

# A7B38UOS Úvod do operačních systémů

## 5. Cvičení

### Regulární výrazy a grep

# Grep

## ***grep [přepínače] vzor [soubory]***

Implicitně vypíše na standardní výstup řádky, které vyhoví zadanému vzoru.

*-i nerozlišuje malá a velká písmena*

*-v vypíše na standardní výstup řádky, které neobsahují zadaný vzor*

*-c vypíše pouze počet odpovídajících řádek*

*-l vypíše pouze jména souborů, které odpovídající řádky obsahují*

*-n vypíše odpovídající řádky a jejich pořadové číslo v souboru*

**Vzor může být definován základním regulárním výrazem**

**grep je zkratka pro skupinu příkazů editoru ex (globally search for regular expression and print result).**

# Grep – jednoduché příklady

```
grep 'The' /etc/ssh/ssh_config
```

```
grep 'the' /etc/ssh/ssh_config
```

```
grep -i 'The' /etc/ssh/ssh_config
```

```
grep -ci 'the' /etc/ssh/ssh_config
```

```
grep -ni 'the' /etc/ssh/ssh_config
```

```
grep -l 'kill' /etc/init.d/*
```

```
grep root /etc/group
```

```
grep -v root /etc/group
```

# Regulární výrazy I

Symbol	Význam
.	právě jeden jakýkoliv znak
znak*	žádný nebo libovolných počet výskytů předchozího znaku
[ ]	jeden znak z množiny/intervalu (např. [a,d,f], [a-h])
[^ ]	jeden libovolný znak mimo znaků z množiny/intervalu
^	začátek řádky
\$	konec řádky
\<	začátek slova
\>	konec slova
\znak	ruší speciální význam následujícího znaku

Nezapomeňte ochránit speciální znaky před zpracováním shellem!  
=> apostrof nebo uvozovky

# Regulární výrazy I

```
ls -l | grep -c '^l'
```

```
ypcat passwd | grep '/bin/bash$'
```

```
grep 'the' /etc/ssh/ssh_config
```

```
grep '\<the\>' /etc/ssh/ssh_config
```

```
grep 'bag' /usr/share/dict/words
```

```
grep '^bag' /usr/share/dict/words
```

```
grep 'bag$' /usr/share/dict/words
```

```
grep '^bag$' /usr/share/dict/words
```

# Regulární výrazy I

```
grep '^b[aeiou]g' /usr/share/dict/words
```

```
grep '^b[^aeiou]g' /usr/share/dict/words
```

```
grep '^b.g$' /usr/share/dict/words
```

```
grep '^woo*' /usr/share/dict/words
```

```
grep '^wood' /usr/share/dict/words
```

```
grep '^wood.*d' /usr/share/dict/words
```

```
grep '^wood.*d$' /usr/share/dict/words
```

# Regulární výrazy I

Výpis řádků obsahujících RE

```
ls /home/* 2>/dev/null | grep novak
```

Výpis řádků neobsahujících RE

```
echo $PATH | tr '[:]' '\n' | grep -v bin
```

Výpis řádků obsahujících RE bez ohledu na velikost písmen

```
man ls | grep -i command
```

Výpis pouze názvů souborů obsahujících RE

```
grep -l start /etc/* 2>/dev/null
```

Výpis počtu řádků obsahujících RE

```
grep -c example /usr/man/man1/grep.1
```

# Regulární výrazy I

Výpis řádků obsahujících stejný znak na začátku i konci řádku

```
grep '^(\.)*\1$' words
```

Výpis pětiznakových palindromů

```
grep '^(\.)(\.)\2\1$' words
```

Výpis řádků obsahujících 2 stejná slova

```
man head | grep -i '\(<[a-z][a-z]*\>\1\>'
```



# Regulární výrazy II – počítáme výskyty

Symbol	Význam
$\text{znak}\{m\}$	právě $m$ výskytů předchozího znaku
$\text{znak}\{m,\}$	nejméně $m$ výskytů předchozího znaku
$\text{znak}\{m,n\}$	$m$ až $n$ výskytů předchozího znaku

# Regulární výrazy II – počítáme výskyty

```
grep '^[A-Z]' /usr/share/dict/words
```

```
grep '^[A-Z][A-Z]' /usr/share/dict/words
```

```
grep '^[A-Z]\{2\}' /usr/share/dict/words
```

```
grep '^[A-Z]\{2,3\}' /usr/share/dict/words
```

# Regulární výrazy II – počítáme výskyty

```
cd /usr/share/dict
```

```
grep o words
```

```
grep oo words
```

```
grep oo$ words
```

```
grep ^oo words
```

```
grep o.o words
```

```
grep 'o.*o' words
```

```
grep '^o.*o$' words
```

```
grep 'o*' words
```

```
grep 'oo*' words
```

```
grep '[^o]' words
```

```
grep -v o words
```

```
grep . words
```

```
grep ^.$ words
```

```
grep ^..$ words
```

```
grep '^[a-z][a-z]$\n' words
```

```
grep '^[^a-z].*[^a-z]\n' words
```

```
grep '^[qxy]...$\n' words
```

```
grep '^[qxy].\{3\}\n' words
```

```
grep '^[qxy].\{12,15\}\n' words
```

```
grep '^[qxy].\{13,\}\n' words
```

```
man grep | grep '<file>'
```

```
man grep | grep '<file'
```

```
man grep | grep 'file>'
```

# Regulární výrazy III+

## *Egrep*

- Implicitně vypíše na standardní výstup řádky, které obsahují zadaný vzor.
- Vzor může být definován rozšířeným regulárním výrazem.
- Nepodporuje znaky: \(\, \), \n, \<, \>, \{, \}
- Navíc podporuje znaky: +, ?, |, (, )
- Přepínače jsou podobné jak u příkazu grep.

# Regulární výrazy III+

Symbol	Význam
znak+	jeden nebo libovolných počet výskytů předchozího znaku
znak?	žádný nebo jeden výskyt předchozího znaku
RE1 RE2	RE1 nebo RE2
(RE)	označení reg. podvýrazu

# Regulární výrazy III+

## *Fgrep*

- Implicitně vypíše na standardní výstup řádky, které obsahují zadaný vzor.
- Vzor může být definován pouze jako obyčejný řetězec.
- Příkaz je rychlejší než grep a egrep.
- Přepínače jsou podobné jak u příkazu grep.

# Regulární výrazy III+

Výpis dnů v týdnu ze souboru words

M	o	n	day			
T	u	e	s	day		
W	e	d	n	e	s	day
T	h	u	r	s	day	
F	r	i	day			
S	a	t	u	r	day	
S	u	n	day			

```
grep '[MTWFS][ouehra][neduit][snru]*day' words
```

```
egrep '(Mon|Tues|Wednes|Thurs|Fri|Satur|Sun)day' words
```

# Regulární výrazy

Nalezněte v souboru  
`/home/predmety/uos/common/vzorky/messages,`  
které MAC adresy získaly 4. 3. 2013 IP adresu 10.12.14.30.



# Regulární výrazy

```
grep "Mar *4.*DHCPACK on 10.12.14.30" messages
```

# Regulární výrazy

Nalezněte v souboru  
`/home/predmety/uos/common/vzorky/access_log`,  
kolikrát se v listopadu 2012 stáhla www stránka  
`/vyuka/paradoxy`.

# Regulární výrazy

```
grep "/Nov/2012.*vyuka/paradoxy " access_log | wc -l
```