



Sed

O co jde
Regulární výrazy
Příkazy
Adresování
Příklady

Michal Trs, trsm1@fel.cvut.cz, www.michaltrs.net/cvut_fel/36unx/sed.pdf

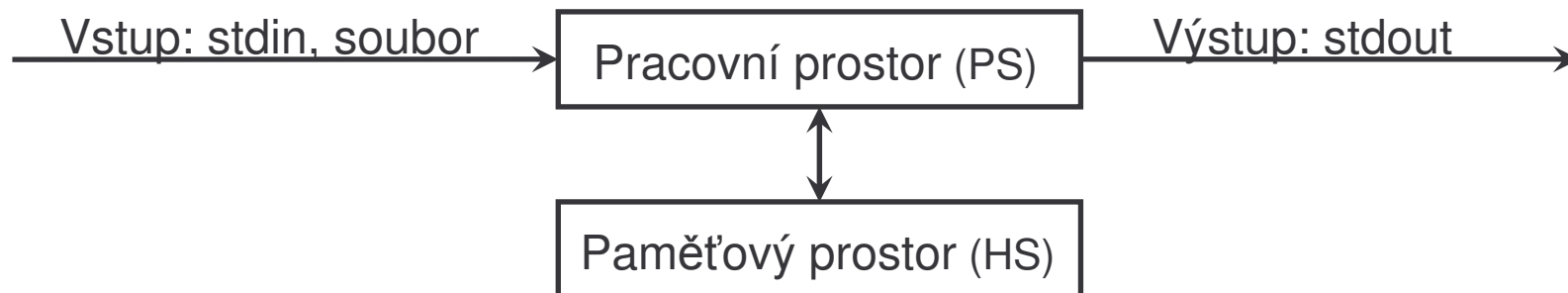


O co jde

- Stream **editor**
- textový editor vhodný pro dávkové zpracování textu
- je to filtr
- vstup: `soubor sed [-n] [-e script] [-f scriptFile] [soubor]`
`stdin cat soubor | sed [-n] [-e script] [-f scriptFile]`
- kde:
 - n potlačení kopírování vstupu na výstup
 - e sedovský příkaz zadaný z řádky
 - f načte soubor se sedovskými příkazy
- výstup: vždy na stdout => **vstupní soubor se nemění**, pozor na přesměrování!
- 2 režimy práce: standardně se vše kopíruje ze vstupu na výstup
–n nic se nekopíruje
- pro práci nutná znalost regulárních výrazů

O co jde (pracovní cyklus)

- 1) načtena řádka ze zdrojového souboru a interně očíslována
- 2) řádka vložena do pracovního prostoru
- 3) na obsah pracovního prostoru aplikován sedscript (posloupnost sedovských příkazů)
- 4) obsah pracovního prostoru poslán na standardní výstup





Regulární výrazy

- vzor porovnáváný s daným textem
- regulární výraz používá speciální znaky v jiném významu než shell ([,], ., *, \$)

Jednoznakové

Konstrukce	se porovná
<i>znak</i>	sám se sebou
<i>\spec_znak</i>	se speciálním znakem (., *, [,], \) jako takovým.
.	s jedním libovolným znakem.
[<i>znaky</i>]	s libovolným jedním znakem z uvedených
[<i>^znaky</i>]	s libovolným jedním znakem různým od uvedených
-	uvedené mezi znaky znamená interval

Regulární výrazy (**Víceznakové**)

Konstrukce	se porovná
$\langle 1nRE \rangle^*$	s libovolným (i nulovým) opakováním předchozího jednoznakového RE. Je-li více možností, vezme se první porovnání zleva co nejdelší.
$\langle 1nRE \rangle \{m\}$	s právě m opakováními jednoznakového RE (m je číslo 1 až 255)
$\langle 1nRE \rangle \{m,\}$	s nejméně m opakováními jednoznakového RE
$\langle 1nRE \rangle \{m,n\}$	s m až n opakováními jednoznakového RE
$\langle \langle RE \rangle$	s odpovídajícím řetězem od začátku slova
$\rangle \langle RE \rangle$	s odpovídajícím řetězem na konci slova
$\wedge \langle RE \rangle$	odpovídajícím řetězem od začátku řádky
$\langle RE \rangle \$$	odpovídajícím řetězem na konci řádky



Příkazy

- d (delete) - vymaže řádku
- p (print) - opíše řádku na výstup
- r soubor - načte uvedený soubor a pošle ho na výstup
- s/RE/náhrada/volby - nahradí text porovnaný s **RE**
v **náhrada** se může vyskytnout:
 - & - text, se kterým se porovnálo **RE**
 - \n (**n** je číslice) - řetězec se kterým se porovnal **n**-tý výskyt
 - \(**REn**\) v **RE**
- : návěští návěští
- **b**návěští - nepodmíněný skok na návěští
- **t**návěští - podmíněný skok, pokud poslední náhrada s byla úspěšná
- Příkazy pro přesouvání mezi PS a HS: h,H,g,G,x
- # komentář (#n na první řádce to samé jako -n)
- spousta dalších



Adresování

- 1) bez adresy příkaz proveden pro každou řádku vstupu
`sed -n -e p t.txt`
- 2) 1 adresa proveden pouze pro tuto řádku
`sed -e /aa/d t.txt`
- 3) Rozsah adres reaguje na řádky od – do (včetně)
`sed -e /aa/world/d t.txt`

Adresou může být:

- číslo řádku
- regulární výraz (RE)
- \$ (poslední řádka souboru)



Příklady

```
$ cat t.txt
hello world!!!
how are you?
aa bbbb
abc
aaaa
aa bbb dd
qwerty
```

```
$ sed -e /a/d t.txt
hello world!!!
qwerty
```

```
$ sed -n -e /a/p t.txt
how are you?
aa bbbb
abc
aaaa
aa bbb dd
```

```
$ cat s.sed
#n
# sed script
s/[ab]$/end/p
```

```
$ sed -f s.sed t.txt
aa bbbend
aaaend
```




Další informace

- unix.felk.cvut.cz/unix/download/predn_05.pdf
- www.root.cz/clanky/neco-malo-o-sedu/
- www.root.cz/clanky/neco-malo-o-sedu-priklady-pouziti/