HPC.NRW

INTRODUCTION TO LINUX

(in an HPC context)

Version 20.09 | HPC.NRW Competence Network

FILES

HPC.NRW Competence Network

INTRODUCTION TO LINUX

- Linux: extensions do not matter
  - But: conventions to help humans
  - Some programs also look at extensions
- Most important: text file or not?
  - Configuration files
  - Scripts
  - System information files
- Binary file: generally not searchable
- Use `file <filename>` to identify file type

FILE MANIPULATION COMMANDS

HPC.NRW

- Simple commands to handle files
  - Most also work on directories
- You already know `ls`
- Rename file/directory: `mv <oldname> <newname>` (move)
- Copy file/directory: `cp <filename> <newname>` (copy)
  - Also needs `-r` for directories

- Create directory: `mkdir <dirname>`
- Create empty file: `touch filename`
  - Updates access time on an existing file
- Remove file/directory: `rm <filename>`
  - Check access permissions!
  - To delete content of subdirectories: `rm -r` (recursive)
  - Common option: `-f` (force) → never prompt for confirmation

- Previous examples: one command, one file
- Select multiple files according to patterns
- Wildcard (placeholder) characters
  - Also called globbing
- Most important
  - `*` zero or more characters
  - `?` exactly one character
  - `[]` range of characters

SEARCHING FOR FILES

HPC.NRW

- Use `find` command
- Syntax: `find <targetdir> <options>`
  - Example: `find . -name "ex1.txt" -type f`
- Allows very complex searches
  - Wildcards
  - Only files modified after X
- Allows executing command for every found file: `-exec`

- Wildcards: common source of problems, especially in scripts
  - Expanded by shell before being given to program
  - Problem not limited to `find` command
- Example: `find` command `-name` option

```
$ find . -type f -name *test*
```

  - The `find` command is handed multiple names, cannot handle this
- Fix: `$ find . -type f -name "*test*"`
  - Now string with wildcards is handed to `find` command