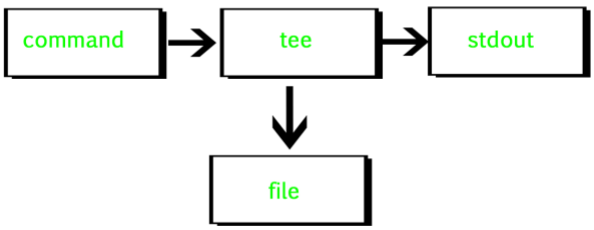
Practical tasks

Part 1. SSH Connect (3 tasks)

1. ssh-keygen
    a. Generate an **ed25519** key with 1024 bits and save it as files **new_key_ed25519** and **new_key_ed25519.pub**
    b. Change the passphrase of this key
2. ssh-copy-id
    a. Do a dry-run of writing the new ssh key to a remote
    b. Run copy id on the remote

Part 2 (2 tasks)

1. Use **nohup** to launch a long running (for example use `bash -c “sleep 10; echo complete”`).
    a. Enter the remote CLI, launch a long running program in the background and write **stdout** to std.out and **stderr** to std.err. Disconnect from the server without waiting for the completion of the background task.
    b. Repeat point a. but execute all the steps in a single terminal command
2.
    a. Redirect standard output stream to standard error, and echo output of a command `ls -lh` to standard error stream.
    b. Write a command that redirects standard output and standard error streams into separate respective files
    c. Use the *tee* program to save standard output in a file while showing it (it reads from the standard input and writes to both standard output and one or more files at the same time)



    d. Use Process Substitution (https://tldp.org/LDP/abs/html/process-sub.html)
    (it creates a FIFO and lets tee listen on it. Then, it uses `>` (file redirection) to redirect the `stdout` of command to the FIFO that your first tee is listening on)

Part 4. Copying and Downloading Data (3 tasks)

1. Copy all files from some local `*/directory_local*` to the remote `/directory_remote` only if their size is between 10MB and 20MB. And show the progress
2. Copy all files from some local `*/directory_local*` to the remote `/directory_remote` only if their names don't start with *“a”* and don’t end with *“z”*. And show the progress
3. Copy the directory structure without copying files from a local `*/directory_local*` to the remote `/directory_remote`. Hint: use -f flag and regular expression pattern

Part 5. Use SSH to resolve the following problems on the remote and retrieve results locally:

Problem 1. From the files in practical4_logs.tar.xz output the top 3 most frequent IP addresses that performed the GET method in files logs_2015-01-13-10, …, logs_2015-01-13-17 alltogether.

The answer for this task will be the top 3 most frequent IP addresses and their frequencies. And commands that you’ve used to obtain it

Problem 2. In the same archive find a list of all files with the .tsv extension that are larger than 5 MB, copy them in a “logs/archive” subdirectory and report number of files in logs/archive.

The answer for this task will be the number of items (.tsv files). And commands that you’ve used to obtain it

Problem 3

The data_for_science.tar.xz directory contains files of the following format: target class, tab, comma-separated list of keywords. Your task is to find unique words for the *bad* class, which are contained in the three largest files. Remember, *DOG* and *dog* are the same word.

The answer for this task will be 10-th,11-th and 12-th entries in a sorted set of unique words. And commands that you’ve used to obtain it