CHALMERS

GÖTEBORGS UNIVERSITET

Introduction to Linux

Lecture 4

Prof. Dr. Philippe Tassin
Chalmers University of Technology

Dr. Evangelos Siminos
Chalmers University of Technology and University of Gothenburg

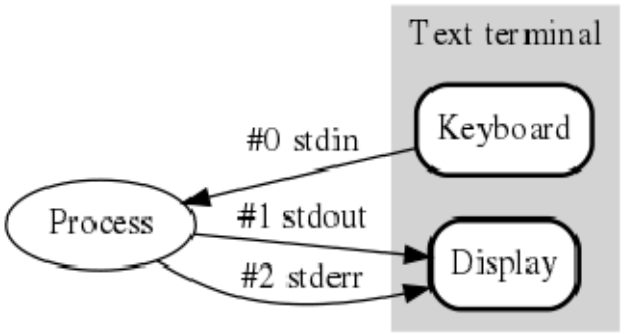
Göteborg, 17 April 2019

Standard I/O streams

- Three standard streams:
  - Standard input (stdin, fd0)
  - Standard output (stdout, fd1)
  - Standard error (stderr, fd2)

- Abstraction of input and output devices

- Default:



- Streams can be redirected to files, other terminals, etc.

Redirecting output and input

- Pipe (|) : redirects standard output to standard input of a following command
- > : redirects standard output to a file or terminal
  - cat loremipsum >loremipsum2
  - cat loremipsum >/dev/pts/5
- >> : append output
- < : redirect input

File descriptors

- Every file opened by a process has a file descriptor
- File descriptors can be used in I/O redirection
- `ls 2>&1 >dirlist`
- `ls 2>/dev/pts/5 >dirlist`

Text editors / Vi

- Many different text editors are available
- “Standard” text editor used on Linux is Vi
- No mouse input; only interaction through the keyboard!
  - Effective, but steep learning curve
- Two modes:
  - insert mode (`i`,`a`,`o`)
  - command mode (`ESC`)
- Many commands (delete, yank, change) need selecting text

  ⇒ Combine a **command** with a *motion*

  Example: `dw` = **delete** *until next word*

Regular expressions

- Often you want to match a pattern, a description of a set of strings
- Example: `lin.+`
- Regular expressions contain characters, wildcards, repetition operators, and/or character classes

Operator	Effect
.	Matches any single character.
?	The preceding item is optional and will be matched, at most, once.
*	The preceding item will be matched zero or more times.
+	The preceding item will be matched one or more times.
{N}	The preceding item is matched exactly N times.
{N,}	The preceding item is matched N or more times.
{N,M}	The preceding item is matched at least N times, but not more than M times.
-	represents the range if it's not first or last in a list or the ending point of a range in a list.
^	Matches the empty string at the beginning of a line; also represents the characters not in the range of a list.
$	Matches the empty string at the end of a line.
\b	Matches the empty string at the edge of a word.
\B	Matches the empty string provided it's not at the edge of a word.
\<	Match the empty string at the beginning of word.
\>	Match the empty string at the end of word.

sed stream editor

- Matches a pattern using a regular expression
- Applies a command to the matching text
  - print
  - delete lines
  - substitute text
  - …
- Commands are similar to those in vi !
- Matching with regular expressions between //

awk

- Stream editor to operate on structured data
  (think of an Excel sheet with columns of data, for example)

- awk ’REGEX { PROGRAM }’ inputfile
- command | awk ’REGEX { PROGRAM }’

- REGEX: select data with regular expressions
- PROGRAM: commands to operate on the data

- $1 - $9 are fields in the structured data file

- You can define variables

- You can have control structures in the program (if…then, etc)

Any other questions?