storage devices (SASD)
- **Direct**: records can be retrieved in any order
  - Devices used are called direct access storage devices (DASDs)

Hard Disks

http://computer.howstuffworks.com/hard-disk.htm/printable

Hard Disks

- Platters
- Read/Write Head
- Actuator Arm

http://www.flickr.com/photos/gek_at2000/485151116/

http://computer.howstuffworks.com/hard-disk.htm/printable

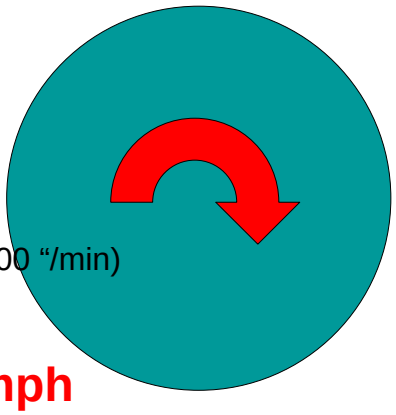
Hard Disks

http://computer.howstuffworks.com/hard-disk.htm/printable

Speed

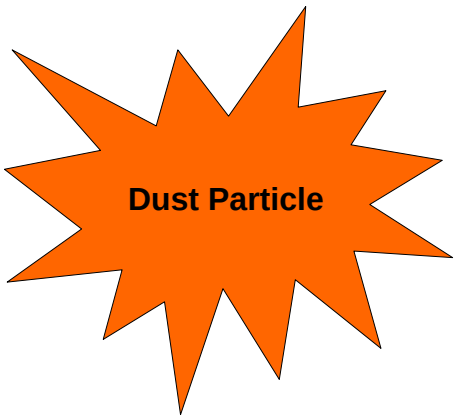
r = 3.5” / 2
diameter = 3.5 “
circumference = 2π r

⇒3.5” * 3.1415 = ca. 11”



⇒7200 rpm
⇒11” * 7,200 inch/minute (79,200 “/min)
⇒11” * 7,200 * 60 inch/hour
⇒4,752,000 inch/hour
⇒4,752,000 / 63,360 => **75 mph**

Dust and Abrasion



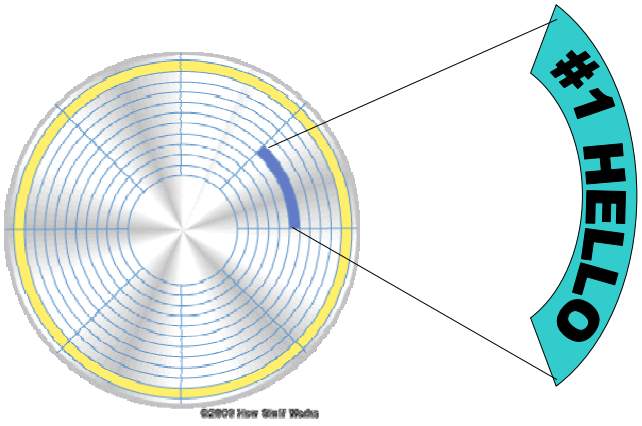
a few micrometers
(ca. 1/1,000,000 yard)

Headcrash

http://www.flickr.com/photos/124330160/89745500/

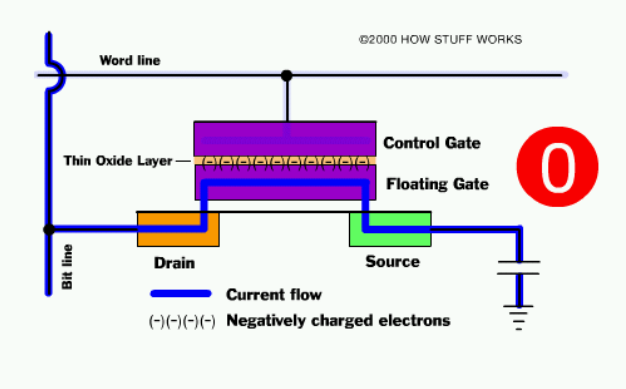
Hard Disks

- **Tracks**
- **Sectors**



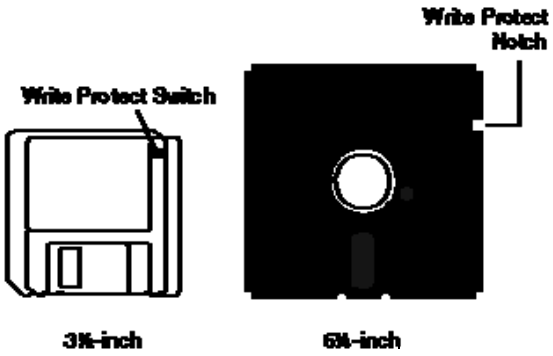
http://computer.howstuffworks.com/hard-disk.htm/printable

