Solution of Exercise Sheet 4

Exercise 1   (Hard Disk Drives)

1. What are sectors (= blocks) in HDDs?

   The tracks are divided into logical units (segments of a circle), which are called blocks or sectors. Sectors are the smallest addressable units of HDDs.

2. What are tracks in HDDs?

   The disc surfaces are magnetized in circular tracks by the heads.

3. What are cylinders in HDDs?

   All tracks on all disks at a specific arm position are part of a cylinder.
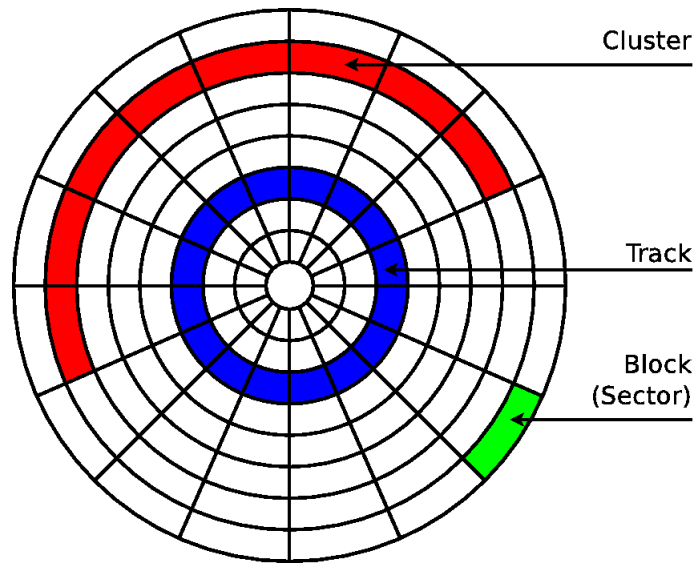
4. What are clusters in HDDs?

   Clusters are groups of sectors with a fixed size, e.g. 4 or 8 kB. In modern operating systems, clusters are the smallest addressable unit of HDDs.

5. Draw the structure of a hard disk drive schematically. Explain with your drawing(s) the meaning of the following terms:

   a) Sector (= Block)

   b) Track

   c) Cylinder

   d) Cluster
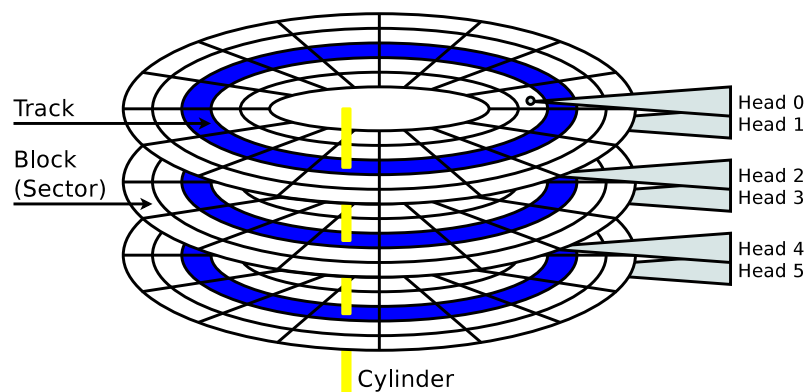
6. Why is it impossible to improve the performance (especially the latency) of HDDs infinitely?

   HDDs contain moving parts. Physical boundaries must be accepted.

7. Which factors influence the access time of HDDs?

   - *Average seek time*
   - *Average rotational latency time*

8. Describe the factors of subtask 7.

   - *The average seek time is the time that it takes for the arm to reach a desired track.*
   - *The average rotational latency time is the delay of the rotational speed, until the required disk sector is located under the head. It depends entirely on the rotational speed of the disks.*

Exercise 2   (Disk Geometry of HDDs)

An old HDD provides these information:

```
Western Digital WD Caviar 64AA           Enheanced IDE Hard Drive
Drive parameters     13328 cyl     15 heads     63 spt     6448.6 MB
S/N: WM653  321 5163    MDL: WD64AA - 00AAA4    DATE: 02 FEB 2000
```

1. Calculate the capacity of one disk of the HDD.
   (Provide the calculation steps!)

   The number of cylinders (CYL) is equal to the number of tracks per disc. The size of the sectors is 512 Byte.

Solution: The capacity of one side of a disk = number of cylinders ∗ Number of sectors per track (`spt`) ∗ bytes per sector (block).

$13,328 * 63 * 512 = 429,907,968$ Byte

Each disc has two surfaces. This means that both surfaces of a disk together have a capacity of $859,815,936$ Byte.

For 15 heads both surfaces are used for 8 discs and only one surface is used for one disc.

2. Calculate the capacity of one track of the HDD.
   (Provide the calculation steps!)

   solution: Size of a track = number of sectors per track (`spt`) ∗ bytes per sector (block).

   $63 * 512 = 32,256$ Byte

3. Calculate the total capacity of the HDD.
   (Provide the calculation steps!)

   The number of `heads` is identical to the number of surfaces of the disks.

   *Solution: Total capacity = capacity of one disk * number of `heads`.*

   $429,907,968 * 15 = 6,448,619,520$ Byte

4. Do the information on the HDD describe the physical disk geometry?
   (Explain your answer!)

   It makes no sense for a hard disk to have 15 heads. With these values a different (logical) geometry (cylinder/head/sector) is used than the physical geometry. These specifications and the logical geometry that has nothing to do with physical geometry result from the limitations of the BIOS.

Exercise 3   (Solid State Drives)

1. Why is it wrong to call SSDs Solid State Disks?

   They do not contain moving parts. Consequently, they do not include disks..

