Zpracování textových souborů v prostředí GNU/Linux

Pro čtení slouží příkazy more či less.

- starší verze více možností využíval less, more dříve postrádal hlavně se vrátit i na předchozí již odscrollované stránky.
- při načtení daného souboru vypíší na monitor právě jednu obrazovku
- „enter“ posune o jednu řádku
- „space“ posune o jednu stránku
- vrácení o stránku zpět klávesa "b"
- vyhledávání textu v souboru stiskneme klávesu "/" a napíšeme dané slovo
- opětovné vyhledání téhož výrazu "n" (less) nebo klávesu "." (more)
- less má dole stavový řádek v němž vypisuje název právě otevřeného souboru acelou řadu parametrů.
- more vypisuje ve stavovém řádku pouze procentuální údaj o poloze v souboru.
- ukončení výpisu souboru dojde po ododscrollování textu až na konec souboru nebo jej můžeme ukončit klávesou „q“.
- klávesa „v“ spustí editor vi a umožní tak daný soubor editovat

editor vi

- editor má dva režimy:
  - příkazový – v tomto režimu se nachází editor po spuštění
  - vkládací – kdy editujeme soubor, do něj se dostaneme pomocí:
    - i - text bude vložen před pozici kurzoru
    - a - text bude vložen za pozici kurzoru
    - o - text bude vložen na novou řádku za pozicí za aktuální řádkou
    - R - text bude přepisovat starý text od pozice kurzoru
    - ukončíme ho klávesou ESC

K ukončení práce v editoru jsou přikazy:

- :w - pro zápis změn na disk
- :w nazev_souboru - uloží pod novým názvem
- :q - pro ukončení práce
- :q! - pro ukončení práce bez uložení změn

Pozor: Příkaz less binární data vypíše, zatímco more je nevypíše, nebo je zobrazí v podobě rozsypaného čaje.

Cat - vypíše kompletně celý text a ukončí se

- tento příkaz čte daný text po řádcích a lze jej s výhodou použít:
  - pro spojování souborů
  - zápis textu do souboru (přesměrování)
  - připsání textu na konec již existujícího souboru (přesměrování)

spojování souborů:
cat soubor1 soubor2 soubor3 > vysledny_soubor

Existuje podobný příkazy **tac**, který vypíše soubor po řádcích od jeho konce do počátku.

head

- vypíše x prvních řádků/znaků ze začátku daného souboru

- pro vypsání prvních 5 řádků tak zadáme:
  head –n 5 a.txt
  parametr n – vypíše n zadaných řádků
- pro vypsání prvních 5 **znaků** tak zadáme:
  head –c 5 a.txt
  parametr c– vypíše c zadaných **znaků**

tail

- vypisuje x posledních řádků/znaků ze souboru
- specialita je parametr -f, který stále vypisuje posledních x řádků. Po přečtení x řádků se neukončí a vypisuje dále nově příchozí data - použití se pro čtení logů.

  tail -f a.txt

sort

- příkaz pro setřídění textu

sort a.txt – seřadí obsah souboru dle abecedy

- parametry:
  - pro neopakování stejných řádků (řetězců)
    -u (uniq)
  - pro seřazení sestupné
    -r
  - pro seřazení podle čísel
    -n (numeric)
  - pro seřazení dle velikostí násobků jednotek
    -h (human)

grep

- umožnuje v obsahu souboru lze vyhledávat konkrétní slova (řetězce)
- parametry:
  - n – očísluje vypsané řádky
  - i – ignoruje velikost abecedy

  grep auto a.txt
  - najde výskyt řetězce auto v souboru a.txt

cut

- extrahuje nějaký text ze souboru
- parametry:
  - -d – tzv. delimiter - oddělovací znak
  - -f - určení polí, která se mají vypisovat
- Např.: ze souboru /etc/passwd se vypíší veškertá uživ. jména a jejich UID
  - Oddělovací znak je „:“ a budeme vypisovat první a třetí sloupec.
  - cut -d: -f1,3 /etc/passwd

Použití bílých znaků v parametrech příkazů. např. cut –d, má-li být oddělovací znak mezi kategoriemi např. mezera, tak ji zapisujeme do apostrofů. paste

- příkaz pro spojení textu z více souborů
- vytvořili jsem si následující dva soubory:
  cat > soubor.txt cat > soubor2.txt

plechovka	rezava
linux	Fedora
windows	microsoft

- sloučení provedeme následovně:
  paste soubor2.txt soubor.txt

join

- umí spojit dva soubory a vytvořit tak jednoduchou „databázi“
- princip příkazu spočívá:
  - soubor obsahuje v prvním sloupci klíč (obvykle číselná hodnota)
  - podle tohoto klíče program join páruje dané položky k sobě
- např: data1.txt data2.txt
  1 pivo 1 25Kc
  2 kofola 2 15Kc
- sloučení provedeme: join data1.txt data2.txt
  1 pivo 25Kc
  2 kofola 15Kc

tr

- příkaz pro záměnu znaků
- má VELIKOU řadu možností viz nápověda k příkazu
- !!!hranaté závorky je třeba správně uvozovat!!!
- např:
  - příkaz změní v souboru malá písmena za velká
    tr '[:lower:]' '[:upper:]' < soubor.txt
  - záměna písmen
    cat soubor.txt | tr "A-Z" "TOSHRKPNELUDCWFYIGJAMQBVZX"

wc

- informativní příkaz
- vypisuje počet řádků, slov a znaků v souboru.

  wc a.txt

  8 8 59 a.txt