Flash Memory



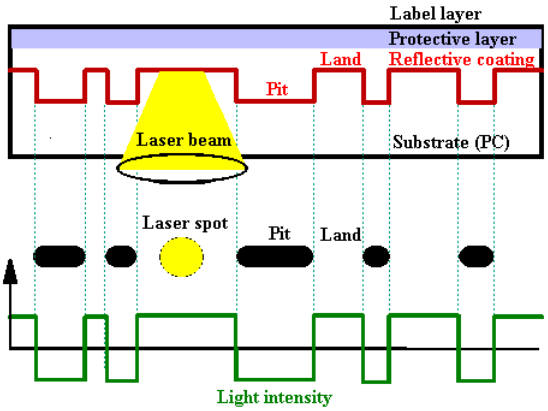
http://computer.howstuffworks.com/flash-memory.htm/printable

Floppy Disk



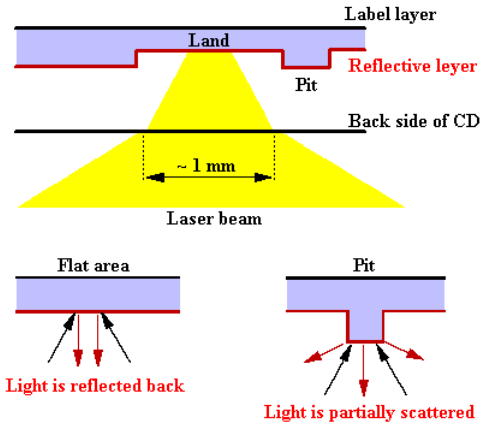
http://www.webopedia.com/TERM/f/floppy_disk.html

CD-ROM



http://www.usbyte.com/common/compact_disk_3.htm

CD-ROM



http://www.usbyte.com/common/compact_disk_3.htm

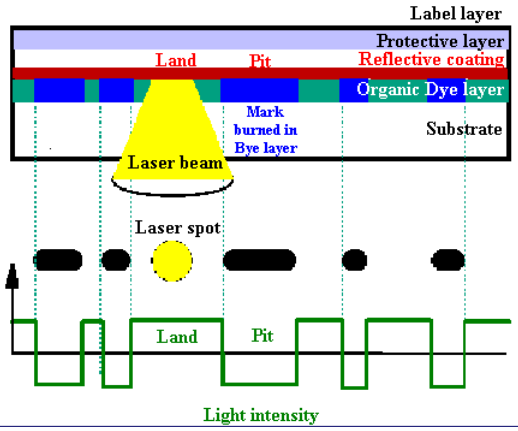
Only ONE Track!

CLV vs. CAV

CLV: Constant Linear Velocity: inches per second remains constant

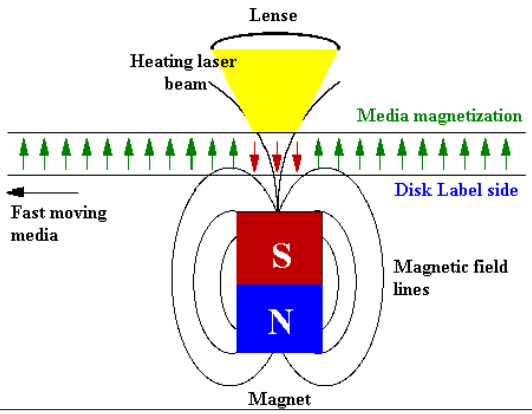
CAV: Constant Angular Velocity: rotations per second remains constant

CD-R



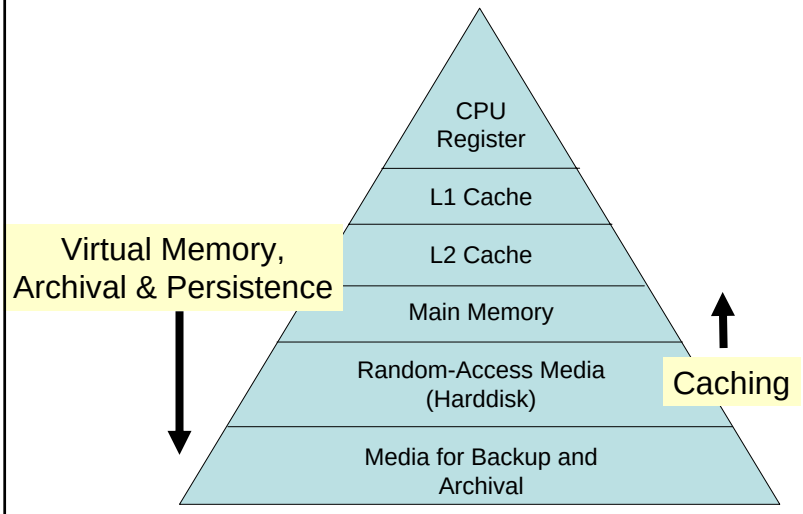
http://www.usbyte.com/common/recordable_CD.htm

Magneto-Optical Drives



http://www.usbyte.com/common/MOsystems.htm

Memory Hierarchy



Assignment for Next Week

- WI2: pp. 213-300
- Review the slides

WI1 = Hansen/Neumann: Wirtschaftsinformatik 1; WI2 = Hansen/Neumann: Wirtschaftsinformatik 2; IBIS = Wigand et al: Introduction to Business Information Systems.

Thank you!

The slides and additional materials will be available at
http://www.heppnetz.de/teaching/gwi/
shortly.

Don‘t forget: Tutorials will start this week!