2. Name four advantages of SSDs over HDDs.

   - Fast access time
   - Low power consumption
   - No noise generation
   - Mechanical robustness
   - Low weight

- The location of data does not matter $\implies$ defragmenting makes no sense

3. Name two drawbacks of SSDs over HDDs.

   - Higher price compared with HDDs of the same capacity
   - Secure delete or overwrite is hard to implement
   - Limited number of program/erase cycles

4. Why are erase operations on flash memory more complex than read operations?

   The memory cells are arranged in groups of pages and blocks. Depending on the structure of a flash memory, a block always contains a fixed number of pages. Write and erase operations can only be carried out for complete pages or blocks.

5. Name an advantage and a drawback of NOR memory.

   - *Benefit:*
     - *Random access for read and write operations $\implies$ Better latency compared with NAND memory*
   - *Drawback:*
     - *More complex ($\implies$ expensive) construction*
     - *Higher power consumption compared with NAND memory*
     - *Typically small capacities ($\leq$ 32 MB)*

6. Name an advantage and a drawback of NAND memory.

   - *Benefit:*
     - *Lesser data lines $\implies$ requires less surface area as NOR memory*
     - *Lower manufacturing costs compared with NOR flash memory*
   - *Drawback:*
     - *No random access $\implies$ Poorer latency compared with NOR memory*
     - *Read and write operations can only be carried out for entire pages*
     - *Erase operations can only be carried out for entire blocks*

7. Describe the difference between NAND memory of the categories Single-Level Cell (SLC), Multi-Level Cell (MLC) and Triple-Level Cell (TLC).

   - *TLC memory cells store 3 bits*
   - *MLC memory cells store 2 bits*
   - *SLC memory cells store 1 bit*

8. What is the objective of wear leveling algorithms?

   Wear leveling algorithms evenly distribute write operations.

Exercise 4   (RAID)

1. Which RAID levels improve the data transfer rate for write?

   ☒ RAID-0   ☐ RAID-1   ☒ RAID-5

2. Which RAID levels improve the reliability?

   ☐ RAID-0   ☒ RAID-1   ☒ RAID-5

3. How many drives are allowed to fail in a RAID 0 array without data loss?

   None.

4. How many drives are allowed to fail in a RAID 1 array without data loss?

   At least a single drive must work properly.

5. How many drives are allowed to fail in a RAID 5 array without data loss?

   A single drive maximum is allowed to fail.

6. Please comment the statement: „A RAID array can be used to replace the regular backup of important data".

   RAID is not a backup replacement. Wrong file operations or virus attacks take place on all drives.

7. Why is it not useful to store all parity information on a single drive, but to distribute the parity information on all drives?

   Each write operation on the RAID causes write operations on the dedicated parity drive $\implies$ bottleneck.

8. What is the net capacity of a RAID 0 array?

   The net capacity is n, if n is the number of drives.

9. What is the net capacity of a RAID 1 array?

   The net capacity is the capacity of the smallest drive.

10. What is the net capacity of a RAID 5 array?

    The net capacity is $n - 1$, if n is the number of drives.

11. How are the parity information of a RAID 5 array calculated?

    P(16-19) = Block 16 XOR Block 17 XOR Block 18 XOR Block 19.

12. Name one advantage and one drawback of software RAID compared with hardware RAID.

Benefit: No cost for additional hardware.

Drawback(s): Operating system dependent, additional CPU load.

Exercise 5 (Character Count, Time and Date, Aliases, Redirecting, Search for Files)

1. Create a file `Quote.txt` with this content by using the command `echo`:

   ```
   Was man nicht weiß,
   das eben brauchte man,
   und was man weiß,
   kann man nicht brauchen.

   Gothe (Faust)
   ```

   ```
   $ echo -e "Was man nicht weiß,\ndas eben brauchte man,\nund was
   man weiß,\nkann man nicht brauchen.\n\nGothe (Faust)" > Quote.txt
   ```

2. Print out the number of characters in the file `Quote.txt` by using the command `wc`.

   ```
   $ wc -m Quote.txt
   ```

3. Print out the number of words in the file `Quote.txt` and redirect the output into the command `wc`.

   ```
   $ wc -w Quote.txt
   ```

4. Print out the calendar of the year 1999 and redirect the output into a new file `Calendar.txt`.

   ```
   $ cal 1999 > Calendar.txt
   ```

5. Use the command `date` to create an output in the shell with the current date and formated like this example:

   ```
   Heute ist Donnerstag, der 24. Oktober 2013.
   Es ist 16:08 Uhr und 07 Sekunden.
   In UNIX-Zeit ist es genau: 1382623687
   ```

   Redirect the output in a way that it is attached at the file `Calendar.txt`.

   ```
   $ date +"Heute ist %A, der %d. %B %Y.%nEs ist %H:%M Uhr und
   %S Sekunden.%nIn UNIX-Zeit ist es genau: %s" >> Calendar.txt
   ```

6. Calculate the number of entries (files and directories) in the directory `/dev` with the command `wc`. Additionally, the processing speed must be measured.

```
$ time ls -l /dev | wc -l
```

7. Print out a list of existing aliases in the shell.

```
$ alias
```

8. Create an alias `zeit`, which produces the output of subtask 5.

```
$ alias zeit="time ls -l /dev | wc -l"
```

9. Remove the alias `zeit`.

```
$ unalias zeit
```

10. Search with an appropriate command all files in your home directory, which match these search criteria:

   - Search only for files and not for directories or links.
   - The file name must contain the string `BTS` (case insensitive)
   - The files must belong to your user account (user ID).
   - The age of the files must be at least 1 day.
   - The last modification must have taken place more than 3 days ago.
   - The file size must be at least 10 kB.

   For each file found, the number of lines must be printed out in the shell.

```
$ find . -iname "*bts*" -user USERNAME -atime -1 -mtime +3
-type f -size +10k -exec wc -l "{}" ";"
```